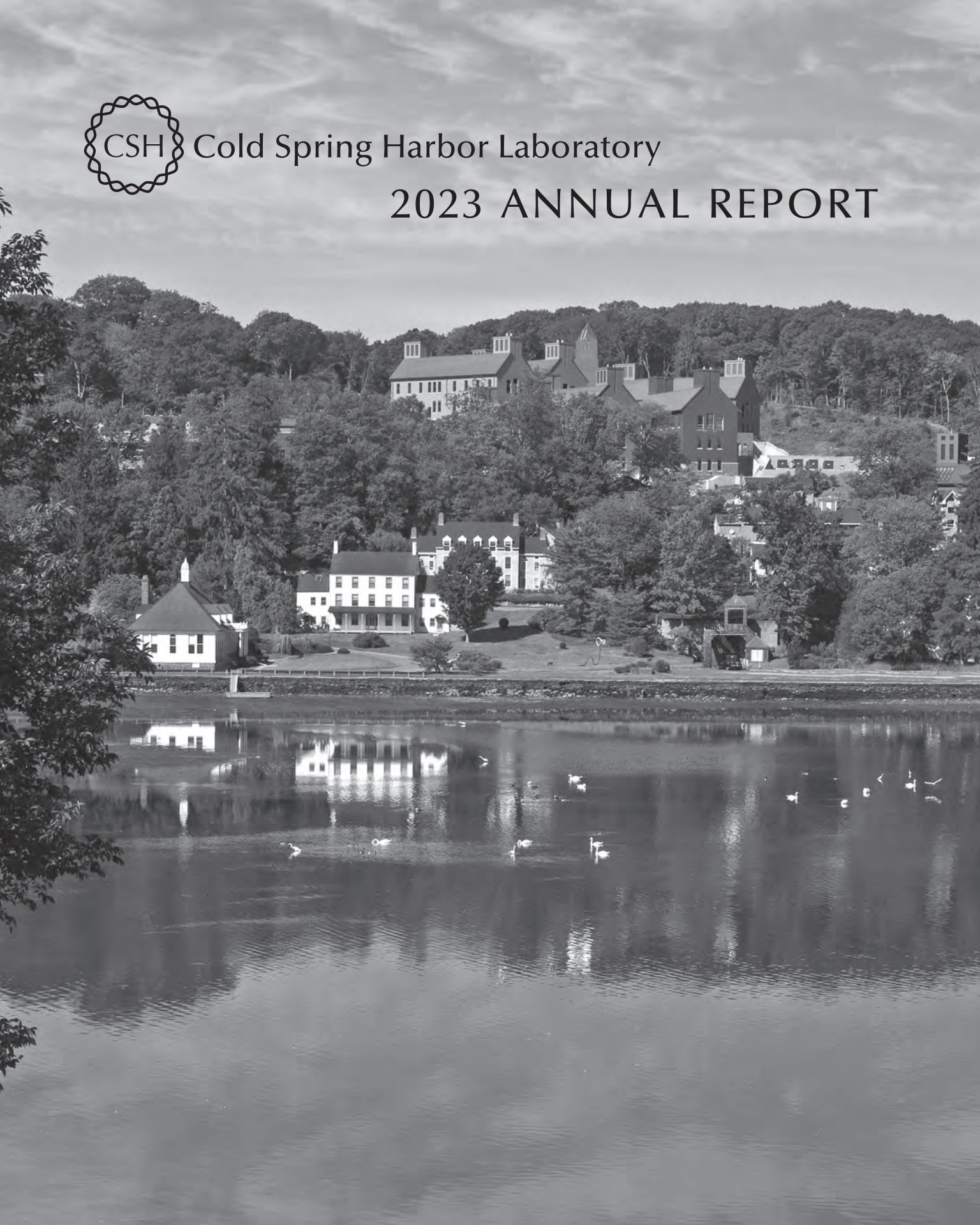
Cold Spring Harbor Laboratory

# 2023 ANNUAL REPORT



**ANNUAL REPORT 2023**



Cold Spring Harbor Laboratory
One Bungtown Road
Cold Spring Harbor, New York 11724

www.cshl.edu

**Managing Editors** Samuel Diamond, Philip Renna, Lisa Cruz
**Production Editor** Kathleen Bubbeo
**Copy Editors** John Schneider, Kathleen Bubbeo, Samuel Diamond
**Proofreader** Joy Jones
**Production Manager** Denise Weiss
**Nonscientific Photography** Constance Brukin, Len Marks Photography
**Cover Designer** Sue Weil-Kazzaz

**Front Cover Photography**: © 2023 Philip Renna

# Contents

# Officers of the Corporation

**Marilyn H. Simons, Ph.D.,** *Chair*
**Charles I. Cogut,** *Vice-Chair*
**Robert D. Lindsay,** *Vice-Chair*
**Paul J. Taubman,** *Vice-Chair*
**Elizabeth McCaul,** *Treasurer*
**Robert W. Lourie, Ph.D.,** *Secretary*
**Bruce W. Stillman, Ph.D.,** *President and CEO*
**John P. Tuke,** *Chief Operating Officer*

# Board of Trustees

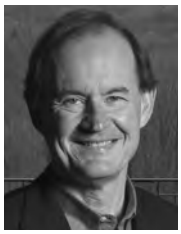
## Life Trustee

**Jamie C. Nicholls**
*Chair Emeritus, New York, NY*

## Individual Trustees

**Christine Anderson**
*Senior Managing Director Global Public Affairs and Marketing, Blackstone*

**Lalit R. Bahl, Ph.D.**
*Senior Research Scientist, Renaissance Technologies Corp.*



**David Boies**
*Chairman, Boies, Schiller & Flexner LLP*

**Michael R. Botchan, Ph.D.**
*Richard and Rhoda Goldman Distinguished Professor of Biochemistry, University of California, Berkeley*

**Mark L. Claster**
*Managing Partner, Carl Marks Advisors, Carl Marks*

**Elizabeth Cogan Fascitelli**
*Goldman, Sachs & Co. (Retired Partner)*

**Charles I. Cogut**
*Retired Partner, Simpson Thacher & Barlett, LLP*

**Elaine Fuchs, Ph.D.**
*Rebecca C. Lancefield Professor/HHMI Investigator, Rockefeller University/Howard Hughes Medical Institute*

**Ellen Futter**
*President Emerita, American Museum of Natural History*

**Mark Hamer**
*President, Harvest Real Estate Services, Inc.*

**Tracy L. Johnson, Ph.D.**
*Professor of Molecular, Cell and Developmental Biology, Cecilia and Keith Terasaki Presidential Endowed Chair, Howard Hughes Medical Institute Professor, Dean, Division of Life Sciences at UCLA*

**Jeffrey E. Kelter**
*Chief Executive Officer, KSH Capital; Chairman, Jack Creek Investment Corp.*

**Laurie J. Landeau, V.M.D.**
*General Manager, Listowel, Inc.*

**Robert D. Lindsay**
*Co-Managing Partner, Lindsay Goldberg LLC*

**Thomas H. Lister**
*Head of Future Research, Renaissance Technologies LLC*

**Robert W. Lourie, Ph.D.**
*Partner, Permira Advisers LLC*

**Elizabeth McCaul**
*ECB Representative, Supervisory Board, European Central Bank*

**Howard L. Morgan, Ph.D.**
*Chairman, B Capital Group; Managing Member, MFCIF, LLC*

## Individual Trustees *(continued)*

**Jeanne Moutoussamy-Ashe**
*New York, NY*

**Lyon Polk**
*Founder, The Polk Wealth Management Group at Morgan Stanley*

**Bruce Ratner**
*New York, NY*

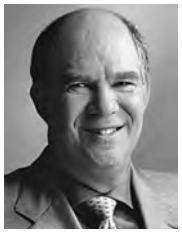
**Geoffrey Robertson**
*Director of Business Assistance, Vermont Sustainable Jobs Fund*

**Douglas Schloss**
*President and CEO, Rexford Management*

**Marilyn H. Simons, Ph.D.**
*President, MJS Foundation; Co-Chair, Simons Foundation*

**Laura Slatkin**
*Co-Founder and Board Chair, NEXT for AUTISM*

**Bruce W. Stillman, Ph.D.**
*President and Chief Executive Officer, Cold Spring Harbor Laboratory*



**James M. Stone, Ph.D.**
*Chairman, The Plymouth Rock Company*

**Karel Svoboda, Ph.D.**
*Vice President and Executive Director, Allen Institute for Neural Dynamics*

**Paul J. Taubman**
*Chairman and CEO, PJT Partners Inc.*

**Diana L. Taylor**
*New York, NY*

**Stuart Weisbrod, Ph.D.**
*Weisbrod Family Office, LLC*

**George D. Yancopoulos, M.D., Ph.D.**
*Chief Scientific Officer & President, Regeneron Pharmaceuticals, Inc.*

## Honorary Trustees

**Bayard Clarkson, M.D.**
*Chair Emeritus; Memorial Sloan Kettering Cancer Center*

**Charles F. Dolan**
*Oyster Bay, NY*

**Lola N. Grace**
*Founder, Middle East Children's Institute*

**Leo A. Guthart**
*Founder and CEO, Topspin Partners*

**Nancy Abeles Marks**
*Nancy Marks Interiors and Carl Marks & Co., Inc.*

**Eduardo G. Mestre**
*Chair Emeritus; Vice Chairman, Evercore Partners*

**William S. Robertson**
*Chairman, Robertson Foundation for Government*

**James H. Simons, Ph.D.**
*Chairman, Simons Foundation*

**Henry Wendt III**
*Friday Harbor, WA*

**Roy J. Zuckerberg**
*Senior Director, Goldman, Sachs & Co.*

# CSHL Scientific Advisory Council

The Scientific Advisory Council (SAC) is an external advisory group that advises the senior management of Cold Spring Harbor Laboratory (CSHL) on matters pertaining to science (both current and future), including the development of a research strategy to maintain CSHL as a world leader. The SAC includes a Chair of Council who is an individual known for scientific breadth and a detailed understanding of research management at the senior management level. The other members are world leaders in their respective fields and as such are able to provide advice on different research areas of the Laboratory.

**Frederick W. Alt, Ph.D. (Chair)**
*Harvard University*

**Joanne Chory, Ph.D.**
*Salk Institute for Biological Studies*

**Carol W. Greider, Ph.D.**
*University of California, Santa Cruz*

**Eve Marder, Ph.D.**
*Brandeis University*

**Markus Meister, Ph.D.**
*California Institute of Technology*

**Steven Salzberg, Ph.D.**
*Johns Hopkins University*

**Kevan M. Shokat, Ph.D.**
*University of California, San Francisco, Berkeley*

**Max S. Wicha, M.D.**
*University of Michigan*

**Wei Yang, Ph.D.**
*National Institutes of Health*

# Governance

The Laboratory is governed by a Board of Trustees of up to 40 members that meets at least three times a year. Authority to act for the Board of Trustees between meetings is vested in the Executive Committee of the Board. The Executive Committee is composed of the Officers of the Board and any other members who may be elected to the Executive Committee by the Board of Trustees. Additional standing and ad hoc committees are appointed by the Board of Trustees to provide guidance and advice in specific areas of the Laboratory's operations.

Representation on the Board of Trustees includes business and community leaders and scientists from major educational and research institutions.

The Laboratory is chartered as an educational corporation by the Board of Regents for and on behalf of the Education Department of the State of New York. It is authorized to operate a graduate program under the name "Cold Spring Harbor Laboratory, School of Biological Sciences" and to confer the degrees of Doctor of Philosophy (Ph.D.), Master of Science (M.S.), and Doctor of Science (Sc.D.), Honorary.

The Laboratory is designated as a "public charity" under Section 501(c)(3) of the Internal Revenue Code.

# Committees of the Board or Laboratory

(As of November 2023)

## BOARD COMMITTEES

### Executive

**Marilyn Simons, Chair**
Casey Cogut
Robert Lourie
Elizabeth McCaul
Howard Morgan
Jamie Nicholls
Bruce Stillman
Paul Taubman
Diana Taylor

### Audit and Risk

**Douglas Schloss, Chair**
Lalit Bahl
Elizabeth Cogan Fascitelli
Leo Guthart (Honorary)
Tom Lister
Elizabeth McCaul

### Nominating

**Paul Taubman, Chair**
Ellen Futter
Jeffrey Kelter
Robert Lindsay
Tom Lister
Elizabeth McCaul
Bruce Ratner
Douglas Schloss
Marilyn Simons
Laura Slatkin
Bruce Stillman

## COMMITTEES OF THE LABORATORY

### Advancement

**Howard Morgan, Chair**
Christine Anderson
David Boies
Casey Cogut
Jeffrey Kelter
Laurie Landeau
Robert Lourie
Charlie Prizzi
Laura Slatkin
Bruce Stillman
Paul Taubman
Diana Taylor
John Tuke

### Commercial Relations

**James Stone, Chair**
Mark Claster
Casey Cogut
Robert Martienssen
Bruce Stillman
John Tuke
Christopher Vakoc
Stuart Weisbrod
Andrew Whiteley

### Compensation Review and Advisory

**Marilyn Simons, Chair**
Ellen Futter
Tracy Johnson
Jamie Nicholls
Douglas Schloss
Diana Taylor

### Education*

**James Stone, Chair**
Lalit Bahl
Michael Botchan
Mark Claster
Casey Cogut
Elaine Fuchs
Ellen Futter
Jacob Goldfield
Tracy Johnson
Laurie Landeau
Robert Lourie
Jeanne Moutoussamy-Ashe
Marilyn Simons
Bruce Stillman
Karel Svoboda

### Facilities

**Jeffrey Kelter, Chair**
Elizabeth Ainslie
Hans Bosch
Elizabeth Cogan Fascitelli
Elaine Fuchs
Mark Hamer
Steven Monez
Frank Sciame
Bruce Stillman
John Tuke

### Finance

**Elizabeth McCaul, Chair**
Elizabeth Cogan Fascitelli
Mark Hamer
Nick Milowski
Jamie Nicholls
Douglas Schloss
Marilyn Simons
John Tuke

*Formerly Academic Affairs and DNA Learning Center Committees.

**Investment**

**Jamie Nicholls, Chair**
Alan Breed
Elizabeth Cogan Fascitelli
Anne Dinneen
Tom Lister
Lyon Polk
Thomas Purcell
Boaz Sidikaro
John Tuke
Stuart Weisbrod

**Robertson Research Fund**
**(Board of Trustees)**

Bruce Stillman (President)
Geoffrey Robertson (Vice President)
Walter Goldschmidts
William Gridley
Leo Guthart
Walter Meier
Lyon Polk
Marilyn Simons

**Robertson Research Fund, Inc.**
**Investment Committee**

William Gridley
Jamie Nicholls
Lyon Polk
Geoffrey Robertson
John Tuke

# Helen Dolan (1926–2023)

The Laboratory lost a great friend and supporter with the death of Helen Dolan on August 19, 2023. Helen, together with Charles, her husband of more than 73 years, made outstanding contributions to Cold Spring Harbor Laboratory (CSHL), and her name will forever be synonymous with public genetics education thanks to the couple's transformational support of the DNA Learning Center.

Helen was born in 1926 in Cleveland. As a student, she was passionately interested in music and art—loves that remained with her throughout her life. At 7, she went to classes at the Cleveland Museum of Art, and at 12, she was awarded a scholarship to attend the Cleveland School of Art. Later in life, she would take up the cello and perform with the Long Island Symphony Orchestra.

Helen attended John Carroll University, and it was there she met Charles, or "Chuck," as we know him. The couple married on July 4, 1950. Working together in their Cleveland home, the Dolans produced and edited short film reels of sports events for syndication to television stations across the country. In the late 1950s, the Dolans moved to New York, where Chuck founded Manhattan Cable Television and HBO before launching Cablevision in 1973. In 1986, Cablevision moved to Melville on Long Island. By then, Helen had already become associated with CSHL. She joined the CSHL Board of Trustees in 1984, serving on the Building Committee. Notably, she would return for a second term in 1996, again serving on the Building Committee.

Helen worked with the Building Committee during a period of great expansion at CSHL and played a significant part in securing crucial funding for the construction of Grace Auditorium. Today, many around the world know Grace as a splendid conference hall for scientists attending meetings at CSHL. However, when it opened in 1986, it was the first major construction at CSHL in more than 30 years. Inevitably, as more scientists came to CSHL to attend these meetings, the issue arose of where these visitors would stay. One usual stomping ground, the rustic, unheated Page "Motel," was already 40 years old and more suited to a summer camp in the Adirondacks than for housing world-class scientists. It was clear that new accommodations were needed. Jim Watson saw an opportunity to use the land above Bungtown Road for a building that would provide both lodging for visitors and research laboratories for CSHL's neuroscience program. The Beckman Foundation contributed to the Neuroscience Center at the north end of the complex, and the Dolan Family Foundation underwrote the cost of a wonderful new building at the south end. With 60 rooms, it was the largest building the Laboratory had built for accommodation, and now, for almost 40 years, visiting scientists have been able to experience the delight of staying on campus in Dolan Hall.

Dolan Hall is not the only CSHL institution bearing the family name. By 1997, the DNA Learning Center (DNALC), which originally operated out of two vans, had outgrown its home in the former school building in Cold Spring Harbor Village. At Helen and Chuck's recommendation, the Dolan Family Foundation again answered the call. The Foundation made a generous pledge to fund an expansion of the DNALC and, in 2000, the 9,000-square-foot BioMedia addition was opened, providing essential space for a computer laboratory, wet labs, museum space, a broadcast media studio, and a conference room. We were honored to name the expanded facility the Dolan DNA Learning Center.

Helen and Chuck have supported many other activities at CSHL as well. In 1986, the Dolans provided Cablevision facilities and staff to help with the production of "Biological Revolution," a video directed by Dave Micklos to celebrate the 100th anniversary of CSHL. During that same year, they helped underwrite one of the most remarkable cultural events in CSHL history—a gala concert performed by Emmanuel Ax, Yo-Yo Ma, and Midori. They also helped underwrite the outdoor sculpture exhibition *Nothing But Steel*, several pieces of which continue to adorn and enliven our grounds.

Helen's support for biomedical science research and education extended beyond CSHL. She was a co-founder of the Lustgarten Foundation for research on pancreatic cancer. I am glad to say that CSHL maintains a very close connection to the Foundation with Dave Tuveson serving as its Director of Science. Elsewhere on Long Island, Helen served as a director of the Oyster Bay Community Foundation. The Dolans were also longtime supporters of Friends Academy, with Helen demonstrating her love of the arts by underwriting construction of the Academy's Helen A. Dolan Center, which includes a performance theater and arts studio. Further afield, at Helen's alma mater, John Carroll University, there is the Charles and Helen Dolan Center for Science and Technology.

In 2003, Helen and Chuck were appointed Honorary Trustees of CSHL, and in 2017, I was delighted to present them with Double Helix Medals in recognition of their outstanding service to the Laboratory. On a personal note, I have known Helen for most of my time at CSHL and though she may have described herself as "introverted," I always found her to be most forthcoming and passionate when it came to her support of CSHL. She was not only a generous philanthropist but a true friend to the Laboratory and the greater scientific community. On behalf of the CSHL Board of Trustees, faculty, and staff, I extend our deepest condolences to Chuck and the entire Dolan family.

**Bruce Stillman**
*Cold Spring Harbor Laboratory*

# PRESIDENT'S REPORT

Cold Spring Harbor Laboratory (CSHL) is a vibrant intellectual community that enables groundbreaking discovery. Our idyllic setting and state-of-the-art facilities offer a uniquely supportive environment for bioscience breakthroughs. However, CSHL's community is not bound by the geography of our campus. Science worldwide has been strengthened by people who spend time on our campus and by the ideas, experiences, and exchanges that they have here. The impact of our work creates waves throughout society.

These waves can take the form of new drugs and technologies as well as advances in knowledge that fuel scientific progress across the globe. They make their way to other research laboratories and educational institutions and from there to business boardrooms and governmental offices. And like oceanic waves, they can have residual effects. Waves come in to the shore of society and back out to the sea of science, helping CSHL expand its capabilities and driving the continuous cycle of discovery.

## Improving Lives around the World

The deep knowledge of biological systems that has been uncovered at CSHL has enabled important medical advances, improving the lives of patients around the world. Take, for example, the fundamental discoveries about cell proliferation that David Beach and his team made here in the 1980s. As a result of their research, drugs like Ibrance (palbociclib), which stops cell growth by blocking the CDK4 and CDK6 enzymes, have become first-line therapies for the treatment of $HR^+$ and $HER2^-$ metastatic breast cancers. Testing for another protein that Beach's laboratory discovered, p16, has become a valuable marker for cancer diagnosis.

Although our investigators focus on basic discovery biology, CSHL's Office of Business Development & Technology Transfer establishes partnerships aimed at translating our discoveries into diagnostic, therapeutic, and technological advances that impact people's lives. One recent success story is the 2016 approval of Spinraza®, the first therapy for spinal muscular atrophy (SMA), using a breakthrough new class of drugs.

*The impact of our work creates waves throughout society.*

Spinraza® restores patients' motor functions using innovative technology called antisense oligonucleotide therapy—co-developed by CSHL Professor Adrian Krainer and a small biotech company called Ionis Pharmaceuticals. The root of this therapy was the Nobel Prize–winning discovery at CSHL of RNA splicing, a mechanism of controlling gene expression. Krainer continued CSHL's studies on RNA splicing and realized that alteration of this process in children with SMA might lead to a new therapy. Thanks to this work, thousands of children who would never have been able to stand or even crawl have now reached these important milestones.

More success stories are written each day. Today, partnerships with venture capital firm Autobahn Labs and the Feinstein Institutes for Medical Research at Northwell Health facilitate clinical trials investigating potential new therapeutics for blood disorders known as myelodysplastic syndromes, based on knowledge uncovered by CSHL Assistant Professor Lingbo Zhang.

## A Global Community of Scientists

CSHL's impact does not stop with discoveries made in our laboratories. We have helped build a strong global community of scientists through the many thousands of researchers who have trained here over our 134-year history. At any given time, our campus community includes hundreds of postdoctoral fellows and graduate students developing skills and ways of thinking they will take with them throughout their careers. Most go on to train students and postdocs of their own, becoming part of a broad intellectual family tree whose branches stem from CSHL. This has happened over and over again, supporting the growth of many new fields in bioscience. Just a few decades ago, the most sophisticated biology research happened almost exclusively in the United States and Europe. That is no longer the case, in part because CSHL scientists, the CSHL Press, and our advanced training courses have helped disseminate knowledge so extensively. Likewise, exceptional scientists come to CSHL from around the world to be part of our community. Our current faculty includes scientists from five continents, and their laboratories are equally diverse.

Approximately 70% of our postdocs and 60% of our students come from outside the United States. Many end up staying in the country, becoming leaders in their fields. That includes people like Supriya Prasanth, who came to CSHL from India in 2001 to do a postdoc in my laboratory and now heads the Department of Cell and Developmental Biology at the University of Illinois Urbana-Champaign. CSHL trainees have found tremendous success outside the United States, too. Gregory Hannon, once a postdoc in David Beach's laboratory, headed a laboratory here for 18 years before leaving CSHL in 2014 to direct the Cancer Research UK Cambridge Institute. Zachary Mainen, who did postdoctoral research in former CSHL neuroscientist Roberto Malinow's laboratory, was also part of the CSHL faculty until 2007. Now he directs the neuroscience program at the Champalimaud Centre for the Unknown in Lisbon, Portugal. Numerous other alumni, including many Nobel laureates, have passed through CSHL.

CSHL's intellectual community becomes even broader and more diverse when we consider the thousands of scientists who visit our campus every year, either virtually or in person. CSHL has been a place for global collaboration since its beginnings. In the 1940s, Max Delbrück and Salvador Luria spent their summers here, taking time away from positions at Vanderbilt and Indiana University, respectively, to study bacteria-infecting viruses called bacteriophage. Their work together, and later with Alfred Hershey, helped launch the field of molecular genetics and led to a 1969 Nobel Prize. Today, CSHL remains a focal point for science's interdisciplinary "supergroups." We are expanding on-campus housing via our Foundations for the Future campaign to support more such collaborations, so scientists from different institutions can come together for extended periods to collaborate just as Delbrück, Luria, and Hershey did.

Some scientific problems demand large-scale collaborations, and we are a hub for several. One example is the International Brain Laboratory, whose experimental and theoretical neuroscientists work together to investigate the brain circuits that control complex behaviors. Those scientists, including CSHL Professor Tony Zador, work at 22 laboratories in the United States and Europe. CSHL is where they come to discuss their progress and plan for the project's future. Likewise, we are the gathering point for the U.S. National Cancer Institute and Cancer Research UK Cancer Grand Challenges–supported team investigating cachexia, the wasting disease that causes extreme weight loss in many people with advanced cancers. The international team is co-led by Eileen White, a former postdoc in my laboratory, now deputy director of the Rutgers Cancer Institute of New Jersey. CSHL Associate Professor Tobias Janowitz is a co-leader of the $25 million project that involves 13 institutions in the United States and United Kingdom.

## Sharing Scientific Knowledge

Our commitment to sharing and disseminating scientific knowledge is perhaps best illustrated through CSHL's world-renowned Meetings & Courses Program. Each year, we host dozens of

courses, drawing attendees from across the globe. Our courses equip scientists with the skills and knowledge needed to explore new territory or delve deeper into their field of study. They have supported massive progress in fields such as neuroimaging, structural biology, genome engineering, and molecular biology—not just from the courses themselves, but via the CSHL Press compiling laboratory protocols and experimental methods and making them readily available to the world's research communities.

Many leading scientists recall courses they took at CSHL as hugely influential on their careers. One prominent example is Venki Ramakrishnan, who came here in 1988 to learn about crystallography and went on to share the 2009 Nobel Prize in Chemistry for determining the structure and mechanism of action of the machinery that makes proteins in cells. Writing about his experience at CSHL, Ramakrishnan states: "The faculty were world-famous and one of them went on to win the Nobel Prize a day after the course. Both I and Rod MacKinnon, a student in a subsequent year, used our newly acquired skills to do the work that won us our own Nobels." To date, 11 scientists who have been awarded the Nobel Prize have taken advanced laboratory courses at CSHL.

By opening our meetings to virtual attendance, we have expanded participation significantly, improving access for far-away scientists who might otherwise be unable to attend. Scientists from Africa and Southeast Asia, in particular, are now much better represented at our meetings. About 13,000 people attended CSHL meetings and courses in 2023, gathering to learn powerful research techniques and discuss the latest findings in a range of fields, from plant genomics to immunology and neurobiology. These events give scientists fresh ideas that they take back with them to laboratories around the world. Moreover, they foster the open exchange of knowledge that is vital for successful science. Our Foundations for the Future project will create additional space for these exchanges—allowing for an even greater societal impact.

Meanwhile, at the Banbury Center, we bring together smaller groups of experts. The format allows them to discuss complex or potentially controversial ideas freely, so Banbury meetings are well suited to consider social issues and policy matters. Critical concerns in medicine, from pain management to Lyme disease, have been on the agenda in recent years, along with topics beyond the biological sciences. In 2021, for example, the Banbury Center convened a meeting on the environmental impact of deep-sea mining, during which participants considered how extracting metals from the seabed might impact global climates and ocean ecosystems. Today, policymakers in the United States and around the globe are discussing these same issues. Again, you can see the ripples of CSHL research and education anywhere you look.

## Strengthening Scientific Literacy

Our DNA Learning Center (DNALC), established in 1988, has always recognized that the science we do at CSHL touches the lives of everyone. Our educational programs for middle and high school students equip children with knowledge of modern genetics and its impact on society. More than 750,000 students have now been exposed to the scientific method through our hands-on approach to learning.

---

*…the science we do at CSHL touches the lives of everyone.*

---

The DNALC program started with outreach to schools in our local community and has expanded dramatically, with a presence in six countries. We operate 13 teaching laboratories in New York. Our methods are also used at licensed centers in Nigeria, China, Austria, Singapore, and the United States, and 10 other cities worldwide have educational programs modeled after the DNALC. Their reach continues to grow. Summer camps at the DNALC in Suzhou also draw students from across China.

Additionally, two new DNALC facilities will open in 2024: one at the Passaic County Technical Institute's Biotechnology Innovation Center in New Jersey, which will serve students of the institute's technical high school and community college, and one in Arecibo, Puerto Rico, as part of the new Arecibo Center for Culturally Relevant and Inclusive Science Education, Computational Skills, and Community Engagement.

The DNALC's programs are for everyone—not just students who have shown an interest in or aptitude for science. Indeed, participation sometimes sparks that interest, setting students off on a path and career they had never considered. To facilitate the careers of teenagers whose families are not experienced in science, the DNA Learning Center's STARS Program—started by DNALC alum-turned-Assistant Director Jason Williams—mentors students through high school, offering them research opportunities as well as guidance on college applications and scholarships. This highly successful program attracts students who otherwise might not have pursued careers in science or medicine. This is another way we can strengthen and support scientific literacy and grow a diverse, global scientific community.

## Connecting Science and Society Worldwide

Indeed, everything we do here at CSHL is for the world we all share. Our programs improve people's lives in myriad ways, from expanding scientific literacy to enabling the development of lifesaving medications and finding new ways to confront climate change. Our bustling intellectual community here on Long Island is deeply integrated with the broader global community, and we take tremendous pride in the ways we support and connect science and society worldwide.

Our institution is a beacon for cutting-edge research and a hub for international collaboration. CSHL research and education programs demonstrate science's ability to break down barriers and drive progress. We represent humanity's greatest ambitions—to explore life, to better understand ourselves, and to improve each other's lives—so it is no wonder those same goals inspire the work we do here every day.

Thank you, CSHL faculty members, postdoctoral fellows, graduate students, employees, trustees, and supporters for contributing to CSHL's vibrant intellectual community and daring to conduct breakthrough bioscience. Your tireless efforts put our institution at the forefront of bioscience research and education. It is my distinct pleasure to serve CSHL alongside you.

**Bruce Stillman, Ph.D., F.R.S.**
*President and Chief Executive Officer*

# Highlights of the Year

## Research

In 2023, hundreds of scientists working in more than 60 Cold Spring Harbor Laboratory (CSHL) research groups published their findings in the world's major scientific journals. Their efforts reflect the full spectrum of CSHL's programs in Cancer, Neuroscience, Plant Biology, Quantitative Biology, and Genomics. The following is a sampling of this year's important findings.

### *A Secret Weapon Exposed*

RNAs are having a moment. The foundation of COVID-19 vaccines are messenger RNAs (mRNAs) that encode protein antigens. They have made their way from biochemistry textbooks into popular discourse. There are many types of RNA in our cells. Unfortunately, when some of them malfunction, it can result in cancer or developmental disorders. Our cells have molecular "machines" that eliminate RNAs at the right time and place. Most come equipped with a "motor" to generate the energy needed to untangle RNA molecules. But one machine, Dis3L2, is an exception. The enzyme can unwind and destroy RNA on its own. This has puzzled scientists for years. Finally, CSHL biochemists have pieced together what is happening.

L. Joshua-Tor

Using state-of-the-art molecular imaging technology, CSHL Professor and HHMI Investigator Leemor Joshua-Tor captured Dis3L2 at work. She fed the molecular machine hairpin snippets of RNA and imaged them getting "eaten" at various stages. After the machine had chewed up the tip of the RNA, it swung open a big arm of its body to peel apart the hairpin and finish the job. Joshua-Tor's team discovered that Dis3L2 has a protruding wedge that enables it to unwind RNA. If the researchers removed the wedge, Dis3L2 could no longer untangle the RNA hairpin.

The findings reveal a surprising new way in which our cells' RNA-controlling machines execute their tasks. Rather than solid structures, these molecular workhorses need to be considered malleable and versatile. "We have to start thinking about these proteins as much more dynamic entities," Joshua-Tor says, "and take that into account when we design therapeutics." This new outlook may help scientists develop better treatments for diseases and disorders caused by RNA gone haywire.

### *Shape-Shifting Antibiotics*

J.E. Moses

In the United States alone, drug-resistant bacteria and fungi infect almost three million people per year and kill about 35,000. Now, CSHL Professor John E. Moses has created a new weapon against these drug-resistant superbugs—an antibiotic that can shape-shift by rearranging its atoms.

Moses came up with the idea of shape-shifting antibiotics while observing tanks with rotating turrets in military training exercises. A few years later, Moses learned of a molecule called bullvalene, whose atoms can swap positions. This gives it a changing shape with over a million possible configurations—exactly the fluidity Moses was after.

Several bacteria have developed resistance to a potent antibiotic called vancomycin, used to treat everything from skin infections to meningitis. Moses thought he could improve the drug's performance by combining it with bullvalene. He turned to click chemistry, a Nobel Prize–winning class of fast, high-yielding chemical reactions that "click" molecules together reliably. Using this technique, Moses and his colleagues created a new antibiotic with two vancomycin "warheads" and a fluctuating bullvalene center.

They found the shape-shifting antibiotic significantly more effective than vancomycin in treating a deadly superbug called vancomycin-resistant *Enterococci* (VRE). Remarkably, the bacteria did not develop resistance to the new antibiotic.

### *Once Sarcoma, Now Muscle*

"Every successful medicine has its origin story. And research like this is the soil from which new drugs are born," says CSHL Professor Christopher Vakoc.

C. Vakoc

A devastating and aggressive type of pediatric cancer, rhabdomyosarcoma (RMS), resembles children's muscle cells. For six years, Vakoc's laboratory has been on a mission to transform sarcomas into regularly functioning tissue cells through a method known as differentiation therapy. To carry out their mission, Vakoc's team created a new genetic screening technique. Using genome-editing technology, they scanned for genes that, when disrupted, would force RMS cells to become muscle cells. That is when a protein called NF-Y emerged.

With NF-Y impaired, the scientists witnessed an astonishing transformation. "The cells literally turn into muscle," Vakoc recalls. "They're switching from a cell that just wants to make more of itself to cells devoted to contraction." The newfound relationship between NF-Y and RMS may set off the chain reaction needed to bring differentiation therapy to patients. The mission does not stop at RMS. Differentiation therapy could be applicable to other cancer types. For example, Vakoc and his team have already succeeded in transforming Ewing sarcoma cells into healthy tissue cells. Notably, both the Ewing sarcoma and RMS discoveries were supported by local families who had lost loved ones to these cancers. Those families and Vakoc's laboratory may yet become the heroes of a new origin story—a scientific breakthrough that could someday help save children's lives and revolutionize cancer treatment as we know it.

### *From Tragedy to Triumph*

A. Krainer

CSHL Professor Adrian Krainer is best known for his groundbreaking research on antisense oligonucleotides (ASOs)—molecules that can control cells' protein levels. His efforts led to Spinraza®, the first FDA-approved treatment for spinal muscular atrophy (SMA).

Following his success with SMA, Krainer started looking into other diseases in which ASOs could make a difference. He soon set his sights on a lethal pediatric brain cancer called diffuse intrinsic pontine glioma (DIPG). "I was contacted by a neurologist and his friend, who had lost her child to DIPG," Krainer explains. Now, Krainer's laboratory has found a way to increase survival rates in mice with DIPG using ASO technology similar to that used for the development of Spinraza®.

The new ASO works by shutting down a mutated protein called H3.3K27M. In DIPG, the dominant mutation blocks closely related proteins from turning many genes on and off, leading to uncontrolled cell growth. When the team used the ASO on mice with DIPG, the genes it affected returned to normal. The tumors stopped growing as fast, and the animals lived longer.

"We could see a lot fewer proliferating cells, and the tumor cells were differentiating into healthy nerve cells," Krainer says. "That tells us DIPG's malignant changes are reversible to an extent."

S. Beyaz

### *Cancer's Multiple Personalities*

Cancer is very complex. Mutations in the same genes can lead to different tumor subtypes in different people. Highly similar proteins produced from the same gene are called isoforms. Different isoforms generate different tumors through a process known as exon skipping. Now, CSHL Assistant Professor Semir Beyaz has developed a new method to model certain liver tumor subtypes using the gene-editing tool CRISPR-Cas9.

Beyaz and his colleagues produced two distinct tumor subtypes by targeting a single section of the mouse gene *Ctnnb1* with CRISPR-Cas9. This tool is mostly used to inhibit gene function. In fact, this is the first time CRISPR-Cas9 has been used to generate different cancer-causing gain-of-function mutations in mice. These mutations enhance protein activity to promote tumor growth. Beyaz's team sequenced each tumor subtype to figure out which isoform was associated with the differences they observed. Next, to confirm that these isoforms actually caused the variances, they produced them in the mouse without using CRISPR. They found that they were indeed able to generate the two different tumor subtypes along with their respective characteristics. Both subtypes are also found in humans.

The mutations Beyaz targeted can lead to colon and liver cancers. Targeting exon skipping has emerged as a potential therapeutic approach for treating cancer and other diseases. Beyaz's method allows researchers to investigate this phenomenon in living mouse cells using CRISPR-Cas9 technology. The platform could someday help researchers develop new therapeutic interventions, such as ASOs that can modulate exon skipping.

### *These Worms Have Rhythm*

Growing from a tiny cluster of cells into an adult organism takes precise timing and control. The right genes must turn on at the right time, for the right duration, and in the correct order.

C. Hammell

CSHL Professor Christopher Hammell found that in the *C. elegans* worm, this genetic orchestra has no single conductor. Instead, a quartet of molecules works in concert to time each developmental stage. Each begins with two proteins, NHR-85 and NHR-23. Together, they spark a pulse of gene expression, switching on the microRNA *lin-4*, which controls stem cell development patterns. The pulse's timing, strength, and duration depend on the short period of time when NHR-85 and NHR-23 interact with each other. The pulse of gene expression of the microRNA *lin-4* is terminated by another protein called LIN-42, which ends each developmental period by shutting off NHR-85.

Hammell teamed with Wolfgang Keil from the Curie Institute in Paris, France, to observe this gene expression cycle in action. Their collaboration resulted in the first-ever video footage to capture active gene expression throughout an entire animal in real time. Understanding how the worm's developmental clock is regulated could help explain how time affects development in other animals, including humans.

### *The Science of Supermoms*

When you pick up a baby, you can't help feeling happy. But where does this feeling come from? How does it influence future behavior? CSHL Professor Stephen Shea has found that, when it comes to maternal care, social contact is its own reward.

S. Shea

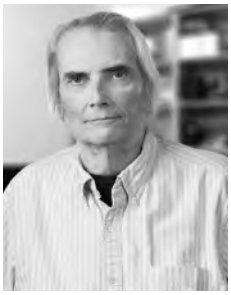
Shea and postdoc Yunyao Xie traced maternal motivation in mice to a region of the brain called the ventral tegmental area (VTA). They found that neurons in the VTA use dopamine to nurture maternal instincts through a process called reinforcement learning. When a mom picks up a crying child, the mother's brain's VTA neurons release dopamine. This creates an expectation of future rewards, driving her to pick the child up again the next time it cries.

To observe reinforcement learning in action, Shea's team built an enclosure with two chambers. They placed a mouse on one side and a pup on the other, behind one of two doors. They then played specific sounds to signal whether the pup was present and if it was behind door number one or two. Each time the mouse heard a sound and retrieved the pup, VTA neurons rewarded it with dopamine.

Deciphering how social contact is encouraged in mice gives researchers a clue about VTA neurons' involvement in rewarding other behaviors. This may lead to a better understanding of neurodevelopmental conditions that affect social interactions, like autism.

### *Holy Immunity, Bat Genomes!*

CSHL Professors W. Richard McCombie and Adam Siepel and postdoc Armin Scheben have sequenced the genomes of the Jamaican fruit bat and Mesoamerican mustached bat. Comparing these sequences to those from other mammals, the team found that rapid evolution has streamlined bat genomes to defend against infection and cancer.



W.R. McCombie

A. Siepel

The Jamaican fruit bat and Mesoamerican mustached bat belong to the world's most ecologically diverse superfamily of mammals. McCombie, Siepel, and Scheben created complete genomes for both bats using new Oxford Nanopore sequencing technology. They then compared these sequences to 15 other bat and mammal genomes, including humans. This analysis revealed an unknown shift in levels of two inflammatory protein–coding genes called interferon-alpha and -omega.

The team also found that compared with other mammals, bat genomes contain more changes in cancer-related genes, including six that repair DNA and 46 that suppress tumors. McCombie, Siepel, and Scheben are currently exploring how bats' immune genes are regulated and how they might be expressed in different parts of the body. They hope their work will provide new insights into the links between immunity, aging, and cancer.

### *Plants Pass on Memories*

CSHL Professors and HHMI Investigators Rob Martienssen and Leemor Joshua-Tor have been researching how plants pass along the markers that keep transposons inactive. To silence them and protect the genome, cells add regulatory marks to specific DNA sites. Martienssen and Joshua-Tor have now shown how a protein called DDM1 makes way for the enzyme that places these marks on new DNA strands. Plant cells need DDM1 because their DNA is tightly packaged. To keep their genomes compact and orderly, cells wrap their DNA around proteins called histones.

R. Martienssen

Martienssen and former CSHL colleague Eric Richards first discovered DDM1 30 years ago. Now, through genetic and biochemical experiments, Martienssen has pinpointed the exact histones DDM1 displaces. Joshua-Tor used cryo-electron microscopy to capture detailed images of the enzyme in action. Together, they saw how DDM1 grabs onto particular histones to remodel packaged DNA. "An unexpected bond that ties DDM1 together turned out to correspond to the first mutation found all those years ago," Joshua-Tor says.

The experiments also revealed how DDM1's affinity for certain histones preserves regulatory controls across generations. The team showed that a histone found only in pollen is resistant to DDM1 and acts as a placeholder during cell division. "It remembers where the histone was during plant development and retains that memory into the next generation," Martienssen says.

Plants may not be alone here. Humans also depend on DDM1-like proteins. The new discovery may help explain how those proteins keep our genomes functional and intact.

### *Autism in the Family Tree*

Scientists long thought that siblings born with autism spectrum disorder (ASD) share more of their mother's genome than their father's. But CSHL Professors Ivan Iossifov and Michael Wigler have now shown that, in many cases, it is dad who might be playing a bigger genetic role.

Over the last two decades, CSHL scientists have led a multimillion-dollar effort to uncover the genetic origins of autism. They discovered thousands of genes that, when damaged, may cause a child to be born with ASD. However, their work was not able to account for all cases of ASD. So, Iossifov and Wigler set out to find the missing sources.

The duo analyzed the genomes of more than 6,000 volunteer families. They found that in families that have two or more children with ASD, the siblings shared more of their father's genome. Meanwhile, in families in which only one sibling had ASD, the children shared less of their father's genome. No one is sure how dad's genome makes its mark on children with ASD, but Iossifov has a couple of interesting ideas. He thinks some fathers may carry protective mutations that fail to get passed on, or fathers may pass down mutations that trigger the mother's immune system to attack the developing embryo.


I. Iossifov


M. Wigler

"If one or two of those theories prove to be true," Iossifov says, "then it opens different treatment strategies, which can, in the future, affect quite a lot of families." This research also offers helpful tools for educators and therapists, as it allows for earlier diagnoses and a better overall understanding of autism.

## Research Faculty

### *Awards and Appointments*

The Society for RNA Therapeutics named Professor Adrian Krainer a member of the Society's charter board of directors. The board consists of 14 individuals from some of the world's most renowned science and health institutions. Krainer is recognized as a leading expert in RNA splicing. His research in this area led to the breakthrough drug Spinraza® as well as potential treatments for cystic fibrosis and a deadly pediatric brain cancer. Krainer was also named an honorary member of the International Association for Dental Research (IADR). The IADR honored Krainer for fundamental and translational research that led to a lifesaving treatment for spinal muscular atrophy (SMA).

Associate Professor Jessica Tollkuhn won the Pershing Square Foundation's inaugural Maximizing Innovation in Neuroscience Discovery (MIND) Prize. The $750,000 award will support her research on how sex hormones affect the brain over the course of a lifetime.


J. Tollkuhn


L. Cheadle

The Howard Hughes Medical Institute (HHMI) named Assistant Professor Lucas Cheadle a Freeman Hrabowski Scholar. Cheadle is one of 31 individuals selected for this new program. The program recognizes early-career faculty with strong potential to become leaders in their fields and advance diversity, equity, and inclusion in science.

Assistant Professor Lingbo Zhang received the National Cancer Institute (NCI) MERIT Award. Short for Method to Extend Research in Time, the MERIT Award provides early-career scientists a source of stable, long-term funding. Zhang's laboratory studies how diet and other environmental factors influence cancer development.


L. Zhang


A. Banerjee

Assistant Professor Arkarup Banerjee was awarded the Klingenstein-Simons Neuroscience Fellowship. The award provides $300,000 in research funding. Banerjee will use the funding to continue his study of "singing" mice, which allows his laboratory to analyze the parts of the brain that drive communication.

Taking a lead role in a crucial nationwide effort, Professor and Cancer Center Director David Tuveson was named the founding chair of the newly launched American Association


D. Tuveson


S. Sun


B. Stillman

for Cancer Research (AACR) Cancer Centers Alliance. The new initiative fosters collaboration and innovation to advance lifesaving scientific discoveries.

The HHMI named Simón(e) Sun, a postdoc in Associate Professor Jessica Tollkuhn's laboratory, a 2023 Hanna Gray Fellow. The fellowship provides Sun up to $1.5 million for up to eight years in support of her studies on the effects of sex hormones on the brain during adolescence.

I received the American Society for Biochemistry and Molecular Biology's (ASBMB's) Earl and Thressa Stadtman Distinguished Scientist Award for outstanding achievement in basic research. I was honored to accept the award in recognition of my laboratory's discovery of the origin recognition complex, which allowed for a new understanding of DNA replication.

### *New Hires/Promotions*


N. Milowski


C.G. Pehle


M. Zaborowski

Nicholas Milowski joined the Laboratory as Chief Financial Officer. Christian Gerno Pehle joined as a NeuroAI Scholar. MaryJo Zaborowski joined as Vice President of Information Technology and Chief Information Officer.

Hannah Meyer was promoted to Assistant Professor. Tobias Janowitz and David McCandlish were promoted to Associate Professor. Florin Albeanu, Christopher Hammell, Ivan Iossifov, and Stephen Shea were promoted to Professor.

Sara Goodwin and Zihua Wang were promoted to Research Assistant Professor.


H. Meyer


T. Janowitz


D. McCandlish

## Education Highlights

### *Meetings & Courses Program*

The CSHL Meetings & Courses Program delivered a comprehensive program of hybrid and in-person sessions across the biological and biomedical sciences. Overall, the program has seen a strong resurgence in attendance, building on the momentum of the second half of 2022. The artificially low "COVID cap" introduced in 2022 was eliminated in 2023, and in-person attendance has now reached pre-pandemic attendance levels.


F. Albeanu


S. Goodwin


Z. Wang

Symposium lunch

In total, 31 advanced courses were offered in 2023, attracting 1,454 applicants, of whom 618 were accepted. Graduate students accounted for 53% of admitted students, and postdocs 28%. Women represented 56% of accepted students. Students from minority groups accounted for 17% of this year's student body—a figure that has more than doubled over the past decade. Altogether, Meetings & Courses enrollment exceeded 10,000 participants in 2023.

Notably, Katalin Karikó and Drew Weissman—winners of the 2023 Nobel Prize in Physiology or Medicine for their foundational work toward the development of COVID mRNA vaccines—both have a history with CSHL. Dr. Weissman discussed mRNA vaccines at our COVID/SARS CoV2 Rapid Research Reports virtual conference in July 2020. Dr. Karikó spoke about the science behind the success of mRNA vaccines at a 2023 conference organized with the CSHL Center for Humanities & History of Modern Biology.

### *Banbury Center*

The Banbury Center hosted 26 CSHL activities in 2023: 15 meetings, six retreats, and five special events.

Banbury meeting themes in 2023 spanned a range of research, education, and policy issues. Research-focused meetings covered cell and plant biology, sepsis, Fragile X syndrome, Kennedy's disease, cancer clinical trials, and environmental contributors to human health. Two meetings brought together CSHL scientists with colleagues at Northwell Health to discuss collaborations in pancreatic cancer and glioblastoma, and a convening of CSHL and Stony Brook University leaders aimed to enhance connections between their institutions. Banbury hosted three external strategy retreats: the Clinical Lab 2.0 Colloquium, Dryad Retreat, and the Lustgarten Foundation's annual scientific meeting. Policy issues were at the center of meetings on evaluation practices related to the U.N. Sustainable Development Goals and the ethics of therapeutic use of psychedelics.

### *Center for Humanities & History of Modern Biology*

For 2023, the Center's major annual meeting was Recombinant DNA: Fifty Years of Discovery and Debates. As with each of our annual meetings, this event gathered participants who shaped a significant field of scientific research along with current practitioners, educators, and historians.

K. Karikó, B. Stillman, and Z.J. Huang

Speakers included five Nobel laureates. The day after the meeting, we awoke to the news that the Nobel Prize had been awarded to yet another presenter, Katalin Karikó, for developing mRNA-based methods of delivering vaccines and therapies.

This year saw the continuation of our project on the history of science and technology on Long Island, supported by the Robert D.L. Gardiner Foundation. The project culminated in the exhibition, "Plant Science, Biotechnology and Agriculture on Long Island, 1900–2020," public lectures, and a workshop for local librarians and historians.

The Center also spearheads a Women in Science initiative that has hosted a public roundtable and added dozens of interviews to CSHL's oral history collection. Other key activities in 2023 included co-organizing our annual Nobel laureate lecture with CUNY, hosting researchers and artists at CSHL through the Sydney Brenner Research Fellowships and the Celia and Walter Gilbert Artist-in-Residence Program, and contributing to frameSHIFT campus events. Recordings of all Center events are available through the Library & Archives website.

### *School of Biological Sciences*

In 2023, the School welcomed its 25th incoming class and graduated its 20th. Forty-six of our graduates have now secured tenure-track faculty positions (although four have since left these positions for industry). Eight have been promoted to associate professor (often conferring tenure), and 11 are full professors. Our graduates have also moved into influential positions in administration, publishing, consulting, science communication, and industry.

During 2023, scientific papers published by students of the School appeared in major journals, bringing the cumulative total to more than 500. As in past years, current and former students won prestigious and highly competitive scholarships, fellowships, and prizes. In August, the School welcomed 12 new students. Members of the Class of 2023 were selected from an extraordinary 971 applicants. Other new graduate students entered as visitors from other institutions, including seven from Stony Brook University and two from Hofstra University.

Established 65 years ago, our Undergraduate Research Program (URP) welcomed 19 students for the summer of 2023. The equally innovative Partners for the Future program brought gifted local high school students to CSHL labs for hands-on research experience. And, in 2023, the School welcomed five scholars to its very first Postbaccalaureate Research Education Program (PREP) class.

URPs 2023

### *DNA Learning Center*

The DNA Learning Center (DNALC) enjoyed a record-breaking year in 2023. Attendance at the Brooklyn facility (DNALC NYC) increased twofold over that of 2022, and 53% of New York City public schools received field trip scholarships—fulfilling the goal of increasing access for low-income and underrepresented minority students. Summer camps were another high point, with attendance increasing to an all-time high of 1,361. Impressively, the Regeneron DNALC was able to reach full occupancy during only its third full year of operations.

A National Science Foundation grant was awarded to update the DNALC bioinformatics education platform, *DNA Subway*. Reimagined mobile and desktop versions of the platform—*DNA Subway 2.0*—will give students the flexibility to collect data and complete assignments in school, at home, or on the road from their preferred device. This will help better serve the needs of low-income, rural, and limited-sight students.

WiSE Fun with DNA

In August, a memorandum of understanding was signed with Oxford Nanopore Technologies. In addition to massive DNA sequencers used at the CSHL Genome Center, Oxford Nanopore produces the MinION sequencer, which is about half the size of a mobile phone. Together, MinION and *DNA Subway 2.0* will offer the first integrated system to support DNA sequencing and analysis anytime, anywhere, by anyone.

### *Cold Spring Harbor Laboratory Press*

CSHL Press achieved its strategic goals for 2023. More institutional read-and-publish agreements were signed, continuing the research journals' momentum toward open access. *Genes & Development, Genome Research,* and *RNA* remained prominent in their fields. The quality and output of the review journals were maintained, and usage of all the Cold Spring Harbor Collection titles remained high, as the open-access transition strategy requires. The born-open-access journal *Life Science Alliance* continued to prosper, publishing more papers and providing its three partner–owners with a first profit share.

Journal advertising had a strong year in a challenging marketplace. Regrettably, the open-access journal *Molecular Case Studies*, which looked promising two years ago, attracted too few submissions to remain viable and was discontinued in December.

The book publishing program generated new laboratory manuals for teams working with bacteria, mice, frogs, and killifish. Two new books on the future of science and the promise of RNA therapeutics aimed for larger, more general audiences. E-book editions were created for 11 strong titles from the book backlist. The year's successes enabled the Press to contribute once again a substantial surplus to support the Laboratory.

### *Preprints in Biology and Medicine*

A preprint is a research manuscript made freely available by its authors without peer review. The Laboratory's preprint platforms—bioRxiv for biology and medRxiv for health sciences—are accelerating biomedical research by transforming how scientific work is shared. Freely accessible and not reliant on the slow pace of peer review, a preprint can prompt other scientists to examine results quickly and, if appropriate, help to advance them.

In 2023, bioRxiv marked its 10th anniversary by posting 39,200 preprints, a record annual total. After the extraordinary spikes of COVID-19 research papers posted in 2020 and 2021, medRxiv posted 11,000 preprints, more than 90% of which were not pandemic-related. Together, the platforms now host over a quarter of a million preprints—communications from hundreds of thousands of scientists around the world that are read millions of times each month.

New features added in 2023 for authors and readers included audio abstracts for neuroscience preprints as well as AI-generated summaries for bioRxiv manuscripts. More journals were integrated into the servers' manuscript transfer networks, and their presence was enhanced on several social media platforms. Efforts continued to build sources of financial support in addition to Laboratory funding and the generous core grant provided by the Chan Zuckerberg Initiative.


The Preprints team celebrates bioRxiv's 10th anniversary.

## Board of Trustees

The Board of Trustees at Cold Spring Harbor Laboratory continues to provide extraordinary leadership and support to CSHL. The Board has been integral to the launch and progressive efforts of CSHL's most ambitious fund-raising campaign, "Foundations for the Future," with the aim of propelling the institution's research and education programs to new heights.

The Board and CSHL team were excited to welcome a new member to the Trustees this year. On June 8, the CSHL Board of Trustees elected Ellen V. Futter to the private not-for-profit institution's governing body. Futter is the president emerita of the American Museum of Natural History. She presided over the museum for three decades, following 13 years as president of Barnard College.

Futter has been a nationally recognized academic leader since 1980, when, at age 30, she was named president of Barnard College. At the time, Futter was the youngest person to hold that position at a major American college. In 1993, she became the first woman to lead the American Museum of Natural History. During her tenure, Futter played an essential role in the museum's expansion. We are delighted to add Futter's expertise and enthusiasm to the Board and look forward to working with her on future initiatives.

## Business Development and Technology Transfer

The Business Development and Technology Transfer Department at CSHL demonstrated remarkable resilience and impact in 2023, navigating a continuously challenging biotech landscape shaped by dramatically reduced investment in early-stage drug development. Despite market volatility, we achieved notable successes, including securing a mediation settlement of more than $10 million tied to patent rights. This strategic win not only contributed to $10.9 million in total licensing revenue but also enabled new investments, such as funding new postdoctoral housing to support CSHL's mission.

Amid the closure of four portfolio companies, the department maintained its momentum by extending funding for key collaborations and forging innovative partnerships. Highlights include a collaboration agreement with Autobahn Labs to advance cancer immunotherapy and an expanded collaboration with Caper Labs on medical food and drug development targeting cancer treatment. Sponsored research funding also rose to $4.25 million, reflecting strong support for CSHL's pioneering science.

Through active portfolio management and a commitment to translating discovery into solutions, the Business Development and Technology Transfer Department remains a driving force in transforming CSHL's cutting-edge research into lifesaving therapies and technologies to advance impactful science for societal benefit.

## Infrastructure

In 2023, the CSHL Facilities Department continued its mission to enhance campus infrastructure, support scientific endeavors, and provide an optimal environment for research and education. Several notable capital projects and improvements were completed, reflecting CSHL's commitment to excellence.

### *McClintock Laboratory Renovations*

Renovations to the McClintock Laboratory were completed, ensuring state-of-the-art facilities for researchers. These upgrades were tailored to meet the evolving needs of the laboratory's scientific community, facilitating innovation and productivity.

McClintock Laboratory

### *Jerome Cottage Faculty Housing*

As part of CSHL's efforts to provide comfortable and accessible accommodations for faculty members, Jerome Cottage was renovated and converted into residential housing. This project underscores CSHL's dedication to supporting its academic staff and fostering camaraderie along with a vibrant community.

### *Hillside Laboratory Faculty Parking Lot*

At the close of 2023, construction of the Hillside Laboratory faculty parking lot was completed. This new facility provides convenient access for faculty members while improving overall campus functionality, especially as the phaseout of pandemic-era travel restrictions and work policies brings more visitors and commuter vehicles to CSHL.

### *Modernizing Campus Infrastructure*

The department continued its programs aimed at modernizing and improving the heating, ventilation, air conditioning, electrical, and plumbing systems throughout campus facilities. These ongoing efforts are crucial to ensuring a comfortable and efficient working environment for researchers and staff while promoting energy efficiency and sustainability across CSHL's infrastructure.

## Community Outreach

CSHL reintroduced all in-person community outreach programming in 2023. Visitors from across Long Island and afar were excited to attend our events, which included both new activities and the return of several beloved programs.

Fourteen public walking tours conducted between April 29 and December 3 brought more than 220 community members to the Laboratory. Each tour brought in groups of up to 25 visitors, who had purchased tickets online. The tours, led by trained graduate students, offer attendees insight into CSHL's history, scientific breakthroughs, Nobel Prize achievements, architecture, art, and more. This year also saw the successful relaunch of *Cocktails & Chromosomes,* our public

Camila dos Santos speaking at a *Cocktails & Chromosomes* event

lecture series hosted at a local bar. The events, presented at Industry bar in Huntington, New York, offer our neighbors the opportunity to learn about science from a CSHL faculty or laboratory member. Each installment features a different scientist speaking about a topic relevant to their research in a way that anyone can understand.

As our audience at these in-person events expanded, our online presence continued to grow. Web and social media content, both paid and organic, allowed us to reach large-scale scientific and nonscientific audiences. Additionally, CSHL's research and education programs continued to attract mainstream media coverage.

### *Public Presentations*

**March 2:** *Dr. John Carpten on Closing the Gaps to Achieve Cancer Health Equity*. Presented as part of the ongoing Roy J. Zuckerberg community engagement series from CSHL's NCI-designated Cancer Center.

Helen Hou and Isabella Rossellini

**March 4–5:** Isabella Rossellini's *Darwin's Smile*. Grace Auditorium hosted two performances of the actress' live one-woman show that reconciles the worlds of art and science. The March 4 performance was followed by a Q&A with Assistant Professor Helen Hou.

**June 23:** "From *HER* to Here: ChatGPT & The New Age of AI Companionship," a movie screening at Cinema Arts Centre in Huntington, NY, featuring a lecture and audience Q&A with Professor Anthony Zador and NeuroAI Scholar Kyle Daruwalla.

**June 29:** *Cocktails & Chromosomes*: "Talking AI" with NeuroAI Scholar Kyle Daruwalla

**July 27:** *Cocktails & Chromosomes*: "The shucking origins of corn" with Professor David Jackson

**August 31:** *Cocktails & Chromosomes*: "Uprooting climate change" with Associate Professor Ullas Pedmale

**September 14:** Concert Series: Internationally known cellist Adrian Daurov and pianist Spencer Myer performed a one-hour concert in Grace Auditorium, followed by a reception.

**October 5:** *Cocktails & Chromosomes*: "Breast cancer awareness edition" with Associate Professor Camila dos Santos

**October 6:** Concert Series: Husband-and-wife piano duo Anna and Dmitri Shelest performed a one-hour concert in Grace Auditorium, followed by a reception.

**October 26:** *Cocktails & Chromosomes*: "The mysterious songs of mice" with Assistant Professor Arkarup Banerjee

**November 30:** *Cocktails & Chromosomes*: "You are what you eat … but how?" with Assistant Professor Semir Beyaz

### *Social Media*

In 2023, CSHL continued to grow its presence on social media by utilizing Facebook, X (formerly Twitter), Instagram, LinkedIn, and YouTube to engage broader scientific and nonscientific audiences. Through the daily promotion of scientific stories, educator highlights, features, and the campus community, CSHL's LinkedIn following grew by 22.3%, Instagram following grew by 16.5%, YouTube by 13%, X by 9.2%, and Facebook by 2.5%.

### *In the News*

**January 30:** "After a decade, CRISPR gene editing is a 'revolution in progress.' What does the future hold?" *USA Today*

**February 1:** "Scientists across the country continue to make headway in the fight against glioblastoma," *WABC-TV*

**February 1:** "Monogamous Prairie Voles Reveal the Neurobiology of Love" and "This Common Aquatic Plant Could Produce Biofuel," *Scientific American*

**March 7:** "CRISPR technology modifies cells to fight cancer: Ultimate tool for manipulating life," *Fox 5 New York*

**April 6:** "Brain scans show how different factors can influence obesity in men and women," *NBC News*

**April 20:** "Cutting-edge advances in cancer treatment are underway. Here are 3 that could change lives." *USA Today*

**August 28:** "We're finally figuring out how plants pass on genetic memories," *Popular Science*

**September 22:** "Bats could play vital role in preventing and treating cancer: 'first step' discovery," *New York Post*

**November 9:** "DNA Explained," *NBC Nightly News: Kids Edition*

**December 13:** "How Tomatoes are Revolutionizing Urban Farming," *PBS*

## Looking Forward

Cold Spring Harbor Laboratory launched the Foundations for the Future campaign to support the Laboratory's vision to meet the needs of the future of science. At the core of this ambitious endeavor lies a visionary seven-acre expansion dedicated to pioneering research in three critical areas that will shape the world of tomorrow: neuroscience, neuroAI, and brain–body physiology. Construction will ramp up in 2024 on Phase I to include three new research buildings.

**Bruce Stillman, Ph.D., F.R.S.**
*President and Chief Executive Officer*

# CHIEF OPERATING OFFICER'S REPORT

The year 2023 was one of many significant developments for CSHL that are worth reflecting on as we continue to advance the frontiers of biology through research and education and embrace the promise and opportunity ahead.

J. Tuke

## Financial Resources

CSHL's philosophy is to budget conservatively, especially with revenue sources that are difficult to project with certainty, like third-party awards and philanthropic support. Despite unanticipated challenges, we were pleased to finish the year with a modest surplus. In 2023, CSHL revenues achieved their second highest level ever at $183 million.

Every major source of income exceeded original budget projections, including third-party awards (+10%), endowment funds, royalties, and interest (+10%), and our education divisions (+2%). For the second year in a row, unrestricted support from our generous donors set a record, capping the year with a very successful Double Helix Medals dinner. However, expenses also exceeded original budget projections as inflation continued to impact salaries, benefits, staff turnover, supplies, and utilities. Thankfully, income growth bested expense changes and our operating results improved, with the modest surplus creating a reserve that will help support 2024 operations.

Our endowment serves as a critical buffer against the uncertainties of federal and private funding. Led by a robust fourth quarter in the financial markets, the endowment returned 14.5% in 2023, its fifth best performance in 30 years. All but one asset class exceeded its benchmark, and the return was 6.5% higher than inflation (as measured by CPI) after a 4.5% draw from the endowment. That means the purchasing power of this critical income source was enhanced. Our 2023 endowment performance put it in the top quartile of a pool of more than 600 endowments and foundations. The endowment started 2023 with a balance of $698 million and—after returns, new gifts, and a $32.5 million operations draw—ended the year at $780 million.

## Operating and Capital Matters

Managing operations across four campuses on 220 acres with 86 buildings requires significant capital and human resources. Our Facilities Department has more than 100 employees working across a wide range of functions. In 2023, $9.1 million was invested in our campus facilities for repairs, upgrades, equipment, and normal life cycle activities. However, this investment must increase, and we are working with the Board's Facilities Committee to create a long-term capital plan to meet that goal. CSHL is fortunate that historically it has been able to supplement its capital budgets with restricted gifts and grants, and 2023 was no exception.

The capital project supported by our Foundations for the Future campaign is now officially underway. At the June 2023 Board meeting, our Trustees approved moving forward with the $240 million Phase I expansion of our Hillside campus. Over the next three years, research buildings focused on Neuroscience, Artificial Intelligence/Quantitative Biology, and Brain–Body Physiology will rise up from an underground superstructure that will include the foundations, parking, and, eventually, a vivarium.

Another multiyear effort, dubbed "Project Helix," will bring significant long-term benefits to CSHL via the implementation of a new enterprise resource planning (ERP) platform. In 2023,

CSHL selected Workday to provide an integrated system for its financial, human resources, procurement, and third-party awards activities. Hundreds of business processes will be streamlined, digitized, and automated, improving the efficiency, accuracy, and timeliness of our data access and analysis.

CSHL's branding project also commenced in 2023 with the ultimate goal of broadening our audiences while maintaining the strong following we enjoy globally within scientific communities. Results of the first phase of this project will be presented at the June 2024 Board meeting, including recommendations to embrace a bolder, new approach on market positioning, messaging, and visuals. Look for more on this in next year's Annual Report.

## Our Global Community

CSHL is an amazing community of 1,000 people working across research laboratories, education divisions, and various support functions. We hail from 77 nations, creating a rich fabric of cultures and traditions all working toward a single mission. Each year we acknowledge the extraordinary work of employees who go "above and beyond" through our Community Recognition Program. Several years ago, we started the frameSHIFT Program, which brings speakers to campus to examine the social, cultural, and historical sides of working at an academic research institution. Last year we welcomed two new members to CSHL's Administrative leadership. MaryJo Zaborowski, our Chief Information Officer, brings decades of experience in the private biotech sector. CSHL CFO Nick Milowski, CPA, joined us late in the year from Fordham University, where he was VP, Finance, and Assistant Treasurer. Both bring experience with digital transformation and are valuable members of the Project Helix Steering Committee.

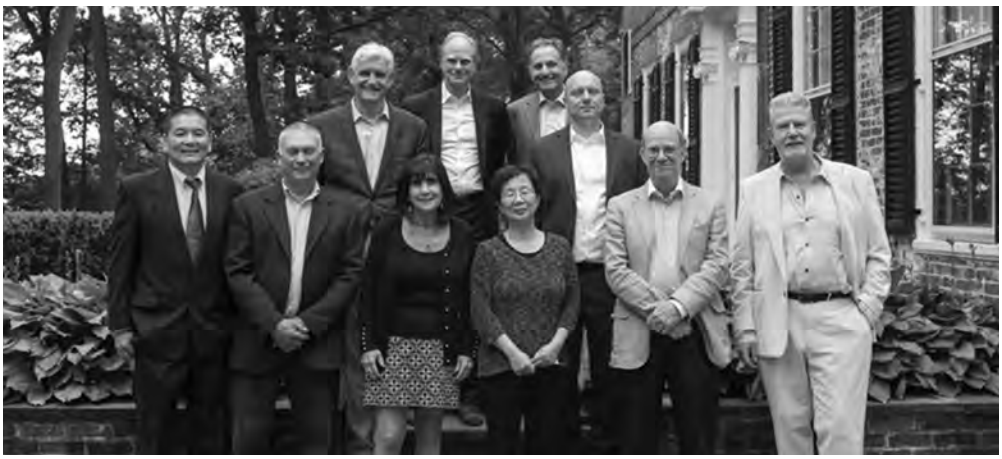
Thank you to the entire CSHL community for contributing to our success in 2023.

**John P. Tuke**
*Chief Operating Officer*

# Long-Term Service

The following employees celebrated milestone anniversaries in 2023:

| | |
|---|---|
| 45 years | Michael Wigler, *Wigler Laboratory* |
| 35 years | Frank Bowdren, *Culinary Services* |
| 30 years | Deborah Aufiero, *Procurement*; Diane Esposito, *Research Operations*; Wayne Pav, *Facilities*; David Stewart, *Meetings & Courses Program* |
| 25 years | Drew Comer, *Facilities*; Paul Edwards, *Facilities*; Louis Hunter, *Facilities*; Jeffrey Klaverweiden, *Facilities*; Oscar Lastra, *Facilities*; Amanda McBrien, *DNA Learning Center*; Susana Roman, *Culinary Services*; Ming Wang, *Joshua-Tor Laboratory*; Lifang Zhang, *Ware Laboratory* |
| 20 years | Eileen Earl, *Laboratory Animal Resources*; Corrisa Farmer, *CSHL Press*; Wayne Hamilton, *Procurement*; Alexei Koulakov, *Koulakov Laboratory*; Partha Mitra, *Mitra Laboratory* |



Front row (*left to right*): Ming Wang, Jeffrey Klaverweiden, Diane Esposito, Lifang Zhang, Jim Stone, David Stewart; back row (*left to right*): Bruce Stillman, John Tuke, David Spector, Alexei Koulakov

15 years

Florin Albeanu, *Albeanu Laboratory*; Nancy Bolanos, *dos Santos Laboratory*; Elmer Canales, *Culinary Services*; Yon Chang, *Mills Laboratory*; Lauren Correri, *DNA Learning Center*; Martin Davis, *Banerjee Laboratory*; Alexander Dobin, *Dobin Laboratory*; Camila dos Santos, *dos Santos Laboratory*; Vlad Drozdoff, *Business Development & Technology Transfer*; Senem Mavruk Eskipehlivan, *McCombie Laboratory*; Brad Frey, *Meetings & Courses Program*; Elena Ghiban, *McCombie Laboratory*; Thomas Gingeras, *Gingeras Laboratory*; Silvia Gonzalez, *Culinary Services*; Wilman Gutierrez, *Culinary Services*; Manzar Hossain, *Stillman Laboratory*; Michael Hutchinson, *Facilities*; Ivan Iossifov, *Iossifov Laboratory*; Jian Lin, *Facilities*; Zachary Lippman, *Lippman Laboratory*; Kristine Murphy, *Finance & Accounting*; Darryl Pappin, *Pappin Laboratory*; Jason Pellegrini, *Laboratory Animal Resources*; Marlenis Romero, *Culinary Services*; Richard Sever, *CSHL Press*; Damianos Skopelitis, *Vakoc Laboratory*; Tara Skopelitis, *Jackson Laboratory*; Nicole Solomon, *Finance & Accounting*; Jessica Toner, *Human Resources*; Christopher Vakoc, *Vakoc Laboratory*; Maria Vilorio, *Culinary Services*

(*Left to right*) Vlad Drozdoff, David Spector, Wilman Gutierrez, Jessica Toner, Nicole Solomon, Damianos Skopelitis, Tara Skopelitis, Elena Ghiban, Senem Mavruk Eskipehlivan, Bruce Stillman, Zachary Lippman, Martin Davis

# RESEARCH

# CANCER: GENE REGULATION AND INHERITANCE

**Camila dos Santos**' laboratory studies the epigenetic regulation of normal and malignant mammary gland development, with an emphasis on the alterations brought by pregnancy. Significant changes mark the pre- and postpubescence mammary developmental stages, but those associated with pregnancy have the greatest effect on cellular function, tissue reorganization, and breast cancer susceptibility. The dos Santos group has recently found that mammary glands react differently to a second pregnancy than they do to the first one, with associated changes in DNA methylation. These findings suggested that pregnancy changes the state of mammary cells, and this may permanently alter how they react to the next pregnancy. In addition, the dos Santos laboratory is exploring how pregnancy-induced epigenetic changes might influence cell transformation and the risk of breast cancer. This research uses genomic and computational approaches to define the pre- and postpregnancy mammary epigenome. An additional objective of dos Santos' laboratory is to use functional genomics to discover novel transcriptional regulators that modulate mammary stem cell self-renewal, lineage specification, and cell transformation. The long-term objective of dos Santos' group is to improve understanding of the mammary epigenome during normal development and use this information to gain insight into new preventive and curative strategies to target breast cancer.

**Christopher Hammell**'s laboratory is interested in understanding gene regulatory processes that give rise to robust phenotypes associated with normal development in animals (specifically, how the timing of developmental processes is controlled) as well as the alterations in these pathways that give rise to diseases such as cancer (as in the alterations in mitogenic pathways in melanoma). Hammell and colleagues approach this elemental problem by using a variety of model organisms and patient-derived cancer cell lines. To directly identify the components that function in controlling normal developmental timing, they use the small nematode *Caenorhabditis elegans*, applying forward and reverse genetic approaches. In contrast to the extreme robustness of cell fate lineage in *C. elegans*, in which specification of developmental programs is hardwired, mutations that alter conserved signaling pathways in melanoma create relatively plastic developmental landscapes that allow these lesions to become aggressive tumors. Notably, the gene regulatory architecture of melanoma cells allows them to acquire resistance to therapeutic agents. Hammell's team is interested in epigenetic mechanisms that contribute to resistance, specifically dramatic changes in gene expression patterns and intracellular signaling pathways. They are performing high-throughput screens to identify cellular factors that allow these rewiring events to occur, with the idea that these components would make ideal therapeutic targets to complement existing clinical strategies.

In **Leemor Joshua-Tor**'s laboratory, researchers study the molecular basis of nucleic acid regulatory processes using the tools of structural biology and biochemistry. One such regulatory process is RNA interference (RNAi), in which a small double-stranded RNA triggers gene silencing. Joshua-Tor and her team offered crucial insight when they solved the crystal structure of the Argonaute protein and identified it as the long-sought Slicer. They then went on to explore the mechanism of the slicing event. The structure of human Argonaute 2 (hAgo2) bound to a microRNA (miRNA) guide allowed Joshua-Tor and her colleagues to understand how mRNA is cleaved during RNAi. Recently, members of the Joshua-Tor laboratory explored the function of a very similar protein, called Argonaute 1, that has no slicing ability, even though it is almost identical in structure to the slicing hAgo2. Using biochemical methods and mutational analysis, they were able to identify key parts of the protein that are required for slicing activity. The laboratory

also studies the generation of PIWI-interacting RNAs (piRNAs), which serve to protect the genome of germ cells. Joshua-Tor's team also helped to determine the structure and function of Zucchini, a key nuclease in the initial generation of piRNAs in fruit flies. In other work, the laboratory is exploring the mechanisms of heterochromatin formation and gene silencing through the study of a protein complex called RNA-induced initiation of transcriptional gene silencing (RITS). Joshua-Tor is also well known for her work on the E1 helicase enzyme, which acts to unwind DNA strands during the DNA replication process.

**Adrian Krainer**'s laboratory studies the mechanisms of RNA splicing, ways in which they go awry in cancer and genetic diseases, and the means by which faulty splicing can be corrected. For example, they study splicing in spinal muscular atrophy (SMA), a neuromuscular disease that is the leading genetic cause of death in infants. In SMA, a gene called *SMN2* is spliced incorrectly, making it only partially functional. The Krainer laboratory found a way to correct this defect using a powerful therapeutic approach. It is possible to stimulate SMN protein production by altering mRNA splicing through the introduction into cells of chemically modified pieces of RNA called antisense oligonucleotides (ASOs). Following extensive work with ASOs in mouse models of SMA, one such molecule, known as nusinersen or Spinraza, was taken to the clinic, and at the end of 2016 it became the first FDA-approved drug to treat SMA, by injection into the fluid surrounding the spinal cord. The Krainer laboratory is currently using antisense technology to develop therapies for other diseases caused by splicing defects, including familial dysautonomia, and to target a cancer-specific alternative splicing event that controls the Warburg effect. In addition, they are applying antisense technology to stabilize mRNAs that are destroyed by a process called nonsense-mediated mRNA decay (NMD), both to learn about the underlying mechanisms and to develop new therapies (e.g., for a nonsense allele in cystic fibrosis). The Krainer laboratory has also worked to shed light on how splicing factors and alternative splicing promote cancer progression in the context of breast, liver, brain, pancreatic, and blood malignancies. Finally, the laboratory continues to study fundamental mechanisms of splicing and its regulation, focusing on the precise recognition of highly diverse intronic and exonic pre-mRNA features by various spliceosome components.

**John Moses**' laboratory specializes in click chemistry, a powerful discovery method that relies on the most robust chemical reactions to synthesize functional molecules. Small molecules are important because nature's machinery, including proteins, enzymes, and receptors, evolved to interact selectively with molecular ligands, similar to how a key fits a lock.

Using click chemistry, they create molecular probes for studying biological systems that may lead to new treatments for deadly diseases, including cancer. For example, they developed a new class of therapeutic DNA binding ligand that interacts selectively with telomeric regions of the genome involved in cellular maintenance. Several of these telomere binding ligands show remarkable selectivity and potency against cancer cells and tumors.

Further developing click chemistry, they recently described a discovery method called diversity-oriented clicking, which exploits a focused group of reliable click chemistry reactions to achieve structural diversity. Their diversity-clicking approach led to the discovery of a new group of antibiotics with activity against multidrug-resistant bacteria, including MRSA.

Through the application of click chemistry, the Moses group is committed to developing chemistry *for* biology at Cold Spring Harbor Laboratory.

The **Andrea Schorn** laboratory investigates how small RNAs identify and silence transposable elements when they become active during development and disease. Transposable elements or "mobile genes," which are closely related to viruses, promote active gene expression in a selfish manner. These elements are usually buried in inactive, condensed DNA by their host to prevent mutagenic

damage. However, both stem cells in the embryo and cancer cells undergo genome-wide reprogramming that reactivates silent transposable elements. The Schorn laboratory is exploring how the host recognizes transposons among thousands of genes and noncoding DNA and specifically restricts transposon mobility.

They found that a highly conserved 18-nt sequence motif is the Achilles' heel of a widespread class of transposable elements that are closely related to retroviruses such as HIV. These retroelements initiate replication at the 18-nt binding site using transfer RNA (tRNA), an essential RNA component of the cell. In turn, cells produce short fragments of tRNAs that they discovered inhibit this class of retroelements. These tRNA fragments are processed from mature tRNAs under yet unknown conditions and potentially protect many cell types in eukaryotes. The Schorn laboratory is investigating under which conditions cells produce this class of small RNAs and assessing their impact on development and pluripotency. tRNA fragments are an ancient link between the "RNA interference" silencing machinery, transposons, and genome stability, with potential roles in transgenerational inheritance and cancer.

**David L. Spector**'s laboratory is focused on characterizing long noncoding RNAs (lncRNAs) that exhibit altered levels of expression in breast cancer progression and during embryonic stem cell differentiation. A major focus of their efforts has been on Malat1 lncRNA, which is one of the most abundant lncRNAs. The Spector laboratory previously identified a novel mechanism of 3′-end processing of this RNA. More recent studies have revealed that increased levels of Malat1 lncRNA impact breast cancer progression and metastasis. Knockout or antisense oligonucleotide knockdown of Malat1 results in the differentiation of mammary tumors and a significant reduction in metastasis. Studies are currently under way to elucidate the mechanism of action of this abundant nuclear-retained lncRNA and to implement innovative therapeutic approaches that can impact its function in vivo. In addition, the Spector laboratory has identified other lncRNAs, termed mammary tumor–associated RNAs, that are up-regulated in breast tumors, and they are currently assessing the function of these lncRNAs using 3D tumor organoids as well as mouse models.

A second area of study in the Spector laboratory is based on their earlier discovery of an increase in random autosomal monoallelic gene expression upon the differentiation of mouse embryonic stem cells to neural progenitor cells. These data support a model in which stochastic gene regulation during differentiation results in monoallelic gene expression, and for some genes, the cell is able to compensate transcriptionally to maintain the required transcriptional output of these genes. Therefore, random monoallelic gene expression exemplifies the stochastic and plastic nature of gene expression in single cells. Ongoing studies are examining the relationship of monoallelic gene expression to lineage commitment.

**Bruce Stillman**'s laboratory studies the process by which DNA is copied within cells before they divide in two. Working with yeast and human cells, Stillman and colleagues have identified many of the cellular proteins that function at the DNA replication fork during the S phase, the portion of the cell-division cycle when DNA synthesis occurs. Among these proteins are those that facilitate the assembly of chromatin, the protein–DNA complexes that form the chromosomes. Current research focuses on the mechanism that initiates the entire process of DNA replication in eukaryotic cells. At the heart of this mechanism is a protein that binds to "start" sites on the chromosomes, called the origin recognition complex (ORC). The Stillman laboratory is part of an ongoing collaboration that determined the cryo-EM structure of ORC proteins in complex with a group of proteins, called a helicase, that unwind DNA during replication. These images offer molecular insights into how the helicase is loaded onto DNA. Stillman's research also focuses on the process by which duplicated chromosomes are segregated during mitosis. They found ORC at centrosomes and centromeres, structures that orchestrate chromosome separation during mitosis. At centromeres, ORC subunits monitor the attachment of duplicated chromosomes to the mitotic

spindle that pulls the chromosomes apart when they are correctly aligned. Stillman's team has discovered that mutations in the Orc1 protein alter the ability of ORC to regulate both DNA replication and centrosome duplication. These mutations have been linked to Meier–Gorlin syndrome, a condition that results in people with extreme dwarfism and small brain size but normal intelligence.

Research in **Chris Vakoc**'s laboratory investigates how transcription factors and chromatin regulators cooperate to control gene expression and maintain the cancer cell state. This work makes extensive use of genetic screens to reveal cancer-specific functions for transcriptional regulators, as well as genomic and biochemical approaches to identify molecular mechanisms. One theme that has emerged from their efforts is that blood cancers are often vulnerable to targeting transcriptional coactivators, such as BRD4 and the SWI/SNF chromatin remodeling complex. Vakoc's team demonstrated that chemical inhibition of BRD4 exhibits therapeutic effects in mouse models of leukemia, a finding that has motivated ongoing clinical trials in human leukemia patients. The Vakoc laboratory has also developed a CRISPR-Cas9 screening approach that can reveal individual protein domains that sustain cancer cells. Their laboratory is now deploying this technology in a diverse array of human cancers to reveal therapeutic opportunities and basic mechanisms of cancer gene control.

The research of the **Lingbo Zhang** laboratory focuses on decoding the role of metabolites and their genetic effectors in the tumor microenvironment of hematologic malignancies. They utilize a combination of functional genomics, metabolomics, circuit mapping, and optogenetics approaches to systematically uncover critical dietary and neuronal activities that regulate stem and progenitor cell development, and to identify key drug targets for treating hematologic malignancies.

Together, the Zhang laboratory recently uncovered a series of critical metabolites in the tumor microenvironment, including acetylcholine and pyridoxal, and their genetic effectors as novel regulators of hematologic malignancies. They identified cholinergic receptor muscarinic 4 (CHRM4) as a novel regulator of early erythroid progenitor self-renewal and a therapeutic target for myelodysplastic syndrome (MDS). Their research uncovered the hematopoietic arc as a novel neuronal activity–based regulatory mechanism of hematopoietic stem and progenitor cell self-renewal. They also identified the vitamin B6 pathway as a nutritional and metabolic dependency in acute myeloid leukemia (AML) that coordinates nucleotides and putrescine metabolism specifically required for leukemia maintenance. Their research uncovered the vitamin B6 pathway as a pharmacologically actionable target for the treatment of leukemia with minimal myelosuppression effect. Through collaborations with medicinal chemists at their spin-off company, they are building pharmacological approaches at the industry level to target these novel regulators and translating their discoveries into first-class therapeutics. Their findings will help treat devastating hematologic malignancies, including refractory anemia, myelodysplasia, and leukemia.

# UNDERSTANDING THE EPIGENETIC REGULATION OF NORMAL AND MALIGNANT MAMMARY GLAND DEVELOPMENT

**C. dos Santos** D. Anandan M. Callaway D. Chatterjee M. Ciccone L. Comfort A.L. Domney S. Henry S. Lewis Y. Li M. Lozada L. Téllez Pérez M. Trousdell S.T. Yang

## Defining a Role for BPTF Inhibition in Blocking Breast Cancer Progression

M. Ciccone, M. Trousdell, D. Chatterjee, D Anandan, S. Lewis, S.T. Yang [in collaboration with A. Siepel and J. Tollkuhn, CSHL]

The goal of this project is to define the effects of bromodomain protein transcription factor (BPTF) inhibition on controlling the development and progression of mammary tumors. To date, we have established distinct roles for the epigenetic factor utilizing BPTF in a series of murine models of mammary tumorigenesis. We have monitored mammary tumor development in a cohort of Brca1 knockout (KO) female mice (basal/triple-negative breast cancer [TNBC] breast cancer model), and observed mammary tumors developed within 2–9 months after tumor-assisted macrophage (TAM) administration in all wild-type (wt) mice ($n = 12$). In marked contrast, Bptf KO mice did not develop mammary tumors during the 1-year observation period, suggesting that Bptf depletion significantly blocked mammary tumorigenesis in Brca1-deficient female mice.

We also found that populations of basal-like lineages gained the expression of estrogen receptor alpha (ERα) after Bptf deletion. Treatment of female mice with tamoxifen, an ERα antagonist, exclusively inhibited the growth of orthotopically transplanted Bptf KO mammary tumors. These results suggest that Bptf inhibition could represent a novel strategy to convert hormone-negative breast cancer subtypes into those that express ERα, and therefore increase their response to hormone-targeting therapies. This hypothesis is currently being addressed in classical models of human TNBC, including cell lines and 3D organoid cultures. By mining the TCGA clinical breast cancer data, we have found that there is a relationship between ESR1 and BPTF expression. Using estrogen supplementation of organoids in a collagen I protrusion assay, we have discovered that estrogen is not sufficient to induce branching and activation of Bptf KO epithelial cells, but this can be reversed through the addition of transforming growth factor beta (TGF-β) to culture.

## Investigating the Relationship between the Systemic Response to Urinary Tract Infection and Brca1 Deficiency on Tumor Progression in the Mammary Gland

S. Lewis

Following up on our finding that urinary tract infection (UTI) drives cellular and tissue-level changes to the mammary gland through systemic response factors such as TIMP1 and CSF3, we set out to understand the role of the systemic response to infection in our Brca1 breast cancer mouse model. Just as in wt UTI-bearing mice, we have found that after 2 weeks of infection, collagen accumulates in the mammary tissue. By characterizing the epithelial cells in the gland of Brca1 KO and Brca1-UTI mice, we have found that this collagen deposition triggered by the infection response leads to an increase in YAP signaling in epithelial cells as well as the emergence of a basal–luminal population of epithelial cells. Based on the unique gene expression of these cells, which we speculate are tumor-initiating cells, we sorted these cells from a viably preserved Brca1 tumor sample and transplanted them into immunocompromised recipients. We discovered that, indeed, these cells have tumor-initiating potential and their growth outpaces that of control cells from the same sample.

Following up on our finding of activated fibroblasts in UTI-bearing wt mice, we are characterizing the phenotypic plasticity of fibroblasts in Brca1-UTI and plan to test their functional potential by leveraging in vitro fibroblast-organoid co-culture models followed by co-injection.

## Impact of Female Hormones on Normal and Malignant Male Mouse Breast Development

L. Téllez Pérez

The mammary gland is a sexually dimorphic organ that is only fully developed and functional in females and is present as a rudimentary structure in males. Nevertheless, the male mammary gland is affected by specific disorders, such as gynecomastia or male breast cancer. These conditions remain largely understudied and treatments for them are inferred from studies of the female mammary gland. It is known that estrogen is essential for the normal development of the female mammary gland, and exposure to exogenous estrogens is related to an increased risk of developing breast cancer. However, the effects of estrogen exposure in the normal male mammary gland (as observed, in practice, during male-to-female gender transition), and the mechanisms by which it could impact breast cancer onset in these populations are yet to be assessed. Beyond the translational implications, the study of the male mammary gland presents an understudied framework to better understand mammary gland development and opens up questions about the mechanisms by which an adult organ is able to fully develop later in life when exposed to the right hormonal milieu.

We have developed both organoid and in vivo models to understand the differences between the male and female mammary gland and their responses to estrogen. Male mammary gland–derived organoids can be derived similarly to female organoids, and they show both the characteristic basal ($Krt5^+$) and luminal ($Krt8^+$) mammary epithelial cell populations. When looking into organoid-derived single-cell RNA sequencing (scRNA-seq) data sets, we observe that all mammary epithelial populations are recapitulated in male-derived samples; however, the levels of expression of canonical markers are generally lower than in females. We have also observed male-characteristic cell populations beyond the described populations present in female, with specific markers and characteristics that resemble a less differentiated mammary gland state seen in prepubescent females. Regarding estrogen response, both male and female mammary gland organoids show a response in their transcriptome, which is different depending on the sex. With our in vivo models, we are able to explore the effects of estrogen and castration long-term in male mice. Adult males exposed to a peak in estrogen show terminal end bud structures on their mammary glands, which are seen in pubescent females undergoing pubertal mammary gland development. We have recently developed a new scRNA-seq data set from mammary glands from wt, castrated, and castrated + estrogen-treated males, to understand the effects of estrogen at different time points throughout the whole organ.

We have also started to develop novel models to study male breast cancer and the effect of female sex hormones in genetically predisposed male models. A pilot study of $BRCA1^{-/-}$ male mice shows that 7% of these animals develop breast tumors. We have ongoing studies with male $BRCA1^{-/-}$ castrated and estrogen-treated animals. These studies will help us better understand the risk of developing breast cancer in males undergoing hormone treatment and the characteristics of the tumors that arise in this scenario.

## Determining the Immunological Basis of Pregnancy-Induced Breast Cancer Protection

D. Anandan, M. Trousdell, M. Lozada, S.T. Yang [in collaboration with J. Yeh, CSHL]

Women who undergo a full-term pregnancy before the age of 25 benefit from lifelong protection against breast cancer. The goal of this project is to define the pregnancy-induced changes to mammary immunity that enable long-term mammary oncoprotection. We have published that there is a stable mammary-specific expansion of natural killer (NK)-like γδ-T cells in the post-pregnancy mammary gland (MG) that is linked to the post-pregnancy prevention of mammary oncogenesis in two mouse models: cMYC overexpression (CAGMYC) and Brca1 KO. scRNA-seq of pre- and post-pregnancy mammary tissue from Brca1 KO mice also reveals an enrichment of B cells—specifically antibody-secreting memory B cells and plasma cells—in post-pregnancy Brca1 KO mammary tissue.

We are now validating our scRNA-seq findings by performing flow cytometry for B-cell populations in pre- and post-pregnancy mammary tissue. To understand whether pregnancy shapes the B-cell receptor (BCR) repertoire and facilitates the acquisition of oncoprotective memory, we have generated BCR-seq data sets from pre- and post-pregnancy wild-type murine mammary tissue. We have also extracted the sequences of the most abundant clonotypes in the post-pregnancy MG and have used them to synthesize monoclonal antibodies (mAbs) in collaboration with the CSHL Antibody & Phage Display Core Facility. We are now working to validate the mAbs we have generated via tissue staining analyses. Future experiments will evaluate the antibody-secreting potential of pre- and post-pregnancy mammary B cells and test whether post-pregnancy mAbs may be used as an immunization strategy to prevent tumorigenesis in Brca1 KO mice.

Considering the role of the microbiome in mediating immunity and epithelial epigenetic regulation in the gut, we are also exploring how the influx of breast milk microbes from the maternal gut to the MG may play a role in oncoprotection. To comprehensively profile pregnancy-induced changes to the mammary microbiome, we performed whole-genome sequencing of MGs from pre- and post-pregnancy mice. Our analyses reveal an enrichment of the core breast milk microbes *Lactobacillus* and *Bifidobacteria* in the post-pregnancy MG and suggest a role for *Lactobacillus* in epigenetic regulation in the post-pregnancy MG. We also performed flow cytometry on mammary tissue from mice that lack *Lactobacillus* and *Bifidobacteria* (VAF Elite mice), and we failed to observe the post-pregnancy expansion of NK-like γδ-T cells, thus suggesting a role for the microbiome in mediating oncoprotective features of the post-pregnancy MG. To further understand the role of the microbiome in the post-pregnancy MG, we are now using VAF Elite and non-Elite mice to test for evidence of pregnancy-induced "epigenetic memory" and further profile pre- and post-pregnancy mammary immune composition. We have also developed two mouse models of microbial modulation in non-Elite mice: (1) antibiotic treatment during pregnancy and lactation and (2) probiotic treatment of nulliparous mice. The first model will identify how disruptions of the microbiome specifically during pregnancy/lactation affect the mammary microbiome and epithelial and immune cells; the second model will elucidate how feeding probiotics rich in *Lactobacillus* and *Bifidobacteria* to nulliparous mice can influence the mammary microbiome and epithelial and immune cells. Future experiments implementing these models in CAGMYC and Brca1 KO mice will reveal how microbial modulation may affect mammary oncogenesis.

## Investigating the Tissue Architecture Plasticity of the Mammary Gland through an Application of Network Theory and Quantification to the Ductal Tree

S. Lewis, S. Henry, L. Téllez-Pérez [in collaboration with S. Navlakha, CSHL]

Despite the long-standing knowledge of the process of ductal genesis in the mammary gland, the dynamics of this ductal proliferation and regression in development and during and after pregnancy require further quantification. By applying a network theory approach, we are manually tracing images of whole mammary tissue from whole mount histology and converting them into networks with features that can be rigorously quantified using NEFI and NetworkX. We are quantifying the number of nodes (alveoli) and edges (ducts), and their distribution pattern throughout the whole organ. We are quantitatively assessing the branching dynamics under different physiological conditions such as before and after puberty to better appreciate the developmental dynamics.

## Elucidating Key Physical and Molecular Drivers of Gestational Breast Cancer

M. Callaway

This project aims to understand key drivers of gestational breast cancer or cancer that arises during or shortly after pregnancy. Because of the high recurrence and poor prognosis of gestational breast cancer, we seek to characterize distinct cellular and tissue alterations occurring during pregnancy that may prime breast tissue to support disease progression and recurrence post-pregnancy.

For the past year, we have been developing approaches to understand how changes in the mammary microenvironment support the development of gestational breast cancers. In doing so, we have optimized

a series of multiplexed tissue staining strategies. These approaches maintain the spatial distribution of immune and stroma cells in relation to the developing cancer while allowing for quantitative analysis of cellular abundances and states. In addition, given the rarity of gestational breast cancer cases, such an approach would enable the investigation of tumor and microenvironmental differences in prospectively isolated tissue specimens from patients. Specifically, we have comprehensively characterized the immune composition of tumors formed in vivo in the presence of pregnancy hormones, revealing distinct increases in intratumoral $CD206^+$ macrophages and tumor-associated fibroblasts (CAFs). This suggests that the onset of pregnancy influences an array of cell types that collectively may regulate tumor growth and progression. In fact, further analysis of the tumor stroma indicated increased extracellular matrix (ECM) deposition in tumors that grow during pregnancy, specifically laminin and collagen fibers—known changes that increase breast cancer metastatic potential.

Given that pregnancy involves a collection of signals that change mammary tissue and whole-body physiology, we next asked whether mammary tumors would alter the pool of circulatory factors in pregnant female mice. In doing so, we collected blood samples from healthy (no tumor) and tumor-bearing nonpregnant and pregnant mice and quantified the amount of plasma circulatory factors. Our approach identified uniquely elevated factors in the plasma of tumor-bearing pregnant female mice, including CXCL12, CXCL13, G-CSF, and CD14—factors known to influence metastasis and immune suppression in cancer. We are interrogating the role of these factors in driving differential immune cell recruitment and tumor cell motility during pregnancy while establishing protocols to assess the prognostic value of these candidates in noninvasive screening modalities for gestational breast cancer diagnosis.

## Defining the Impact of Tissue Aging on Mammary Gland Development

S. Henry

We have previously defined mechanisms by which pregnancy provides protection against mammary tumor development, but the protective effects of pregnancy are age-dependent, as women who undergo a first pregnancy after the age of 35 have a 30%–40% increase in breast cancer risk. Previously, we have found that mammary organoids derived from aged mice are observed to have insignificant increase in size and formation of branching structures compared to untreated aged mucoepidermoid carcinomas (MECs) and CSN2 protein expression, indicating that aged MECs are less responsive to pregnancy hormones.

To further explore this altered response to pregnancy hormones with age, we have utilized a mouse model in which we implant 21-day slow-release estrogen progesterone (EP) pellets to simulate pregnancy. Mammary glands were collected across early- (D4, D6), mid- (D12) and end- (D21) pregnancy time points and preserved in histology and whole mounts. Utilizing these complementary methods, we were able to gain insight into differences in branching/ductal structures across various pregnancy time points with age. In summary, visually, during early–mid pregnancy time points there is an observed delay in mammary gland development with age. By D21 both young and aged look fairly similar, indicating that deficiency in development occurs at earlier stages of mammary gland development with age.

These findings on delay in response to pregnancy were further validated using our scRNA-seq data sets derived from organoid cultures, where down-regulation in hormonal associated pathways in the pregnancy hormone treated aged MECs when compared to young MECs. Specifically, in the aged setting it was seen that luminal hormone sensing cells had a decrease in estrogen receptor expression, suggesting that estrogen, a key hormone in pregnancy-induced mammary gland development, function may be dysregulated with age. Current work is focused on assessing aged MECs' response to estrogen alone, examining receptor expression through immunofluorescence, and exploring changes in estrogen motif accessibility and interaction with co-factors as potential mechanisms that determine this lack of response to pregnancy hormones with age.

## Investigating the Cellular Origin of Basal-Like Breast Cancers at Single-Cell Resolution

D. Chatterjee

To date, we have aimed to define the role of the NURF complex subunit BPTF in blocking tumor onset by

delaying cancer progression in several models of mammary-gland tumorigenesis. Monitoring the development of mammary tumor in a cohort of Brca1 KO female mice (K5A model), we noted the development of mammary tumors within 2–9 months after Brca1 KO induction in all mice with wt Bptf activity ($n = 18$), while the double-KO (Brca1 KO + Bptf KO K5BA model) mice did not develop mammary tumors ($n = 9$) within the 1-year observation period. This observation therefore suggested that Bptf depletion may block mammary-gland tumorigenesis in the K5A female mice. At single-cell resolution, however, all the expected cells of the K5A mammary gland lineage were detected in the K5BA mice, including the putative tumor-initiating cells (TICs)—an abnormal Brca1 loss–induced, mixed-lineage population. Compared with the K5A TICs, the K5BA TIC population displayed changes consistent with those reported in studies involving other models (see other projects in the laboratory)—mainly, the increase in Esr1 expression as well as an increase in the hallmark signaling downstream of estrogen receptor activation. Trajectory inference analysis of these TICs further detected a transcriptomic deviation in the fates of Brca1 KO cells with Bptf KO and Bptf wt activity. Modules (303 genes) associated with the Bptf wt lineage were down-regulated in the Bptf KO TICs and the high expression of their human orthologs was found to be associated with poor prognosis in patients with $ER^-$ breast cancer. Similarly, high expression of differentially up-regulated genes active in the Bptf KO lineage were also associated with better relapse-free survival in patients with $ER^+$ breast cancer. These results, along with those observed in other models, indicate that Bptf inhibition could represent a novel strategy to convert hormone-negative breast cancer subtypes into ERα-expressing subtypes that are more responsive to hormone-targeted therapies.

To further describe the fate of Bptf KO TICs and other epithelial cells in the K5BA mice where tumors were ultimately not detected, we further explored the likely role of the immune population in mediating the clearance of abnormal cells. Using single-cell transcriptomics, we detected a relative increase in communication (Bptf KO vs. wt comparison) received by $CD8^+$ T cells, as well as a decrease in that associated with the periductal fibroblasts, neutrophils, and the plasma B cells. In contrast, plasmacytoid dendritic cells (pDCs) were found to differentially increase communication with those T cells and the periductal fibroblasts and the mixed-lineage epithelial-cell population in the Bptf KO samples. Using immunofluorescence, the presence of Siglec-$H^+$ $Tlr9^+$ pDCs was validated both within the lymph node as well as adjacent to the epithelial ducts. Given that pDCs respond to extracellular unmethylated CpG DNA, we will next validate these differential communications as well as investigate whether hypomethylated mitoDNA is released as a result of increased Bptf KO-induced cellular injury or because of changes to the chromatin in epithelial cells resulting directly from Bptf KO.

## To Define the Role of Ezh2 in Pregnancy-Induced Epigenomic Reprogramming and cMYC Overexpression

Y. Li

Women who undergo a full-term pregnancy before the age of 25 benefit from lifelong protection against breast cancer. Our goal is to define genes/programs that play an essential role in post-pregnancy mammary oncogenesis inhibition. We have previously shown that inhibition of p300, known to control the expression of senescence genes, increased abnormal branching of post-pregnancy organoid cultures in response to cMYC overexpression, thus supporting that specific perturbations to post-pregnancy MECs can revert their ability to respond to cMYC overexpression and engage in malignant transformation. More recently, we found that small-molecule inhibition (UNC1999, 5 μM, nonlethal dose) of the histone methyltransferase Ezh2 specifically blocked the response of post-pregnancy organoids to pregnancy hormones, suggesting a role for Ezh2 in controlling cellular and molecular responses of organoids previously exposed to pregnancy hormones. We hypothesize that Ezh2 may represent an epigenetic switch in MECs that controls the dynamics between the nonsenescence, proliferating state of MECs (during pregnancy) to a presenescence state, which is further augmented in response to oncogenic stress (in response to cMYC overexpression). Therefore, we will investigate the cellular and molecular perturbations brought to MECs during Ezh2 inhibition and their contributions to mammary oncogenesis.

## The Effects of Gestational Breast Cancer on the Fetoplacental Unit in a Mouse Model

L. Comfort

Cancer during pregnancy has been associated with fetal growth restriction, especially after exposure to antepartum chemotherapy. Yet, there is little information regarding the risks of breast cancer itself on the pregnancy or information about the placental effect of breast cancer, despite it being one of the most commonly diagnosed reproductive-age cancers. Breast cancer cells themselves do not commonly pass to the placenta or fetus, but may have an effect on the placenta and developing fetus through indirect mechanisms. Cancer progression involves multiple signaling events, including immunosuppressive cytokines. Cytokines have many effects on prenatal health, with increased immune response associated with risk of miscarriage, preeclampsia, gestational diabetes, and preterm labor. In fact, elevated cytokines IL-6 and IL-1β in the amniotic fluid as well as placental inflammation are predictors of brain injury in premature infants. We hypothesize that there is similarly increased pro-inflammatory cytokine production in tumor-affected mice, with subsequent increased cytokine production in mouse pups, thus contributing to gestational breast cancer associated fetal growth restriction.

## Evaluating the Effects of Standard Fertility Medications on Normal Breast Tissue and Breast Cancer Development

A.L. Domney

Many factors that modify breast cancer risk are associated with estrogen exposure, including age of menarche, age at first full-term birth, parity, obesity, and postmenopausal hormone therapy. Over the past decades, numerous studies investigating the association between fertility drugs and breast cancer risk have yielded conflicting or inconclusive results. Heightened estrogen levels during fertility treatment raise concerns in specific subsets of patients, including those with a history of certain cancers, fertility preservation due to active cancer, history of blood clots, BRCA status, and more. In these types of patients, an additional medication called letrozole is also used. Letrozole is an aromatase inhibitor that has been shown to reduce estrogen exposure during ovarian stimulation when combined with gonadotropins. With this rapid rise in gonadal hormone levels, the direct and indirect effects of these agents on estrogen-sensitive tissues, especially breast tissue, and consequently on an individual's risk of having breast cancer, remain unclear and of significant clinical importance. Our goal is to utilize breast organoid cultures as a unique strategy to study the effects of fertility drug treatment on normal breast development and potential understanding of malignancy formation and progression.

### PUBLICATIONS

Callaway MK, dos Santos CO. 2023. Gestational breast cancer—a review of outcomes, pathophysiology, and model systems. *J Mammary Gland Biol Neoplasia* **28:** 16. doi:10.1007/s10911-023-09546-w

Ortiz JR, Lewis SM, Ciccone MF, Chatterjee D, Henry S, Siepel A, dos Santos CO. 2023. Single-cell transcription mapping of murine and human mammary organoids responses to female hormones. bioRxiv doi:10.1101/2023.09.28.559971

*In Press*

Henry S, Lewis SM, Cyrill SL, Callaway MK, Chatterjee D, Somasundara AVH, Jones G, He X, Caligiuri G, Ciccone MF, et al. Host response during unresolved urinary tract infection alters mammary tissue homeostasis through collagen deposition and TIMP1. bioRxiv doi:10.1101/2024.02.05.578977

# CONTROL OF CELL FATE SPECIFICATION DURING ANIMAL DEVELOPMENT

**C.M. Hammell** O. Huiwi P. Wu
B. Kinney H. Zhang
J. Wang C. Zhao

Dynamic changes in gene expression are a hallmark of developmental biology, in which cell fates are determined in specific orders to produce functional tissues and organs. Our laboratory studies how genes are turned on and off during development. We are specifically focused on understanding how the precise timing and sequence of developmental events are organized and how the correct expression levels of essential regulatory proteins are established. One-half of the laboratory studies how oscillatory gene expression during the development of *Caenorhabditis elegans* larvae contributes to the establishment and stability of postembryonic cell fate specification. The second half of the laboratory studies how protein translation is regulated by controlling—and specifically how two conserved families of RNA-binding proteins, UBAP2/2L and G3BP1/2, control—the subcellular localization of PTEN and AKT/PI3K downstream signaling.

## Development of the LlamaTag System to Visualize Transcription Factor Dynamics during *C. elegans* Development

J. Wang, H. Zhang, C. Zhao

Our laboratory's primary focus is directly monitoring gene expression dynamics during development. With this information, we aim to identify and understand how temporal information is generated and used to define sequential cell fates in the cells and tissues of a developing animal. We have pioneered the use of destabilized green fluorescent protein (GFP) and the MS2 system to quantify the dynamics of the *C. elegans* heterochronic pathway. This has enabled us to determine how gene dosage within this pathway is controlled and what mechanisms and strategies are used to determine distinct cell fates. This led to the discovery that the transcriptional activation of key microRNA genes that posttranscriptionally regulate gene expression in the developmental timing pathway are expressed in a highly pulsatile manner in which short, periodic bursts of mRNA production are followed by long periods of transcriptional inactivity. In this fashion, microRNAs sequentially repress a suite of transcription factors and RNA-binding proteins needed to execute cell fate division programs specific to individual stages.

For many years, our laboratory has attempted to directly monitor the dynamic expression of the targets of the heterochronic microRNAs through conventional means, including immunofluorescence and the use of GFP tags. Although conventional immunofluorescent approaches indicate that most heterochronic microRNA targets are easily detectable and dynamically expressed, quantifying GFP fusions of endogenous genes for microRNA targets has proven extremely difficult. A common feature of these experiments is that the endogenously tagged alleles of these genes generate fusion proteins that are very dim. This contrasts with their ability to be visualized and measured using fixed samples and antibody staining.

This discrepancy may reflect an often overlooked feature of fluorescent tags universally used to monitor protein expression in vivo. GFP is a typical tag derived from a fluorescent protein from jellyfish. Although this tag is easily transferable to any gene of interest when fused in-frame with an endogenous gene, the polypeptide generated by the GFP gene must go through a series of isomerization events that enable the protein to fluoresce when physically excited by the correct wavelength of light. This takes an often overlooked amount of time, which, at best, is calculated to be 35 minutes to an hour at the temperature at which *C. elegans* is cultivated. Therefore, proteins with short half-lives or, more specifically, proteins with half-lives shorter than 30 minutes, can be generated at levels

that should be easily visualized with antibody-based approaches but are destroyed before the average maturation time of GFP in vivo.

We used single-chain antibodies developed in vitro that recognize GFP to overcome this limitation. Specifically, we took advantage of variants of camelid antibodies (nanobodies) directed against GFP that stimulate GFP expression by disaggregating AvGFP. In our strategy, we reversed that conventional approach by encoding these single-chain antibodies (LlamaTag) as the epitope on our gene of interest using CRISPR genomic editing. We then ubiquitously expressed a variant of GFP known to be stabilized by the nanobody. As a test case, we compared HBL-1::GFP expression levels to those measured with the HBL-1::LlamaTag plus the ubiquitously expressed AvGFP. As shown in Figure 1, this leads to a greater than fivefold increase in relative expression. This suggests that the half-life of HBL-1 is relatively short, and the lack of visibility of HBL-1::GFP transgenes in *C. elegans* embryos is likely due to the rapid turnover of HBL-1. We are now adapting this system to visualize other heterochronic temporal identity genes.

## LIN-42 Functions as a Scaffold to Coordinate the Activity of Multiple Classes of Transcription Factors

P. Wu, B. Kinney, J. Wang

Mutations in *lin-42* result in precocious developmental phenotypes. Our laboratory's central, long-standing project is to determine how LIN-42 controls periodic gene expression during *C. elegans* larval development and to understand how LIN-42 functions directly to dampen transcription. In addition to impacting temporal cell fate specification, *lin-42*(*lf*) mutants also cause animals to exhibit molting phenotypes where animals fail to molt periodically or molting patterns are inappropriately reiterated during adulthood. To understand how LIN-42 coordinates these cellular and behavioral features, in the last two years we have focused on the physical association between LIN-42 and various transcription factors. Prior experiments demonstrated that one large class of transcription factors bound by LIN-42 is the expanded nuclear hormone receptor (NHR) class. Specifically, LIN-42 interacts with 66 of the 283 *C. elegans* NHRs.

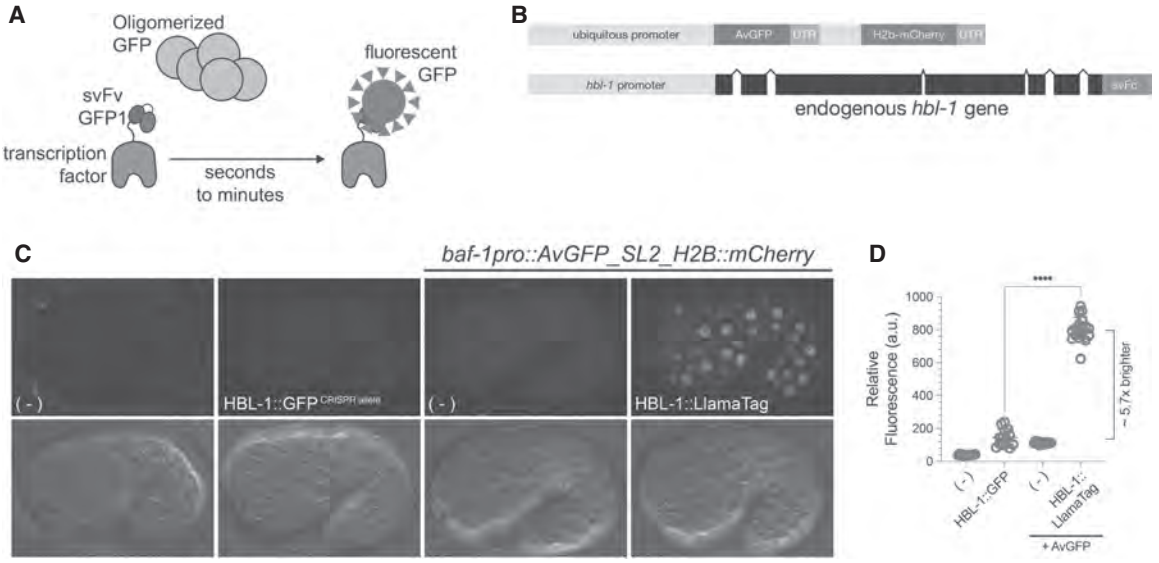



**Figure 1.** LlamaTags and the visualization of short half-life proteins. (*A*) Depiction of the LlamaTag strategy in which a target protein of interest harbors an encoded single-chain antibody (svFV-GFP1) that can recognize oligomerized and aggregated green fluorescent protein (GFP) molecules. Once the svFC-GFP epitope (LlamaTag) recognizes the aggregated, nonfluorescent GFP molecule, it becomes soluble and fluorescent. (*B*) An illustration of the transgenes used to establish the LlamaTag system in *C. elega*ns. (*C*) Tagging of the *hbl-1* gene with the LlamaTag epitope increases fluorescent signal compared with conventional GFP-tagging. (*D*) Quantification of signal intensities from multiple animals of the indicated genotypes.

We have recently found several additional transcription factors (non-NHRs) that interact with LIN-42. In both instances, we hypothesize that these interactions directly modulate the function of these DNA-binding proteins.

A significant focus in the last year has been mapping the domains of LIN-42 that mediate transcription factor interactions and then generating mutations that prevent these interactions. We initially focused on defining the interaction surfaces of LIN-42 that interact with NHR-85. Using yeast two-hybrid strategies, we have mapped a nine-amino acid region of LIN-42 that interacts with NHR-85 (Fig. 2). This region of LIN-42 harbors a conserved α-helical element that binds to a conserved pocket within NHR-85. Further experiments indicate that this region is essential for binding to other nuclear hormone receptors. This domain's deletion leads to mild heterochronic phenotypes and a failure to cease molting after the larval-to-adult temporal transition.

As noted above, we tested interactions between LIN-42 and other classes of transcription factors and identified a new, conserved MYRF-1 transcription factor as a robust LIN-42 binder. This transcription factor is expressed cyclically in all somatic tissues once per larval stage. Mutations of the *myrf-1* gene cause larval arrest phenotypes, indicating an essential role in development. We mapped the regions of LIN-42 that bind MYRF-1 to a conserved carboxyl-terminal motif that likely forms a β-sheet that intercalates into the β-sheet structures of the MYRF-1 trimer. We are deleting this region from the endogenous *lin-42* gene

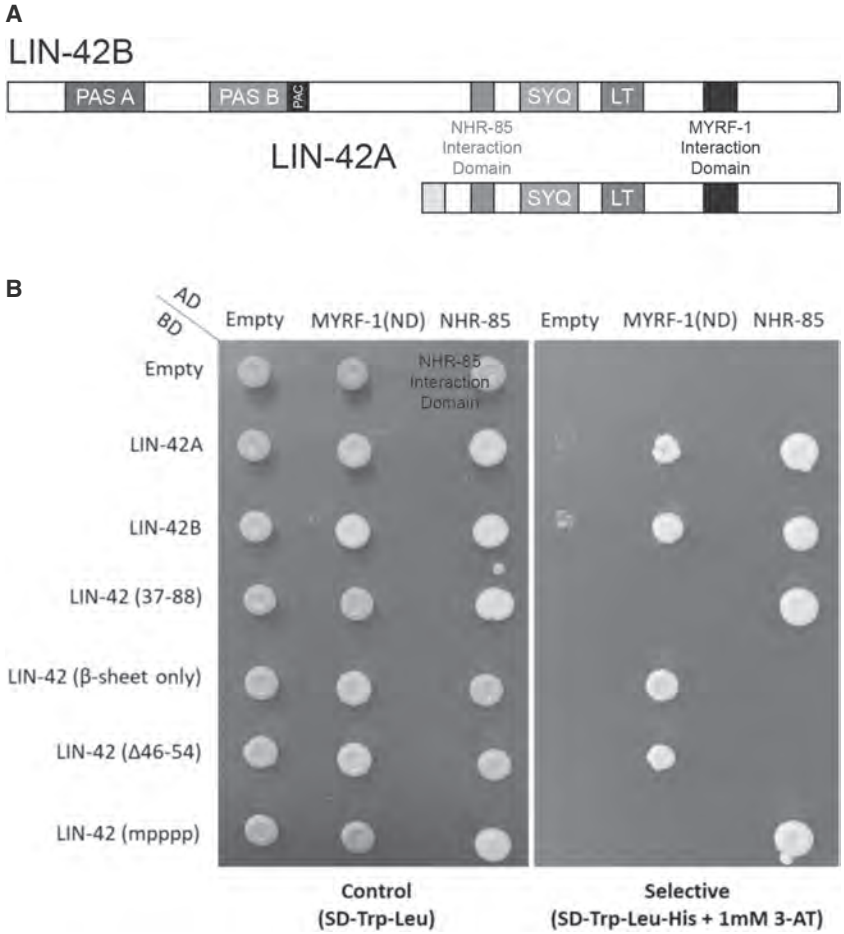



**Figure 2.** LIN-42 interacts with several classes of transcription factors to coordinate gene expression. (*A*) A schematic of the major isoforms of LIN-42 with defined interaction domains highlighted. (*B*) A summary of the yeast two-hybrid used to determine LIN-42 interactions with NHR-85, a conserved nuclear hormone receptor, and MYRF-1, a ubiquitously expressed trimeric transcription factor.

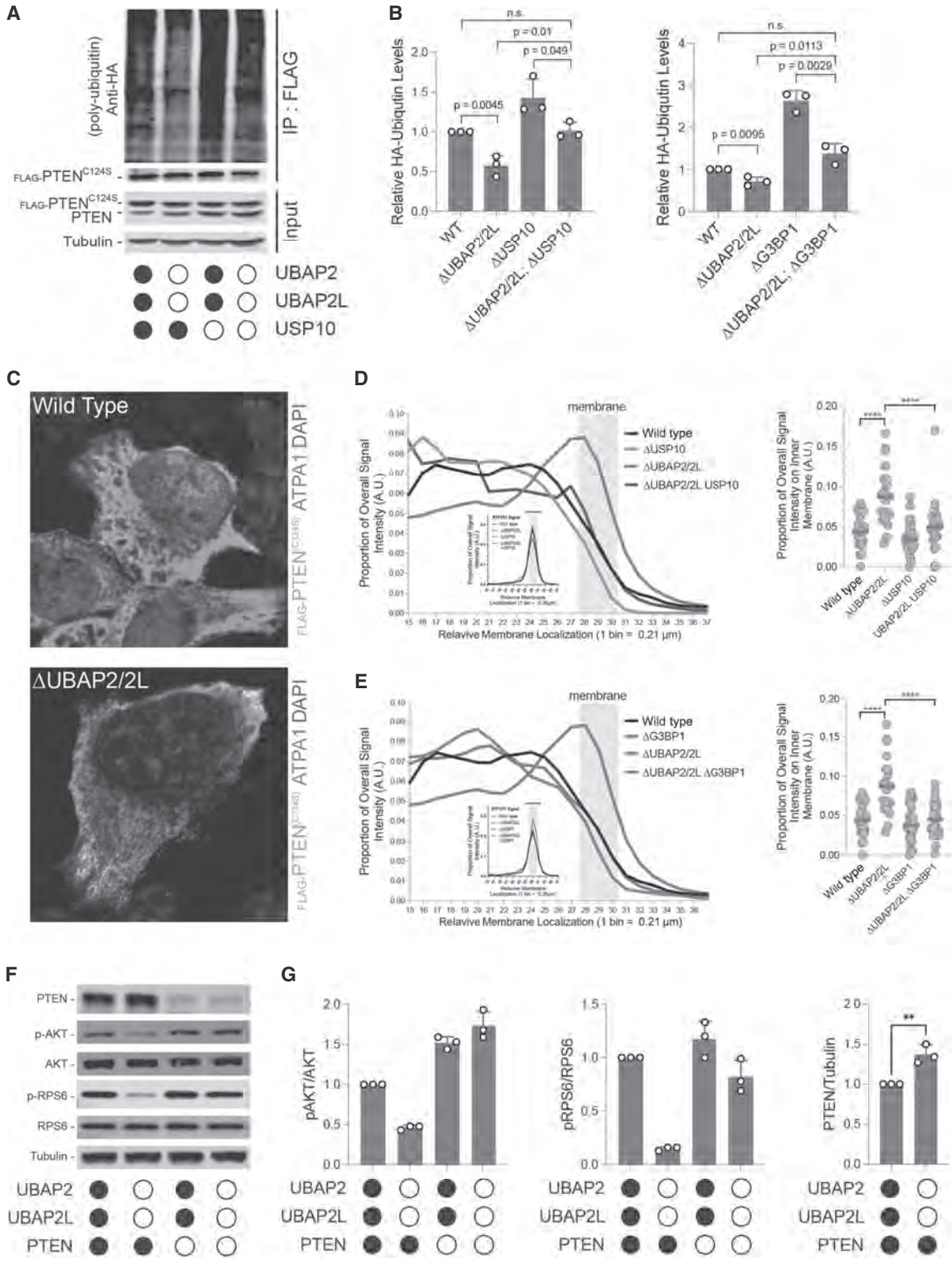



**Figure 3.** UBAP2/2L modulates PTEN localization and activity. (*A*) Deletion of USP10 in ΔUBAP2/2L cells restores normal ubiquitination of PTEN. (*B*) Quantification of PTEN ubiquitin levels in various genetic backgrounds. (*C*) Localization of PTEN in wild-type and ΔUBAP2/2L cells. (*D,E*) The quantification of PTEN membrane localization in ΔUBAP2/2L cells and deleted cells for either G3BP1 or USP10 indicate that the deletion of USP10 or G3BP1 suppresses the constitutive localization of PTEN in ΔUBAP2/2L cells. ATP1A1 was used as a membrane marker in these experiments. (*F,G*) Deletion of PTEN in ΔUBAP2/2L cells restores normal cellular signaling downstream of mTORC1.

to determine whether these interactions are essential for LIN-42 and MYRF-1 regulation and to identify the transcriptional targets of both proteins.

## UBAP2 and UBAP2L Regulate PTEN Localization and Activity

O. Huiwi, J. Wang

Previous experiments from our laboratory indicate that UBAP2 and UBAP2L function redundantly to control mTORC1 signaling. We have used CRISPR to probe the genetics of this circuit. We had previously used proteomics to identify UBAP2- and UBAP2L-interacting proteins. We identified two proteins that functionally interact with both proteins. The first component is a highly conserved de-ubiquitinating protein called USP10. The second protein, G3BP1, is a direct USP10 binding protein that appears to be a cofactor for the USP10 deubiquitinase. Deleting either protein suppresses the ΔUBAP2/2L mTORC1 signaling defects, restoring downstream mTORC1 signaling and translational activation. USP10 is known to modulate PTEN localization and activity by deubiquitinating the C2 domain of PTEN. Deubiquitination of the C2 domain of PTEN restores PTEN localization to the membrane, where it can attenuate mitogenic intracellular signaling. We hypothesized that UBAP2/2L function normally limits this activity, and in their absence, PTEN becomes hypo-ubiquitinated and overactive. Consistent with these results, we find that PTEN is hypo-ubiquitinated in UBAP2/2L cells and constitutively localized to the membrane fraction of the cells by immunofluorescence (Fig. 3). We also find that deleting either USP10 or G3BP1 alleviates the constitutive localization of PTEN to the membrane and increases the level of PTEN ubiquitination (Fig. 3). This model suggests that PTEN deletion should also suppress UBAP2/2L deletion phenotypes. As shown in Figure 3, this is the case. This model may explain why UBAP2/2L overexpression in cancer cells is associated with more aggressive tumors and worse patient survival.

### PUBLICATIONS

Kinney B, Sahu S, Stec N, Hills-Muckey K, Adams DW, Wang J, Jaremko M, Joshua-Tor L, Keil W, Hammell CM. 2023. A circadian-like gene network programs the timing and dosage of heterochronic miRNA transcription during *C. elegans* development. *Dev Cell* **58:** 2563–2579.e8. doi:10.1016/j.devcel.2023.08.006

Medwig-Kinney TN, Kinney BA, Martinez MAQ, Yee C, Sirota SS, Mullarkey AA, Somineni N, Hippler J, Zhang W, Shen K, et al. 2023. Dynamic compartmentalization of the pro-invasive transcription factor NHR-67 reveals a role for Groucho in regulating a proliferative-invasive cellular switch in *C. elegans*. *Elife* **12:** RP84355. doi:10.7554/eLife.84355

Xiao Y, Yee C, Zhao CZ, Martinez MAQ, Zhang W, Shen K, Matus DQ, Hammell C. 2023. An expandable FLP-ON::TIR1 system for precise spatiotemporal protein degradation in *Caenorhabditis elegans*. *Genetics* **223:** iyad013. doi:10.1093/genetics/iyad013

# MECHANISMS OF GENE REGULATION AND NONCODING RNAs

**L. Joshua-Tor** D. Adams, A. Axhemi, J. Bauer, L. Braviner, O.P. Chouhan, A. Garg, S. Goldsmith, S. Gonuguntla, B. Farhi, J. Ipsaro, M. Licht, Y. Nishino, K. On

We study the molecular basis of nucleic acid regulatory processes: gene expression and RNA interference (RNAi), as well as DNA replication. We use the tools of structural biology, biochemistry, and biophysics to study protein complexes associated with these processes and elucidate how they work. Cryo-electron microscopy, X-ray crystallography, and other structural techniques enable us to obtain the three-dimensional structures of these molecular machines. Biochemistry, biophysics, and molecular biology allow us to study properties that can be correlated to their function and biology.

## Epigenetic Inheritance of DNA Methylation

D. Adams and J. Ipsaro [in collaboration with R. Martienssen, CSHL]

DDM1 (Decrease in DNA Methylation 1) is a Snf2-like chromatin remodeler that serves as a master regulator in the epigenetic silencing of heterochromatin in *Arabidopsis*. In *ddm1* mutants, DNA methylation and H3K9me silencing marks are down-regulated, leading to a derepression of transposable element transcription. DDM1 colocalizes with H3.1 and H3.3 during the cell cycle and DDM1 also promotes histone variant exchange. Genetic mutations in *ddm1* show a loss of histone variant H3.1 (associated with heterochromatin) deposition, resulting in increased and ectopic histone H3.3 (associated with active transcription) deposition.

To establish the molecular specificity of physical interactions between DDM1 and nucleosomes, we determined the molecular structure of DDM1 with a variant nucleosome at 3.2 Å resolution by single-particle cryogenic electron microscopy (cryo-EM) (Fig. 1). Both the DEXD ATPase and HELICc domains of DDM1 engage with the nucleosome at the SHL2 position, making contacts with both DNA gyres of the nucleosome, while also making specific contacts with histone H3.3 and the unmodified tail of histone H4. The helical structure of DNA was notably unwound where the HELICc domain of DDM1 binds to the nucleosome. We showed that a disulfide bond that we discovered within the HELICc domain is essential for enzymatic activity. Curiously, *ddm1-1*, the first allele of *ddm1* that was characterized and that has strong defects in DNA methylation, harbors a C615Y substitution that would disrupt this disulfide bond.

To probe our structural observations further, peptide-binding, ATPase-activity, and nucleosome sliding assays were performed using various recombinant DDM1 constructs. Through microscale thermophoresis binding affinity assays, DDM1 was shown to bind to unmodified H4 and methylated H4K20 peptides with a $K_D$ of ~100 μM, whereas fully acetylated H4 peptides failed to bind. These assays demonstrated a specific affinity for deacetylated H4 tails, consistent with the state of heterochromatic nucleosomes, which are strongly deacetylated in *Arabidopsis*. Nucleosome sliding and ATPase assays show that the amino terminus of DDM1 is inhibitory, with the enzymatic activity of wild-type DDM1 being substantially lower than DDM1(Δ1-132). Interestingly, the disulfide bond mutant of DDM1(C615S) had no measurable enzymatic activity, hinting at a layer of redox regulation of chromatin remodeling and DNA methylation.

The molecular structure and subsequent experimental validation of the DDM1-nucleosome complex reveal many factors that regulate the action of DDM1 to maintain heterochromatin through DNA methylation. Dynamic characteristics of chromatin, such as the histone variant composition, the epigenetic state of histone tails, and the local redox environment, all impact the silencing of transposable elements via DDM1 remodeling.

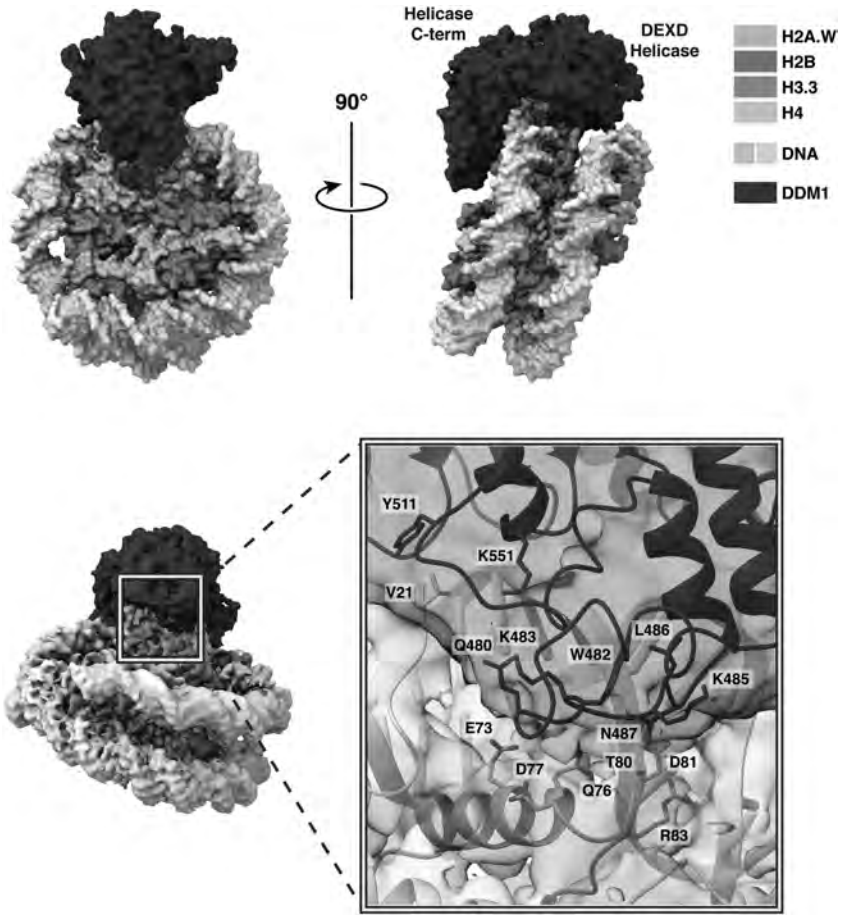



**Figure 1.** (*Top*) Overview of the molecular structure of the DDM1-nucleosome complex as determined by cryogenic electron microscopy (cryo-EM). DDM1 domains, corresponding to the two lobes, are labeled on the side view. The final structure spans residues 200-435 of DDM1 including the DEXD ATPase domain, and residues 442-673, which includes the helicase superfamily carboxyl-terminal (HELICc) domain. DDM1 clasps the nucleosome on the outside of superhelical location-2 (SHL-2), making contact with both gyres of DNA. (*Bottom*) DDM1-histone interactions. The experimental density map of the complex shows that DDM1 interacts with histones H3.3 and H4. The majority of the DDM1-histone interface is formed by a loop in DDM1 (residues 480-487) that makes contact with H3. (*Inset*) A cartoon representation with a partially transparent cryo-EM map is shown. Amino acids along the DDM1-histone interface (6 Å cutoff) are displayed as sticks and include T80 and D81 of histone H3. In this region, DDM1, DNA, histone H3.3, and histone H4 all contact one another.

## A Circadian-Like Gene Network Regulating the Expression of a Developmental miRNA

D. Adams [in collaboration with C. Hammell, CSHL]

The development of an organism from the embryonic to the adult stage is dependent on the precise organization, communication, and progression of cellular responses over time. The coordination and regulation of gene dosages, temporal expression patterns, and localization of genetic expression are important for the proper progression of an organism through its life stages. In *Caenorhabditis elegans*, a heterochronic gene regulatory network controls the progression through developmental stages via the precise timing and dosage of regulatory microRNAs (miRNAs), including the miRNA, *lin-4*, that is the subject of this study.

Previous work in the Hammell laboratory showed that the timing of *lin-4* transcription correlates with the phased expression of two transcription factors, NHR-23 and NHR-85, and the cyclical expression of an antagonistic factor, LIN-42. Critical to *lin-4* transcription is a brief temporal overlap in the expression

of NHR23 and NHR85, leading to the cooperative binding of these two NHR proteins to a DNA pulse control element (PCE).

We used electrophoretic mobility shift assays (EMSA) to characterize the binding of NHR-23 and NHR-85 to the PCE. Although NHR-23 and NHR-85 both bind the wild-type PCE fragment, combining the two proteins enhances the binding of the PCE. Furthermore, multi-angle light scattering experiments show that the binding of the PCE is dependent on the availability of the "GGTCA" repeat within the PCE. We also showed that the two proteins interact with high affinity through microscale thermophoresis assays.

Understanding how these proteins interact at the molecular level will give insights into the cooperativity of NHR23/NHR85 binding of the PCE and how LIN-42 might antagonize this interaction.

## PUBLICATIONS

Gao Y, He X-Y, Wu XS, Huang Y-H, Toneyan S, Ha T, Ipsaro JJ, Koo PK, Joshua-Tor L, Bailey KM, et al. 2023. ETV6 dependency in Ewing sarcoma by antagonism of EWS-FLI1-mediated enhancer activation. *Nat Cell Biol* **25:** 298–308. doi:10.1038/s41556-022-01060-1

Kinney B, Sahu S, Stec N, Hills-Muckey K, Adams DW, Wang J, Jaremko M, Joshua-Tor L, Keil W, Hammell CM. 2023. A circadian-like gene network programs the timing and dosage of heterochronic miRNA transcription during *C. elegans* development. *Dev Cell* **58:** 1–17. doi:10.1093/doi:10.1016/j.devcel.2023.08.006

Lee SC, Adams DW, Ipsaro JJ, Cahn J, Lynn J, Kim H-S, Berube B, Major V, Calraco JP, LeBlanc C, et al. 2023. Chromatin remodeling of histone H3 variants by DDM1 underlies epigenetic inheritance of DNA methylation. *Cell* **186:** 4100–4116. doi:10.1016/j.cell.2023.08.001

Meze K, Axhemi A, Thomas DR, Doymaz A, Joshua-Tor L. 2023. A shape-shifting nuclease unravels structured RNA. *Nat Struc Mol Biol* **30:** 339–347. doi:10.1038/s41594-023-00923-x

Qian Z, Song D, Ipsaro JJ, Bautista C, Joshua-Tor L, Yeh JT-H, Tonks NK. 2023. Manipulating PTPRD function with ectodomain antibodies. *Genes Dev* **37:** 743–759. doi:10.1101/gad.350713.123

# RNA SPLICING

**A.R. Krainer** J. Apollo, C. Cizmeciyan, M. Daley, L. Han, Y. Ishigami, L. Jia, A. Kral, K-T. Lin, Y.H. Liu, L. Manche, E.E. Nakagaki-Silva, D. Segovia, D. Voss, L. Wan, L. Yang, Q. Zhang

## Mechanisms of Constitutive and Alternative Pre-mRNA Splicing

RNA splicing is required for expression of most eukaryotic protein-coding genes. The spliceosome selects authentic splice sites with very high fidelity, relying on limited sequence information present throughout introns and exons. In humans, >90% of genes are expressed via alternative splicing, giving rise to multiple protein isoforms. The choice of alternative splice sites is commonly regulated to alter gene expression, either tissue-specifically or in response to a developmental program or to signaling pathways. The fact that multiple protein isoforms can be expressed from individual genes demonstrates that the classical "one gene–one enzyme" paradigm is no longer valid and provides an explanation for the unexpectedly small number of protein-coding genes uncovered by genome-sequencing projects.

Both constitutive and alternative splicing mechanisms involve numerous protein components, as well as five noncoding RNA components that are part of small nuclear ribonucleoprotein (snRNP) particles. These components are sequentially assembled with a pre-mRNA substrate into a spliceosome, which catalyzes the two transesterification steps of splicing. The work in our laboratory focuses on the identification and molecular characterization of protein factors and sequence elements that are necessary for the catalysis and fidelity of splicing and/or for the regulation of alternative splice-site selection. We are interested in how the spliceosome correctly identifies the exons on pre-mRNA and how certain point mutations in either exon or intron sequences cause aberrant splicing, leading to various human genetic diseases. Related areas of interest include the remodeling of mRNP architecture as a consequence of splicing, which influences downstream events, such as nonsense-mediated mRNA decay (NMD); the various roles of alternative splicing dysregulation in cancer; and the development of effective methods—particularly antisense technology—to correct defective splicing or modulate alternative splicing or for gene/allele-specific inhibition of NMD, in various disease contexts. A summary of five of our recently published studies is provided below.

## Experimental and Computational Analyses of Splicing

SRSF1 is the founding member of the SR protein family. It is required—interchangeably with other SR proteins—for pre-mRNA splicing in vitro, and it regulates various alternative splicing events. Dysregulation of SRSF1 expression contributes to cancer and other pathologies (see below). To further explore the functions and mechanisms of action of SRSF1, we characterized its interactome using proximity labeling and mass spectrometry. This approach yielded 190 proteins enriched in the SRSF1 samples, independently of the amino- or carboxy-terminal location of the biotin-labeling domain. The detected proteins reflect established functions of SRSF1 in pre-mRNA splicing and reveal additional connections to spliceosome proteins, in addition to other recently identified functions. We validated a robust interaction with the spliceosomal RNA helicase DDX23/PRP28 using bimolecular fluorescence complementation and in vitro binding assays. The interaction is mediated by the amino-terminal RS-like domain of DDX23 and both RRM1 and the carboxy-terminal RS domain of SRSF1. During pre-mRNA splicing, DDX23's ATPase activity is essential for the pre-B to B spliceosome complex transition and for release of U1 snRNP from the 5′ splice

site. We showed that the RS-like region of DDX23's amino-terminal domain is important for spliceosome incorporation, whereas larger deletions in this domain alter the protein's subnuclear localization. Finally, we showed that DDX23 also interacts with several other members of the SR protein family.

In collaboration with Justin Kinney (CSHL), we previously generated and analyzed complete libraries of all possible 5′-splice-site sequences in three different minigene contexts, using barcoding and massively parallel RNA sequencing after transient transfection. Among other findings, these massively parallel splicing assay (MPSA) data revealed pairwise epistatic interactions between certain nucleotides of the 5′ splice site, such as a positive epistatic interaction between G at the –1 position and G at the +5 position. Continuing these combined experimental/computational collaborative efforts with the Kinney and McCandlish laboratories, we used the MPSA approach, together with transcriptomics, biophysical modeling, mutational analysis, and quantitative drug-dose response analysis in cell culture, to study the mechanism of action of two splice-modifying small-molecule drugs for spinal muscular atrophy (SMA), risdiplam and branaplam. This analysis quantitatively defined the specificities of these two drugs for 5′-splice-site sequences, suggested that branaplam recognizes 5′ splice sites via two distinct interaction modes, and disproved a previous two-site hypothesis for risdiplam activity at *SMN2* exon 7. More generally, our results showed that single-drug cooperativity and multidrug synergy are widespread among both small-molecule drugs and antisense-oligonucleotide drugs that promote exon inclusion.

## Alternative Splicing as Driver and Therapeutic Target in Solid Tumors

Inflammation is strongly associated with pancreatic ductal adenocarcinoma (PDAC), a highly lethal malignancy. Dysregulated RNA splicing factors have been widely reported in tumorigenesis, but their involvement in pancreatitis and PDAC is not well understood. We found that the splicing factor SRSF1 is highly expressed in pancreatitis, PDAC precursor lesions, and tumors. Using pancreas-specific inducible overexpression of SRSF1 in a PDAC mouse model, we showed that increased SRSF1 is sufficient to induce pancreatitis and accelerate $KRAS^{G12D}$-mediated PDAC. Mechanistically, SRSF1 activates MAPK signaling, in part by up-regulating interleukin 1 receptor type 1 (IL1R1) through alternative-splicing-regulated mRNA stability. Additionally, we found that SRSF1 protein is destabilized through a negative feedback mechanism in phenotypically normal epithelial cells expressing $KRAS^{G12D}$, both in mouse pancreas and in pancreas organoids acutely expressing $KRAS^{G12D}$, buffering MAPK signaling and maintaining pancreas-cell homeostasis. This negative feedback regulation of SRSF1 can be overcome by hyperactive MYC, facilitating PDAC tumorigenesis. Our findings thus implicate SRSF1 in the etiology of pancreatitis and PDAC and point to SRSF1-misregulated alternative splicing as a potential therapeutic target.

In addition to the role of splicing misregulation in PDAC initiation, we contributed to a study of PDAC metastatic spread, in collaboration with former laboratory member Rotem Karni (Hebrew University, Jerusalem). Analysis of RNA-splicing data from a large cohort of primary and metastatic PDAC tumors identified differentially spliced events that correlate with PDAC progression. De novo motif analysis of these events detected enrichment of motifs with high similarity to the RBFOX2 binding site, UGCAUG. Overexpression of RBFOX2 in a patient-derived xenograft metastatic PDAC cell line drastically reduced its metastatic potential in vitro and in vivo, whereas depletion of RBFOX2 in primary PDAC cell lines increased their metastatic potential. These findings indicate that RBFOX2 functions as a potent metastatic suppressor in PDAC. RNA-seq and splicing analysis of RBFOX2 target genes revealed enrichment of genes in RHO GTPase pathways, suggesting a role of RBFOX2 targets in cytoskeletal organization and focal-adhesion formation.

## Development of Antisense Oligonucleotide Therapeutics

Diffuse midline gliomas (DMGs) are lethal pediatric high-grade brain tumors in the thalamus, midbrain, or pons; the latter subgroup is termed diffuse intrinsic pontine gliomas (DIPG). The brainstem location of these tumors limits the clinical management of DIPG, resulting in very poor outcomes. A heterozygous, somatic point mutation in one of two

genes coding for the noncanonical histone H3.3 is present in most DIPG tumors. This dominant mutation in the *H3-3A* gene results in replacement of lysine 27 with methionine (K27M) and causes a global reduction of trimethylation on K27 of all wild-type histone H3 proteins, which is thought to be a driving event in gliomagenesis. We designed and systematically screened 2′-*O*-methoxyethyl phosphorothioate antisense oligonucleotides (ASOs) that direct RNase H–mediated knockdown of *H3-3A* mRNA. We identified a lead ASO that effectively reduced *H3-3A* mRNA and H3.3K27M protein and restored global H3K27 trimethylation in patient-derived neurospheres. We then tested the lead ASO in two mouse models of DIPG: an immunocompetent mouse model using transduced mutant human *H3-3A* cDNA, and an orthotopic xenograft with patient-derived cells implanted in immunodeficient mice. In both in vivo models, ASO treatment restored K27 trimethylation of histone H3 proteins and reduced tumor growth; promoted neural-stem-cell and tumor-cell differentiation into astrocytes, neurons, and oligodendrocytes; and increased survival. These results demonstrate the involvement of the H3.3K27M oncohistone in tumor maintenance, confirm the reversibility of the aberrant epigenetic changes it promotes, and provide preclinical proof of concept for DMG antisense therapy.

## PUBLICATIONS


Hastings ML, Krainer AR. 2023. RNA therapeutics. *RNA* **29:** 393–395.

Jbara A, Lin KT, Stossel C, Siegfried Z, Shqerat H, Amar-Schwartz A, Elyada E, Mogilevsky M, Raitses-Gurevich M, Johnson JL, et al. 2023. RBFOX2 modulates a metastatic signature of alternative splicing in pancreatic cancer. *Nature* **617:** 147–153.

Kim YJ, Krainer AR. 2023. Antisense oligonucleotide therapeutics for cystic fibrosis: recent developments and perspectives. *Mol Cells* **46:** 10–20.

Wan L, Lin K-T, Rahman MA, Wang Z, Jensen MA, Park Y, Tuveson DA, Krainer AR. 2023. Splicing factor SRSF1 promotes pancreatitis and $KRAS^{G12D}$-mediated pancreatic cancer. *Cancer Discov* **13:** 1678–1695.

Zhang Q, Yang L, Liu YH, Wilkinson JE, Krainer AR. 2023. Antisense oligonucleotide therapy for H3.3K27M diffuse midline glioma. *Science Transl Med* **15:** eadd8280.


### *In Press*


Ecker DJ, Aiello CD, Arron JR, Bennett CF, Bernard A, Breakfield XO, Broderick TJ, Callier SL, Canton B, Chen JS, et al. 2024. Opportunities and challenges for innovative and equitable healthcare. *Nat Rev Drug Discov* **23:** 321–322.

Ishigami Y, Wong MS, Martí-Gómez C, Ayaz A, Kooshkbaghi M, Hanson S, McCandlish DM, Krainer AR, Kinney JB. 2024. Specificity, synergy, and mechanisms of splice-modifying drugs. *Nat Commun* **15:** 1880.

Maia-Silva D, Cunniff PJ, Schier AC, Skopelitis D, Trousdell MC, Moresco P, Gao Y, Kechejian V, He XY, Sahin Y, et al. 2024. Interaction between MED12 and ΔNp63 activates basal identity in pancreatic ductal adenocarcinoma. *Nat Genet* **56:** 1377–1385.

Segovia D, Adams DW, Hoffman N, Safaric Tepes P, Wee TL, Cifani P, Joshua-Tor L, Krainer AR. 2024. SRSF1 interactome determined by proximity labeling reveals direct interaction with spliceosomal RNA helicase DDX23. *Proc Natl Acad Sci* **121:** e2322974121.

Wan L, Kral AJ, Voss D, Schäfer B, Sudheendran K, Danielsen M, Caruthers MH, Krainer AR. 2024. Screening splice-switching antisense oligonucleotides in pancreas-cancer organoids. *Nucleic Acid Ther* **34:** 188–198.

Xiang X, Bhowmick K, Shetty K, Ohshiro K, Yang X, Wong LL, Yu H, Latham PS, Satapathy SK, Brennan C, et al. 2024. Mechanistically based blood proteomic markers in the TGF-β pathway stratify risk of hepatocellular cancer in patients with cirrhosis. *Genes Cancer* **15:** 1–14.

Yang X, Bhowmick K, Rao S, Xiang X, Ohshiro K, Amdur RL, Hassan MI, Mohammad T, Crandall K, Cifani P, et al. 2024. Aldehydes alter TGF-β signaling and induce obesity and cancer. *Cell Rep* **43:** 114676.

# CLICK CHEMISTRY, SYNTHESIS, AND CHEMICAL BIOLOGY

**J.E. Moses** J. Homer R. Johnson A. Moorhouse M. Rufrano D. Vishwakarma
Q. Huang R. Koelln S. Pati S. Sun Z. Wang

The Moses laboratory focuses on developing novel click chemistry methodologies and applying these in cancer and infectious disease research, leveraging two key advantages of click reactions. First, click reactions exhibit high selectivity and robustness, making them particularly well suited for late-stage functionalization. This is especially crucial when manipulating the complex chemical functionalities often present in biologically active scaffolds and drug molecules. The direct functionalization facilitated by click chemistry can also enable the facile synthesis of chemical probes for application in protein enrichment assays, an essential component of mechanism of action studies. Moreover, click chemistries, including sulfur fluoride exchange (SuFEx) and phosphorus fluoride exchange (PFEx), serve as powerful discovery tools. These reactions enable the rapid generation of diverse chemical libraries for screening in functional assays, identifying lead structures for further development. The fidelity of click reactions allows for the swift exploration of chemical space, and the resulting click-derived linkers can function as valuable pharmacophores while simultaneously offering useful stability profiles. Summaries of our major developments follow.

## Building the Click Chemistry Toolbox—Novel Reaction Development

This work was done in collaboration with K.B. Sharpless (Scripps).

Click chemistry has been applied successfully in our laboratory's pursuit of discovering functional molecules. By expanding the repertoire of available click transformations, the possibility of accessing hit and lead molecules through array synthesis is increased.

Until recently, phosphorus-centered connections were notably absent from the click chemistry toolbox. Some of Nature's most vital molecular connectors contain phosphate esters and anhydrides. These linkages are essential in the biosynthesis of nucleic acids, nucleotide coenzymes, nucleoside triphosphates (e.g., ATP), metabolic intermediates, and intermediates in many biochemical processes. They are also found in chemotherapy agents and antiviral drugs such as (–)-remdesivir. Drawing inspiration from nature's phosphate connectors, we recently co-developed PFEx as the latest advance in connective click reaction technology (Fig. 1A). PFEx allows for the dependable linking of P(V)–F-loaded hubs with aryl alcohols, alkyl alcohols, and amines, yielding stable, multidimensional products connected through P(V)–O and P(V)–N bonds. The reactivity profile of P–F hubs surpasses that of their P–Cl counterparts in reaction performance, rate, and outcome, qualifying PFEx as a true click reaction.

The use of Lewis amine base catalysis (e.g., 1,5,7-triazabicyclo[4.4.0]dec-5-ene [TBD]), significantly enhances the rate of PFEx transformations, and the judicious selection of such bases allows for selective serial exchange reactions that result in complex final products in up to four steps. The reactivity profile of PFEx hubs also allows for a series of click reactions to be performed (e.g., SuFEx-PFEx-CuAAC) to rapidly generate complex multidimensional molecules, making PFEx an excellent new addition to the click chemistry toolbox.

Although incredibly useful as connective hubs, compounds containing electrophilic P–F bonds can be toxic. To help align the PFEx reaction more closely with the safety profiles of other click reactions, we have endeavored to enhance the user-friendliness of this transformation. By delivering the reagents in premade solutions and treating all glassware and consumables with appropriate quenching protocols, the risk of substrate and reagent exposure to the operator is minimized. To enhance the accessibility of PFEx to the wider scientific community, we have published a detailed step-by-step protocol alongside a video demonstration (https://www.youtube.com/watch?v=mM2B1vDs0PI).

In addition to PFEx, the Moses laboratory reported the synthesis of the novel SuFEx hub

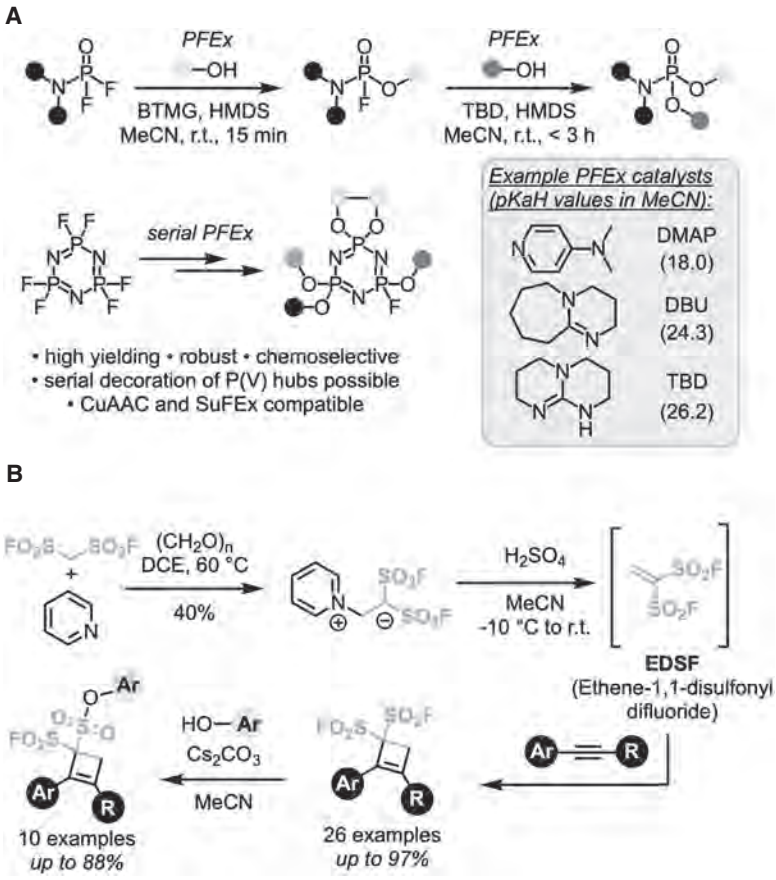



**Figure 1.** (*A*) Catalytic phosphorus (V) fluoride exchange (PFEx). (*B*) Synthesis and application of the EDSF SuFEx hub.

ethene-1,1-disulfonyl difluoride (EDSF). The preparation of this hub was realized via a bench-stable 1,1-bis(fluorosulfonyl)-2-(pyridin-1-ium-1-yl)ethan-1-ide precursor (Fig. 2). The EDSF SuFEx reagent was subsequently utilized to synthesize 26 distinct 1,1-bissulfonylfluoride substituted cyclobutenes via cycloaddition reactions. This cycloaddition process was rapid, straightforward, and highly efficient—aligning with click chemistry principles. The highly functionalized four-membered ring carbocycles produced by this reaction are important structural motifs present in numerous bioactive natural compounds and pharmaceutically relevant small molecules. The cyclobutene cores were further diversified through selective $Cs_2CO_3$-activated SuFEx click chemistry between a single S–F group and aryl alcohol, resulting in the synthesis of sulfonate ester products. Mechanistic insights into the reaction pathway were provided following density functional theory calculations.

## Click Chemistry Discovery of Hit Molecules

This work was done in collaboration with D. Tuveson, S. Lyons, and M. Lukey (CSHL), M. Jackson (Colorado State University), W.D. Fairlie (La Trobe University), S. Haider (University College London), A. Prota (Paul Scherrer Institute), and Mi Steinmetz (Paul Scherrer Institute).

The adaption of click reactions to array format is one way to expedite the discovery of functional molecules via compound library synthesis. Accelerated SuFEx click chemistry (ASCC) (previously developed by the Moses laboratory) is a potent method for coupling aryl and alkyl alcohols with SuFEx-compatible functional groups and is characterized by favorable kinetics and high product yields. ASCC enhances synthetic workflows, simplifies purification processes, and is particularly effective for discovering functional molecules.

Utilizing commercially available aryl alcohols and various SuFExable fragments, including aryl and alkyl sulfonyl fluorides, aryl fluorosulfates, and sulfuryl fluoride

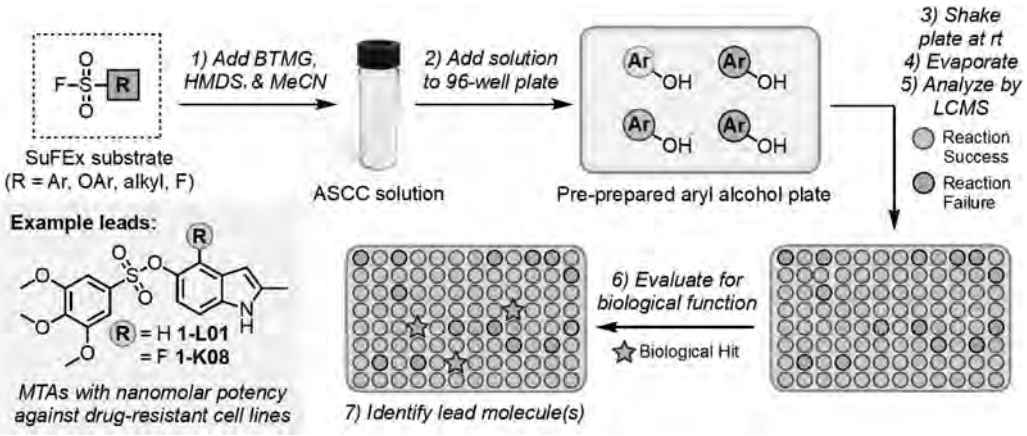



**Figure 2.** Schematic of the use of the accelerated sulfur fluoride exchange (SuFEx) click chemistry (ASCC) reaction in drug discovery.

gas, we synthesized discrete libraries of molecules systematically. The resulting compound libraries were evaluated for biological function. Lead molecules, including those with anticancer and antibiotic properties, were identified, with some demonstrating significant activity against aggressive cancers like pancreatic ductal adenocarcinoma (PDAC) and triple-negative breast cancer (TNBC) (specifically, **1-L01** and **1-K08** depicted in Fig. 2).

Overall, these results highlight the effectiveness and versatility of click derivatization in rapidly identifying biologically active molecules.

## Shapeshifting Antibiotics

In collaboration with P. Williams (University of Nottingham), J. Bolla, C. Robinson (University of Oxford), T. Fallon (The University of Newcastle), and T. Soares da Costa (La Trobe University).

The rapid emergence of multidrug-resistant bacteria is recognized as a major global health crisis, prompting the need for the rapid development of new therapies. However, the development of new antibiotic agents can be prohibitively expensive and time-consuming, reducing the number of innovators in the field. The Moses group re-engineered existing U.S. Food and Drug Administration (FDA)-approved antibiotics of last resort to evade resistance mechanisms in an effort to expedite the introduction of new therapies to the clinic (Fig. 3A,B). This strategy pivoted around the incorporation of a "shapeshifting" core, known as bullvalene, into the structure of vancomycin dimers (Fig. 3C). The lead candidates from this study displayed similar potency to vancomycin against susceptible strains of bacteria in biological evaluation against a panel of Gram-positive bacteria. The novel shapeshifting vancomycin dimers (SVDs) were not disadvantaged by the common mechanism of vancomycin resistance resulting from the alteration of the carboxy-terminal d-Ala-d-Ala dipeptide with the corresponding d-Ala-d-Lac depsipeptide, retaining potency against bacteria that were resistant to the parent drug, including vancomycin-resistant *Enterococcus* (VRE), methicillin-resistant *Staphylococcus aureus* (MRSA), and vancomycin-resistant *S. aureus* (VRSA) (Fig. 3D). Further, evidence suggested that the shapeshifting ligands destabilized the complex formed between flippase MurJ (an enzyme pivotal in bacterial cell wall assembly) and lipid II, implying access to a potentially new mode of action for polyvalent glycopeptides. The SVDs showed little propensity for acquired resistance by enterococci and additionally demonstrated limited cytotoxicity against HEK293 human embryonic kidney cells (Fig. 3D). The Moses group continues to expand their second-generation library of shapeshifting antibiotics to aid in the fight against drug-resistant pathogens.

## Jerantinine-A—Sustainable Synthesis, Analog Generation, and Anticancer Properties

This work was done in collaboration with D. Tuveson, S. Lyons, M. Lukey, and J. Yeh (CSHL) and Kevan Shokat (UCSF).

(–)-Jerantinine A (**JA**) is a versatile polypharmaceutical with multiple biological actions, including

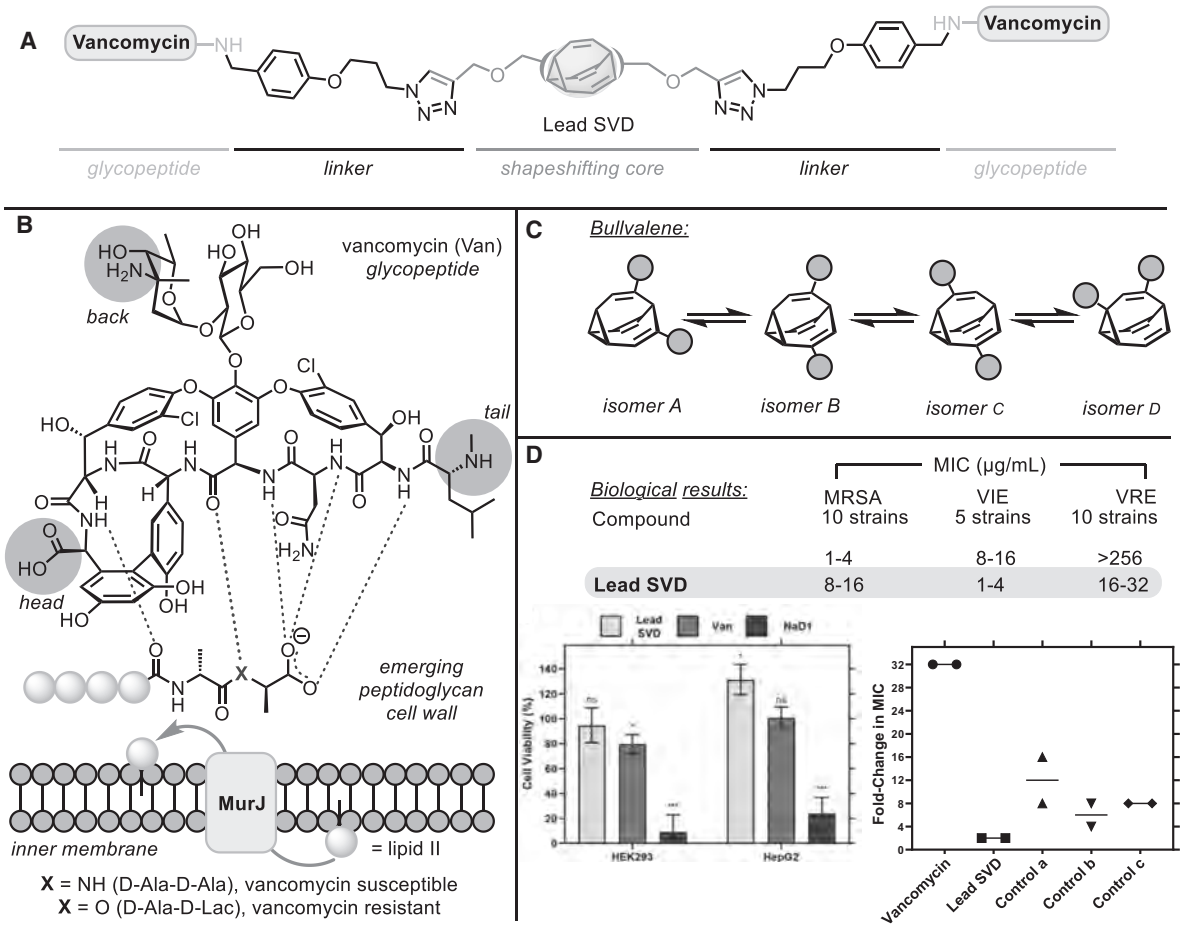



**Figure 3.** (*A*) Shapeshifting antibiotic dimer design. (*B*) Vancomycin binding to lipid II. (*C*) Bullvalene, a shapeshifting molecule. (*D*) Antibacterial panels against multiple strains of MRSA, VIE, and VRE. (MSSA) Methicillin-sensitive *Staphylococcus aureus* (ATCC® 9144™), (MRSA) methicillin-resistant *S. aureus* (ATCC® BAA-1720), (VSE) vancomycin-sensitive *Enterococcus faecium* (ATCC® BAA-2127™), (VIE) vancomycin-intermediate E. gallinarum (ATCC® 49608™); and VRE: vancomycin-resistant *E. faecium* (ATCC® 700221™), (VSSA) SH1000, (VRSA) vancomycin-resistant *S. aureus* (VRS1/HIP11714). Viability studies of HEK293 cells and HepG2 cells treated with 500 μg/mL lead shapeshifting vancomycin dimer (SVD) (light gray) relative to 500 μg/mL vancomycin (dark gray), and 50 μM of the cytotoxic peptide NaD1 (black), normalized to a 1% (v/v) DMSO vehicle control. Error bars, S.D. ($n = 3$); ns $p > 0.05$, $*p < 0.05$, $***p < 0.001$; and propensity of *E. faecium* to develop resistance. *E. faecium* (ATCC® BAA-2127™) was exposed to increasing concentrations of vancomycin, lead SVD, control a, control b, and control c, up to 4× minimum inhibitory concentration (MIC). Data plotted as the fold-change in MIC value relative to a DMSO vehicle control ($N = 2$).

induction of $G_2$/M cell cycle arrest and apoptosis in cancer cells by disrupting microtubule polymerization, inhibiting polo-like kinase 1 (PLK1), and targeting the spliceosome through up-regulation of SF3B1 and SF3B3 proteins. Its multifaceted mode of action positions **JA** as a promising anticancer agent, particularly for treating tumors with high heterogeneity and drug resistance. The Moses laboratory has developed a streamlined semisynthesis of **JA**, its likely biosynthetic precursor (–)-melodinine P (**MP**), and derivatives thereof (Fig. 4A), demonstrating potent cytotoxicity against aggressive TNBC cells (Fig. 4B). Additionally, live-cell imaging assays revealed **JA**'s ability to induce oxidative stress (Fig. 4C), and metabolomics analysis identified its role in inhibiting nucleotide metabolism, specifically pyrimidine biosynthesis, highlighting its potential therapeutic efficacy and mechanism of action in TNBC cells.

A significant offshoot of this work stems from the discovery that **JA** and **MP** bind covalently to

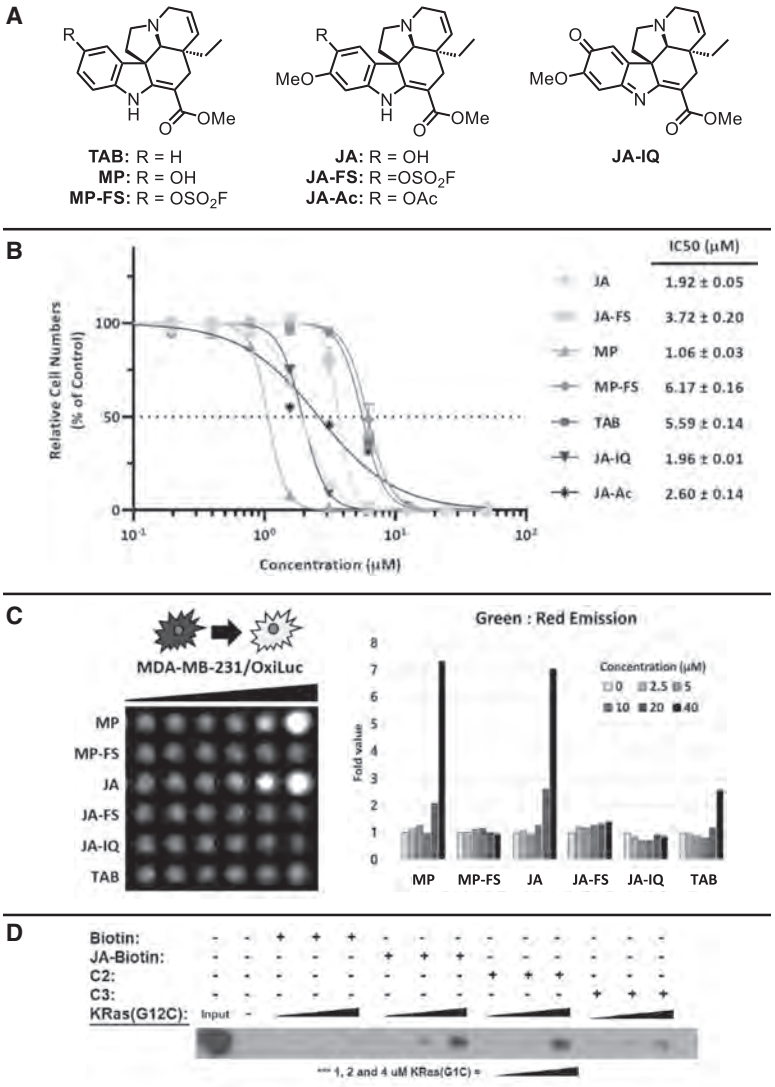



**Figure 4.** (*A*) Jerantinine-A and derivatives tested. (*B*) Growth inhibitory effect of JA and its derivatives against the MDA-MB-231 breast cancer cell line. Cell numbers were quantified by CyQUANT assay after 5 days of treatment at indicated concentrations (error bars indicate SEM; $n = 4$). $IC_{50}$ values were calculated using GraphPad Prism. (*C*) In vitro imaging assays detecting oxidative stress. Lighter color indicates reactive oxygen species (ROS) generation. (*D*) JA-KRas(G12C) pulldown assay results.

K-Ras(G12C). In collaboration with the Shokat laboratory, we designed, synthesized, and tested a **JA-Biotin** probe using click chemistry. Whole-cell pulldown assays using the **JA-Biotin** probe and two control molecules **C2** and **C3** confirmed the reproducible interaction between K-Ras(G12C) and the **JA-Biotin** pulldown probe (Fig. 4D). In an ongoing collaboration with the Shokat laboratory, whole-protein mass spectrometry was used to confirm the covalent labeling of K-Ras(G12C) by both **JA** and **MP**. **JA** and **MP** did not label K-Ras(wt) or K-Ras(G12D). The **JA-FS** derivative and Tabersonine (**TAB**), another member of the *Aspidosperma* alkaloids, did not label K-Ras(wt), K-Ras(G12C), or K-Ras(G12D). In **JA-FS**, the phenolic oxygen of **JA** at C-15 (also present in MP) cannot participate in any potential oxidation of the phenol-indole moiety to iminoquinone, a known process of *Aspidosperma* alkaloids. **TAB** does not possess the C-15 phenolic oxygen. Further, the covalent adducts of K-Ras(G12C) and both **JA** and **MP** lost two mass

units. These labeling preferences and adduct masses led us to hypothesize that **JA** and **MP** may label K-Ras(G12C) through the iminoquinone derivatives of **JA** and **MP**. To test this hypothesis, we conducted an experiment to determine whether K-Ras(G12C) labeling was oxidation-dependent. The effect of the treatment of **JA** and **MP** with sodium ascorbate on the percentage of labeling of K-Ras(G12C) by **JA** and **MP** was determined, and the results are consistent with an oxidation-based method of covalent bond formation. The Shokat laboratory went on to confirm that cancer cell death was not a direct result of K-Ras(G12C) inhibition by using a Ba/F3 cell model system to examine on-target inhibition.

## Developing Bromodomain 8–Selective Inhibitors

This work was done in collaboration with A. Mills (CSHL).

In human cancer development, the inhibition of the tumor-suppressive function of p53 is a crucial factor. Despite this, p53 remains unmutated in most cases of glioblastoma (GBM). Recent research from the Mills laboratory has revealed that the chromatin regulator bromodomain 8 (BRD8), through the EP400 histone acetyltransferase complex, maintains H2AZ occupancy at p53 target loci, consequently promoting GBM proliferation by hindering p53 transactivation. Targeting BRD8 with small-molecule inhibitors presents a promising strategy to enhance or reactivate p53 activity, thus restoring tumor suppression in GBM.

However, the development of BRD8-selective small-molecule inhibitors is challenging because of difficulties in expressing the dual bromodomains of BRD8 in crystallographic quality and quantity. The lack of structural data further complicates inhibitor design. Recently, the Forrester group reported the first small-molecule inhibitors of BRD8, discovered serendipitously during an AlphaScreen assay. Despite promising selectivity, the efficacy of these inhibitors in cells was not studied. The Moses laboratory evaluated a literature-based BRD8 inhibitor against GBM cell lines and found it lacking in both selectivity and potency.

Inspired by this, the Moses Laboratory has designed and synthesized a series of triazole-based small molecules as selective BRD8 inhibitors. Their synthesis choices were informed by docking studies using a homology model of BRD8. Initial screening identified one derivative as slightly more active in BRD8-sensitive GBM cell lines. Further refinement through a SuFEx-based discovery platform has yielded several more potent derivatives, with ongoing work aiming to increase selectivity between BRD8-sensitive and -nonsensitive GBM cell lines.

## Developing Nonaddictive Morphine Derivatives

This work was done in collaboration with B. Li (CSHL).

Although opioids have revolutionized pain management, the associated risks, including addiction and overdose, are significant societal challenges. Developing nonaddictive opioid-like molecules would have profound benefits for public health, societal welfare, and economic burden reduction. By modifying morphine with ASCC to covalently capture a novel amino acid residue in the structure of the μ-opioid receptor, the project seeks to permanently block the receptor, thus preventing activation by additional opiate molecules. Preliminary studies in mice have shown promising results, indicating well-tolerated morphine derivatives with beneficial pain relief profiles and nonhabit-forming properties.

### PUBLICATIONS


Gialelis TL, Wang Z, Homer JA, Yang W-H, Chung T, Hu Q, Smedley CJ, Pawar NJ, Upadhyay NS, Tuveson DA, et al. 2023. Inhibition of mitochondrial metabolism by (–)-jerantinine A: synthesis and biological studies in triple-negative breast cancer cells. *RSC Med Chem* **14:** 710–714. doi:10.1039/D3MD00049D

Homer JA, Xu L, Kayambu N, Zheng Q, Choi EJ, Kim BM, Sharpless KB, Zuilhof H, Dong J, Moses JE. 2023. Sulfur fluoride exchange. *Nat Rev Methods Primers* **3:** 59. doi:10.1038/s43586-023-00241-y

Moorhouse AD, Homer JA, Moses JE. 2023. The certainty of a few good reactions. *Chem* **9:** 2063–2077. doi:10.1016/j.chempr.2023.03.017

Ottonello A, Wyllie JA, Yahiaoui O, Sun S, Koelln RA, Homer JA, Johnson RM, Murray E, Williams P, Bolla JR et al. 2023. Shapeshifting bullvalene-linked vancomycin dimers as effective antibiotics against multidrug-resistant Gram-positive bacteria. *Proc Natl Acad Sci* **120:** e2208737120. doi:10.1073/pnas.2208737120

Smedley CJ, Giel M-C, Fallon T, Moses JE. 2023. Ethene-1,1-disulfonyl difluoride (EDSF) for SuFEx click chemistry: synthesis of SuFExable 1,1-bissulfonylfluoride substituted cyclobutene hubs. *Angew Chem Int Ed* **62:** e202303916. doi:10.1002/anie.202303916

Sun S, Homer JA, Smedley CJ, Cheng Q-Q, Sharpless KB, Moses JE. 2023. Phosphorus fluoride exchange: multidimensional catalytic click chemistry from phosphorus connective hubs. *Chem* **9:** 1–16. doi:10.1016/j.chempr.2023.05.013

*In Press*

Homer JA, Koelln RA, Barrow AS, Gialelis TL, Boiarska Z, Steinohrt NS, Lee EF, Yang W-H, Johnson RM, Chung T, et al. 2024. Modular synthesis of functional libraries by accelerated SuFEx click chemistry, *Chem Sci* **15:** 3879. doi:10.1039/d3sc05729a

Homer JA, Sun S, Koelln RA, Moses JE. 2024. Protocol for producing phosphoramidate using phosphorus fluoride exchange click chemistry. *STAR Protoc* **5:** 102824. doi:10.1016/j.xpro.2023.102824

Moorhouse AD, Homer JA, Moses JE. 2024. Click chemistry: a catalyst for the democratization of synthesis, *Chem* **10:** 2615. doi:10.1016/jchempr.2024.07.030

# SMALL RNA PATHWAYS PROMOTE GENOME STABILITY

**A.J. Schorn** M. Peacey J.I. Steinberg J. Xie
H. Sertznig J. Wilken

Transposable elements and retroviruses are threats to genome stability. They are usually embedded in inactive chromatin, but epigenetic reprogramming during development and disease erases repressive chromatin marks. This is when small RNA–mediated silencing mechanisms become crucial to regulate their expression, inhibit their mobility, and maintain genome integrity. Small RNAs derived from the 3′ end of transfer RNAs (tRNAs) (3′-tRFs) strongly inhibit long terminal repeat (LTR) retroelements by targeting their highly conserved tRNA primer binding site (PBS). The use of host tRNAs as a primer for reverse transcription and replication is a hallmark of LTR retroelements, which include endogenous retroviruses (ERVs) in mammals but also closely related infectious retroviruses such as HIV.

## Biogenesis of tRNA-Derived Small RNA and RNA Interference

J.I. Steinberg, H. Sertznig, M. Peacey, J. Wilken, A.J. Schorn

tRNA-derived fragments (tRFs) are a novel class of small RNAs that potentially protect many eukaryotes. They are processed from full-length, mature tRNAs under yet unknown conditions. 3′-tRFs come in two distinct sizes: 22-nucleotides (nt)-long tRF3b fragments posttranscriptionally silence coding-competent ERVs, whereas ~18-nt tRF3a fragments specifically interfere with reverse transcription and retrovirus mobility. We quantified 3′-tRFs and their "parental" full-length, mature tRNA molecules from several cell types using hydrolysis-based tRNA sequencing. tRNA levels are not directly proportional to tRF3a or tRF3b, indicating selective processing or half-life times (Fig. 1A). Interestingly, specific tRFs are highly enriched in a pattern conserved between mouse and human (Fig. 1B). Moreover, tRFs from certain tRNA isoacceptors that differ by only a few nucleotides are much more abundant than others (Fig. 1C). This is unlikely because of selective endonucleolytic cleavage of tRFs from tRNAs, but rather suggests selective stability and turnover mechanisms.

Distinct mechanisms regulate small RNA targeting and turnover depending on their sequence complementarity to their targets, their terminal RNA modifications, and the clade of AGO/PIWI proteins they are binding to. We found that tRFs are 2′-O-methylated at their 3′ ends, protecting them from degradation and promoting ERV silencing. 2′-O methylation of 3′-tRFs is mediated by HENMT1, the same enzyme that methylates PIWI-interacting RNAs (piRNAs) in the germline and stabilizes small RNAs with extensive target complementarity in other organisms. In the absence of HENMT1, nontemplated tailing of 3′-tRFs occurs by terminal nucleotidyl transferases (TENTs) known to regulate small RNA stability and decay.

Because of the perfect sequence complementarity of 3′-tRFs to endogenous retroviral sequences, they have tens of thousands of targets in mammalian genomes. In collaboration with the Kinney laboratory (CSHL), we performed a massively parallel reporter assay to determine target site rules for 3′-tRFs. We tested ~2,000 sequence variations of the Lysine$^{TTT}$-PBS of *Mus musculus* particle type D (MusD), a highly active murine ERV targeted by Lysine$^{TTT}$ 3′-tRFs. Surprisingly, 3′-tRFs did not follow microRNA (miRNA) target rules that are dominated by a so-called "seed" region. Instead, MusD reporter expression was strongly silenced without sequence complementarity to 3′-tRF in the seed region, but disrupted by mutations outside the seed region. This is reminiscent of PIWI-mediated silencing of transposons in the germline. Both terminal RNA modifications of 3′-tRFs and their target site rules suggest that 3′-tRFs are an ancient substrate of the RNA interference machinery that can distinguish self (genes) from nonself (transposons and retroviruses).

## Pseudouridine Guides Germline Small RNA Transport and Epigenetic Inheritance

A.J. Schorn [in collaboration with R.A. Martienssen, CSHL; T. Kouzarides, Gurdon Institute, Cambridge, UK]

Epigenetic modifications that arise during plant and animal development, such as DNA and histone modifications, are mostly reset when gametes are formed,

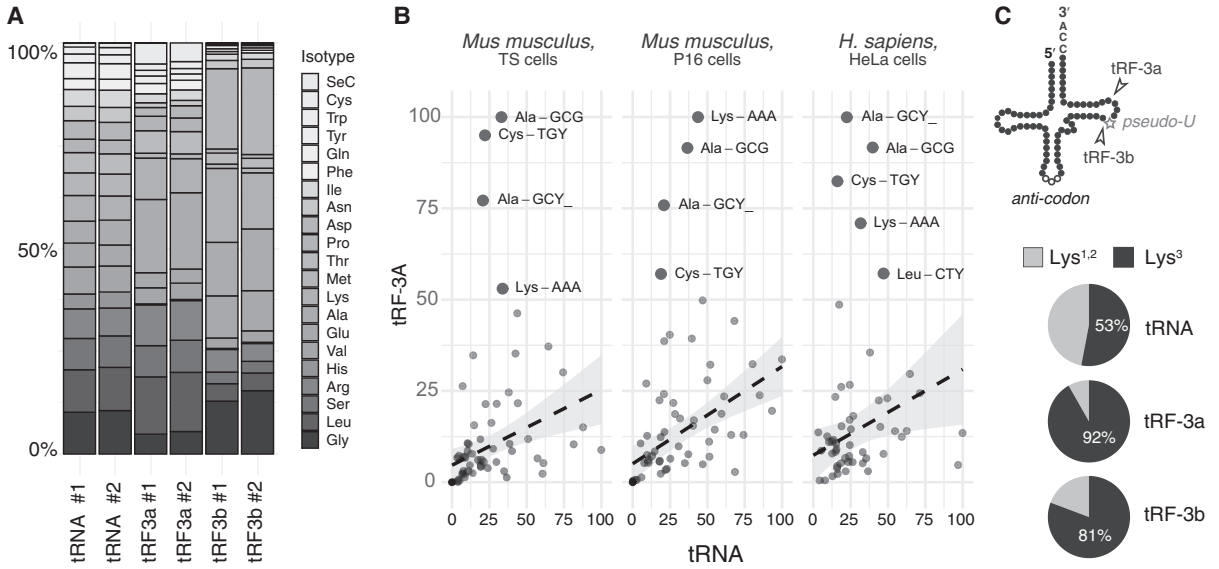



**Figure 1.** Selective enrichment of small RNAs from the 3′-end of tRNAs compared to full-length tRNAs in human and mouse cell lines. (*A*) Quantification of full-length tRNAs by hydrolysis-based sequencing side-by-side with amounts of tRF3a (18 nt) and tRF3b (22 nt) fragments determined by small RNA sequencing. Percent tRNA, tRF-3a, and tRF-3b by amino acid isotype, two replicates each, in a human cancer cell line. (*B*) Enrichment of specific isoacceptors is conserved across mouse and human. (*C*) The cell processes tRF3a and tRF3b fragments from mature tRNAs that contain a posttranscriptional CCA tail. The position of the RNA modification pseudo-uridine (star) is indicated. Lysine isoacceptor ratios differ significantly between tRNAs and tRF.

but some are inherited by the next generation. Small RNAs guide epigenetic modifications and are often transmitted to the next generation. In worms (*Caenorhabditis elegans*), inherited small RNAs have poly(UG) nucleotide tails, but it was unknown how small RNAs are marked for transmission or inheritance in other animals and plants. Pseudo-uridine (pseudo-U) is the most abundant RNA modification in the cell but had not been explored in small RNAs. In this collaborative study, novel assays to detect pseudo-U in short RNA sequences revealed the presence of pseudo-U in miRNAs, their precursors, and germline small RNAs of mouse and the model plant *Arabidopsis.* piRNAs in mouse testes as well as epigenetically activated siRNAs (easiRNAs) in *Arabidopsis* sperm cells show high enrichment for this RNA modification. In *Arabidopsis,* pseudo-U is required for transport of the marked small RNAs into sperm cells and therefore the next generation. The transporter PAUSED/HEN5 (PSD) is the plant homolog of Exportin-t, a nuclear pore protein that transports tRNAs into the cytoplasm. tRNAs and tRF3b fragments are heavily pseudo-uridylated. This suggests that tRF3b and pseudo-uridylated piRNAs in mammals, both of which control transposon activity, might be selected for inheritance and "inform" the next generation or at least be stabilized for intercellular transmission to enforce genome stability in germline tissues.

## PUBLICATION

Herridge RP, Dolata J, Migliori V, de Santis Alves C, Borges F, Van Ex F, Parent JS, Lin A, Bajczyk M, Leonardi T, et al. 2023. Pseudouridine guides germline small RNA transport and epigenetic inheritance. bioRxiv doi:10.1101/2023.05.27.542553

# REGULATION OF GENE EXPRESSION

**D.L. Spector** D. Aggarwal M. Gandhi B. Liu S. Russo W. Xu
B. Balasooriya R. Hazra P. Naik Y. Sahin

Most cellular processes can trace their beginnings to the nucleus, where a gene is activated, resulting in the production of an RNA molecule—some of which encode for proteins, whereas others function as non-protein coding RNAs. Although much biochemical information is available regarding many of the factors involved in gene expression, the spatial and temporal parameters that influence gene expression and the role of noncoding RNAs in regulating this multifaceted process are just beginning to be elucidated. Over the past year our research has continued to focus on identifying and characterizing the role of long noncoding RNAs (lncRNAs) in breast cancer progression, and in nuclear organization and gene expression. Further, a new project has identified lncRNAs up-regulated in glioblastoma multiforme. Following is an overview of some of our accomplishments over the past year.

## Identification of lncRNAs Involved in Breast Cancer Progression

D. Aggarwal, M. Gandhi, B. Liu, S. Russo, Y. Sahin, W. Xu

Large-scale genome-wide studies have revealed that thousands of RNAs that lack protein-coding capacity are transcribed from mammalian genomes. A subset of these noncoding RNAs are greater than 200 nucleotides in length and are referred to as long noncoding RNAs (lncRNAs). With breast cancer being the most frequent malignancy in women worldwide, we aim to identify lncRNAs that play roles in breast cancer progression and to evaluate their mechanism of action and potential as therapeutic targets.

The lncRNA *Metastasis Associated Lung Adenocarcinoma Transcript 1* (*MALAT1*) has been shown to be overexpressed in more than 20 different cancer types, including breast cancer. Over the past year Disha Aggarwal and Suzanne Russo have continued to focus on the role of *MALAT1* by implementing breast cancer organoid take-rate experiments in NOD scid gamma (NSG) mice with the goal of establishing patient-derived-organoid-xenograft (PDO-X) models to examine the impact of *MALAT1* knockdown in vivo on tumor state and metastasis. Thus far, we have achieved a 70% success rate in establishing PDO-X tumors from 11 different validated, fast-growing PDO lines. Further, we have screened 26 different PDO lines for *MALAT1* knockdown (KD) efficiency in vitro using antisense oligonucleotides (ASOs). This data allowed us to select PDO lines with excellent ASO uptake efficiency for the xenograft *MALAT1* KD experiments. In a pilot study in which we injected 50,000 organoids of the PDO line NH85TSc into the mammary glands of NSG mice, we achieved a 100% take-rate. Upon treatment with ASOs (50 mg/kg 2×/week, targeting *MALAT1*), we achieved a 70% *MALAT1* KD and a subsequent ~58% reduction in lung metastatic burden ($n = 3$). Additionally, we observed 1,599 differentially expressed genes and, in collaboration with Adrian Krainer and Kuan-Ting Lin, 815 alternatively spliced genes in the primary tumors. Based on these preliminary results, we are expanding our cohorts to study the impact of knocking down *MALAT1* in multiple PDO-X models. For these experiments, the NSG mice were injected with two to four million cells per mammary gland depending on the in vitro growth rate of the organoid lines. We generated PDO-X models for PDO lines NH85TSc, DS115T, and NH048T with 96%–100% take-rate. These mice were treated with *MALAT1*-targeting ASOs as described above. We achieved an average of 73%, 58%, and 76% KD efficiency in each of the 3 PDO-X in vivo models, respectively ($n = 6–8$ mice per cohort). We are currently evaluating potential changes in gene expression in the primary tumors and changes in lung metastasis. In addition, we are using an affinity-based Cas9-mediated enrichment (ACME) method, in collaboration with Shruti Iyer, Sara Goodwin, Melissa delaBastide, and Richard McCombie, to identify potential mutations and methylation changes in the *MALAT1* gene sequence between tumor and normal organoids.

A second lncRNA that Bodu Liu has been investigating is *Mammary Tumor Associated RNA 42* (*MaTAR42*), which was found to be up-regulated in

MMTV-PyMT and MMTV-Neu-NDL mammary tumors compared with normal mammary epithelial cells. We previously determined that *MaTAR42* was only expressed in a distinct $CD106^+/CD61^+$ population (1%–5%) of mammary tumor cells. This is the result of down-regulation of its transcriptional repressor Runx3 in this population. Elevated *MaTAR42* rendered the cells hypersensitive to hypoxia-induced epithelial–mesenchymal transition (EMT), therefore promoting these cells to metastasize. To study the underlying molecular mechanism of action, we performed RNA antisense purification coupled with mass spectrometry (RAP-MS) analysis. The results identified Casein Kinase 2 alpha (CK2α) to be specifically associated with *MaTAR42*, which was verified by immunoblot as well as RNA immunoprecipitation. Because CK2 is a kinase that can phosphorylate a spectrum of proteins, which function in almost every cellular process, it is crucial to identify the specific target that may be regulated by *MaTAR42*. To do this, we analyzed our single-cell RNA-seq results comparing vector versus *MaTAR42*-overexpressing 4T1 cell mammospheres. We performed a transcription factor enrichment assay to analyze the top 100 genes whose expressions are positively correlated with that of *MaTAR42*. We have identified a c-Myc-transcribed gene signature significantly enriched in these genes (odds ratio = 6.73, $p$ value = $2.09 \times 10^{-9}$). c-Myc is a known master regulator of tumorigenesis, metastasis, and EMT in breast cancer, which resonates with our previous findings. Next, we performed immunoblot analysis that showed a significant up-regulation of c-Myc protein, but not mRNA as assayed by quantitative real-time polymerase chain reaction (qRT-PCR), in 4T1 and 4T07 mouse mammary tumor cells, as well as NMuMG normal mouse mammary epithelial cells, upon ectopic overexpression of *MaTAR42*—suggesting that *MaTAR42* may posttranslationally up-regulate c-Myc. Given that it was previously found by others that c-Myc is a phosphorylation target of CK2 (Yaylim et al., *Asian Pacific J Cancer Prev* **13:** 5233 [2012]), and CK2 activity is required for c-Myc stability in lymphoma cells (Channavajhala and Seldin, *Oncogene* **21:** 5280 [2002]), we hypothesized that *MaTAR42* may enhance c-Myc stability through regulating phosphorylation of c-Myc by CK2. Indeed, inhibition of CK2 activity in *MaTAR42*-overexpressing 4T1 cells abrogated the effect of enhanced c-Myc stability by *MaTAR42*. On the other hand, inhibition of protein degradation by the proteosome inhibitor MG-132 also maintained the level of c-Myc. Both results suggest that *MaTAR42* stabilizes c-Myc protein by regulating its phosphorylation by CK2. To verify this, we are currently performing a phosphorylation assay of c-Myc by CK2 in vitro. Taken together, we have found that *MaTAR42* may regulate the phosphorylation of c-Myc by CK2, thus stabilizing it from protein degradation. The stabilized c-Myc can transcriptionally activate a spectrum of genes, leading to hypersensitivity to hypoxia-induced EMT and metastasis of breast cancer cells.

PDOs provide a very innovative platform to study cancer, as they can recapitulate many aspects of the disease with high fidelity. We previously analyzed the transcriptome of several triple-negative breast cancer (TNBC) and luminal B organoid lines in order to identify lncRNAs that are up-regulated and that represent potential therapeutic targets. Among these, *LINC01235* showed high expression in luminal-progenitor cells of several TNBC organoids and normal organoids. Wenbo Xu showed that down-regulation of *LINC01235* expression, via ASO knockdown or CRISPR-Cas9 knockout in TNBC MDA-MB-468 cells, resulted in a significant decrease in cellular proliferation. Similarly, ASO-mediated KD of *LINC01235* in NH93 and NH95 TNBC organoids results in a reduced rate of organoid formation. We have observed that the expression of *NFIB*, a neighboring gene of *LINC01235*, exhibits a positive correlation in expression with *LINC01235*. Down-regulation of *LINC01235* results in a significant decrease in *NFIB* expression in both MDA-MB-468 cells and NH95 and NH93 organoids. The Notch pathway has been previously identified as being regulated by *NFIB* (Granit et al., *Cell Rep* **24:** 3237 [2018]. We observed that the down-regulation of *LINC01235* correlates with a decrease in the expression of Notch genes, specifically *NOTCH1* and *NOTCH3*. Consequently, our focus is now to unravel the mechanism by which *LINC01235* influences Notch signaling via *NFIB*.

After observing that *LINC01235* is markedly enriched in the chromatin compartment, and that it regulates the expression of *NFIB*, we were interested in determining whether it did so by directly binding to the *NFIB* promoter region. To do so, we employed chromatin isolation by RNA purification (ChIRP) to isolate RNA/DNA complexes, using specific biotin-labeled antisense oligonucleotides targeting

*LINC01235*. The efficacy of RNA pulldown was then evaluated through qRT-PCR analysis, revealing retrieval rates of ~42% for odd probes and 64% for even probes, thus validating the efficiency of the pulldown method.

Subsequently, the DNA samples procured from the pulldown were analyzed using qPCR to assess the enrichment at the target site, specifically focusing on the *NFIB* promoter region, which is 5′ of the *LINC01235* gene. For this analysis, seven pairs of qPCR primers were designed to target various regions within the *NFIB* promoter. Additionally, four pairs of primers were designed to target the *MPDZ* promoter region, a gene situated directly 3′ of *LINC01235*, thereby serving as a control for binding specificity. Two sets of primers targeting the transcription start-site (TSS) of *LINC01235*, within a range of ±200 base pairs, were also designed to serve as a positive control for binding specificity. *GAPDH* was used as a negative control. As expected, the TSS of *LINC01235* emerged as a hotspot for its association. Subsequent qPCR analysis revealed that, relative to *GAPDH* enrichment, two pairs of primers targeting the *NFIB* promoter region (PP3 and PP4) demonstrated ~10-fold higher enrichment. In contrast, the primers targeting the *MPDZ* promoter region exhibited enrichment levels similar to those of the negative control *GAPDH*. These results demonstrate that *LINC01235* RNA specifically binds to the *NFIB* promoter region but not to the *MPDZ* promoter region, thereby highlighting its potential role in the regulation of *NFIB* expression. Based on these findings, we propose a model whereby *LINC01235* directly interacts with the *NFIB* promoter, positively influencing *NFIB* transcriptional expression. Subsequently, NFIB modulates the downstream NOTCH pathway, impacting the proliferation of luminal progenitor (LP)-like cells during breast cancer progression.

To identify additional lncRNAs associated with TNBC progression, using our RNA-seq expression data from TNBC and normal PDOs, Yunus Sahin developed a new analysis pipeline. This pipeline begins with differential gene expression analysis using TNBC organoids, the Cancer Genome Atlas (TCGA) TNBC patient data, and their respective normal counterparts. We utilized the latest genome annotation (GENCODE v43) to identify up-regulated lncRNAs in TNBC compared with normal tissue. From this analysis, we identified 853 and 2,107 differentially expressed lncRNAs in the TNBC organoids and TCGA TNBC patient data sets, respectively, with a $\log_2$-fold change of >1 and an adjusted $p$ value of 0.05. By overlapping these two lncRNA lists, we obtained 245 common lncRNAs that are up-regulated in both TNBC organoids and TCGA TNBC patient samples. We classified the common lncRNA list based on their genomic location, resulting in 138 intergenic lncRNAs among the 245 identified.

Our goal is to perform a CRISPR screen of the 138 intergenic lncRNAs. We have chosen the CRISPR interference system for this screening approach because of its low off-target ratio and high knockdown efficiency without altering chromatin structure. Currently, we are generating organoid lines stably expressing ZIM3 KRAB-dCas9 and testing the system's efficacy using control gRNAs (essential genes, lncRNAs, positive controls, and negative controls). Simultaneously, we are working on preparing a gRNA library. After verifying the CRISPR system's stability, we will perform a viability screen using this library.

In addition. we are currently collaborating with Peter Koo and Kaeli Rizzo to generate a prognostic lncRNA signature using machine-learning algorithms. We plan to test this signature using the SCAN-B TNBC data set and perform further functional wet lab experiments.

Invasive lobular carcinoma (ILC) is the second most frequently diagnosed histologic subtype of breast cancer, representing 10%–15% of diagnosed invasive breast cancers, and it is hormonally driven by estrogen and progesterone. Our goal is to identify novel estrogen-responsive lncRNAs and to investigate their molecular function and their therapeutic potential in ILC. As a first step, Minakshi Gandhi performed transcriptome profiling using validated ILC PDOs ($n = 12$) and normal breast organoid models ($n = 14$) to identify differentially expressed lncRNAs in ILC. We have identified 3,395 differentially expressed lncRNAs, of which 1,111 were significantly ($p$ value < 0.05) up-regulated (>1.5 fold change [FC]) and 1,824 were significantly ($p$ value < 0.05) down-regulated (<0.5 FC). We are in the process of designing and using CRISPR-Cas13 screening modality to silence the up-regulated lncRNAs to identify lncRNAs essential for viability in ILC using our established and validated PDO lines (in collaboration with the Sanjana laboratory at The New York Genome Center).

In addition, we have designed and expanded a drug panel consisting of estrogen receptor (ER) signaling modulators from $n = 8$ drugs last year to $n = 25$

standard-of-care drugs used for treating ILC patients in the clinical setting, which includes different classes of drugs such as estrogen antagonists, selective estrogen receptor modulators (SERMs), and selective estrogen receptor down-regulators (SERDs)—including some recently approved targeted therapy drugs. As a first step, we tested different drug concentrations at multiple time points, validated the down-regulation of ESR1 using qRT-PCR, and chose the lowest possible concentration of each respective drug for treating ILC organoids. Next, as proof of concept, we treated a validated ILC organoid line, LNS055, with the ER signaling modulators and isolated total RNA for RNA-seq analysis to identify lncRNAs responsive to modulation of estrogen signaling. We identified multiple estrogen-responsive lncRNAs and one lncRNA, ENSG00000223808, stood out as the most promising candidate with expression only in breast cancer data sets from TCGA and no expression in normal breast samples. We have now received $n = 30$ ASOs from Ionis Pharmaceuticals for screening to silence this lncRNA to further probe its cellular function and impact on hallmark cancer processes such as cell viability/cell proliferation, etc. We are now in the process of validating the knockdown efficiency and ultimately hope to identify at least two to three ASOs with optimum silencing efficiency from this ASO cohort.

In addition, we have been establishing ILC patient-derived organoid xenograft (PDO-X) models in NOD scid gamma mice. We have now concluded an initial pilot experiment using 2 ILC models to determine the take-rate of different ILC organoids in mice and to optimize the number of organoids per injection. We also initially identified promising results using ultrasound, in collaboration with Scott Lyons and Joseph Merrill; we noticed development of cysts and small tumors in the mammary glands of a small subset of mice while the tumors were still unpalpable. We have validated the initial ultrasound findings ex vivo using a histological approach by immunolabeling with an antihuman mitochondrial antibody that selectively differentiates injected human organoids from mouse tissues. Other than optimizing the injection volume and number of organoids, we have also established that the injected ILC organoids recapitulated the single-file growth pattern (displayed due to loss of the *CDH1* gene in ILC patients) of parent ILC tumors in mice. Next, we will focus on identifying additional ILC PDOs that can recapitulate the single-file growth pattern of ILC. We propose to utilize these established models in combination with advanced imaging modalities (positron emission tomography computed tomography [PET-CT]/contrast-enhanced magnetic resonance imaging [MRI]), in collaboration with Scott Lyons and Joseph Merrill, to assess the impact of lncRNA silencing in vivo on tumor formation and metastasis using a previously established ASO-based lncRNA targeting approach.

## Identification of Glioblastoma Stem Cell–Associated lncRNAs Using Single-Cell RNA-Sequencing Data Sets

R. Hazra, P. Naik [in collaboration with R. Utama and A. Dobin, CSHL]

Glioblastoma multiforme (GBM), a World Health Organization grade IV glioma, ranks as the deadliest primary malignant brain cancer, with nearly 25,700 new cases diagnosed each year in the United States. The currently available treatment regimen comprises maximal surgical resection and a combination of radiotherapy and chemotherapy, which extends patient survival to a median of only 14.6 months. Despite recent progress in understanding the tumor's biology and multimodal therapy options, GBM is one of the most treatment-resistant malignancies, and even after successful treatment, the tumors eventually recur. Glioblastoma stem cells (GSCs) initiate the tumor and are known culprits of therapy resistance. Rasmani Hazra, in collaboration with Raditya Utama and Alexander Dobin, carried out a comprehensive analysis of the GSC transcriptome using scRNA-seq databases of GBM tumors, tumor-derived organoids, and normal developing human brain. We identified a total of 1,426 expressed lncRNAs in GBM tumors, among which 374 were uncharacterized lncRNAs associated with GSCs. The expression and subcellular localization of four selected lncRNAs were validated by qRT-PCR, single-molecule RNA fluorescence in situ hybridization (FISH), and subcellular fractionation. The radial glia and olfactory progenitor cell (OPC) GSC-enriched lncRNA *GIHCG* was found to be highly expressed in GSC lines and localized to both the cytoplasmic and nuclear fractions. In contrast, the neuronal GSC-enriched lncRNA *LINC01563* is highly enriched in the nuclear fraction of GSCs. In addition, as proof of principle, we independently

depleted the expression of two lncRNAs (*GIHCG* and *LINC01563*) and revealed that they each promote GSC proliferation, migration, and maintenance of stemness. In summary, these analyses identified a large number of lncRNAs associated with the GSCs' transcriptome, which represent candidates to be pursued at the functional level and prioritized as potential therapeutic targets in GBM.

## A Subset of Guanine- and Cytosine-Rich Genes Are Actively Transcribed at the Nuclear Lamin B1 Region

B. Balasooriya

Chromatin organization in the mammalian cell nucleus plays a vital role in the regulation of gene expression. The lamina-associated domain at the inner nuclear membrane has been shown to predominantly harbor heterochromatin, whereas the nuclear interior has been shown to contain the majority of the euchromatin. Gayan Balasooriya carried out immunolabeling of mouse embryonic stem cells (mESCs) and in vitro–derived OPCs for Lamin B1 and RNA Pol II-pSer2, a signature for actively transcribing genes. In addition to being distributed at internal nuclear regions, surprisingly a considerable amount of RNA Pol II pSer2 signals were also localized at the Lamin B1 region in mESCs and OPCs. To identify and characterize the genes that are uniquely expressed at the Lamin B1 region and identify the molecular characteristics of these genes, we designed a sequential chromatin immunoprecipitation (ChIP) assay using Lamin B1 and RNA Pol II pSer2 antibodies. First, Lamin B1–associated chromatin was immunoprecipitated and then the transcriptionally active chromatin in the Lamin B1 fraction was immunoprecipitated using an antibody to RNA Pol II pSer2. DNA sequencing (150PE) using the Illumina HiSeq sequencing platform revealed that ~2.7% of the actively transcribed genes in mESCs are transcribed at the Lamin B1 region. In OPCs, we found that the total number of actively transcribing genes declined by nearly 28%. Surprisingly, in OPCs, ~52% of the total actively transcribing genes were expressed solely at the Lamin B1 region. These nuclear periphery–associated actively transcribing genes primarily represent housekeeping genes, and their gene bodies are significantly enriched with guanine and cytosine compared to genes actively transcribed at the nuclear interior. We found the promoters of these genes to also be significantly enriched with guanine and to be predominantly regulated by zinc finger protein transcription factors.

### PUBLICATIONS

Aggarwal D, Russo S, Naik P, Bhatia S, Spector DL. 2023. Establishment and culture of patient-derived breast organoids. *J Vis Exp* **192:** e64889. doi:10.3791/64889

Hazra R, Utama R, Naik P, Dobin A, Spector DL. 2023. Identification of glioblastoma stem cell–associated lncRNAs using single-cell RNA-sequencing datasets. *Stem Cell Reports* **18:** 2056–2070.

Mattick JS, Amaral PP, Carninci P, Carpenter S, Chang HY, Chen L-L, Chen R, Dean C, Dinger ME, Fitzgerald KA, et al. 2023. Long non-coding RNAs: definitions, functions, challenges and recommendations. *Nat Rev Mol Cell Biol* **24:** 430–447.

Mousset A, Lecorgne E, Bourget I, Lopez P, Jenovai K, Cherfils-Vicini J, Dominici C, Rios G, Girard-Riboulleau C, Liu B, et al. 2023. Neutrophil extracellular traps formed during chemotherapy confers treatment resistance via TGFβ activation. *Cancer Cell* **41:** 757–775. doi:10.1016/j.ccell.2023.03.008

*In Press*

He X-Y, Gao Y, Ng D, Michalopoulou E, George S, Adrover JM, Sun L, Albrengues J, Dassler-Plenker J, Han X, et al. 2024. Chronic stress increases metastasis via neutrophil-mediated changes in the microenvironment. *Cancer Cell* **42:** 474–486.

# DNA REPLICATION AND CHROMATIN INHERITANCE

**B. Stillman** K. Bhalla, A. Cocozzelli, K. Hanes, M. Hossain, Y. Hu, G. McDermott, S. Naik, M. Quyang, Y-J. Sheu, N. Zali

Our studies of DNA replication in eukaryotic cells have primarily focused on using the budding yeast *Saccharomyces cerevisiae* as a model organism, but it became apparent that this yeast and a small group of related budding yeasts determine the location of origins of DNA replication via DNA sequence–specific binding of the origin recognition complex (ORC) to a defined DNA sequence located in the intergenic regions of the genome. *S. cerevisiae* has about 6,500 genes and very little DNA other than the promoter regions that determine the control of gene expression. Thus, it has a gene-dense genome. Based on prior studies, we were convinced that nearly all eukaryotes—including plants, animals, and the vast majority of fungi, including yeasts—specified the location of origins of DNA replication very differently from *S. cerevisiae*. In addition to our long-standing studies on the initiation of DNA replication in human cells, we selected a budding yeast that has a common ancestor with *S. cerevisiae* that existed ~300 million years ago (equivalent to the evolutionary distance between modern humans and primitive reptile-like animals in the Permian period). We initially chose to study the yeast *Yarrowia lipolytica* because we predicted that it would have a very different mechanism of specifying the location of origins of DNA replication in its genome. The genome in *Y. lipolytica* is ~1.8× larger than the genome of *S. cerevisiae* but has the same number of protein-coding genes present in six chromosomes. Furthermore, most genes in *Yarrowia* have introns, unlike genes in *S. cerevisiae*, and therefore, *Yarrowia* has large tracks of noncoding DNA.

Yeasts do not have the capacity to convert exogenous thymidine or thymidine analogs into the triphosphate nucleotide because they lack the thymidine kinase enzyme. To label *Yarrowia* with the thymidine analogs 5-ethynyl-2′-deoxyuridine (EdU) or bromodeoxyuridine (BrdU), we expressed the herpes simplex thymidine kinase (TK) gene and the human ENT1 nucleoside transporter genes (called the BrdU-Inc cassette) in the MatA and MatB strains of *Yarrowia* we are using. We integrated the BrdU-Inc cassette into the *Yarrowia* genome at the nonessential *APX* gene locus using selected homologous recombination. A detailed analysis of the location of origins of DNA replication and the temporal patterning in *Yarrowia* has been completed. The most informative approach has been to starve cells to arrest them in $G_1/G_0$ and then release them into fresh media containing high glucose. This avoids converting the cells into hyphal growth, which makes it difficult to study replication patterning. Cells were also released into media containing hydroxyurea (HU) at levels that prevent extensive replication fork elongation but allow a slow activation of all origins of DNA replication over time. This analysis has revealed that *Yarrowia* has 634 origins, more than the ~400 that are present in *S. cerevisiae* (Fig. 1).

Analysis of the DNA sequences that lie under the peaks of EdU incorporation show that the origins are very G·C rich, much more than the A·T rich origins present in *S. cerevisiae* and more like the G·C rich origins of DNA replication in human cells. In addition to EdU-Seq, we constructed a plasmid that contains a *Yarrowia* centromere sequence and a selectable auxotrophic maker into which short fragments of the *Yarrowia* genome were cloned to search for sequences that confer high-frequency transformation and hence are potential origins of DNA replication *(ORI)*. Without an origin, the plasmid does not transform *Yarrowia* cells. Unlike *S. cerevisiae,* in which origins of DNA replication are sufficient to confer autonomous replication *(ARS)* activity on a plasmid, *Yarrowia ARS* sequences require both a *CEN* and *ORI.* We selected a number of sequences under the initiation peaks found in whole-genome analysis of origins of DNA replication and, when cloned into a plasmid containing a *CEN* sequence, the plasmid transformed yeast at high frequency, indicative of a functional *ORI*. In contrast, insertions of *Yarrowia* DNA sequences from the genome that were not under an EdU peak failed

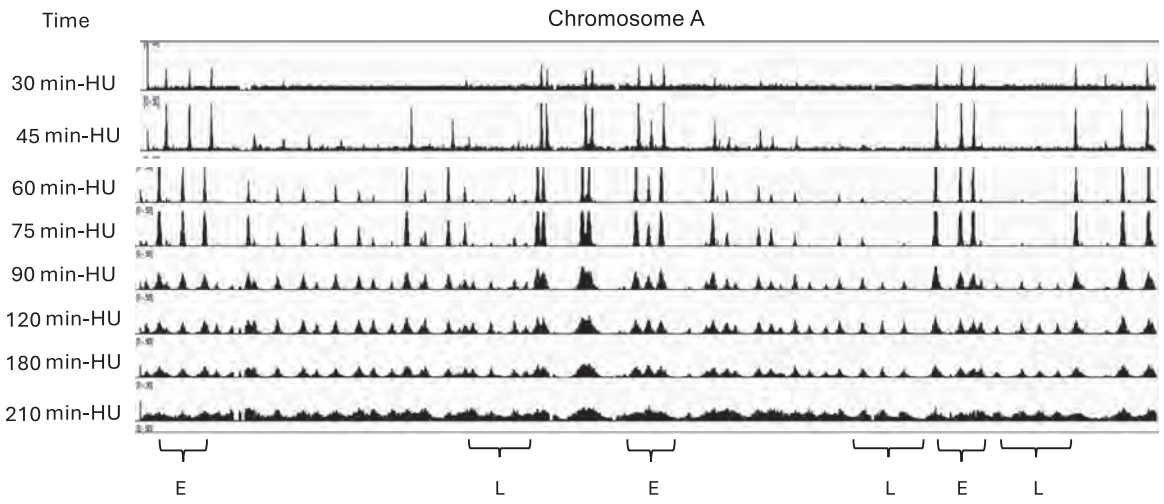



**Figure 1.** Time course of DNA replication in the budding yeast *Yarrowia lipolytica* chromosome A. Cells were arrested in $G_1/G_0$ by starvation and released into complete medium in the presence of 5 mM hydroxyurea that slows the DNA replication forks. Cells were continuously labeled with EdU, and then the DNA was sheared, the EdU containing DNA was isolated by click chemistry, and the DNA was sequenced. The replication pattern in chromosome A is shown. (HU) Hydroxyurea.

to transform yeast at high frequency when cloned into a *CEN* plasmid. Thus, the *ARS* assay faithfully reflects active origins of DNA replication.

The absence of an obvious conserved sequence in the *Y. lipolytica ORI*s resulted in a systematic genetic analysis of multiple *ORI*s using DNA linker-scan mutagenesis. The scan of two origins, one near a centromere (*ORI-C-061*) and another selected from the left arm of chromosome A (*ORI-A-006*) (the *ORI*s were systematically named with the chromosome number [in this case Chromosome A] and the number of the *ORI* from the left telomere). The linker scans revealed that both *ORI-C-061* and *ORI-A-006* have an essential region of ~30 base pairs that are required to support the high-frequency transformation of yeast with the plasmid. These results suggest that there is an essential element in *Y. lipolytica ORIs*. In cases in which transformation yielded colonies, the stability of the plasmid was determined by measuring the percentage of cells retaining a selectable marker after growth for four days under selection for an auxotrophic marker. All linker-scan mutations outside of the essential region retained wild-type plasmid stability, further emphasizing that the ~30-base-pair region is necessary. Indeed, a 50-base-pair DNA sequence containing the essential region was sufficient for high-frequency transformation in the presence of a centromere. With these results, a comparison of essential sequences was possible that revealed a weak, short consensus sequence present in the two origins. However, most single point mutations in these base pairs did not affect the high-frequency transformation of the plasmid, and double or triple mutations were required to inactivate the *ORI*.

Next, in collaboration with Justin Kinney and Kaiser Loell, each base pair in a 90-base-pair region surrounding the essential region in both *ORI-A-006* and *ORI-C-061* was subjected to insertion of mutations at a 15% substitution frequency across the entire 90 base pairs and then the resulting library of mutations was subjected to selection in *Yarrowia* for multiple generations of cell division. The plasmid DNAs that were selected, which select for efficient origins, were sequenced and analyzed for conserved motifs. A weak consensus sequence $YATR_{NNNNNN}C$ $.AWTT_{NNNNNN}Y_NYAA$ (N = any nucleotide, R = purine, W = A or T, Y = pyrimidine) was identified and subjected to point mutation analysis. Mutations in some of the highly selected base pairs either affected high-frequency transformation or plasmid stability.

In parallel, biochemical and structural studies were performed in collaboration with Leemor Joshua-Tor and Jack Bauer. The interaction of purified ORC and Cdc6 with each origin was studied. ORC and Cdc6 were expressed in recombinant baculovirus-infected insect cells and both DNA binding and cryo-EM

structure analysis of the proteins with and without DNA were undertaken. Structures of *Y. lipolytica* ORC in the absence of DNA and structures of ORC and Cdc6 bound to DNA were obtained, using the origin DNA sequences described above. The results show differences between *Yarrowia* and *Saccharomyces* ORC structures, with *Yarrowia* more similar to the human ORC that we have undertaken previously.

Very little basic biology has been performed using *Yarrowia*, and therefore we have spent considerable time in making reagents for the study of DNA replication. Monoclonal or polyclonal antibodies were produced that target the five ORC (1–5) subunits of the origin recognition complex and Cdc6. Polyclonal antibodies have been made against the Orc2 and Orc6 subunits, and antibodies are in preparation targeting Cdt1 and three subunits (Mcm2, Mcm3, and Mcm7) of the Mcm2–7 complex, so that we have antibodies against the proteins that form the pre-replicative complex.

Since we do not have methods to synchronize cells in $G_1$ phase, we have starved cells to arrest them in $G_1/G_0$ and then released them into media with high glucose (also see Fig. 1). So far, preliminary analysis of the time course as cells progress from $G_1/G_0$ into the cell cycle suggests that the Orc1 subunit may be degraded after entry into S phase and resynthesized just before mitosis. This is similar to the pattern of the ORC1 subunit in human cells that we identified many years ago. We are following up on this observation to determine the temporal dynamics of ORC, Cdc6, and Cdt1 in *Y. lipolytica* and examine how and when these proteins bind to DNA, including using protein localization studies. We are also investigating methods of arresting *Yarrowia* cells in $G_1$ phase or synchronizing them at other phases of the cell division cycle.

A retrospective comparison of how the mechanism of initiation of DNA replication was discovered in bacteria and eukaryotes showed that although it was possible to establish cell-free extracts from bacteria that could support the initiation of DNA replication from the bacterial *OriC* origin, cell extracts from eukaryotic cells could not support the initiation of DNA replication from yeast origins. The principal difference is that in eukaryotic cells, the initiation of DNA occurs via a two-step process during which the first step, formation of pre-RCs at each origin, occurs first in an environment that lacks the cyclin-dependent protein kinases (CDKs). This assembly of pre-RCs is inhibited by CDKs, but then CDKs activate the pre-RC to promote initiation of DNA synthesis. Thus, a cell extract that has active CDKs cannot perform the first step. This dual process ensures accurate and even duplication of the entire genome from the multiple origins of DNA replication present on chromosomes in eukaryotic cells. In contrast, bacteria have a single origin of DNA replication, and it is therefore possible to establish origin recognition and initiation of DNA syntheses in a single cell extract.

## PUBLICATIONS

Hu Y, Stillman B. 2023. Origins of DNA replication in eukaryotes. *Mol Cell* **83:** 352–372. doi:10.1016/j.molcel.2022.12.024

*In Press*

Stillman BW. 2024. Establishing a biochemical understanding of the initiation of chromosome replication in bacteria. *Proc Natl Acad Sci* **121:** e2400667121. doi:10.1073/pnas.2400667121

Wilson MC, Cheeger J, Lawson B, Lourie R, Bluestein Simons B, Neuwirth L, Patterson N, Kra I, Phillips T, et al. Jim Simons 1938–2024. *Not Am Math Soc* **72:** 32–47.

# CANCER DRUG TARGET DISCOVERY AND MECHANISMS

**C.R. Vakoc** L. Almeida, A. Alpsoy, P. Cunniff, S. Espinosa, C. Fitzpatrick, Y. Gao, V. Kechejian, O. Klingbeil, P. Kumar, J. Liverpool, C. Lopez-Cleary, D. Maia-Silva, J. Milazzo, B. Nalbant, S. Pal, L. Shanley, D. Skopelitis, M. Sroka, K. Tam, K. Taneja, Y. Wang, X. Wu, T. Yoshimoto

Massive genome-wide reprogramming of transcription is critical for malignant transformation. As a consequence, cancer cells are vulnerable to perturbations of the transcriptional apparatus, which includes targeting of DNA-binding transcription factors/cofactors and chromatin regulatory machineries. Over the past decade, our laboratory has taken a genetic screening approach to identify transcriptional dependencies in cancer cell lines. Upon identifying cancer-specific patterns of essentiality, we have pursued detailed molecular mechanisms that underpin these cellular phenotypes. By understanding transcriptional dependencies in cancer, we have revealed fundamental mechanisms of gene control, novel processes that drive cancer formation, and new therapeutics that reprogram transcription to eliminate cancer cells. The broad goals of our current research are (1) to identify novel cancer-specific dependencies and evaluate underlying mechanisms, (2) to reveal detailed molecular mechanisms of lineage master regulator transcription factors that drive cancer cell growth, (3) to develop chemical probes that modulate the function of lineage master regulators, and (4) to explore how lineage cell-of-origin and trans-differentiation processes contribute to the pathogenesis and therapy of human tumors.

## ASCL1, POU2F3, and NRF2 as Onco-Transcription Factor Dependencies in Lung Cancer

A. Alpsoy, S. Espinosa, C. Fitzpatrick, P. Kumar, S. Pal, L. Shanley, D. Skopelitis, K. Taneja, X. Wu

Tuft cells belong to a rare chemosensory lineage that coordinates immune and neural responses to foreign pathogens in mucosal tissues. Recent work from our laboratory revealed tuft-cell-like human tumors, particularly as a variant of small-cell lung cancer. Both normal and neoplastic tuft cells share a genetic requirement for the transcription factor POU2F3, although the transcriptional mechanisms that generate this cell type are poorly understood. In a recent study published in 2022, we showed that binding of POU2F3 to the uncharacterized proteins C11orf53 and COLCA2 (renamed here OCA-T1/POU2AF2 and OCA-T2/POU2AF3, respectively) is critical in the tuft cell lineage. OCA-T1 and OCA-T2 are paralogs of the B-cell-specific coactivator OCA-B; all three proteins are encoded in a gene cluster and contain a conserved peptide that binds to class II POU transcription factors and a DNA octamer motif in a bivalent manner. We demonstrated that binding between POU2F3 and OCA-T1 or OCA-T2 is essential in tuft-cell-like small-cell lung cancer (SCLC). Moreover, we generated OCA-T1-deficient mice, which are viable but lack tuft cells in several mucosal tissues. These findings reveal that the POU2F3–OCA-T complex is the master regulator of tuft cell identity and a molecular vulnerability of tuft-cell-like small-cell lung cancer. In ongoing research, we have solved the X-ray crystal structure of POU2F3 bound to OCA-T1 and to OCA-T2. We have also generated OCA-T2 knockout mice and characterized their tuft cell phenotypes. Finally, we are generating a genetically engineered mouse model of tuft-cell-like human tumors by expressing Cre recombinase in tuft cells together with floxed alleles of *Trp53, Rb1, Rb2, Myc,* and *Pten.*

In the classical neuroendocrine form of SCLC, tumors express and depend on ASCL1 for their viability. Thus, we hypothesize that molecular mechanisms of ASCL1 might inform novel targeted therapy approaches for the neuroendocrine subtype of SCLC. In a new project, we have devised a genetic screening strategy that seeks to expose ASCL1 cofactors in an unbiased manner. For this purpose, we have developed an intracellular fluorescence-activated cell sorting

(FACS) stain for DLL3 expression as a molecular reporter, a gene highly expressed with ASCL1 in SCLC and highly expressed in normal embryonic neuroendocrine cells. We have carried out genomewide CRISPR screens to identify all possible knockouts that cause reduced DLL3 expression in SCLC. Notably, the top hit from this screen in addition to ASCL1 was the transcription factor POU2F1. In our ongoing work, we are investigating the mechanism by which POU2F1 cooperates with ASCL1 at distal enhancer elements using epigenomics. In addition, we are characterizing the differentiation phenotypes of POU2F1 knockout SCLC cells and, potentially in the future, neuroendocrine cells and neural progenitors.

We recently initiated a project that seeks to investigate the function of NRF2 as a transcription factor dependency in KEAP1-mutant lung adenocarcinoma (LA). We have performed a NRF2 reporter screen in search of novel regulators of NRF2 function. These screens validated the known roles of mSWI/SNF and Mediator as NRF2 coactivators. Unexpectedly, these screens nominated the CUL3 ubiquitin ligase as supporting NRF2 function in a KEAP1-independent manner. In a parallel effort, we have uncovered an absolute requirement for the NEH3 domain of NRF2 to support lung cancer proliferation.

## Pediatric Sarcomas Caused by Fusion Oncoproteins

L. Almeida, Y. Gao, C. Lopez-Cleary, M. Sroka, A. Vasudevan, K. Tam, T. Yoshimoto

Recurrent chromosomal rearrangements found in rhabdomyosarcoma (RMS) produce the PAX3-FOXO1 fusion protein, which is an oncogenic driver and a dependency in this disease. One important function of PAX3-FOXO1 is to arrest myogenic differentiation, which is linked to the ability of RMS cells to gain an unlimited proliferation potential. Here, we developed a phenotypic screening strategy for identifying factors that collaborate with PAX3-FOXO1 to block myo-differentiation in RMS. Unlike most genes evaluated in our screen, we found that loss of any of the three subunits of the nuclear factor-Y (NF-Y) complex leads to a myo-differentiation phenotype that resembles the effect of inactivating PAX3-FOXO1. Although the transcriptomes of NF-Y- and PAX3-FOXO1-deficient RMS cells bear remarkable similarity to one another, we found that these two transcription factors occupy nonoverlapping sites along the genome: NF-Y preferentially occupies promoters, whereas PAX3-FOXO1 primarily binds to distal enhancers. By integrating multiple functional approaches, we map the *PAX3* promoter as the point of intersection between these two regulators. We show that NF-Y occupies CCAAT motifs present upstream of *PAX3* to function as a transcriptional activator of PAX3-FOXO1 expression in RMS. These findings reveal a critical upstream role of NF-Y in the oncogenic PAX3-FOXO1 pathway, highlighting how a broadly essential transcription factor can perform tumor-specific roles in governing cellular state.

In addition, we are investigating the mechanism of MYOD1 lineage dependency in RMS. Using a marker-based genetic screen, we revealed the E proteins (TCF12, TCF3, TCF4) as performing a redundant function in support of MYOD1 in this tumor context. This contradicts prior work showing that MYOD1 is nonfunctional in RMS, and leads to a novel hypothesis that RMS cells are addicted to the MYOD1-E protein complex for tumor viability.

The EWS-FLI1 fusion in Ewing sarcoma is also of interest to our laboratory. We are attempting to establish a molecular reporter of these fusion proteins to be assessed by CRISPR screening. In addition, we are pursuing mechanistic studies of MyoD, which has a powerful lineage dependency in rhabdomyosarcoma cells. The EWS-FLI1 fusion oncoprotein deregulates transcription to initiate the pediatric cancer Ewing sarcoma. Here, we used a domain-focused CRISPR screen to implicate the transcriptional repressor ETV6 as a unique dependency in this tumor. Using biochemical assays and epigenomics, we show that ETV6 competes with EWS-FLI1 for binding to select DNA elements enriched for short GGAA repeat sequences. Upon inactivating ETV6, EWS-FLI1 overtakes and hyperactivates these *cis*-elements to promote mesenchymal differentiation, with *SOX11* and *NRTK1* being key downstream targets. We show that squelching of ETV6 with a dominant-interfering peptide phenocopies these effects and suppresses Ewing sarcoma growth in vivo. These findings reveal a strategy for neutralizing the EWS-FLI1 oncoprotein by reprogramming its genomic occupancy.

Using a marker-based genetic screen, we have identified the cohesion loader complex NIPBL/MAU2 as a coactivator of EWS-FLI1 in Ewing sarcoma. Our epigenomic studies indicate that EWS-FLI1 is

necessary and sufficient to direct NIPBL/MAU2 to microsatellite sequences. In addition, Ewing sarcoma cells appear to be hypersensitive to genetic targeting of this complex. Our ongoing work seeks to map the biochemical interface that connects these regulators.

## Pancreatic Cancer Dependencies KLF5, p63, and MARK2/3

P. Cunniff, O. Klingbeil, D. Maia-Silva

One limitation of our previous CRISPR screening strategy is that only single genes are inactivated in our pooled genetic screens. Importantly, evolution often produces novel genes via duplication events, which can produce gene pairs that function redundantly to support cellular functions. We have been concerned about whether redundancy conceals essential gene functions in our essentiality screens. To address this issue, we developed a CRISPR screening strategy in which two single-guide RNAs (sgRNAs) are expressed from a single lentiviral vector backbone. This allows us to produce single and double knockouts within a single genetic screen. We generated sgRNA libraries that co-target homologous kinase, phosphatase, and chromatin-modifying enzymes. These studies led us to make the discovery that several carcinoma cell lines are dependent on MARK2/MARK3, which function in a redundant manner to support cancer cell line growth. Notably, several hematopoietic and neuroendocrine lineage tumor lines do not require MARK2/MARK3 for survival. We have recently found that MARK2/3 are critical to sustain the function of YAP/TAZ in human cancer and have shown that NF2 and YAP/TAZ are critical substrates of MARK2/3 that underlie this co-dependency relationship. Underlying this observation is direct MARK2/3-dependent phosphorylation of NF2 and YAP/TAZ, which effectively reverses the tumor-suppressive activity of the Hippo module kinases LATS1/2. To simulate targeting of MARK2/3, we adapted the CagA protein from *Helicobacter pylori* as a catalytic inhibitor of MARK2/3, which we show exerts antitumor activity in vivo. Together, these findings reveal MARK2/3 as powerful co-dependencies of YAP/TAZ in human cancer—targets that may allow for pharmacology that restores Hippo pathway–mediated tumor suppression.

The presence of basal lineage characteristics signifies hyperaggressive human adenocarcinomas of the breast, bladder, and pancreas. However, the biochemical mechanisms that maintain this aberrant cell state are poorly understood. Here we performed marker-based genetic screens in search of factors needed to maintain basal identity in pancreatic ductal adenocarcinoma (PDAC). This approach revealed MED12 as a powerful regulator of the basal cell state in this disease. Using biochemical reconstitution and epigenomics, we show that MED12 carries out this function by bridging the transcription factor p63, a known master regulator of the basal lineage, with the Mediator complex to activate lineage-specific enhancer elements. Consistent with this finding, the growth of basal-like PDAC is hypersensitive to MED12 loss when compared with classical PDAC. Taken together, our comprehensive genetic screens have revealed a biochemical interaction that sustains basal identity in human cancer, which could serve as a target for tumor lineage–directed therapeutics.

An additional target in PDAC under investigation is KLF5, a zinc finger transcription factor. Using a combined KLF5 reporter screen coupled with proteomic analysis of KLF5 complexes, we have nominated the RUVBL1/RUVBL2 complex as a coactivator of KLF5 in PDAC. We have effectively applied epigenomics methods to demonstrate that RUVBL1/2 is a catalytic coactivator of KLF5 in PDAC. Leveraging existing small molecules, we have found that RUVBL1/2 can be effectively targeted with small molecules to suppress KLF5 function in PDAC. Together, our findings suggest that KLF5 is a targetable dependency in PDAC.

## SCP4-STK35/PDIK1L Complex Is a Dual Phospho-Catalytic Signaling Dependency in Acute Myeloid Leukemia

A. Alpsoy, J. Liverpool, Y. Wang

Acute myeloid leukemia (AML) cells rely on phospho-signaling pathways to gain unlimited proliferation potential. In a study published in 2022, we used domain-focused CRISPR screening to identify the nuclear phosphatase SCP4 as a dependency in AML. Importantly, we have shown that this enzyme is dispensable in normal human hematopoietic progenitor cells. Using CRISPR exon scanning and gene complementation assays, we found that the catalytic function of SCP4 is essential in AML. Through mass spectrometry analysis of affinity-purified complexes, we identified

the kinase paralogs STK35 and PDIK1L as binding partners and substrates of the SCP4 phosphatase domain. We showed that STK35 and PDIK1L function catalytically and redundantly in the same pathway as SCP4 to maintain AML proliferation and to support amino acid biosynthesis and transport. We obtained evidence that SCP4 regulates STK35/PDIK1L through two distinct mechanisms: catalytic removal of inhibitory phosphorylation and by promoting kinase stability. Our findings reveal a phosphatase-kinase signaling complex that supports the pathogenesis of AML.

## Lineage Dependencies SOX10 in Melanoma and PAX8 in Ovarian Cancer

V. Kechejian, B. Nalbant

We recently initiated projects aimed at validating and genetically interrogating some of the most powerful lineage dependencies in cancer nominated by genetic screening. By analyzing the DepMap database, we prioritized SOX10 in melanoma and PAX8 in ovarian cancer as compelling opportunities for therapeutic targeting. We have established gene complementation assays for both genes in the relevant cancer context, and are now fine-mapping the critical surfaces of these molecules that drive cancer proliferation.

### PUBLICATIONS

Fiskus W, Mill CP, Birdwell C, Davis JA, Das K, Boettcher S, Kadia TM, DiNardo CD, Takahashi K, Loghavi S, et al. 2023. Targeting of epigenetic co-dependencies enhances anti-AML efficacy of Menin inhibitor in AML with MLL1-r or mutant NPM1. *Blood Cancer J* **13:** 53. doi:10.1038/s41408-023-00826-6

Gao Y, He XY, Wu XS, Huang YH, Toneyan S, Ha T, Ipsaro JJ, Koo PK, Joshua-Tor L, Bailey KM, et al. 2023. ETV6 dependency in Ewing sarcoma by antagonism of EWS-FLI1-mediated enhancer activation. *Nat Cell Biol* **25:** 298–308. doi:10.1038/s41556-022-01060-1

Hur SK, Somerville TDD, Wu XS, Maia-Silva D, Demerdash OE, Tuveson DA, Notta F, Vakoc CR. 2023. p73 activates transcriptional signatures of basal lineage identity and ciliogenesis in pancreatic ductal adenocarcinoma. bioRxiv doi:10.1101/2023.04.20.537667

Li BE, Li GY, Cai W, Zhu Q, Seruggia D, Fujiwara Y, Vakoc CR, Orkin SH. 2023. In vivo CRISPR/Cas9 screening identifies Pbrm1 as a regulator of myeloid leukemia development in mice. *Blood* **7:** 5281–5293. doi:10.1182/bloodadvances.2022009455

Maia-Silva D, Schier AC, Skopelitis D, Kechejian V, Alpsoy A, Liverpool J, Taatjes DJ, Vakoc CR. 2023. Marker-based CRISPR screening reveals a MED12-p63 interaction that activates basal identity in pancreatic ductal adenocarcinoma. bioRxiv doi:10.1101/2023.10.24.563848

Pomella S, Cassandri M, D'Archivio L, Porrazzo A, Cossetti C, Phelps D, Perrone C, Pezzella M, Cardinale A, Wachtel M, et al. 2023. MYOD-SKP2 axis boosts tumorigenesis in fusion negative rhabdomyosarcoma by preventing differentiation through $p57^{Kip2}$ targeting. *Nat Commun* **14:** 8373. doi:10.1038/s41467-023-44130-0

Sroka MW, Skopelitis D, Vermunt MW, Preall JB, El Demerdash O, de Almeida LMN, Chang K, Utama R, Gryder B, Caligiuri G, et al. 2023. Myo-differentiation reporter screen reveals NF-Y as an activator of PAX3-FOXO1 in rhabdomyosarcoma. *Proc Natl Acad Sci* **120:** e2303859120. doi:10.1073/pnas.2303859120

Sun X, Klingbeil O, Lu B, Wu C, Ballon C, Ouyang M, Wu XS, Jin Y, Hwangbo Y, Huang YH, et al. 2023. BRD8 maintains glioblastoma by epigenetic reprogramming of the p53 network. *Nature* **613:** 195–202. doi:10.1038/s41586-022-05551-x

Wu XS, Vakoc CR. 2023. Made to order tuft cells by an OCA-T1 isoform switch. *Sci Immunol* **8:** eadh3123. doi:10.1126/sciimmunol.adh3123

# CHARACTERIZATION OF TUMOR MICROENVIRONMENT IN HEMATOLOGICAL MALIGNANCIES

**L. Zhang** D. Dubey T. Kaushik B. Li S. Mishra T. Morales N. Rai S. Rakshit K. Shah P. Shenoy R. Tang

The long-term goal of our laboratory is to dissect the regulation of hematological malignancies by environmental signals. Our research specifically focuses on important but understudied classes of environmental signal modulators, including micronutrients and neurotransmitters. We aim to understand how environmental signals such as diet and neural activity regulate stem and progenitor cell development and cancer. We use a combination of functional genomics, metabolomics, circuit mapping, and optogenetic approaches to systematically reveal key dietary and neural activities that regulate stem and progenitor cell development and identify key drug targets for the treatment of hematological malignancies. Our laboratory recently discovered a series of key metabolites in the tumor microenvironment, including pyridoxal and acetylcholine, and their genetic effectors, that are novel regulators of hematological malignancies. We discovered that the vitamin B6 pathway is a nutritional and metabolic dependency in acute myeloid leukemia (AML), orchestrating nucleotide and putrescine metabolism required for leukemia maintenance. We also identified cholinergic receptor muscarinic 4 (CHRM4) as a novel regulator of early erythroid progenitor differentiation and a therapeutic target for myelodysplastic syndrome (MDS). By collaborating with medicinal chemists at our spin-off company, we are developing pharmacological approaches at the pharmaceutical industry level to target these novel modulators and translate our discoveries into first-in-class therapeutics. Our findings will help treat devastating hematological malignancies, including refractory anemias, myelodysplasias, and leukemias.

## Focusing on Nutrient Availability in Hematological Malignancies

Micronutrients are important risk factors for various cancer types. Our recent studies have shown that the vitamin B6 pathway contributes directly to the proliferation of malignant cells in AML. In line with our findings, multiple independent studies have shown that vitamin dysregulation also directly regulates the proliferation of malignant cells in solid tumors, including pancreatic ductal adenocarcinoma, colorectal cancer, and breast cancer, further highlighting the importance of cancer micronutrient status in cancer regulation.

AML is one of the most common hematological cancers. It is a very aggressive disease with a very low five-year relative survival rate calculated by year of diagnosis. AML is particularly dangerous for the elderly, as patients have a higher incidence and a poorer prognosis. As the first-line treatment for AML, standard induction chemotherapy has limited application because of poor patient health and poor tolerability. Therefore, new therapeutic approaches are urgently needed for AML treatment, especially those with the potential to form combination therapies with existing drugs to reduce their doses and therefore reduce side effects.

One of the characteristics of AML is the abnormal micronutrient status in the leukemic blasts and peripheral blood of patients. My laboratory is systematically identifying selective micronutrient genetic pathway vulnerabilities in AML. Our recent studies have revealed the functions of vitamin B6–dependent enzymes in supporting AML growth. We have further validated the potential of inhibiting the vitamin B6 pathway as a new therapeutic strategy for AML proliferation and have verified the antileukemic effects of vitamin B6 inhibitors in multiple AML cell lines and AML mouse models. Together, our studies have identified the vitamin B6 pathway as a pharmacologically actionable target for the treatment of AML.

## Targeting PDXK and Its Functional Interaction with BCL-2 for the Treatment of AML

Epidemiological investigations have shown that plasma vitamin B6 levels are reduced in leukemia patients,

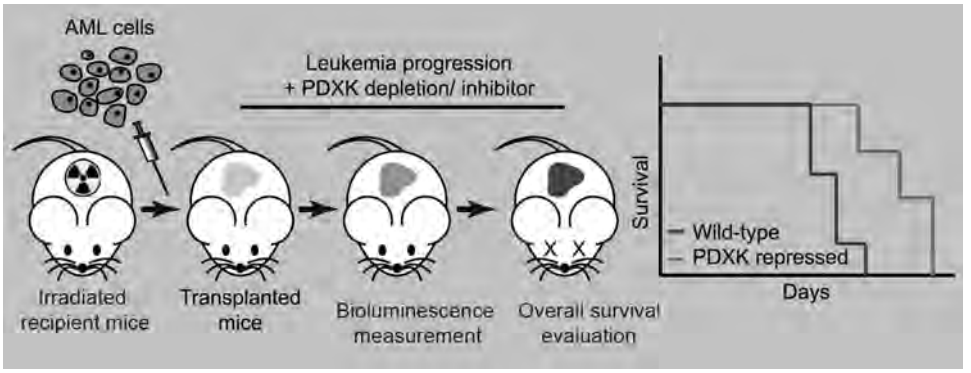



**Figure 1.** Nutrient availability is a critical regulatory aspect of leukemogenesis. Abnormal micronutrient status supports leukemic cell growth and provides potential therapeutic opportunities. Therapeutic targeting of vitamin B6 addiction provides novel therapeutic approaches for leukemia.

indicating that vitamin B6 metabolism is abnormal during leukemogenesis. Accordingly, we have identified PDXK as a priority requirement for AML cell proliferation, whereas many other cell types are not. We further validated the specific dependence of PDXK, but not PNOP or PDXP, in leukemia cells through genome-wide CRISPR-Cas9 screening data in human cancer cell lines. PLP is a cofactor for many enzymes involved in cell proliferation, and PDXK disruption reduces the intracellular concentration of key metabolites required for cell division, whereas disruption of the PLP-dependent enzymes ornithine decarboxylase (ODC1) or glutamate-oxaloacetate transaminase 2 (GOT2) also selectively inhibits AML cell proliferation. Our work identifies the vitamin B6 pathway as a pharmacologically actionable dependency in AML.

PDXK kinase activity is required for PLP production and proliferation in AML cells, and pharmacological inhibition of PLP using the antituberculosis drug isoniazid or its direct inhibitor 4′-O-methylpyridoxine recapitulates the effects of PDXK gene disruption. Notably, PDXK depletion did not trigger leukemic cell differentiation, but instead impaired cell cycle progression and increased apoptosis, resulting in cytostatic and cytotoxic effects. We further show that blockade of the vitamin B6 pathway with isoniazid (INH) exhibits promising therapeutic effects in multiple AML mouse models and significantly improves the survival of diseased mice. We found that leukemic cells are dependent on the vitamin B6 pathway such that inhibition of the vitamin B6 pathway selectively impairs the proliferation of leukemic cells compared with other normal and cancer cell types. Accordingly, although INH reduced intracellular PLP levels in bone marrow hematopoietic stem and progenitor cells (HSPCs), it had no effect on the proliferation of primary HSPCs from normal bone marrow in mice and humans. In conclusion, our findings highlight the therapeutic potential of the vitamin B6 pathway as an antileukemic target with minimal myelosuppression (Fig. 1).

Standard induction chemotherapy for AML is greatly limited by drug side effects. Therefore, new therapies are urgently needed to treat AML that are free of on-target toxicity and have the potential to form combination therapies with small molecules or immunotherapies to reduce their doses and, therefore, side effects. BCL-2 is the first member of the pro-survival subfamily to be identified. Although BCL-2 is found to be overexpressed in cancers including AML, the transient responses of BCL-2 inhibitors (ABT-199/venetoclax) have limited their clinical promise as single agents and prompted the exploration of combination strategies. Our analyses showed that genome-wide CRISPR-Cas9 screens demonstrated similar requirements for PDXK and BCL-2 for AML proliferation, suggesting that they may have synergistic roles in AML. Our further experiments revealed a synergistic role for PDXK and BCL-2 in AML. In conclusion, our study provides a very promising combination therapy targeting nutrient availability.

## Importance of Early Erythroid Progenitors in the Treatment of Hematologic Malignancy Myelodysplastic Syndrome

MDS is a heterogeneous group of hematopoietic malignancies, with anemia being the main manifestation

of MDS, caused by progressive bone marrow failure. Currently, treatment options for MDS are very limited. For patients with MDS who are not dependent on red blood cell (RBC) transfusions, treatment is usually initiated with erythropoietin (EPO) as standard of care. Importantly, serum EPO levels vary among patients with MDS, which determines the response to EPO therapy. Patients with lower serum EPO levels are more likely to respond, whereas patients with higher serum EPO levels are not eligible for EPO therapy. Although EPO has been a first-line treatment option for patients with MDS who have symptomatic anemia, most patients with MDS already have elevated serum EPO levels that result in a lack of response to EPO.

The alternative option for patients who do not respond to EPO (and lenalidomide and other treatments) is RBC transfusion. The dose and frequency of RBC transfusions depend on the severity of anemia and vary from patient to patient. Importantly, transfusion exposes patients to insufficient correction of anemia, alloimmunization, and organ failure secondary to iron overload. The most obvious of these is iron overload toxicity, which occurs because the excess iron from each transfusion accumulates in vital organs. Although iron chelation therapy is often used to treat iron overload in patients with MDS, the clinical response is mixed and the main reason is that most MDS patients are elderly and have other comorbidities that affect the safety and tolerability of this therapy. Together, there is an unmet demand for patients and clinicians to have novel therapeutics to treat MDS.

The burst-forming unit-erythroid (BFU-E) is the first lineage-determined erythroid progenitor, which has unique expansion and differentiation properties and can produce thousands of red blood cells. Clinically, bone marrow BFU-E cell levels predict the response of MDS patients to EPO therapy. MDS patients with BFU-E numbers similar to normal individuals respond to EPO, whereas patients with insufficient BFU-E do not. This highlights BFU-E as a key cell type for treating EPO-resistant anemia in MDS. Given the extremely limited understanding of the molecular mechanisms of BFU-E differentiation, our studies are directed toward genetic modulators and corresponding small compounds targeting these modulators, allowing us to target this process for the treatment of MDS.

## Targeting the Neurotransmitter Receptor CHRM4 for the Treatment of Myelodysplastic Syndromes

Our studies performed a functional screen for modulators of BFU-E expansion and differentiation and unexpectedly identified the muscarinic acetylcholine receptor (CHRM4) pathway as a novel regulator of BFU-E differentiation. Consistently, pharmacological inhibition of this pathway improved erythropoiesis in MDS mouse models and patient samples. Using purified BFU-E, we employed a functional genomics approach and found that inhibition of CHRM4 triggers BFU-E expansion and differentiation. CHRM4 is abundantly expressed in neural tissues and the retina, and analysis of gene expression data from human tissues and cell types showed that CHRM4 expression in BFU-E is second only to that in neural tissues and the retina. Although the role of CHRM4 in nervous system is known, the role of the receptor and its signaling in hematopoiesis has never been evaluated before.

We further showed that inhibition of CHRM4 promotes BFU-E expansion and differentiation and increases erythroid cell production. Genetic deletion of CHRM4 promoted erythroid expansion and differentiation in primary mouse BFU-E cells and in human $CD34^+$ HSPC erythroid expansion culture system. A small compound inhibits CHRM4 signaling, leading to elevated intracellular cyclic AMP levels and altered activity of the downstream transcription factor CREB. These data suggest that the CHRM4-CREB pathway promotes BFU-E differentiation by up-regulating the expression of genes important for maintaining the BFU-E progenitor state. We further revealed that inhibition of CHRM4 triggers BFU-E expansion and differentiation and corrects anemia in vivo. Injection of CHRM4 inhibitor corrects anemia in a genetically engineered MDS mouse model based on expression of the mutant RNA splicing factor SRSF2 and extends the survival of these mice to that of littermate wild-type mice. Importantly, injection of CHRM4-selective inhibitor also corrects the deficiency of BFU-E and reduces abnormal plasma EPO levels in MDS mice to levels comparable to those of wild-type control mice. These results suggest that CHRM4 is a therapeutic target for anemia in MDS and that targeting CHRM4 may overcome early erythroid progenitor deficiency and EPO resistance (Fig. 2A,B).

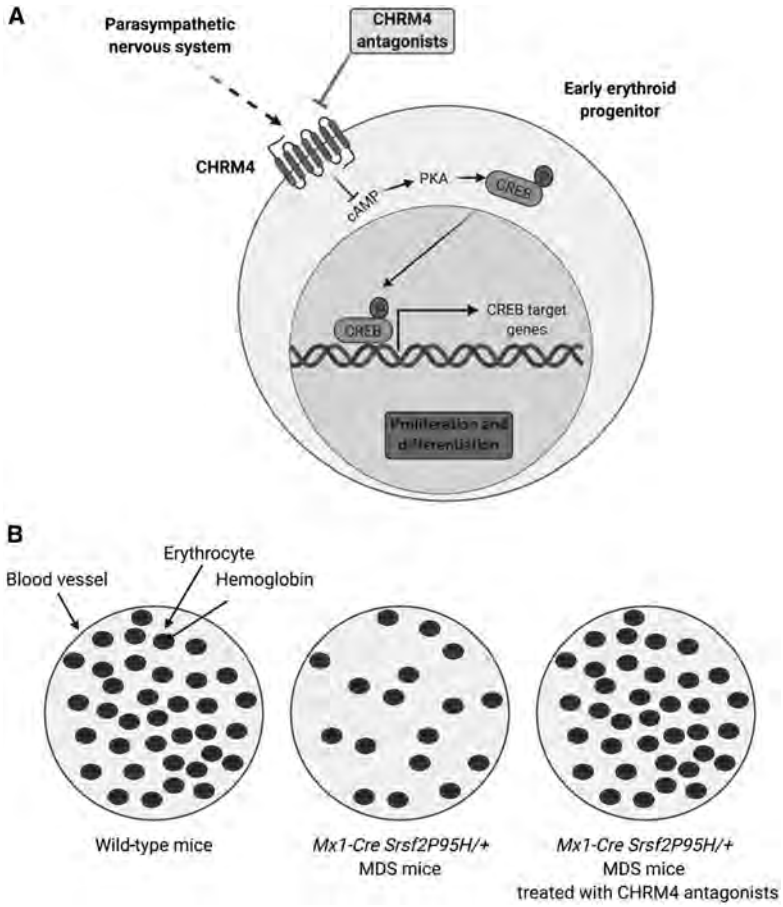



**Figure 2.** Muscarinic acetylcholine receptor is a key regulator of early erythroid progenitor differentiation, mediating neuronal activity through the "Hemopoietic Arc" to enable hematopoietic stem and progenitor differentiation (*A*). Therapeutic targeting of muscarinic acetylcholine receptor provides new treatment approaches for myelodysplastic syndromes and refractory anemias (*B*).

## Discovering New Treatment Strategies for Hematologic Diseases

Insufficiency of early erythroid progenitors leads to treatment resistance. Our study targets treatment-resistant MDS and uses small-molecule drugs to promote the differentiation of early erythroid progenitor BFU-E. Our study shows that pharmacological inhibition of CHRM4 corrects MDS anemia with sustainable long-term efficacy in vivo. Our research is further developing pharmacological approaches to target this pathway and help advance CHRM4 inhibitors into clinical development to treat this hematologic malignancy.

Our research will provide a new treatment for MDS. We have shown that pharmacological inhibition of CHRM4 overcomes MDS resistance to current therapies. Our ongoing research is focused on improving CHRM4-selective inhibitors and their corresponding toxicological and pharmacological profiles. These results will form the basis for an IND application and clinical trials for refractory MDS. We have demonstrated that CHRM4-selective inhibitor significantly improved erythroid progenitor cell expansion and differentiation in MDS patient samples. Our ongoing studies targeting MDS subtypes or molecular subtypes will search for MDS that are more sensitive to CHRM4-selective inhibitor treatment and identify genetic markers or combinations of markers that are responsive to the treatment. Taken together, our study will provide information on genetic molecular markers

of response to CHRM4 inhibitor treatment to facilitate patient participation in clinical trials, and this new therapy will significantly improve the clinical management of MDS.

## PUBLICATION

Mishra S, Millman S, Zhang L. 2023. Metabolism in acute myeloid leukemia: mechanistic insights and therapeutic targets. *Blood* **141:** 1119–1135.

# CANCER: GENETICS

**Alea Mills**' laboratory is studying genetic pathways important in cancer, aging, and autism; identifying the genetic players; and determining how aberrations in their functions culminate in human disease. Through innovative use of a technique called "chromosome engineering," the Mills group discovered that one of the most common genetic alterations in autism—deletion of a 27-gene cluster on chromosome 16—causes autism-like features in mice. These autism-like movement impairments can be identified just days after birth, suggesting that these features could be used to diagnose autism. Mills has also used chromosome engineering to identify a tumor suppressor gene that had eluded investigators for three decades. The gene, called *Chd5*, was shown by Mills to regulate an extensive cancer-preventing network. This year, the Mills laboratory uncovered how *Chd5* acts as a tumor suppressor: It binds to a protein found within chromatin to turn specific genes on or off, halting cancer progression. The epigenetic role of *Chd5* in development, cancer, and stem-cell maintenance is currently being investigated. The Mills laboratory is also studying p63 proteins, which regulate development, tumorigenesis, cellular senescence, and aging, in vivo. They succeeded in halting the growth of malignant tumors by turning on production of one of the proteins encoded by the *p63* gene, called TAp63. TAp63 also exerts other protective effects. The Mills laboratory generated a mouse model that allowed them to find that TAp63 is required to prevent a genetic disorder, known as EEC (ectrodactyly-ectodermal dysplasia cleft lip/palate syndrome), which is characterized by a cleft palate and major deformities of the skin and limbs in infants. In addition, they recently discovered that a different version of *p63*, called ΔNp63, reprograms stem cells of the skin to cause carcinoma, the most prevalent form of human cancer. Modulation of these proteins may offer new ways to treat human malignancies in the future.

The **Peter Westcott** laboratory is developing powerful new strategies to study how the immune system and cancer coevolve, with the goal of expanding the curative potential of immunotherapy to more patients. Westcott's group studies how the immune system shapes tumor evolution during initiation, malignant transformation, and metastasis—the major cause of cancer mortality. Cancer progresses through sequential genetic and epigenetic aberrations, following Darwinian principles of natural selection. During clonally selective sweeps—population explosions of the most "fit" cancer cells—the molecular and cellular architecture of the tumor and its tissue niche are restructured. Using a multidisciplinary approach at the interface of cancer genomics and immunology, the Westcott laboratory aims to understand how immunosurveillance fails at these critical junctures. Central to their approach are next-generation mouse models that capture the mutational complexity and immune interactions of human cancer. The team engineers into these models powerful genetic switches that allow them to control the precise timing of tumor selective events and immune perturbations in vivo. Combined with longitudinal tumor sampling, advanced imaging, single-cell and spatial omics, and other emerging technologies, these models enable the laboratory to deconstruct the spatiotemporal dynamics of tumor-immune cross talk.

**Michael Wigler**'s work provides a new paradigm for understanding and exploring human disease. The Wigler laboratory studies human cancer and the contribution of new mutation to genetic disorders. The cancer effort (with James Hicks, Alex Krasnitz, and Lloyd Trotman) focuses on breast and prostate cancers. It involves collaborative clinical studies to discover mutational patterns predicting treatment response and outcome and the development of diagnostics to detect cancer cells in bodily fluids such as blood and urine. The major tools are single-cell DNA and RNA analysis. The single-cell methods, which are in development, are also being applied to problems in neurobiology to characterize neuronal subtypes, somatic mutation, and monoallelic expression.

The Wigler laboratory's genetic efforts are a collaboration with Ivan Iossifov and Dan Levy, and this team focuses on determining the role of new mutations in pediatric disorders. In a large-scale population sequencing project with W. Richard McCombie and the Genome Sequencing Center at Washington University in St. Louis, and supported by the Simons Foundation, the team has proven the contribution of this mechanism to autism. The work further suggests a relationship between the mutational targets in autism and the process of neuroplasticity that lies at the heart of learning. Smaller-scale population studies of congenital heart disease and pediatric cancer (collaborations with scientists at Columbia University and Memorial Sloan Kettering Cancer Center, respectively) also point to new mutation as a causal factor in these disorders.

# GENETIC AND EPIGENETIC BASIS OF CANCER

**A.A. Mills** S. Balinth, C. Ballon, J. Boockvar, Y. Chang, M. Fisher, H. Liang, P. Shrestha, S. Sun, C. Wu

The overall goal of our work is to elucidate genetic and epigenetic mechanisms that impact cancer so that more effective treatments can be developed in the future. We discover genes that either promote or prevent cancer, figure out how their encoded proteins normally work, and determine how their perturbation contributes to malignancy. Our findings have impacted how clinicians diagnose and treat patients with these syndromes.

**Major discoveries:**

- Revealing p63 as a regulator of development, cancer, and aging
- Discovering *CHD5* as a tumor suppressor
- Identifying BRD8 as a novel target for glioblastoma

## p63 in Development, Cancer, and Aging

Our team has had a long-standing interest in p63, a p53-related protein Dr. Mills discovered as a master regulator of development. We have focused on how p63 normally works, as well as how its deregulation leads to disease. Although p63 looked very similar to the p53 tumor suppressor—which is mutated in the majority of human cancers—its function was completely unknown when it was first discovered. The Mills group revealed that the absence of p63 causes premature aging, and that p63 is required for replenishing stem cells of stratified epithelia such as the skin. Indeed, loss of p63 in mice causes features of aging such as curvature of the spine, hair loss, and severe skin lesions, and mutation of p63 in humans causes premature aging. Levels and types of p63 proteins are crucial, as we discovered that too much expression of one version of p63 (ΔNp63α) causes carcinoma—the most prevalent type of human cancer, whereas too little expression of a different version of p63 (TAp63) causes mesenchymal cancers such as sarcoma. Our finding that TAp63 effectively inhibits tumor growth in cells lacking p53 was novel, as it showed that TAp63 can prevent cancer completely independently of p53.

Dr. Mills had initially discovered that p63 is essential for embryonic development, as its loss causes depletion of stem cells and birth defects of the limbs, skin, and palate. This finding led others to interrogate p63 in humans, revealing that its mutation causes seven different syndromes in which children are afflicted with birth defects similar to those we found in the mice we engineered to lack p63. By generating mouse models for one of these human syndromes, ectrodactyly, ectodermal dysplasia, clefting (EEC) syndrome, we found out why some children with EEC have symptoms that are very severe and even life-threatening, whereas other children with EEC—even those in the same family with the same p63 mutation—have symptoms that are barely noticeable.

During the past year we have advanced our work on p63 by discovering that cancer-associated fibroblasts promote malignancy by inducing expression of the chromatin-modifying protein CBX4 in squamous cell carcinoma (Fisher et al. 2023a). Our findings have important therapeutic implications for the treatment of squamous cell carcinoma (Fisher et al. 2023b). To advance our findings toward the bedside, we have joined forces with clinical oncologists at Northwell Health, allowing us to study how p63 impacts human carcinoma of the head and neck, cervix, and salivary gland—tissues in which we showed p63 to be essential.

## *CHD5*, a New Tumor Suppressor

We discovered *CHD5* as a tumor suppressor mapping to 1p36—a region of our genomes frequently

deleted in a variety of human cancers. Although ample evidence that a tumor suppressor was located in this region had been known for more than three decades, the gene responsible remained a mystery. Using chromosome engineering to generate mice with deletions and duplications of the genomic region corresponding to human 1p36, we discovered *CHD5* as the tumor suppressor gene and found that its product turns on a network of cancer-preventing proteins. In addition, we showed that *CHD5* is frequently deleted in human glioma.

We advanced this work by defining the role of *CHD5* in chromatin dynamics and deciphering how dysregulation of *CHD5* and the pathways it regulates leads to disease. We found that *CHD5* uses its plant homeodomains to bind histone H3, and that this interaction is essential for tumor suppression. This paved the way for further discoveries, and *CHD5* is now known to be mutated in human cancers of the breast, ovary, and prostate, as well as in melanoma, glioma, and neuroblastoma. Indeed, recent reports show that patients with high levels of *CHD5* respond better to treatment than those with low levels. We found that *CHD5* is essential for packaging DNA, and that loss of *Chd5* (the mouse version of human *CHD5*) leads to improperly packaged DNA that is prone to damage. *Chd5*'s absence is particularly important during the process of sperm maturation—an event in which the DNA is first unpackaged and then repackaged using an elaborate series of steps—and that deficiency of *CHD5* causes men to be infertile. We collaborated with clinicians at Northwell Health to reveal that *CHD5* mutations occur in infertile men, providing an intriguing link between cancer and infertility and highlighting the role of epigenetic processes in these syndromes. We also found *CHD5* to be highly expressed in neurons and showed that *CHD5* plays a pivotal role in the brain, suggesting that inappropriate DNA packaging contributes to neurodevelopmental syndromes such as autism. We discovered that *CHD5* regulates a ribosome biogenesis switch that dictates neuronal cell fate and that *CHD5* deficiency leads to an excessive number of astrocytes at the expense of neurons. Over the past year we have gone deeper into the mechanisms whereby CHD5-mediated regulation of chromatin affects gene expression programs that regulate neuronal stem cells and how deregulation of these processes sets the stage for neurodevelopmental syndromes and cancer. We recently discovered that *CHD5* regulates the transcription factor SIX3 to promote neuronal differentiation (Shrestha et al. 2023), elucidating new principles of chromatin biology that have important implications for brain cancer (Sun et al. 2023a).

## BRD8, a New Target for Treating Glioblastoma

During the past year, we published our recent discovery that the bromodomain-containing chromatin regulator BRD8 is essential to glioblastoma lacking p53 mutations, which make up ~71% of cases (Sun et al. 2023b). We found that BRD8 maintains malignancy by crippling p53-mediated tumor suppression in a way distinct from previously described mechanisms: It reprograms the p53 network through the EP400 histone acetyltransferase complex and, by bromodomain-directed occupancy of the histone variant H2AZ at p53-induced targets, it enforces a repressive chromatin state that prevents p53-mediated transactivation.

Importantly, targeting BRD8 in glioblastoma remodels chromatin by evicting H2AZ and enhancing chromatin accessibility, enabling p53 to bind and transactivate its targets. This chromatin remodeling cascade (referred to as the "BRD8/p53 epigenetic switch") re-establishes p53 activity, thereby normalizing gene expression, evoking cell cycle arrest, inhibiting cancer development, and prolonging survival in xenograft models of glioblastoma. These findings present a previously unappreciated mechanism by which cancer cells sidestep p53 to maintain malignancy and indicate that targeting BRD8 re-establishes p53-mediated tumor suppression. We are currently working with neuro-oncological surgeons at Northwell Health to grow organoids from patients with glioblastoma. We believe that BRD8 holds significant promise for more successful treatments for patients with glioblastoma—a deadly cancer for which there are few successful treatment options.

### PUBLICATIONS

Fisher ML, Balinth S, Hwangbo Y, Wu C, Ballon C, Goldberg GL, Mills AA. 2023a. Cancer-associated fibroblasts promote cancer stemness by inducing expression of the chromatin-modifying protein CBX4 in squamous cell carcinoma. *Carcinogenesis* **44:** 485–496. doi:10.1093/carcin/bgad048

Fisher ML, Balinth S, Mills AA. 2023b. ΔNp63α in cancer: importance and therapeutic opportunities. *Trends Cell Biol* **33:** 280–292. doi:10.1016/j.tcb.2022.08.003

Shrestha P, Jaganathan A, Huilgol D, Ballon C, Hwangbo Y, Mills AA. 2023. *Chd5* regulates the transcription factor Six3 to promote neuronal differentiation. *Stem Cells* **41:** 242–251. doi:10.1093/stmcls/sxad002

Sun X, Shrestha P, Mills AA. 2023a. CHROMO domain readers: a rainbow of opportunities. In *Chromatin readers in health and disease*. Elsevier Science, New York.

Sun X, Klingbeil O, Lu B, Wu C, Ballon C, Ouyang M, Wu XS, Jin Y, Hwangbo Y, Somerville TDD, et al. 2023b. BRD8 maintains glioblastoma by epigenetic reprogramming of the p53 network. *Nature* **613:** 195–202. doi:10.1038/s41586-022-05551-x

# IMMUNE SURVEILLANCE IN COLON CANCER EVOLUTION AND THERAPEUTIC OPPORTUNITIES

**P. Westcott** Z. Aminzada, E. Arboleda, A. Cicala, E. Gazzara, S. Gholami, S. Han, A. Liu, C. McLaughlin, N. Persaud, Y. Qin, S. Shui, J. Zhang

The Westcott laboratory studies the role of the immune system in cancer evolution, with a particular focus on (1) early transition to malignancy and (2) metastasis in colon cancer. Understanding the former is important for cancer prevention and early detection, whereas understanding the latter is critical for improving treatment, as metastasis to the liver is the primary cause of mortality in colorectal cancer. Our approach is centered on development and application of powerful genetically engineered mouse models that are faithful to the human disease and enable mechanistic dissection of processes that are challenging to capture in humans. We apply these models to studies of the fundamental biology of stepwise evolution of cancer, early dysfunction of tumor-specific cytotoxic T cells, and preclinical studies of emerging immunotherapies for colorectal cancer liver metastasis—in close collaboration with surgeon oncologist, Dr. Sepideh Gholami, at Northwell Health.

## Intratumoral Heterogeneity of Mutations Undermines Immunity in DNA Mismatch Repair Deficient Cancers

We recently developed genetically engineered mouse models of lung and colon cancer with targeted disruption of *Msh2* and other essential DNA mismatch repair genes (Westcott et al. 2023) with the goal of establishing preclinical platforms that recapitulate the heterogeneity of immune checkpoint blockade (ICB) responses in the clinic. This was motivated by clinical trials showing that tumors with very high levels of somatic mutations, especially those that are DNA mismatch repair deficient (MMRd), have the highest response rates to ICB. Nevertheless, most patients with these tumors do not benefit from long-term, durable responses. Surprisingly, we found that DNA MMRd did not confer increased T-cell infiltration or sensitivity to ICB. We reasoned that this was due to concomitant initiation of tumor-driving events with MMR gene loss and the absence of clonal neoantigens, which is supported by the high intratumoral heterogeneity of mutations of tumors in our model. Using in vivo retransplantation studies, we showed that single-cell clones from MMRd tumors are more responsive to ICB than the parental line or a mixture of clones, and the magnitude and quality of the T-cell response increases as a function of neoantigen clonal fraction. These results demonstrate that intratumoral heterogeneity of mutations is confounding to the T-cell response in MMRd tumors.

Intriguingly, we also found that immune surveillance—the process of lymphocytes patrolling the body and eliminating infected or abnormal cells—failed to restrain the outgrowth of tumor cells with high tumor mutation burden and only eliminated clonal neoantigens—that is, those neoantigens present in the majority of tumor cells. Specifically, the overall burden of neoantigens was not significantly changed in MMRd lung tumors that arose in mice with no T cells, but the clonal architecture of neoantigens was very different, with significantly more clonal tumor neoantigens arising in the absence of immune surveillance. Finally, we reanalyzed two clinical trials of ICB in MMRd gastric and colon cancer and found that clonal, but not subclonal, neoantigen burden is a better biomarker than tumor mutation burden alone, significantly associated with overall response and survival. Altogether, this work has important implications for (1) understanding the fundamental biology of how immune surveillance shapes cancer evolution and (2) developing improved clinical strategies for matching patients with the right immunotherapies and predicting prognosis.

## Deconstructing the Stepwise Evolution of Cancer Using New In Vivo Models

Building on the models and findings above, we have developed and begun to benchmark innovative new mouse models of colon cancer that (1) enable fine spatial and temporal control of stepwise progression to malignancy and (2) faithfully recapitulate neoantigen-specific T-cell dynamics of human cancer. A major goal of this project is to apply these models to mechanistic studies of tumor-immune evolution during benign-to-malignant progression. Stepwise progression is a hallmark of most human cancer, but the immune implications of this process are poorly understood, and few mouse models are optimized to its study. Given that ICB resistance in our MMRd cancer models is driven by extensive intratumoral heterogeneity, *we hypothesize that selective sweeps in tumor progression—population explosions of the most fit cells—reshape neoantigen clonal architecture and engage immunosurveillance, ultimately driving adaptation of immune evasive mechanisms.* Although we and others have identified neoantigen clonality as a critical determinant of antitumor immunity, it is unclear how malignant progression reshapes clonal dynamics and immunosurveillance. We will deeply interrogate our central hypothesis in two complementary model systems of colon cancer that enable precise spatiotemporal induction of adenoma to adenocarcinoma transition in the context of MMRd or defined neoantigens. Focusing on early pre-adenocarcinoma expansions, we will determine (1) how cancer becomes immunogenic and (2) when and how immunosurveillance is dismantled during benign-to-malignant transition.

To enable temporally controlled cell type–specific activation of secondary oncogenic events within established colon adenomas, we developed the first split-Cre-recombinase cancer model (unpublished), wherein the amino-terminal half of Cre (CreN) is expressed endogenously from the epithelial-specific gene, *Epcam*, and the carboxyl-terminal half of Cre (CreC) is delivered via lentivirus. CreN and CreC splice via bacterial intein to form full-length Cre only in the infected epithelial cells. This reaction is controlled by tamoxifen-inducible and trimethoprim (TMP)-stabilized destabilization domain (dd)-CreN at the *Epcam* locus, enabling temporal regulation of secondary events. Mice also harbor constitutive *Cas9* ($R26^{Cas9}$) and Cre-inducible drivers ($Kras^{LSL\text{-}G12D}$; $Trp53^{flox/flox}$) and a fluorescent reporter ($R26^{LSL\text{-}tdTomato}$). The initiating lentivirus also delivers single-guide RNAs (sgRNAs) to knock out the tumor suppressor *Apc* and a DNA MMR gene (e.g., *Msh2*) via CRISPR-Cas9. Colonoscopy-guided injection of lentivirus with the same sgRNAs into the distal colons of $R26^{Cas9}$ mice is sufficient to induce MMRd adenomas, as we previously reported. We have established the feasibility of this system to induce rare secondary cancer driver events in established adenomas, and we continue to optimize and characterize the utility of this system for tracking cancer clonal evolution.

## SOX17 Drives a Fetal Intestinal Program to Evade Immune Surveillance in Early Colon Cancer

In collaboration with the laboratory of Omer Yilmaz at MIT, we discovered a novel mechanism of early immune evasion in colon cancer progression (Goto et al. 2024). We leveraged colon organoid models to define the transcriptional and epigenetic changes associated with early tumor establishment in the distal colons of mice. Even though the key cancer-driving mutations were engineered into the organoids in vitro, in vivo transplantation was associated with profound gene expression and chromatin changes, including up-regulation of *Sox17*, a transcription factor involved in development of the fetal intestine. This was somewhat surprising, as *Sox17* is a repressor of WNT signaling, a key stem cell growth pathway that is hyperactivated in most colon cancer. Notably, deletion of *Sox17* had no impact on organoid growth in vitro, but severely impaired the ability of these organoids to form tumors in vivo.

Mechanistically, we found that *Sox17* binds to the promoter of interferon gamma receptor 1 (*Ifngr1*), a critical sensor of inflammation, and suppresses its expression. This was associated with insensitivity of the organoids to interferon gamma (IFN-γ) in the tumor microenvironment, failure to up-regulate major histocompatibility complex I (MHC-I) and to present antigens, and markedly reduced sensitivity to killing by cytotoxic $CD8^+$ T cells, enabling early evasion of immunosurveillance. *Sox17* expression in tumors was also associated with the fetal intestinal program and loss of the homeostatic stem cell program of the normal colon crypts, marked by *Lgr5*. It is now well appreciated that the fetal regenerative program and

stem cell plasticity are critical drivers of colon cancer progression and metastasis. This work provides further nuance to our understanding of this process and how plasticity can undermine immune surveillance to enable cancer progression.

## PUBLICATIONS

Ely ZA, Mathey-Andrews N, Naranjo S, Gould SI, Mercer KL, Newby GA, Cabana CM, Rideout WM, Jaramillo GC, Khirallah JM, et al. 2023. A prime editor mouse to model a broad spectrum of somatic mutations in vivo. *Nat Biotechnol* **42:** 424–436.

Westcott PMK, Muyas F, Hauck H, Smith OC, Sacks NJ, Ely ZA, Jaeger AM, Rideout WM 3rd, Zhang D, Bhutkar L, et al. 2023. Mismatch repair deficiency is not sufficient to increase tumor immunogenicity. *Nat Genet* **55:** 1686–1695.

*In Press*

Goto N, Westcott PMK, Goto S, Imada S, Taylor MS, Eng G, Braverman J, Deshpande V, Jacks T, Agudo J, Yilmaz OH. 2024. SOX17 enables immune evasion of early colorectal adenomas and cancers. *Nature* **627:** 636–645.

# CANCER AND HUMAN GENETICS


**M. Wigler** M. Abramson, J. Alexander, D. Bradford, B. Debmalya, A. Gruet, I. Hakker, J. Kendall, S. Kim, S. Li, J. Luo, A. Moffitt, P. Morris, S. Negi, A. Peyser, N. Ranade, M. Riggs, M. Ronemus, J. Rosenbaum, D. Stauder, A. Stepansky, A. Tandon, D. Trimboli, Z. Wang, Z. Yu


In its research, the Wigler group combines technology development, computational analysis, and clinical and population discovery. Our work is concentrated in three major areas: autism, cancer, and genomics. We collaborate with the groups of Ivan Iossifov (in autism), Alexander Krasnitz (in cancer), and Dan Levy (in all three) at CSHL. Much of the clinical aspect of our cancer work involves collaborations with physicians at Northwell Health.

## Cancer Genetics


M. Abramson, J. Alexander, A. Gruet, I. Hakker, J. Kendall, S. Li, A. Moffitt, M. Riggs, M. Ronemus, J. Rosenbaum, D. Stauder, A. Stepansky, D. Trimboli, Z. Wang, Z. Yu
[in collaboration with S. Allen, J. Kolitz, N. Chiorazzi, G. Goldberg, and M. Frimer, Northwell Health; R. Levine, Memorial Sloan Kettering Cancer Center; M. Kemeny and R. Parsons, Icahn School of Medicine at Mount Sinai; D. Levy and A. Krasnitz, CSHL]


Our technology development continues within the healthcare arena as we translate research into clinical applications. We continue our strong partnership with Northwell Health, particularly in gynecological cancers and acute myeloid leukemia. We have also strengthened collaborations with several academic medical centers in the metropolitan area to extend our cohorts and expand to other cancer types. With the establishment of a robust logistical pipeline for receiving samples from our clinical collaborators, we have been able to handle an ever-increasing number of specimens covering a wide range of sample types.

In solid cancers, using the MASQ (multiplex accurate sensitive quantitation) method developed in our laboratory, we assess tumor load in two components derived from blood (circulating epithelial cells and cell-free DNA from plasma). Application of MASQ to these specimens has demonstrated the clinical utility of MASQ to contribute to prognosis, treatment efficacy, and potentially early detection, focusing mainly on endometrial and ovarian cancers. Our results show that the range of tumor-specific signal can represent a significant proportion of total cell-free DNA, and this signal is currently detectable in 80% of cases. There is a significant correlation between the levels of detection and months of relapse-free survival, even when accounting for other variables such as grade, size, and type of cancer. Our ongoing efforts will continue our focus on endometrial cancers in a longitudinal study with Northwell Health over the next 5 years, for which we received substantial funding in 2023. In this new study, we will also evaluate the state of microsatellite instability utilizing STORM (stabilizing technique of random mutagenesis; see below).

As with solid cancers, we assess burden over time in acute myeloid leukemia, examining peripheral blood cells, bone marrow, and cell-free DNA from blood and urine. We continue to recruit new patients for this ongoing longitudinal study. The additional patients and time points have reinforced our earlier findings that the high sensitivity of MASQ can accurately assess remission, monitor for minimal residual disease, and detect relapse before other available clinical approaches. We have also initiated a new project to study chronic lymphocytic leukemia and its relationship to monoclonal B-cell lymphocytosis using MASQ.

We have continued to develop our alternative method for tracking cancer genomes: STORM. We apply this to variation in length of microsatellites, short tracts of DNA in which certain motifs (ranging in length from 1 to 6 base pairs) are repeated. Instability within these regions is a hallmark of cancer. Based on the muSeq technology we previously developed, we can accurately assess instability of these regions. Building on our earlier work using synthetic templates, we have applied STORM to approximately 100 pairs of tumor and blood samples from clinical studies. We found that in virtually any tumor, we can

identify a subset of microsatellite loci with unique lengths not observed in the blood. Moreover, some of the endometrial cancers were diagnosed as deficient for mismatch repair (dMMR). This often results in rampant changes in microsatellite lengths, a phenomenon called microsatellite instability (MSI). Whereas traditional MSI testing by gel electrophoresis is an excellent indicator of dMMR status in colorectal cancer, it has an unacceptably high false negative rate in endometrial cancers. Our preliminary results suggest that STORM determines dMMR status with higher sensitivity and specificity than all other methods, and we continue to extend these studies with additional clinical research in a variety of solid cancers.

## Autism Genetics

J. Kendall, M. Ronemus [in collaboration with I. Iossifov, CSHL and the New York Genome Center; D. Levy, A. Krasnitz, H. Meyer, and T. Janowitz, CSHL; K. Baldwin, Columbia University; K. Ye, Albert Einstein College of Medicine; A. Buja and A. Krieger, Wharton School of the University of Pennsylvania]

In collaboration with Kenny Ye (Albert Einstein College of Medicine) and Andreas Buja and Abba Krieger (Wharton School of the University of Pennsylvania), we have continued to measure the extent to which siblings—concordant and discordant for autism—share their parental genomes. We previously applied the method to approximately 5,800 pairs of discordant siblings from the Simons Simplex Collection (SSC), Autism Genetic Research Exchange (AGRE), and Simons Powering Autism Research (SPARK) collections of autistic families, and we have extended this analysis using additional families. As before, we continue to observe that concordant siblings share more from their parents than expected by chance, consistent with transmission of causal determinants in at least half of the multiplex families. We also continue to observe more sharing from fathers than from mothers. The more extensive sharing of paternal than maternal genomes has contradicted our long-established expectations that maternal genomes are the primary source of damaging variants in high-risk families, and we are considering new models for the genetic contribution to autism based on these data.

We continue to investigate the effects of autism-linked mutations on gene expression to explore the molecular mechanisms that may underlie the disorder. In collaboration with Kristin Baldwin (Columbia University) and the New York Genome Center, we previously conducted pilot studies on the effects of new mutations on gene expression in cultured blood cell lines from autistic individuals, as well as in the neuronal cell cultures into which they can be induced to differentiate. Building on the success of these efforts, Simons Foundation Autism Research Institute (SFARI) funded a large-scale project to generate 4,000 RNA sequencing (RNA-seq) profiles from all children in the SSC. We are assessing this large amount of gene expression data to find novel genetic variants associated with perturbed expression or altered splicing of nearby genes.

The paternal bias in transmission that we observe has led us to test the hypothesis that this effect arises from maternal–fetal antigenic incompatibility. Following two separate paths, in an effort led by Tobias Janowitz, we are developing mouse models of maternal–fetal incompatibility, and in collaboration with Alexander Krasnitz and Hannah Meyer, we are applying computational approaches to identify traces of maternal–fetal incompatibility. Initial data from the mouse models indicates that in mothers undergoing stimulation of the immune system, placental weight is lower and mouse pups are deficient in normal social interactions. This promising work has now received independent funding from SFARI, and we continue to collaborate with Tobias Janowitz on assessment of the mouse models.

## Genomics

B. Debmalya, I. Hakker, S. Kim, S. Li, A. Moffitt, A. Peyser, A. Tandon, Z. Wang [in collaboration with D. Levy, CSHL]

Balls of acrylamide gel sequencing (BAG-seq) is a high-throughput single-cell technology that we developed to simultaneously capture both DNA and RNA from the same cell or nucleus. To enhance the analysis of BAG-seq data, we developed a multinomial algorithm that quantitatively measures the distance between expression clusters. This serves as a complement to commonly used 2D visualization tools such as uniform manifold approximation and projection (UMAP) and t-distributed stochastic neighbor embedding (t-SNE), which do not provide this specific information. Using our algorithm, we effectively cluster RNA identities while determining the number

of unique templates required to confidently assign a nucleus to an RNA cluster. These methods have enabled us to identify unique stromal subtypes that are exclusive to certain tissues or individuals, thereby offering new insights into the tumor microenvironment.

Another advance in BAG-seq technology is the development of sequence-specific capture, an improvement over the previous version's random DNA capture approach. Sequence-specific capture enables researchers to selectively focus on genomic regions shared among cells rather than sequencing the entire genome to seek random overlaps. This approach facilitates studies of single-cell phylogeny through single-nucleotide mutations or microsatellite length variations among cells. We have demonstrated the effectiveness of sequence-specific capture in BAGs using human *Alu* repeats. To achieve this, we developed a new method to initiate polymerization of the BAGs postenzymatic reactions using blue light. We have recently found that *Alu* repeats are often followed by a poly(A) microsatellite region, which allows for the study of microsatellite length variations between single cells. This combines the features of varietal tags (unique molecular barcodes) with each molecule for error correction.

## PUBLICATIONS

Li S, Alexander J, Kendall J, Andrews P, Rose E, Orjuela H, Park S, Podszus C, Shanley L, Ma R, et al. 2023. High-throughput single-nucleus hybrid sequencing reveals genome-transcriptome correlations in cancer. bioRxiv doi:10.1101/2023.10.04.560973

Wroten M, Yoon S, Andrews P, Yamrom B, Ronemus M, Buja A, Krieger AM, Levy D, Ye K, Wigler M, Iossifov I. 2023. Sharing parental genomes by siblings concordant or discordant for autism. *Cell Genom* **3:** 100319. doi:10.1016/j.xgen.2023.100319

# CANCER: CELLULAR COMMUNICATION IN CANCER

The **Corina Amor Vegas** laboratory studies cellular senescence. Senescence is a stress response program that is triggered in damaged cells and leads to their elimination by the immune system. If uncleared, the accumulation of senescent cells generates a chronic proinflammatory microenvironment conducive to aging and tumor development. Amor Vegas' laboratory seeks to understand how the immune system recognizes and targets these cells in physiological conditions and how senescent cells evade immune clearance in disease. They aim to leverage their findings to develop immune-based therapeutic approaches to target senescent cells in cancer, aging, and age-related pathologies. In their studies, they develop and combine novel somatic mouse models of cancer and cell-based therapeutics such as their recently developed senolytic CAR T cells.

**Semir Beyaz**'s research interrogates the functional consequences of diets for immune recognition and response pathways that play crucial roles in cancer immunity. Cells respond and adapt to the signals that they receive from their environment. Environmental factors such as nutrients affect cellular states by altering cell state–specific gene expression or metabolic programs. The Beyaz group investigates the causal cellular and molecular mechanisms that link nutrition to organismal health and disease. For example, diets that lead to obesity, such as high-fat diets, are significant environmental risk factors that influence cancer incidence and progression in several tissues. Although the interactions between tumor cells and the immune system play a significant role in tumorigenesis, little is known about how dietary perturbations impact immunity against cancer. By identifying the altered gene expression and metabolic programs in the immune system in response to dietary perturbations, Beyaz aims to uncover mechanistic links that can be therapeutically exploited for the treatment of diseases associated with immune dysfunction such as cancer.

Why do patients with cancer (irrespective of cancer type) frequently experience systemic symptoms like pain, cognitive impairment, deficits in appetite, and disrupted sleep/wake cycles? What is the underlying biology governing these phenomena, and how can this biology be leveraged to improve people's lives? The **Jeremy Borniger** laboratory investigates bidirectional communication between the brain and periphery in the context of cancer. The laboratory aims to determine how tumors disrupt neural circuit function, how aberrant cellular activity promotes cancer-associated systemic dysfunction, and how reciprocal outputs from the brain regulate cancer growth and metastasis. Specifically, the Borniger laboratory uses techniques from systems neuroscience (e.g., optogenetics, calcium imaging, circuit mapping, electrophysiology, and behavioral assays) to dissect how factors in the tumor microenvironment alter host physiology and behavior. Recent work has focused on how central neuromodulator populations participate in cancer-associated sleep and metabolic disruption. The laboratory discovered that nonmetastatic mammary tumors distally alter immune and endocrine signaling to aberrantly activate lateral hypothalamic hypocretin/orexin (HO) neurons. This resulted in disrupted sleep and hepatic glucose metabolism, the latter being driven by the sympathetic nervous system (Borniger et al., *Cell Metab* **28:** 118 [2018]). This research, in combination with clinical work, will facilitate the development of novel treatments to improve outcomes for patients with cancer.

**Mikala Egeblad** and colleagues study cancer and, in particular, the microenvironment in which the cancer cells arise and live. Solid tumors are abnormally organized tissues that contain not only cancer cells, but also various other stromal cell types and an extracellular matrix, and these latter components constitute the microenvironment. Communications between the different components of the tumor influence its growth, its response to therapy, and its ability to metastasize.

Among the tumor-associated stromal cells, the laboratory's main focus is on myeloid-derived immune cells, a diverse group of cells that can enhance angiogenesis and metastasis and suppress the cytotoxic immune response against tumors. Egeblad is interested in how different types of myeloid cells are recruited to tumors and how their behaviors—for example, their physical interactions with cancer cells and other immune cells—influence cancer progression, including metastasis. The Egeblad laboratory studies the importance of the myeloid cells using mouse models of breast and pancreatic cancer and real-time imaging of cells in tumors in live mice. This enables them to follow the behaviors of and the interactions between cancer and myeloid cells in tumors during progression or treatment. This technique was instrumental when the laboratory showed that cancer drug therapy can be boosted by altering components of the tumor microenvironments, specifically reducing either matrix metalloproteinases (enzymes secreted by myeloid cells) or chemokine receptors (signal receptors on myeloid cells). Most recently, the Egeblad laboratory has shown that when a specific type of myeloid cell, called neutrophil, is activated during inflammation, it can awaken sleeping cancer to cause cancer recurrence. The neutrophils do so by forming so-called neutrophil extracellular traps—structures of extracellular DNA—and these alter the extracellular matrix surrounding the sleeping cancer cells to provide a wake-up signal.

The **Douglas Fearon** laboratory studies the interaction between cancer and the immune system. Their underlying premise is that the tumor microenvironment is immune suppressive because cancer cells elicit responses characteristic of wound healing and tissue regeneration. This approach has led to the finding that activated fibroblasts in the tumor stroma mediate immune suppression in several mouse models of cancer, including the autochthonous model of pancreatic ductal adenocarcinoma of the Tuveson laboratory. Their understanding of the basis of immune suppression is evolving, but they know that it involves the production of the chemokine CXCL12 by the fibroblastic stromal cells, binding of this CXCL12 by pancreatic cancer cells, and exclusion of T cells from the vicinity of the cancer cells. T-cell exclusion, which also occurs in several types of human adenocarcinomas, causes antagonists of T-cell checkpoints to be ineffective despite the presence of cancer-specific $CD8^+$ T cells. This immune suppression is interrupted by administering AMD3100, an inhibitor of CXCR4, the receptor for CXCL12, which leads to the rapid accumulation of T cells among cancer cells, thereby uncovering the efficacy of anti-PD-L1 and eliminating cancer cells. Because human pancreatic cancer has certain immunological characteristics of the mouse model, a phase 1 clinical trial of AMD3100 in patients with pancreatic cancer was initiated in 2015. Some of the Fearon laboratory's next steps are to determine the biological process that causes cancer cells to express nonmutated, shared antigens, and the means by which dormant metastases escape immune elimination.

How do tumors interact with the biology of the host system? What can be learned from studying the physiology and biochemistry of the host system in the context of cancer? These are the principal questions that drive research in the **Tobias Janowitz** laboratory. For example, they investigate the convergence of systemic metabolic stress, endocrinology, and suppressed anticancer immunity to discover mechanism-based strategies for combination therapy for patients with cancer. They have shown that interleukin-6-induced metabolic stress is sufficient to down-regulate hepatic ketogenesis. This causes significant systemic stress during periods of caloric deficiency that are often part of the cancer care pathway. The resulting elevation of glucocorticoids suppresses antitumor immunity in model systems of pancreatic cancer. Using clinical samples and data, they have shown correlative findings of weight loss, reduced ketogenesis, and elevated glucocorticoids in patients with pancreatic cancer. Their work therefore confirms that cancer cannot be understood and probably cannot be treated by investigating tumors in isolation. They use findings like these to develop strategies for interventional studies with the aim of improving outcomes for patients with cancer.

The **Michael James Lukey** laboratory aims to understand the nature and regulation of metabolic adaptations in the different stages of cancer, and then to develop therapeutic strategies that target resulting vulnerabilities. Proliferative signals in mammalian cells drive biosynthetic programs that support cell growth and replication. In healthy cells, this process is tightly regulated by growth factors, but in cancer cells, oncogenic lesions can result in continuous signaling to the metabolic machinery. Oncogene-driven metabolic reprogramming supports tumorigenesis but renders cells sensitive to specific metabolic stresses, a phenomenon that is exploited for cancer therapy. Because the distribution of nutrients varies markedly between organs, cancer cells growing at different sites in the body—and in different regions of the tumor microenvironment—must use a range of metabolic strategies to fuel their growth. The Lukey laboratory is especially interested in the biochemical processes underlying nutrient sensing and metabolic/redox homeostasis, including regulation of the protein posttranslational modification landscape by reactive metabolites. They are also exploring the reciprocal connections between tumor metabolism and host physiology, recognizing that metabolic therapies must be designed to synergize with, and not to antagonize, the antitumor immune response.

**Nicholas Tonks** and colleagues study a family of enzymes called protein tyrosine phosphatases, or PTPs, which remove phosphate groups from proteins and other signaling molecules, such as lipids, in cells. Disruption of PTP function is a cause of major human diseases, and several of the PTPs are potential therapeutic targets for such diseases. Tonks' group seeks to fully characterize the PTP family, understanding how PTP activity is controlled and how PTPs modify signaling pathways. In addition, they are working to determine how those pathways are abrogated in serious illnesses, including cancer, diabetes, and Parkinson's disease. The overall goal is to identify new targets and strategies for therapeutic intervention in human disease. Tonks and colleagues have defined new roles for PTPs in regulating signaling events in breast cancer, identifying three PTPs as novel potential tumor suppressors. They have characterized the regulation of PTP1B by reversible oxidation, demonstrating that it is regulated by covalent modification of the active site by hydrogen sulfide ($H_2S$) under conditions of endoplasmic reticulum stress that are linked to protein-folding-related pathologies, such as Parkinson's and Alzheimer's. In addition, they have generated recombinant antibodies that selectively recognize the oxidized conformation of PTP1B; these antibodies display the ability to promote insulin signaling in cells and suggest novel approaches to therapy for diabetes. Finally, they have also discovered a novel mechanism for allosteric regulation of PTP1B activity, offering the possibility of developing small-molecule drugs that could inhibit the phosphatase and thereby modulate signaling by insulin and the oncoprotein tyrosine kinase HER2, potentially offering new ways to treat insulin resistance in type 2 diabetes and breast cancer.

The **Lloyd Trotman** laboratory developed the RapidCaP model system to study metastatic prostate cancer in vivo. This approach allows them to visualize and probe metastasis biology in the fully native and immune competent setting. They combine this with 3D whole-organ imaging at single-cell resolution to define the stepwise progression from primary tumor escape to colonization of metastatic sites such as bone, liver, and lung. A major challenge in metastasis research is the identification of factors that control progression of non-life-threatening tumors to lethal metastatic disease. In spite of much knowledge on the role of tumor–neuron interactions in primary prostate cancer, the role of neurons in metastatic prostate cancer remains to be defined. The Trotman laboratory uses their whole-organ imaging platform on metastatic tumors to map and manipulate interactions between the peripheral nervous system and metastatic lesions in collaboration with neuroscientists at CSHL. The Trotman laboratory also uses the RapidCaP model to understand the evolution of metastasis and therapy resistance in order to advance a 70-year-old standard of care, antihormonal therapy, which is not curative. They combine three cutting-edge approaches to test novel therapeutic concepts in vivo: single-cell analysis of

metastatic therapy resistance, dietary control of prostate cancer progression, and discovery of therapeutic targets in genome-wide screens.

**David Tuveson**'s laboratory uses murine and human models of pancreatic cancer to explore the fundamental biology of malignancy and thereby identify new diagnostic and treatment strategies. The laboratory's approaches run the gamut from designing new model systems of disease to developing new therapeutic and diagnostic approaches for rapid evaluation in preclinical and clinical settings. The laboratory's studies make use of organoid cultures—three-dimensional cultures of normal or cancerous epithelia—as ex vivo models to probe cancer biology. Current projects in the laboratory explore changes in redox metabolism associated with pancreatic cancer tumorigenesis, dissect signaling by the Ras oncogene, discover new biomarkers of early pancreas cancer, and identify mechanisms of cross talk between pancreatic cancer cells and the tumor stroma. Novel treatment approaches suggested by these studies are then tested by performing therapeutic experiments in mouse models. To dissect molecular changes associated with pancreatic tumorigenesis, the Tuveson laboratory has generated a large collection of human patient–derived organoid models. By measuring the therapeutic sensitivities of patient-derived organoids, the laboratory is working to identify novel strategies to treat patients as well as markers of therapeutic response. The Tuveson Laboratory maintains strong links to clinical research, and the ultimate goal is the confirmation of preclinical findings in early-phase trials. Collectively, the laboratory's bench-to-bedside approach is codified as the "Cancer Therapeutics Initiative," and this initiative will provide these same approaches to the entire CSHL cancer community.

Dr. Tuveson serves as Director of the Cold Spring Harbor Laboratory Cancer Center and as Chief Scientist for the Lustgarten Foundation.

**Linda Van Aelst's** laboratory studies how aberrations in intracellular signaling involving enzymes called small GTPases can result in disease. They are particularly interested in Ras and Rho GTPases, which help control cellular growth, differentiation, and morphogenesis. Alterations affecting Ras and Rho functions are involved in cancer and various neurodevelopmental disorders. Van Aelst's team has extended its prior study of mutations in a Rho-linked gene called oligophrenin-1 (*OPHN1*), part of an effort to connect the genetic abnormalities associated with mental retardation to biological processes that establish and modify the function of neuronal circuits. In addition to a role for OPHN1 in activity-driven glutamatergic synapse development, laboratory members have obtained evidence that *OPHN1* has a crucial role in mediating mGluR-LTD (long-term depression), a form of long-term synaptic plasticity, in CA1 hippocampal neurons. Their findings provide novel insight not only into the mechanism and function of mGluR-dependent LTD, but also into the cellular basis by which mutations in *OPHN1* could contribute to the cognitive deficits observed in patients. Defects in cortical neurogenesis have been associated with cerebral malformations and disorders of cortical organization. The Van Aelst team discovered that interfering with the function of the Rho activator DOCK7 in neuronal progenitors in embryonic cerebral cortices increases the number of proliferating neuronal progenitors and defects in the genesis of neurons. In an extension of these studies, the Van Aelst team showed that DOCK7 has a central regulatory role in the process that determines how and when a radial glial cell progenitor "decides" to either proliferate (i.e., make more progenitor cells like itself) or give rise to cells that will mature, or "differentiate," into pyramidal neurons. These lines of research provide novel insight into mechanisms that coordinate the maintenance of the neural progenitor pool and neurogenesis.

# NUTRITIONAL REGULATION OF REGENERATION, CANCER, AND IMMUNITY

**S. Beyaz** M. Barbi O. Eskiocak T. Maher
P. Bunk A. Garipcan V. Shah
C. Chung M. Gorman B. Yueh

The Beyaz laboratory is focused on understanding the mechanistic basis for nutritional and metabolic regulation of cell fate. Although it is becoming increasingly evident that nutrients may directly regulate cell fate and function, little is known about the underlying causal mechanisms. During the past year, we made significant progress on understanding how nutrients and metabolism affect tissue regeneration, cancer, and immunity. For example, we developed new genetic (Mou et al. 2023) and patient-derived organoid models (Katcher et al. 2023) to study cancer. Furthermore, we provided new insights into regulation of cancer cell plasticity and MHC class II–mediated immune responses in cancer progression and metastasis (Lei et al. 2023). Detailed progress on research activities is outlined below.

## Dietary Arachidonic Acid Induces Regenerative Stem Cell Plasticity without Promoting Tumorigenesis

Intestinal epithelial cells acquire stem cell activity by dedifferentiation to promote regeneration in response to tissue damage, but it remains unclear whether such plasticity occurs under normal physiological conditions (Hu et al. 2023). Furthermore, this plasticity response is exploited by cancer cells, posing a challenge for developing regenerative therapeutics to treat diseases with intestinal epithelial degeneration. Here, we uncovered that a single dietary nutrient, arachidonic acid (AA), can induce stem cell plasticity under homeostatic conditions and boosts regenerative capacity of intestinal epithelium against various damaging insults without increasing risk of cancer. Dietary AA emulates the regenerative response to tissue injury by activating an adaptive and conserved epithelial prostaglandin E2 (PGE2)–Ptger4–cAMP–PKA signaling in mice and humans. Mechanistically, AA elicits epigenetic reprogramming by transcriptional regulators Creb1 and Yap1 to promote stem cell plasticity and regenerative capacity. A key target gene of this transcriptional circuit is S100a6, which is crucial for damage-induced regenerative response. These results unveil a nutrient-triggered epigenetic mechanism for inducing regenerative stem cell plasticity.

This discovery uncouples the processes of physiological regeneration from oncogenic transformation, offering significant potential for cancer patients suffering from mucositis—a debilitating side effect of cytotoxic cancer treatments that causes painful inflammation and ulceration of the digestive tract. Given the substantial economic burden of mucositis in the United States, there is a pressing need for interventions that enhance stem cell regeneration without elevating cancer risk. In light of these findings, we have submitted a patent application and secured venture capital funding to test this dietary regimen in clinical settings. Our goal is to improve the tolerability of cytotoxic therapies, alleviate mucositis, and ultimately enhance the quality of life for cancer patients undergoing treatment.

## Reciprocal Interactions between Cancer, Host Metabolism, and Antitumor Immunity

We made fundamental discoveries that define how cancer impairs host immune surveillance against tumor through metabolic reprogramming. We hypothesized that the competition between cancer cells and immune cells in the tumor microenvironment (TME) is a fitness battle and the shared demands of these cells in the TME might necessitate dependence on similar limiting factors that determine the overall fitness of immune cells against cancer. We sought to understand whether enhancing immune cell fitness through metabolic interventions is sufficient to promote immunogenicity of immune-evasive cancers such as metastatic colorectal cancer (mCRC) and improve immunotherapy efficacy.

We focused on PPARδ, which we previously defined as a lipid-sensing transcription factor (TF) that drives tumor-initiating stem cell fitness in the intestine, but its reported roles in cancer and immune cell function have been conflicting (Beyaz et al., *Nature* **531:** 53 [2016]; Beyaz et al., *Cell Stem Cell* **28:** 598 [2021]). In addition, PPARδ orchestrates cellular lipid catabolism, which is associated both with tumorigenesis and antitumor immune cell activity such as memory T-cell differentiation. However, cell- and context-specific roles of PPARδ and lipid catabolism in antitumor immunity remain incompletely understood.

In this work, we found that activating the transcription factor PPARδ selectively in immune cells eradicates tumors regardless of their antigenicity, type, or stage, without causing host toxicity. We discovered that activating PPARδ selectively in immune cells enhances the clonal expansion, infiltration, and fitness of $CD8^+$ T cells in immune-evasive cancers such as mCRC by promoting mitochondrial fatty acid oxidation (FAO). Conversely, dampening FAO in immune cells by ablating Cpt1a, the rate-limiting enzyme of FAO, reduces $CD8^+$ T-cell infiltration and fitness in a highly antigenic, microsatellite instable (MSI) CRC model. Therapeutically, PPARδ activity in $CD8^+$ T cells boosted the effectiveness of checkpoint blockade immunotherapy against orthotopically established microsatellite-stable mCRC and blunted metastasis in mice. Finally, human CAR-T cells with active PPARδ showed enhanced efficacy in restricting tumor growth. This work defines PPARδ activity and lipid catabolism in $CD8^+$ T cells as a causal determinant of tumor immunogenicity and immunotherapy responsiveness, independent of antigenicity. Thus, immunotherapy modalities that incorporate strategies to elevate the metabolic fitness of immune cells hold potential for treating immune-evasive cancers. We have submitted a patent application on promoting the fitness of antitumor immunity by activating metabolic switches, and a manuscript reporting these findings is currently under review.

## Defining the Mechanisms of Reciprocal Interactions between Cancer and Microbiome

Cancer is associated with alterations in the intestinal microbiome, and the intestinal microbiome appears to influence immune response and therapeutic efficacy in cancer. An important causal mechanism through which microbiome influences cancer risk and outcomes is the modulation of inflammation and immunity. Although there are several factors that can alter microbiome composition and function, dietary and metabolic alterations are the strongest influencers. Because cancer alters systemic metabolism of the host, it is important to interrogate the cancer-induced microbiome alterations, together with metabolic state of the organism. We comprehensively dissected the interactions between cancer and host microbiome to define causal microbe-induced mechanisms that influence antitumor immunity. We found that a lard-based pro-obesity high-fat diet dampens the expression levels of genes involved in MHC-II antigen presentation pathway in tumor-initiating stem cells in the intestine and enhances tumor initiation through escape from immune surveillance (Beyaz et al., *Cell Stem Cell* **28:** 1922 [2021]). Recently, we comprehensively investigated the impact of diet on intestinal microbiome and anticancer immune surveillance by performing extensive metagenomics characterization using the cohort described above. We defined specific gut microbes that sustain the expression of MHC-II antigen presentation pathway genes in epithelial cells and colorectal cancer cells. Strikingly, the microbes that elevate MHC-II expression restrict tumor progression and metastasis, which results in longer overall survival in three different orthotopic colorectal cancer models (including immune-evasive mCRC and two different mouse strains). Importantly, we found that these gut microbes promote antitumor T-cell immunity by enhancing T-cell priming and tumor infiltration in a nonimmunogenic and metastatic model of colorectal cancer. Using loss-of-function and gain-of-function models, we found that cancer MHC-II expression is necessary and sufficient to mediate microbe-induced restriction of colorectal cancer progression and metastasis. To test the therapeutic significance of this microbe–MHC-II axis in metastatic colorectal cancer, we first orthotopically established nonantigenic immune-evasive metastatic AKPS tumors in distal colon. Once the primary tumors were confirmed, we inoculated these mice with microbes that elevate MHC-II expression and started combination immunotherapy treatment with anti-PD1 and anti-CTLA4. Although the tumors in control mice did not respond to immunotherapy and continued to grow and metastasize to distant organs

such as liver, colonization with microbes that elevate MHC-II expression synergized with immunotherapy to restrict tumor progression and blocked metastasis. To explore the human relevance of these findings, we overexpressed CIITA, the transactivator that regulates the expression of MHC-II pathway genes, in colorectal cancer patient–derived organoids. We found that CIITA-mediated elevation of MHC-II promotes antitumor immunity in autologous patient-derived organoid–immune cell co-cultures. These findings suggest that gut microbes that enhance MHC-II antigen presentation promote antitumor immunity and may be utilized to enhance efficacy of immunotherapy in treatment-resistant metastatic colorectal cancer.

Although this work defines the elevation of MHC-II antigen presentation in cancer cells as a causal driver of tumor immunogenicity in colorectal cancer, the expression of MHC-II in other cancer types exhibits divergent outcomes on antitumor immunity. For example, we recently published our findings on how breast cancer MHC-II expression without costimulatory molecule expression promotes lymph node metastasis by eliciting Treg responses (Lei et al. 2023). These results indicate contextual and complex roles of MHC-II antigen presentation in cancer cells during immune surveillance of cancer in different tissue types and stages, which warrant further investigations.

## Development of Human Organoid Models to Study Nutrient–Gene Interactions in Humans

To study the mechanistic basis of how perturbations in nutrient metabolism influence human physiology and disease states, we established comprehensive human tissue and patient-derived organoid (PDO) biobanks. Given their significant association with cancer in the context of obesity, we focused on endometrium (between a two- to sevenfold increase in cancer risk in obesity) and colon (~40% increase in cancer risk in obesity). We formed a collaborative network between scientists and clinicians at CSHL, Northwell Health, and the New York Genome Center (NYGC) to contribute to the NYGC Polyethnic-1000 initiative with the aim to understand the mechanistic basis of cancer health disparities in colorectal cancer and endometrial cancer, including potential genetic or environmental risk factors such as obesity. In addition, our collaborative team recently became part of the Chan Zuckerberg Initiative Human Cell Atlas Network to develop the most comprehensive cellular map of endometrium across ancestries with lean and obese states. To study the mechanisms of nutrient–gene interactions in endometrial cancer, we developed PDO models (of more than 200 patients) across endometrial cancer subtypes, including rare but very aggressive cancers such as carcinosarcoma (Katcher et al. 2023).

## CRISPR-Induced Exon Skipping of β-Catenin Reveals Tumorigenic Mutants Driving Distinct Subtypes of Liver Cancer

CRISPR-Cas9 is widely used to define the significance of cancer oncogenes and genetic dependencies in loss-of-function studies. However, how CRISPR-Cas9 influences gain-of-function oncogenic mutations is elusive. Here, we demonstrate that single-guide RNA targeting exon 3 of *Ctnnb1* (encoding β-catenin) results in exon skipping and generates gain-of-function isoforms in vivo. CRISPR-Cas9-mediated exon skipping of *Ctnnb1* induces liver tumor formation in synergy with $YAP^{S127A}$ in mice. We define two distinct exon skipping–induced tumor subtypes with different histological, transcriptional, and metabolic features. Notably, ectopic expression of two exon-skipped β-catenin transcript isoforms together with $YAP^{S127A}$ phenocopies the two distinct subtypes of liver cancer. Moreover, we identify similar *CTNNB1* exon-skipping events in patients with hepatocellular carcinoma. Collectively, our findings advance our understanding of how different β-catenin isoforms influence tumorigenesis (Mou et al. 2023).

### PUBLICATIONS

Burr AHP, Ji J, Ozler K, Mentrup HL, Eskiocak O, Yueh B, Cumberland R, Menk AV, Rittenhouse N, Marshall CW, et al. 2023. Excess dietary sugar alters colonocyte metabolism and impairs the proliferative response to damage. *Cell Mol Gastroenterol Hepatol* **16:** 287–316. doi:10.1016/j.jcmgh.2023.05.001

Cable J, Rathmell JC, Pearce EL, Ho P-C, Haigis MC, Mamedov MR, Wu M-J, Kaech SM, Lynch L, Febbraio MA, et al. 2023. Immunometabolism at the crossroads of obesity and cancer—a Keystone Symposia report. *Ann NY Acad Sci* **1523:** 38–50. doi:10.1111/nyas.14976

Hu T, Allam M, Cai S, Henderson W, Yueh B, Garipcan A, Ievlev AV, Afkarian M, Beyaz S, Coskun AF. 2023. Single-cell spatial

metabolomics with cell-type specific protein profiling for tissue systems biology. *Nat Commun* **14:** 8260. doi:10.1038/s41467-023-43917-5

Katcher A, Yueh B, Ozler K, Nizam A, Kredentser A, Chung C, Frimer M, Goldberg GL, Beyaz S. 2023. Establishing patient-derived organoids from human endometrial cancer and normal endometrium. *Front Endocrinol (Lausanne)* **14:** 1059228. doi:10.3389/fendo.2023.1059228

Lei P-J, Pereira ER, Andersson P, Amoozgar Z, Van Wijnbergen JW, O'Melia MT, Zhou H, Chatterjee S, Ho WW, Posada JM, et al. 2023. Cancer cell plasticity and MHC-II-mediated immune tolerance promote breast cancer metastasis to lymph nodes. *J Exp Med* **220:** e20221847. doi:10.1084/jem.20221847

Mou H, Eskiocak O, Özler KA, Gorman M, Yue J, Jin Y, Wang Z, Gao Y, Janowitz T, Meyer HV, et al. 2023. CRISPR-induced exon skipping of β-catenin reveals tumorigenic mutants driving distinct subtypes of liver cancer. *J Pathol* **259:** 415–427.

Tierney BT, Kim J, Overbey EG, Ryon KA, Foox J, Sierra M, Bhattacharya C, Damle N, Najjar D, Park J, et al. 2023. Viral activation and ecological restructuring characterize a microbiome axis of spaceflight-associated immune activation. *Res Sq* doi:10.21203/rs.3.rs-2493867/v1

Zilbauer M, James KR, Kaur M, Pott S, Li Z, Burger A, Thiagarajah JR, Burclaff J, Jahnsen FL, Perrone F, et al. 2023. A roadmap for the human Gut Cell Atlas. *Nat Rev Gastroenterol Hepatol* **20:** 597–614. doi:10.1038/s41575-023-00784-1

*In Press*

Rogava M, Aprati TJ, Chi W-Y, Melms JC, Hug C, Davis SH, Earlie EM, Chung C, Deshmukh SK, Wu S, et al. 2024. Loss of *Pip4k2c* confers liver-metastatic organotropism through insulin-dependent PI3K-AKT pathway activation. *Nat Cancer* **5:** 443–447. doi:10.1038/s43018-023-00704-x

# LOCAL AND SYSTEMIC INTERACTIONS AMONG THE NERVOUS SYSTEM AND CANCER

**J.C. Borniger** A. Battison L. Boyd A.M. Gomez N. Holland M. Sherman A. Tiwari Y. Wu

## Neuronal Basis for Disrupted Glucocorticoid Circadian Rhythms in Breast Cancer

A.M. Gomez, Y. Wu, L. Boyd, J.C. Borniger [in collaboration with C. Zhang, L. Cheadle, and C. Amor Vegas, CSHL]

Patients with breast cancer often exhibit disrupted or "blunted" circadian rhythms in circulating glucocorticoids (e.g., cortisol). This is associated with reduced quality of life and increased mortality. Adrian Gomez, Yue (May) Wu, and several other laboratory members have now made significant progress in unraveling the neural basis behind this phenomenon and demonstrated that reinforcing robust glucocorticoid rhythms attenuates tumor progression in breast cancer mouse models. We observed that mice with mammary tumors exhibit diminished circadian glucocorticoid rhythms before tumor palpability, concomitant with elevated activity of paraventricular hypothalamic neurons expressing corticotropin-releasing hormone (i.e., $PVN^{CRH}$ neurons) (Figs. 1 and 2). RNA-sequencing these neurons demonstrated that tumors promoted dysregulated expression of genes contributing to neuronal excitatory/inhibitory (E/I) balance. In line with this, $PVN^{CRH}$ neurons from tumor-bearing mice had reduced numbers of inhibitory synapses (VGAT/$Gephyrin^{+}$) contacting them (Fig. 3). Adrenal and pituitary functions remained unaltered, implicating hypothalamic alterations as the primary cause of HPA-axis dysfunction in breast cancer. Chemogenetic stimulation of $PVN^{CRH}$ neurons at the light-to-dark transition reestablished robust endogenous glucocorticoid rhythms, and this was associated with reduced tumor growth and a more robust antitumor immune response. Our findings reveal the impact of cancer on circadian glucocorticoid rhythms and provide novel insights into the potential therapeutic targeting of subcortical neurons regulating these rhythms to mitigate tumor progression.

## Glycan Landscape in Glioblastoma and Development of a GCaMP-BioID2 Fusion Protein

A. Battison, J.C. Borniger [in collaboration with S. George and L. Van Aelst, CSHL]

The overarching goal of this project is to elucidate the role of glycosylation in cancer and address the therapeutic potential of modifying cancer-associated glycans to ameliorate cancer severity. Despite being one of the most prevalent and complex protein posttranslational modifications, the role of glycans in combatting or promoting cancer has not yet been elucidated. Glioblastomas (GBs) present a formidable challenge in oncology because of their aggressive nature and resistance to therapy, leading to poor patient outcomes. Despite advances in treatment strategies, the intricate molecular and cellular landscape of GBs impedes effective therapeutic interventions. Alix Battison's project aims to investigate the role of glycosylation in GB pathogenesis and investigate the therapeutic potential of targeting glycans. This will be achieved using a mouse model of GB created by a CRISPR-Cas9 gene knockdown delivered via in utero electroporation in transgenic mouse strains. A map of glycan localization in and around the tumor will first be obtained using spatial transcriptomics and spatial proteomics. Subsequently the surface glycoproteome of GB cells will be determined using the iPEEL mouse strain to identify and characterize surface glycans on tumor cells, shedding light on their role in tumor behavior. Finally, a system will be developed for manipulating specific glycoforms using CRISPR-Cas9, and this will be applied to evaluate the impact of glycans on tumor growth and neuronal circuit activity relating to glioma-associated epilepsy.

In a separate project, Alix is developing a novel tool to link neuronal activity to protein expression. GCaMP (a genetically encoded calcium indicator [GECI]) is a widely used tool in the field of systems neuroscience

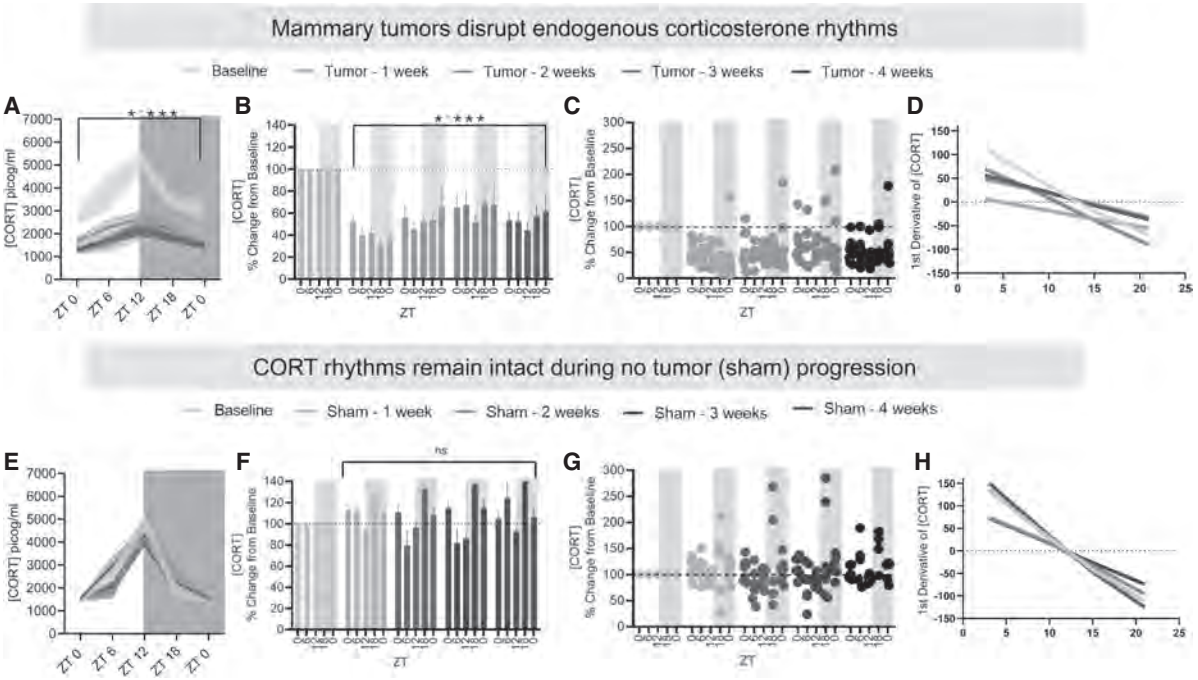



**Figure 1.** Mammary tumors disrupt endogenous glucocorticoid circadian rhythms. (*A*) Corticosterone concentrations over the day at baseline (before tumor cell injection) and weekly during cancer progression. Note the strong rhythm at baseline that reduces significantly by week 1 postinjection. (*B,C,D*) Different ways of looking at the data as a percentage change from baseline and the instantaneous slope of the line (first derivative). (*E–H*) The same as *A–D* for control (non-tumor-bearing) mice, demonstrating the specific effect is driven by tumor presence and not other extraneous factors. ($n = 9–11$ mice/group; error bars represent standard error of the mean.)

and cell biology. GCaMP is a genetically engineered protein that detects changes in intracellular calcium levels, serving as a fluorescent indicator of neuronal activity or other cellular processes. GCaMP is typically used to analyze the behavior of whole cells or large ensembles of cells in vitro or in vivo. Although GCaMP can provide information about temporal and spatial dynamics, it cannot convey information about the channels, proteins, and pathways mediating these dynamics. Proximity-labeling proteomics is a powerful technique used to identify and study proteins that interact with a specific target protein within a cell or subcellular compartment. This method involves the fusion of the protein of interest (the "bait") with a biotin-ligase enzyme. A new tool that would elevate our understanding of the molecular mediators of cellular activity would allow for functional imaging of calcium dynamics followed by proteomic analysis of the exact same cells to determine what proteins are altered in different experimental conditions that then lead to the differing calcium dynamics. This would help bridge the knowledge gap between systems and molecular neuroscience and provide a direct readout of the molecular drivers of function and physiology. Alix has now generated several GCaMP-BioID2 and GCaMP-APEX2 constructs and is testing them in cell lines to determine which to use in subsequent validation in vitro and in vivo.

## Horizontal Mitochondria Transfer between Neurons and Breast Cancer Cells

A. Tiwari, Y. Wu, J.C. Borniger

Mitochondria are powerhouses of the cell, but our understanding of mitochondria has expanded beyond their role in ATP generation. Mitochondria play a significant role in biosynthetic reactions, stress responses, cell death, and cell differentiation. Recent studies have provided insights that mitochondria are significant for intercellular communication by lateral mitochondria transfer. The nervous system regulates organogenesis and oncogenesis. Recent discoveries

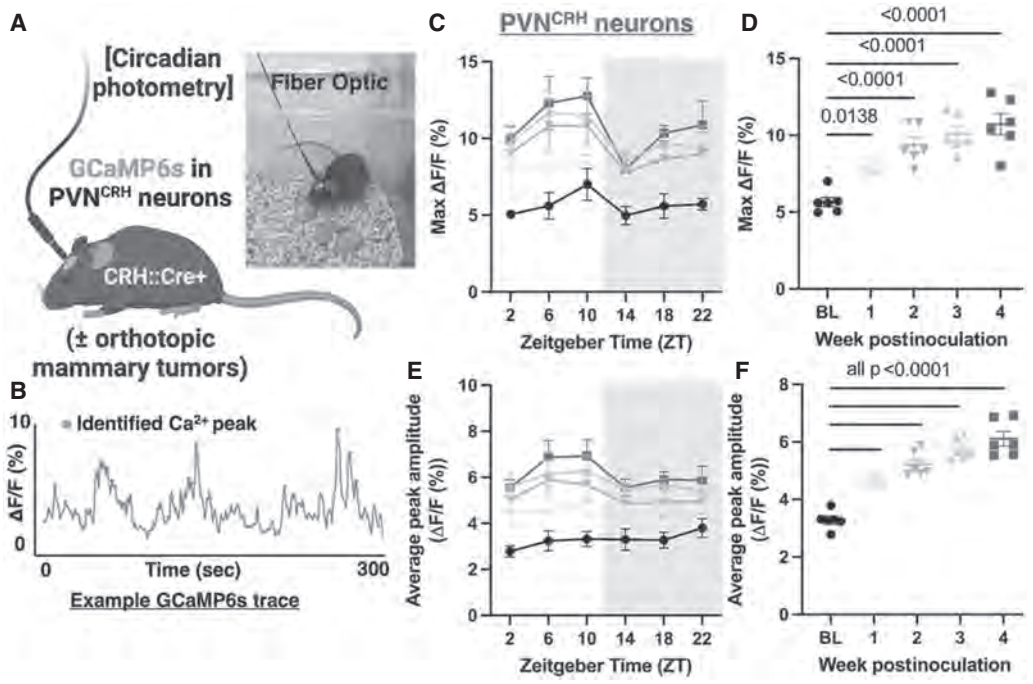



**Figure 2.** Mammary tumors aberrantly activate PVN$^{CRH}$ neurons. (*A*) Experimental design showing fiber-optic implantation into a CRH::Cre mouse that was injected with a GCaMP6s-expressing adeno-associated virus (AAV). Following surgery, these mice were allowed to recover before receiving sham or tumor cell injections. (*B*) Example 5-min $Ca^{2+}$ trace showing identified peaks (dots). (*C,D*) The maximum GCaMP6s signal across the day at baseline (darkest color) and during tumor progression (other lines). (*E,F*) The average peak amplitude of the GCaMP6s signal at baseline and during tumor progression. ($n = 6$ mice/group; error bars represent SEM.)

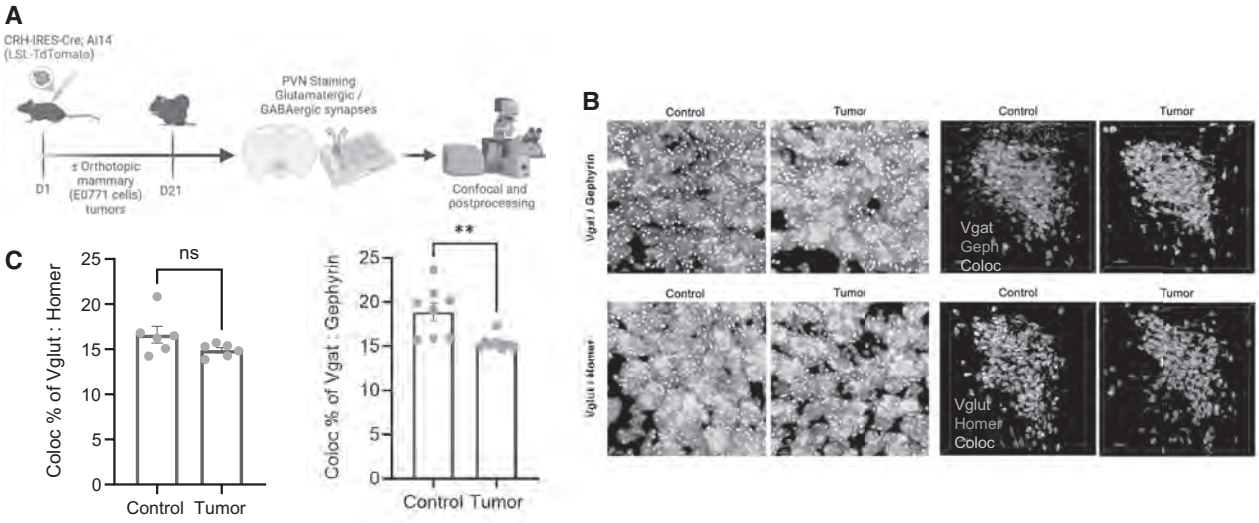



**Figure 3.** PVN$^{CRH}$ neurons from tumor-bearing mice have reduced inhibitory input. (*A*) Experimental design to label and quantify excitatory (VGLUT2/Homer1+) and inhibitory (Vgat/Gephyrin$^{+}$) synapses on cortico-releasing hormone (CRH) neurons in the paraventricular nucleus of the hypothalamus. (*B*) Example confocal images showing pre- (Vglut2 or Vgat) and post- (Homer1 or Gephyrin) synaptic proteins and their colocalization, indicating fully assembled synapses. (*C*) Quantification of these data showing no change in excitatory input but a reduction in inhibitory input to PVN$^{CRH}$ neurons in tumor-bearing mice. ($n = 7$ mice/group; error bars represent SEM.)

have elucidated cross talk between the nervous system and tumors, at both systemic and direct (paracrine or electrochemical) levels, with subsequent effects on tumor growth. Our laboratory recently discovered that sensory neurons are a reservoir of mitochondria that transfer to breast cancer cells. To our knowledge, this is the first evidence of neuron-to-cancer cell mitochondrial transfer. Beyond the lateral mitochondria transfer, mitochondria play a critical role in systemic communication between different organ systems. We are working on deciphering how mitochondria function in the context of cancer is altered in the brain. Further, we are investigating how these alterations in mitochondrial network activity within the brain have systemic effects on whole-body physiology.

## Lateral Proteome Transfer between Sensory Neurons and Breast Cancer

Y. Wu, J.C. Borniger

Recent research indicates a crucial involvement of the nervous system in cancer development, with bidirectional interactions occurring via direct membrane contacts or within the tumor microenvironment. Although investigations into nervous system interactions with gliomas in the brain are extensive, studies exploring the role of the peripheral nervous system in extracranial tumor microenvironments are still limited. Yue (May) is now exploring the interactions between peripheral neurons and breast tumor cells in vivo and in vitro.

Our findings from breast tumor tissue clearing reveal sensory nerves as the predominant nerve type in the tumor microenvironment across various stages of tumor progression in mouse breast tumor models. Therefore, we established a co-culture system in which breast cancer cells directly interact with primary sensory neurons, resulting in increased proliferation and migration rates, particularly in the presence of stromal cells. RNA-sequencing analysis also identified up-regulation of key signaling pathway genes involved in adhesion, migration, and proliferation in co-cultured breast cancer cells, including neuroligin-2, nerve growth factor (NGF) and nerve growth factor receptor (NGFR), GDF-15, syndecan-3, semaphorin, and plexin-B. Furthermore, using bio-orthogonal labeling, stable isotope labeling by amino acids in cell culture (SILAC) and immunostaining, we demonstrated intercellular protein and mitochondria trafficking between sensory neurons and breast cancer cells, suggesting potential implications in cellular homeostasis, damaged cell repair, tumor progression, and immunoregulation.

To further investigate the mechanisms underlying this sensory neuron–cancer cell interaction, May is now focused on identifying the specific proteins being trafficked, investigating the structural aspects of the cellular connections involved, and understanding their functional implications. This will be achieved through techniques such as mass spectrometry, electron microscopy, whole-cell patch-clamp recording, microelectrode array, and intravital imaging. By studying the interactions between the peripheral nervous system and breast cancer within the tumor microenvironment, we aim to gain a comprehensive understanding of the molecular and cellular mechanisms driving breast cancer progression and potentially identify novel therapeutic targets.

## Neuronal Circuit Mechanisms of Stress-Induced Insomnia and Metabolic Dysfunction

M. Sherman, A.M. Gomez, J.C. Borniger

Marina Sherman's project examines stress, and specifically CRH neurons, in the paraventricular nucleus of the hypothalamus (PVN) and their effect on metabolic function and insomnia. This is based on ample evidence that insomnia is associated with the development of metabolic disease, and prior research from our group and others has demonstrated a role for PVN CRH neurons in stress-induced insomnia. So far, we have confirmed the association between the stress response and activation of these CRH PVN neurons in a mouse model. We also assessed the dysregulation of many metabolic functions during stress or chemogenetic CRH PVN neuron activation and found increased energy expenditure and breathing rate across both groups as compared with control mice. Additionally, genes associated with glycolysis and gluconeogenesis were generally up-regulated or unchanged in the liver of stress or activated CRH PVN groups. Further work is being conducted to evaluate how these neurons affect glucose homeostasis through tolerance tests for glucose, insulin, lactate, and pyruvate.

## Role of Locus Coeruleus Noradrenergic Neurons in Breast Cancer

N. Holland, L. Boyd, J.C. Borniger [in collaboration with C. Amor Vegas, CSHL]

Graduate student Nik Holland and technician Leah Boyd have been leading a project to determine how neuromodulation of the brainstem locus coeruleus influences breast cancer outcomes. This project stems from recent advances demonstrating that noradrenergic (NE) signaling within the tumor microenvironment powerfully drives tumor growth via direct actions on tumor cells, modifying endothelial cell function, or polarizing immune cells toward a pro-tumor phenotype. However, no study has examined the upstream neurons (in the locus coeruleus) that drive peripheral norepinephrine release. Leah and Nik are using daily optogenetic stimulation of dopamine-β-hydroxylase expressing LC neurons to determine how changing their activity influences tumor growth. One hour per day (at 3 Hz) of optogenetic stimulation was sufficient to significantly increase primary tumor growth in the EO771 mammary carcinoma model. Currently, Nik and Leah are examining the proximate mechanisms for this by analyzing tumor immune cell infiltrate and determining whether and how cancer in the mammary gland modifies the activity of $LC^{NE}$ neurons. Using pseudorabies virus (PRV) polysynaptic tracing, they identified a subpopulation of LC neurons that project to the developing tumor in the mammary gland. The importance of this connection is currently under investigation.

## Other Projects

L. Boyd, N. Holland, M. Sherman, J.C. Borniger [in collaboration with T. Janowitz, CSHL]

In addition to the main projects described above, there are other research efforts that are still in early stages. These include research by Leah Boyd to elucidate how cancer and chemotherapy affect sleep and circadian rhythms in mice, and, reciprocally, whether correcting altered sleep patterns in tumor-bearing mice improves tumor progression and survival. This is work done in collaboration with Dr. Erin Gibson's group at Stanford University. We recorded longitudinal EEG/EMG signals from mice during and following several cycles of methotrexate (MTX) chemotherapy treatment. This work is based on prior results from the Monje and Gibson laboratories demonstrating that MTX promotes glial cell dysfunction to drive cognitive impairments in mice. So far, the data suggest that MTX causes excessive wakefulness through both increased numbers of microwake episodes (<5 sec) during the normal sleep phase and increased long wakefulness bouts during the active phase. Separately, Leah is collaborating with Tobias Janowitz's laboratory to examine the role of sleep in regulating the onset and development of cancer cachexia. She is using chemogenetics to silence dopaminergic neurons in the mesencephalic ventral tegmental area ($VTA^{DA}$ neurons), which induces natural sleep in mice. She will then determine whether correcting sleep disruption improves outcomes in a novel murine model of cancer cachexia using an IL-6-overexpressing cancer cell line developed in the Janowitz laboratory.

Nik and Marina in the laboratory are spearheading an effort to understand how cancer alters the brain's response to stressful stimuli. Taking advantage of TRAP2::Ai14 mice, in which transiently active neurons can be labeled permanently with TdTomato upon injection of 4-hydroxytamoxifen (4-OHT), they are probing how neuronal activity changes in response to restraint stress with and without the presence of breast cancer. Preliminary data strongly indicate that several subcortical neuronal populations become aberrantly active in response to cancer during stress, indicating that indeed the brain does respond to stressful stimuli in the context of breast cancer. We are now expanding this work to create a whole-brain atlas of stress-induced neuronal activity, and additionally breeding TRAP2::Sun1-GFP mice to isolate 'trapped' neurons for single-cell sequencing.

### PUBLICATION

Wu Y, Borniger JC. 2023. Systemic interactions between cancer and the nervous system. In *Cancer neuroscience* (eds. Amit M, Scheff NN), pp. 147–168. Springer, New York.

# THE MICROENVIRONMENT: HOST CONTRIBUTIONS TO METASTASIS

**M. Egeblad** J. Adrover X. Han S. Shiu
E. Gazzara X-Y. He L. Sun
S. Gholami N. Sivetz

Solid tumors are aberrant tissues. Like organs, solid tumors are composed of cells and stroma. The stroma is the supportive framework of the organs and includes the extracellular matrix (ECM) as well as immune cells, fibroblasts, adipocytes, and cells of the vascular system. Interactions between epithelium and stroma are essential for normal organ development but become dysregulated during disease—particularly cancer.

In solid tumors, the stroma is also known as the tumor microenvironment. The Egeblad laboratory studies how the tumor microenvironment influences cancer cells in the context of tumor initiation, growth, drug resistance, and metastasis. We use mouse models of primarily breast, ovarian, and pancreatic cancer together with real-time spinning disk confocal and multiphoton microscopy in living mice, a technique known as intravital imaging. This allows us to study how cancer cell proliferation, migration, and survival are influenced by the microenvironment in real time. During the last years, our focus has been the bidirectional interactions between cancer and the host—how tumors alter host biology and how altered host biology affects cancer progression.

## Stress-Induced Metastasis

X. He, J. Adrover, X. Han, S. Shiu, L. Sun [in collaboration with D. Ng, Rockefeller University; L. Van Aelst, C. Vakoc, and D. Spector, CSHL]

Stress is a multifaceted physiological reaction to environmental or psychological stimuli that triggers a series of systemic effects, beginning with information processing in the central nervous system. The stress response is characterized by the hypothalamus releasing corticotropin-releasing hormone, which leads to the secretion of adrenocorticotropic hormone by the anterior pituitary, ultimately resulting in the release of glucocorticoids from the adrenal glands. Although there is substantial epidemiological and clinical evidence connecting chronic stress, depression, and social isolation with a risk of cancer onset and recurrence, the precise mechanisms through which stress contributes to metastasis specifically are not well understood.

To explore how stress influences metastasis, a novel model of disseminated breast cancer was developed. In this model, breast cancer cells derived from MMTV-PyMT mice were transplanted into the mammary fat pads of other mice. After the primary tumors reached a manageable size—allowing for some dissemination but before metastases developed—the tumors were surgically removed, and the mice were exposed to chronic stress via restraint. Remarkably, the stressed mice developed a fourfold increase in metastatic burden compared with the control group.

An elevated neutrophil-to-lymphocyte ratio in the blood is linked to poor outcomes in breast and other cancers and is also seen in both animals and humans exposed to stress. Given prior findings of neutrophils' role in metastasis, we explored connections between stress, neutrophils, and metastasis. The results showed that stress caused a significant increase in neutrophil recruitment to the lungs, and depleting neutrophils in the stressed mice reduced metastasis. Neutrophils primarily combat harmful microorganisms, often using a unique mechanism involving the formation of neutrophil extracellular traps (NETs), which are DNA-based structures that trap and kill pathogens. However, research has shown that NETs can also foster metastasis in breast, ovarian, and lung cancers. In our study, chronically stressed mice exhibited elevated NET levels, and daily injections of DNase I, an enzyme that degrades NETs, significantly lowered metastasis. These NETs also altered the lung microenvironment by increasing fibronectin (a protein that supports metastasis) and reducing cytotoxic T-cell infiltration and activation.

The next question was how stress leads to NET formation. Stress hormones, such as glucocorticoids and adrenaline, are produced by the adrenal glands. We observed that adrenalectomy (removal of adrenal glands) in stressed mice prevented NET induction.

Further testing revealed that glucocorticoids, but not adrenaline, triggered NET formation in vitro. Using mice with neutrophil-specific knockouts of the glucocorticoid receptor, we demonstrated that glucocorticoid signaling is necessary for stress to induce NETs and metastasis.

Although the exact signaling pathways responsible for NET formation are not fully understood, the process often involves protein arginine deiminase 4 (PAD4), an enzyme that initiates chromatin decondensation, which is an important step in NET formation. Interestingly, our findings indicate that glucocorticoid-induced NET formation does not rely on PAD4, suggesting alternative pathways. Recent studies suggest that neutrophils may repurpose molecules involved in cell proliferation, such as cyclin-dependent kinase (CDK) 4/6, to form NETs. We found that CDK4/6 inhibitors, specifically palbociclib and abemaciclib, which are used in metastatic breast cancer treatment, inhibited glucocorticoid-induced NET formation in vitro.

Although previous studies have highlighted correlations between chronic stress and cancer progression, little has been known about the cellular and molecular mechanisms driving stress-induced metastasis. Our work provides valuable insights into the role of NETs in this process, and ongoing research aims to further elucidate the pathways through which glucocorticoids promote NET formation and how stress-induced NETs facilitate the metastasis of disseminated cancer cells.

## A Vascular-Restricted, Tumor-Induced Neutrophil Population Drives Vascular Occlusion, Pleomorphic Necrosis, and Metastasis

J. Adrover, E. Gazzara, S. Gholami, X. Han, X-Y. He, N. Sivetz, L. Sun [in collaboration with J. Daßler-Plenker; T. Fujii, National Cancer Institute; Jessica Peters, George Raptis, and C. Evans, Northwell Health; Scott Powers, Stony Brook University; and Won Jin Ho, Johns Hopkins University]

As part of our efforts to characterize pro-metastatic immune responses that regulate the growth of disseminated tumor cells, we identified a previously unknown neutrophil subpopulation that appears exclusively in mice with tumors characterized by large, pleomorphic necrotic regions. These neutrophils are defined by high expression of the Ly6G marker and low expression of Ly6C, a marker found on both neutrophils and monocytes. Our findings revealed that the tumors induce this specific $Ly6G^{High}$ $Ly6C^{low}$ neutrophil population by heavily skewing the hematopoietic system toward granulopoiesis, in part through secretion of CXCL1. Unlike typical $Ly6G^{High}$ $Ly6C^{Medium}$ neutrophils, this newly discovered subpopulation exhibited distinct transcriptional and phenotypic characteristics, including an enhanced ability to form NETs and an inability to exit the bloodstream and infiltrate tissues in response to inflammatory stimuli, such as tissue injury or injection with yeast components. This "vascular-restricted" neutrophil gene expression signature was also detected in neutrophils isolated from breast cancer patients.

We hypothesized that the presence of pleomorphic tumor necrosis in mice with $Ly6G^{High}$ $Ly6C^{low}$ neutrophils was linked to the unique properties of this vascular-restricted neutrophil population. In mouse models with these neutrophils, we observed that NET-containing aggregates formed within the tumor vasculature, leading to vascular occlusion and impaired blood flow, particularly in vessels deficient in pericytes and displaying exposed basal membrane components along with fibrin.

Interestingly, inhibiting NET formation in these models—using conditional knockout mice lacking *Pad4* (a gene essential for NET formation) in neutrophils, or using treatment with DNase I to digest NETs or disulfiram to inhibit NET formation—significantly reduced tumor necrosis. Additionally, lung metastases were diminished across various cancer mouse models. Further analysis using spatial transcriptomics and immunofluorescence revealed that cancer cells near necrotic areas underwent epithelial-to-mesenchymal transition (EMT)-like changes at the transcriptomic and protein levels, a process linked to increased metastatic potential in tumors with pleomorphic necrosis.

For decades, tumor necrosis has been associated with poor prognosis, though it has been thought of as a passive byproduct of tumor growth and largely nontargetable. Our data challenge this view by showing that genetic or pharmacological inhibition of NET formation prevents pleomorphic necrosis, providing crucial evidence that (1) tumor necrosis is not an inevitable consequence of tumor growth and (2) targeting NETs can reduce tumor necrosis and metastasis.

## Sepsis-Induced Inflammation Alters Natural Killer Cell–Mediated Surveillance of Liver Metastasis

N. Sivetz [in collaboration with P.J. Cunniff, C.R. Vakoc, and S. Beyaz, CSHL]

A majority of pancreatic ductal adenocarcinoma (PDAC) patients who undergo surgery for primary tumor resection will develop lethal metastases, despite having no evidence of metastasis at the time of their initial diagnosis. Liver metastases constitute nearly half of recurrences detected within the first six months following PDAC resection. Several clinical and experimental studies provide evidence that perioperative inflammation—including infectious complications like sepsis—positively correlates with liver metastasis across several malignancies. Despite these observations, it is still unknown whether there is a causal link between perioperative inflammation and PDAC liver metastasis.

To gain mechanistic insight into the regulation of PDAC liver metastasis in the context of perioperative inflammation, we established mouse models that combine experimental PDAC liver metastasis with either the cecal ligation and puncture (CLP) model of polymicrobial sepsis or the lipopolysaccharide (LPS) model of acute sterile endotoxemia. Mice challenged with CLP prior to seeding PDAC cells in the liver had increased metastatic burden at end point compared with sham laparotomy controls. In contrast, mice challenged with LPS prior to PDAC cell seeding had reduced metastatic burden at end point compared with PBS-injected controls. We hypothesized that these opposing pro- and antimetastatic effects of CLP and LPS, respectively, were due to qualitatively distinct inflammatory responses elicited in the liver. Our profiling of the host response to either CLP or LPS revealed shared systemic changes, such as the acute-phase protein response. However, LPS-associated inflammation uniquely up-regulated a distinct set of interferon-regulated genes in the liver important for modifying immune cell activity (i.e., *Ccl2*, *Ccl5*, *Cxcl9*, and *Cxcl10*). Accordingly, immunostaining and flow cytometry of the liver revealed that LPS uniquely increased the abundance of activated natural killer (NK) cells and $MPO^+$ myeloid cells. We utilized $Tlr4^{-/-}$ mice, which lack the primary receptor for LPS, to determine that the antimetastatic effect of LPS required a host-mediated response downstream of TLR4-mediated signaling. Through a series of genetic models and short-term depletion experiments, we determined that NK cells restrict PDAC cell seeding and outgrowth in our experimental liver metastasis model, and that this protective effect is further enhanced by activation of type-I interferon signaling in the host.

Our findings confirm prior work identifying NK cell–mediated surveillance of disseminated cancer cells as an important mechanism in protecting against liver metastasis and show that sterile endotoxemia and its associated inflammation establish a protective "antimetastatic niche" marked by enhanced NK cell activity. Further insights into the establishment of this niche may reveal potential therapeutic strategies to reduce the incidence of liver metastasis among resected PDAC patients.

### PUBLICATIONS

Adrover JM, McDowell SAC, He X-Y, Quail DF, Egeblad M. 2023. NETworking with cancer: the bidirectional interplay between cancer and neutrophil extracellular traps. *Cancer Cell* **41:** 505–526. doi:10.1016/j.ccell.2023.02.001

Chen H-A, Ho YJ, Mezzadra R, Adrover JM, Smolkin R, Zhu C, Woess K, Bernstein N, Schmitt G, Fong L, et al. 2023. Senescence rewires microenvironment sensing to facilitate antitumor immunity. *Cancer Discov* **13:** 432–453. doi:10.1158/2159-8290.CD-22-0528

Gao Y, He X-Y, Wu XS, Huang Y-H, Toneyan S, Ha T, Ipsaro JJ, Koo PK, Joshua-Tor L, Bailey KM, et al. 2023. ETV6 dependency in Ewing sarcoma by antagonism of EWS-FLI1-mediated enhancer activation. *Nat Cell Biol* **25:** 298–308. doi:10.1038/s41556-022-01060-1

Hirschhorn D, Budhu S, Redmond D, Gigoux M, Kraehenbuehl L, Schroder D, Chow A, Ricca JM, Gasmi B, De Henau O, et al. 2023. T cell immunotherapies trigger neutrophil activation to eliminate tumor antigen escape variants. *Cell* **186:** 1432–1447.e17. doi:10.1016/j.cell.2023.03.007

Mousset A, Lecorgne E, Bourget I, Lopez P, Dominici C, Jenovai K, Cherfils-Vicini J, Terp MG, Egeblad M, Gaggioli C, Albrengues J. 2023. Neutrophil extracellular traps formed during chemotherapy confers treatment resistance via TGF-β activation. *Cancer Cell* **41:** 757–775.e10. doi:10.1016/j.ccell.2023.03.008

Wang G, Li J, Bojmar L, Chen H, Li Z, Tobias GC, Hu M, Homan EA, Lucott S, Zhao F, et al. 2023. Tumour extracellular vesicles and particles induce liver metabolic dysfunction. *Nature* **618:** 374–382. doi:10.1038/s41586-023-06114-4

*In Press*

He X-Y, Gao Y, Ng D, Michalopoulou E, George S, Adrover JM, Sun L, Albrengues J, Daßler-Plenker J, Han X, et al. 2024. Chronic stress causes metastasis via neutrophil-mediated changes to the microenvironment. *Cancer Cell* **42:** 474–486.e12.

# THE IMMUNOLOGY OF PANCREATIC DUCTAL ADENOCARCINOMA

**D.T. Fearon** J. Li R. Yan
P. Moresco J-I. Yang
J. Pearson M. Yao

## The Plasticity of Cancer-Associated Fibroblasts

R. Yan

Stromal fibroblasts in cancer, also known as cancer-associated fibroblasts (CAFs), have been informative to the Fearon laboratory's view of cancer immunology since 2010 and 2013 when conditional deletion of CAFs resulted in intratumoral immune activation, suppression of tumor growth, and new sensitivity to immune checkpoint blockade. We asked whether this immune-suppressive phenotype can be modulated to support immune activation—which would indicate that the phenotype of CAFs is "plastic." Single-cell RNA sequencing (scRNA-seq) of CAFs from immune-activated mouse pancreatic ductal adenocarcinomas (PDACs) identified a subpopulation with decreased expression of the immune-suppressive chemokine CXCL12 and increased expression of the T-cell-attracting chemokine CXCL9. Recombinant interferon-gamma (IFN-γ) and tumor necrosis factor alpha (TNF-α), cytokines that were present in the immune-activated PDAC, acted together to augment CXCL9 expression, whereas TNF-α alone suppressed CXCL12 expression. This study demonstrates that CAFs have a phenotypic plasticity that allows their adaptation to contrasting immune tissue microenvironments (Yan et al. 2023).

## The Occurrence of Auto-Antibody-Producing Plasma Cells in PDAC

M. Yao [in collaboration with D.A. King, K. Raphael, A. Rishi, D. Sejpal, and M.J. Weiss, Northwell Health and North Shore University Hospital]

Our second study addressed the question of whether PDAC is inherently unable to support antigenic stimulation of lymphocytes, as is generally thought. We sought to discover the antigens driving the intratumoral response of B lymphocytes by using scRNA-seq and immunoglobulin (Ig) sequencing of tumor-infiltrating immune cells from seven primary surgical PDAC samples. Ig sequencing identified plasma cell (PC) clones expressing isotype-switched and hypermutated Igs, indicating occurrence of T-cell-dependent B-cell responses in the PDAC tissue. We assessed the reactivity of recombinant antibodies that represented the products of 235 PCs and observed staining of normal intracellular self-antigens, three of which were identified: F-actin, the nuclear protein RUVBL2, and the mitochondrial heat shock protein member HSPD1. Antibody titers against F-actin and HSPD1 were elevated in the plasma of patients with PDAC compared with healthy donors. Thus, PCs in PDAC produce autoantibodies, which may result from promotion of preexisting, autoreactive B-cell responses. Thus, the inflammatory microenvironment of human PDAC supports the adaptive immune response (Yao et al. 2023).

### PUBLICATIONS

Yan R, Moresco P, Gegenhuber B, Fearon DT. 2023. T cell–mediated development of stromal fibroblasts with an immune-enhancing chemokine profile. *Cancer Immunol Res* **7:** OF1–OF11. doi:10.1158/2326-6066.CIR-22-0593

Yao M, Preall J, Yeh JT, Pappin D, Cifani P, Zhao Y, Shen S, Moresco P, He B, Patel H, et al. 2023. Plasma cells in human pancreatic ductal adenocarcinoma secrete antibodies against self-antigens. *JCI Insight* **8:** e172449.172449.

# WHOLE-BODY CONSEQUENCES OF CANCER PROGRESSION

**T. Janowitz** E. Davidson, B. Demestichas, E. Ertel, M. Ferrer Gonzalez, S. Kleeman, C. Koza, H. Lee, T. Thakir, A. Wang, X. Zhao

The Janowitz laboratory investigates the whole-body response to disease, with a particular interest in cancer. We currently focus on understanding the mechanisms underlying the cancer-induced multiorgan alterations that lead to the development of paraneoplastic syndromes and specifically cachexia, a highly debilitating wasting syndrome that is experienced by many patients with colorectal, lung, and pancreatic cancer. We use basic and translational research approaches to study the interconnection between metabolism, immunology, and neuroendocrinology, emphasizing the importance of studying cancer as a disease that affects the whole body, not just the organ where the tumor is localized.

## Ketogenic Diet Promotes Tumor Ferroptosis but Induces Relative Corticosterone Deficiency That Accelerates Cachexia

M. Ferrer Gonzalez, E. Davidson, S. Kleeman, T. Janowitz

Cancer is, at least in part, a metabolic disease. Cancer cells rely on increased glucose consumption to maintain growth and proliferation. At a whole-body level, cancer can alter the metabolism of the individual and lead to a systemic wasting syndrome named cachexia. Glucose dependency of cancer cells can be targeted with a high-fat, low-carbohydrate ketogenic diet (KD). However, the utilization of KD as energy for survival is hindered when the liver's ability to engage in ketogenesis is suppressed because of cancer-induced alterations. In an experimental approach using models of cancer cachexia, we describe delayed tumor growth but accelerated wasting (cachexia) in the context of KD feeding (Fig. 1). Mechanistically, we found that this uncoupling results from the biochemical interaction of two pathways that causes (1) the accumulation of toxic lipid peroxidation products within the tumor and ultimately death of cancer cells, and (2) impaired synthesis of corticosterone, a glucocorticoid that is essential for survival and regulates the stress response. We then demonstrated that administration of synthetic corticosterone (dexamethasone) delays cachexia and extends survival while preserving small tumor sizes in our model (Fig. 2). Our study highlights the importance of investigating the effects

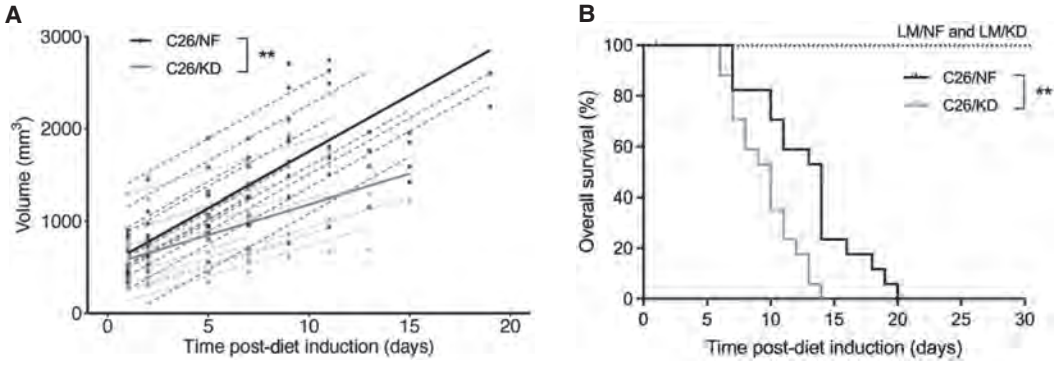



**Figure 1.** Ketogenic diet delays tumor growth but accelerates cachexia and shortens survival. (*A*) Tumor volume over time, analyzed using mixed random effect modeling. (*B*) Overall survival, analyzed using log rank testing. (LM) Littermate, (C26) colon-26 tumor–bearing mice, (NF) normal food (chow diet), (KD) ketogenic diet. **, $p < 0.01$.

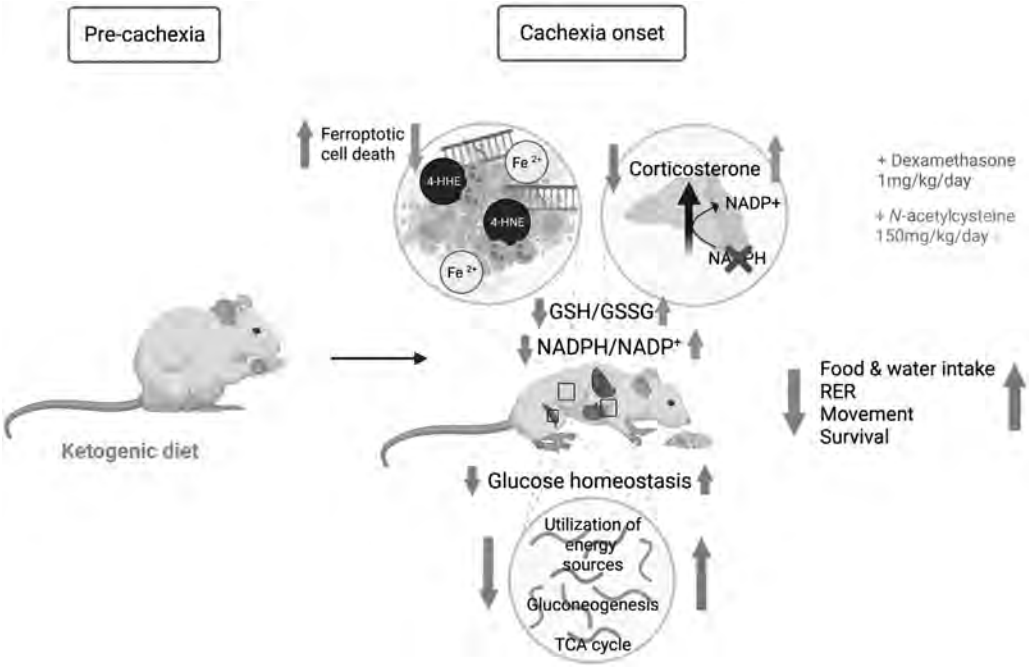



**Figure 2.** Schematic summary of ketogenic diet effects in a murine model of interleukin-6-driven cancer cachexia. Lipid peroxidation depletes reduced glutathione (GSH) and NADPH, which induces ferroptosis in the tumor and a relative corticosterone deficiency that can be treated with synthetic glucocorticoid administration. (4-HNE) 4-hydroxynonenal, ($Fe^{2+}$) ferrous iron.

of therapies on both the tumor and the body to accurately assess their potential. These findings may be relevant to clinical research on nutritional interventions in patients with cancer.

## Cachexia: A Systemic Consequence of Progressive, Unresolved Disease

M. Ferrer Gonzalez, T. Janowitz [in collaboration with the Cancer Cachexia Action Network (CANCAN)]

The CANCAN team, funded by a Grand Challenges (CGC) grant and co-led by Tobias Janowitz, Eileen White (Rutgers University), and Marcus Goncalves (Weill Cornell Medicine), focuses on identifying targetable mechanisms that cause cachexia, a highly debilitating condition suffered by patients with cancer. Cachexia is a systemic wasting syndrome developed as the ultimate consequence of unresolved diseases such as infections, chronic inflammatory diseases, and cancers (Fig. 3). In a consensus review paper, we summarized the knowledge gaps and emerging evidence in the mechanistic understanding of cachexia. The robust research foundation on the late stages of cachexia and the ability to longitudinally track specific molecules across multiple organs open the door to expanding our understanding of the processes that lead to the development of cachexia. The combination of preclinical and clinical research and a comprehensive understanding of the early stages of cachexia will enable better clinical intervention in the years to come.

## Sensitive, Nonimmunogenic In Vivo Imaging of Cancer Metastases and Immunotherapy Response

B. Demestichas, T. Janowitz [in collaboration with S.K. Lyons, CSHL]

Immunotherapy is now an established first-line treatment for multiple metastatic cancers, such as melanoma, colorectal, lung, or bladder cancer. However, not all patients are responsive to immunotherapy, and more research is needed. Noninvasive tumor imaging in murine preclinical models is a key element for studying tumor response to therapy; however, the current most used reporter transgenes are

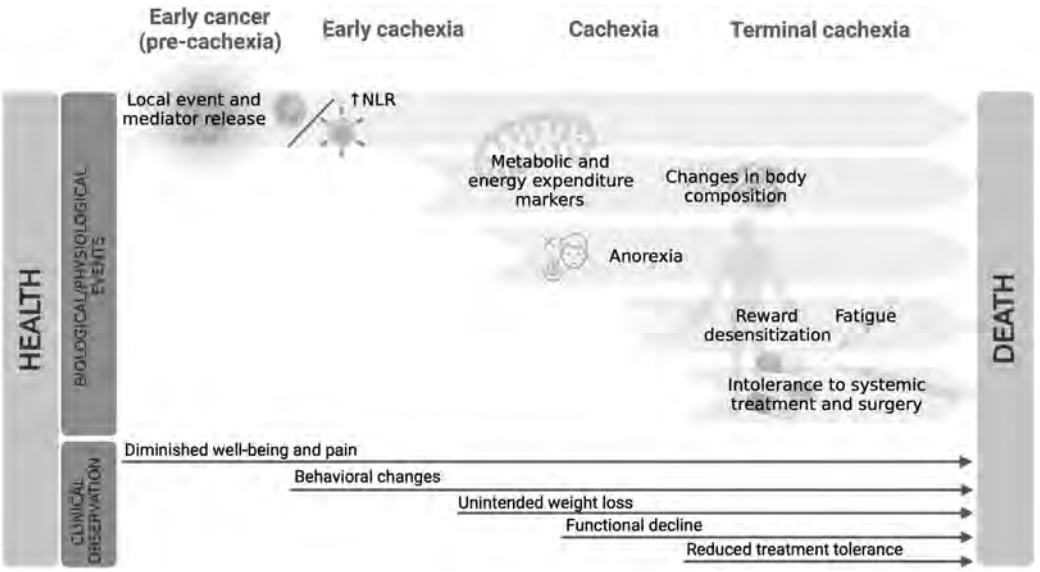



**Figure 3.** Longitudinal progression of biological phenomena and clinical observations from early cancer to advanced cachexia.

immunogenic (nonmurine origin) and therefore impact the biological response to cancer. In a study aimed at developing a sensitive, high-resolution, and nonimmunogenic tumor imaging approach, we established a tumor-cell-labeling viral vector that delivers stable expression of an endogenous transgene of murine origin. The methodology is transferrable to any murine cancer model system based on implanted cells and may aid the future development of cancer immunotherapy strategies.

## PUBLICATIONS


Ferrer M, Anthony TG, Ayres JS, Biffi G, Brown JC, Caan BJ, Cespedes Feliciano EM, Coll AP, Dunne RF, Goncalves MD, et al. 2023. Cachexia: a systemic consequence of progressive, unresolved disease. *Cell* **186:** 1824–1845. doi:10.1016/j.cell.2023.03.028

Ferrer M, Mourikis N, Davidson EE, Kleeman SO, Zaccaria M, Habel J, Rubino R, Gao Q, Flint TR, Young L, et al. 2023. Ketogenic diet promotes tumor ferroptosis but induces relative corticosterone deficiency that accelerates cachexia. *Cell Metab* **35:** 1147–1162.e7. doi:10.1016/j.cmet.2023.05.008

Kleeman SO, Thakir TM, Demestichas B, Mourikis N, Loiero D, Ferrer M, Bankier S, Riazat-Kesh YJRA, Lee H, Chantzichristos D, et al. 2023. Cystatin C is glucocorticoid responsive, directs recruitment of Trem2$^+$ macrophages, and predicts failure of cancer immunotherapy. *Cell Genom* **3:** 100347. doi:10.1016/j.xgen.2023.100347

Langer HT, Ramsamooj S, Dantas E, Murthy A, Ahmed M, Hwang SK, Grover R, Pozovskiy R, Liang RJ, Queiroz AL, et al. 2023. Restoring adiponectin via rosiglitazone ameliorates tissue wasting in mice with lung cancer. bioRxiv doi:10.1101/2023.07.31.551241

Liu Y, Dantas E, Ferrer M, Liu Y, Comjean A, Davidson EE, Hu Y, Goncalves MD, Janowitz T, Perrimon N. 2023. Tumor cytokine-induced hepatic gluconeogenesis contributes to cancer cachexia: insights from full body single nuclei sequencing. bioRxiv doi:10.1101/2023.05.15.540823

Merrill JR, Inguscio A, Chung T, Demestichas B, Garcia LA, Habel J, Lewis DY, Janowitz T, Lyons SK. 2023. Sensitive, non-immunogenic in vivo imaging of cancer metastases and immunotherapy response. *Cell Stress* **7:** 59–68. doi:10.15698/cst2023.08.288

Mou H, Eskiocak O, Özler KA, Gorman M, Yue J, Jin Y, Wang Z, Gao Y, Janowitz T, Meyer HV, et al. 2023. CRISPR-induced exon skipping of β-catenin reveals tumorigenic mutants driving distinct subtypes of liver cancer. *J Pathol* **259:** 415–427. doi:10.1002/path.6054

Sun Q, van de Lisdonk D, Ferrer M, Gegenhuber B, Wu M, Tollkuhn J, Janowitz T, Li B. 2023. Area postrema neurons mediate interleukin-6 function in cancer-associated cachexia. bioRxiv doi:10.1101/2023.01.12.523716

# THE CAUSES AND CONSEQUENCES OF METABOLIC REPROGRAMMING DURING CANCER PROGRESSION

**M. Lukey** J. de Ruiter-Swain Y. Ji O. Stamatatos
Q. Hu Y. Qiu W-H. Yang

A century ago, Otto Warburg made the seminal discovery that tumors process nutrients differently from adjacent healthy tissue. Metabolic reprogramming is now known to be a hallmark of tumorigenesis and is essential for cancer cells to fulfill the biosynthetic, bioenergetic, and redox requirements of malignant growth. Within the tumor microenvironment (TME), during passage in the circulation, upon metastatic colonization, and as a consequence of therapeutic intervention, cancer cells experience fluctuating and limited nutrient supplies and must possess metabolic flexibility to adapt to these changes. Moreover, tumors interact bidirectionally with the broader metabolism of the host and consequently can be impacted by dietary changes and can induce systemic metabolic changes. Our laboratory aims to understand the biochemical underpinnings of metabolic reprogramming during tumorigenesis and metastasis, with the goal of identifying metabolic essentialities that can be selectively targeted for cancer therapy while sparing the host from adverse side effects.

## Breast Cancer Adaption to Amino Acid Starvation

Certain features of tumor metabolism, such as increased consumption of glucose and glutamine and a shift toward fermentative metabolism alongside respiration, are broadly conserved across many cancer types. Glutamine is the most abundant amino acid in blood plasma, and its critical importance for cancer cells has been known since the establishment of the first cancer cell lines in the 1950s. However, recent clinical trials of drugs targeting glutamine metabolism have yielded largely disappointing results. By exploring how highly glutamine-addicted triple-negative breast cancer (TNBC) cells adapt to glutamine starvation or pharmacological blockade of glutamine catabolism, we discovered that the de novo serine biosynthesis pathway becomes strictly essential under these conditions. In this context, serine production itself is not the key function of the pathway, but instead we found that a pathway side product is able to compensate for the loss of glutamine-derived carbon. Because serine biosynthesis is pharmacologically tractable, we have applied our findings to develop a synergistic combination therapy that blocks cancer cell glutamine catabolism and simultaneously prevents an adaptive metabolic response. In murine models of TNBC, this combination therapy sharply suppresses tumor growth. In addition to glutamine's role as a major carbon source for cancer cells, the amide and amine groups are obligate nitrogen donors for several biosynthetic pathways. In a parallel study, we have traced a nutrient scavenging mechanism that allows breast cancer and pancreatic cancer cells to bypass the direct need for glutamine-derived nitrogen. Targeting this process restricts the metabolic plasticity of TNBC cells and markedly inhibits tumor growth in vivo.

## Metabolic Reprogramming in Brain-Metastatic Breast Cancer

Different organs in the body present metastatic cancer cells with distinct nutrient profiles, and metabolic colonization of distant sites is dependent on cancer cells adapting to the available fuel supplies. The development of brain metastases is typically an end-stage event in breast cancer progression, with median survival after diagnosis measured in months regardless of the disease subtype. The brain presents a unique metabolic environment to cancer cells arising from peripheral tumors, with much lower availability of nutrients such as glucose and glutamine than in other organs. This is thought to be one reason why brain metastasis is a biologically inefficient process, with cancer cells that traverse the blood–brain barrier showing a high rate of attrition. Reflecting strong selection for a specific metabolic phenotype, metastatic cancer cells appear to undergo convergent evolution in the brain

such that brain metastases arising from diverse primary tumor types exhibit a conserved metabolic phenotype. To study the metabolic adaptations that are required for metastatic growth of breast cancer cells in the brain, we have employed animal models coupled with mechanistic ex vivo studies using a culture medium that models the nutrient composition of the brain interstitial fluid. This has allowed us to conduct a whole-genome CRISPR-Cas9 functional genomics screen aimed at identifying genes that are selectively essential for the survival and proliferation of breast cancer cells following metastatic spread to the brain. The screen has pinpointed a conditional essentiality of genes clustered around two pathways of mitochondrial redox metabolism and has indicated that several genes of unknown function might be involved in these pathways. We have used the information obtained to develop therapeutic interventions targeting the metabolic adaptations required for metastatic growth in the brain.

## Metabolic Regulation of Microglia-Mediated Immunosuppression in Brain Tumors

Beyond cancer cell–intrinsic metabolic adaptation, we have found that metastatic and primary brain cancer cells co-opt other cell types in the brain to provide metabolic support and/or to mediate immunosuppression within the TME. Using metabolomics and stable isotope tracing approaches, in a collaboration with the Van Aelst group we have identified a major metabolic interaction between glioblastoma cells and microglia, the resident immune cells of the brain. This axis serves the dual functions of inducing immunosuppressive polarization of microglia downstream of a metabolite-sensing receptor and providing glioblastoma multiforme (GBM) cells with metabolic support. Using genetic approaches to disrupt a critical step in the intercellular metabolic pathway, we have found that microglia-mediated immunosuppression can be reversed, leading to a robust anticancer immune response. We are now following up on these discoveries, aiming to define the molecular mechanisms by which the GBM cell–derived metabolites activate immunosuppressive signaling pathways in microglia and developing strategies that combine the inhibition of this metabolic process with radiation and/or immunotherapy to maximize the therapeutic effect.

## Intersections of Cellular Signal Transduction and Metabolism

Many of the signaling pathways that are dysregulated in cancer drive downstream changes in cellular metabolism such that the tumor oncogenotype is a key determinant of its ultimate metabolic phenotype. Reciprocally, numerous metabolites are able to influence cellular signal transduction pathways—for example, by providing precursors for protein prenylation, by altering gene expression by influencing the activity of enzymes that catalyze epigenetic changes, or by direct allosteric modulation of signaling proteins. Phospholipase enzymes lie at the intersection of signal transduction and metabolism, hydrolyzing phospholipids into bioactive signaling lipids that regulate diverse physiological processes. We have identified a new role for phospholipase signaling in the inner mitochondrial membrane (IMM) whereby the lipid product of the phospholipase reaction directly binds and activates the metabolic enzyme glutaminase (GLS), which is associated with the IMM. This interaction occurs with extremely high affinity and shows exquisite specificity regarding the length and structure of the fatty acid tail. Moreover, we have identified lipid species that are able to compete with and displace clinical GLS inhibitors, potentially constituting a resistance mechanism to this class of drug. We are now working to define the molecular mechanism of this lipid-mediated regulation of glutamine metabolism, focusing on the formation of enzyme filaments during the activation process. Additionally, we are developing and evaluating combination therapies that target GLS and phospholipase signaling, thereby shutting down a key metabolic pathway for cellular proliferation and simultaneously blocking a potential resistance mechanism.

### PUBLICATIONS

de Ruiter Swain J, Michalopoulou E, Noch EK, Lukey MJ, Van Aelst L. 2023. Metabolic partitioning in the brain and its hijacking by glioblastoma. *Genes Dev* **37:** 681–702.

Ferrer M, Mourikis N, Davidson EE, Kleeman SO, Zaccaria M, Habel J, Rubino R, Gao Q, Flint TR, Young L, et al. 2023. Ketogenic diet promotes tumor ferroptosis but induces relative corticosterone deficiency that accelerates cachexia. *Cell Metab* **35:** 1147–1162.e7.

*In Press*

Homer JA, Koelln RA, Barrow AS, Gialelis TL, Boiarska Z, Steinohrt NS, Lee EF, Yang W-H, Johnson RM, Chung T, et al. 2024. Modular synthesis of functional libraries by accelerated SuFEx click chemistry. *Chem Sci* **15:** 3879–3892.

# PROTEIN TYROSINE PHOSPHATASES AND THE CONTROL OF SIGNAL TRANSDUCTION

**N.K. Tonks** J. Anduaga, Y. Cen, L. Christensen, C. Felice, D.S. Kao, T.C. Kuo, V. Mandati, P. Venkataramani, I. Zubiete Franco

As cells encounter stimuli, such as growth factors, cytokines, and hormones, receptors on the cell surface modulate the activities of protein kinases and phosphatases. The functions of these enzymes, which promote the addition and removal of phosphate groups, respectively, are coordinated in signal transduction pathways that mediate the cellular response to environmental stimuli. These pathways are of fundamental importance in the control of cell function and their disruption frequently underlies major human diseases. Consequently, the ability to modulate such signal transduction pathways selectively with drugs holds enormous therapeutic potential. To date, drug discovery efforts in tyrosine phosphorylation–dependent signal transduction have emphasized the protein tyrosine kinases (PTKs). Since the approval of the first PTK inhibitor drug, imatinib, for treatment of chronic myelogenous leukemia, 80 small-molecule kinase inhibitors have been approved as therapeutics by the U.S. Food and Drug Administration (FDA). In fact, kinase inhibitors represent a multi-billion-dollar industry and are a major component of modern cancer therapy. Nevertheless, drug resistance, both acquired and intrinsic, remains a major challenge to the effectiveness of kinase inhibitors in the clinic. Considering the reversibility of protein tyrosine phosphorylation, there is the potential to manipulate signal transduction pathways at the level of both PTKs and the family of protein tyrosine phosphatases (PTPs), which Dr. Tonks discovered more than 30 years ago. Although the PTPs have been garnering attention as potential therapeutic targets, they remain largely an untapped resource for drug development. The long-term objectives of the Tonks laboratory are to characterize the structure, modes of regulation, and physiological function of members of the PTP family of enzymes. Through this basic research, the Tonks laboratory is trying to devise creative new approaches to exploiting these enzymes as targets for therapeutic intervention in major diseases, including cancer, diabetes, obesity, and neurodegenerative disease.

During the last year, there were several changes in the group. Zhe (Changer) Qian was awarded his Ph.D. and moved to a Staff Scientist position at Merck. Dongyan Song, having been awarded her Ph.D. in 2022, completed her studies and moved to a postdoctoral position at MD Anderson Cancer Center. Der-Shyang (Leo) Kao joined the laboratory as a postdoctoral fellow and Javier Anduaga (Pharmacology Program) and Tan-Chun (John) Kuo (Genetics Program) joined the laboratory as graduate students from Stony Brook.

## Targeting the Receptor-Like PTPRD (RPTPδ) with Antibodies to Its Extracellular Segment

One of the central tenets of the control of signal transduction is that ligand binding brings two molecules of a receptor protein tyrosine kinase (RPTK) together on the cell surface in a dimer, promoting trans-autophosphorylation within the dimer, and activation of the signaling pathway. Variations on this theme apply across the whole family of RPTKs. Twenty-one of the *PTP* genes in humans encode receptor-like transmembrane proteins that also have the potential to regulate signaling pathways directly in response to environmental cues; however, it has been a challenge to define how this signaling function may be regulated. Crystal structures suggested that ligand binding to RPTPs may induce dimer formation and inactivation of the phosphatase, the reciprocal of receptor PTK regulation. An intriguing example is the opposing effects of heparin (HSPG) and chondroitin (CSPG) sulfate proteoglycans on neuronal extension, which are mediated through their interaction with the receptor phosphatase RPTPσ. HSPG-induced oligomerization inactivates PTP function to favor neurite outgrowth.

In contrast, CSPG leaves RPTPσ in a monomeric state that favors phosphatase function and antagonizes growth. This concept of regulation of RPTP function by dimerization also prompted the development of antibodies to engage the extracellular segment of receptor phosphatases to manipulate their function therapeutically. There are several advantages to such approaches, including the high affinity and specificity of antibody–antigen interactions, the fact that they act on the cell surface and so avoid the requirement for the drug to cross the plasma membrane to reach its target, the capacity to exploit the intrinsic bivalency of antibodies to promote dimerization, and the fact that such approaches avoid the problems that have been encountered developing small-molecule drugs that target the charged PTP active site. In 2023, we published the characterization of an array of antibodies against RPTPδ/PTPRD, an enzyme that can elicit tumor-promoting effects in certain contexts. PTPRD dephosphorylates the inhibitory carboxy-terminal phosphorylation site of the PTK SRC, relieving autoinhibition of SRC and activating downstream signals. In cancer cells that feature loss of missing-in-metastasis/MTSS1, the levels of PTPRD are increased; treatment of such cells, including breast cancer models, with antibody to the extracellular segment of PTPRD induced rapid RPTP dimerization and inactivation, inhibition of the SRC signal, and suppression of cell invasion. Ultimately, antibody treatment induced the degradation of PTPRD protein. The extent to which such antibodies are functional in vivo now awaits further study and will determine their utility in a therapeutic context.

## Characterization of PTPN23 as a Cell Death Checkpoint in Restraining Apoptosis, Necroptosis, and Pyroptosis

Cell death plasticity is crucial for modulating tissue homeostasis and immune responses to diverse stimuli; however, our understanding of the molecular components that regulate cell death pathways to determine cell fate remains limited. We conducted a CRISPR screen of acute myeloid leukemia cells and identified PTPN23 as an essential requirement for survival. Loss of PTPN23 activated NF-κB, apoptotic, necroptotic, and pyroptotic signaling pathways, by causing the accumulation of death receptors and TLRs (Toll-like receptors) in endosomes. These effects were recapitulated by depletion of PTPN23 co-dependent genes in the ESCRT pathway that regulates membrane integrity. Through BioID analysis, we showed that the protein NAP1 interacted with PTPN23 to promote endosomal sorting of TNFR1, thereby sensitizing cells to TNF-α-induced cytotoxicity. Collectively, our findings reveal PTPN23-dependent ESCRT machinery as a novel cell death checkpoint that regulates the spatiotemporal distribution of death receptors and TLRs to restrain multiple cell death pathways (Fig. 1). Over the last year, we have completed additional studies, including the demonstration that PTPN23 plays a pivotal role in restraining multiple cell death modalities in primary myeloid cells, thereby indicating its broader relevance beyond transformed cancer cell contexts. A revised manuscript describing these studies is currently under review.

## Famotidine and Redox Regulation of PTP Function

In early 2020, the novel COVID-19 disease, caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), triggered a rush to repurpose drugs for the emergency treatment of patients while effective vaccines were approved and produced. One of the proposed drugs was famotidine, sold under the brand name PEPCID, which is a well-tolerated and safe over-the-counter drug used to treat peptic ulcers and gastroesophageal reflux. Famotidine acts as a histamine 2 receptor (H2R) antagonist in gastric parietal cells and thereby reduces the secretion of acid into the stomach. Histamine 2 receptors are widely expressed in a range of different organs, including lung, and belong to the G-protein-coupled receptor (GPCR) family of signaling proteins that can induce changes in phosphorylation of downstream targets and changes in cellular calcium levels. The U.S. Department of Veterans Affairs, which provides healthcare to military veterans, conducted an observational study of hospitalized patients with COVID-19 infections, which suggested that there was a survival benefit for those patients undergoing famotidine treatment at the time of infection. In addition, several molecular docking studies to examine the structure of viral proteases suggested that famotidine could be an inhibitor of the early viral proteases that governed processing of the

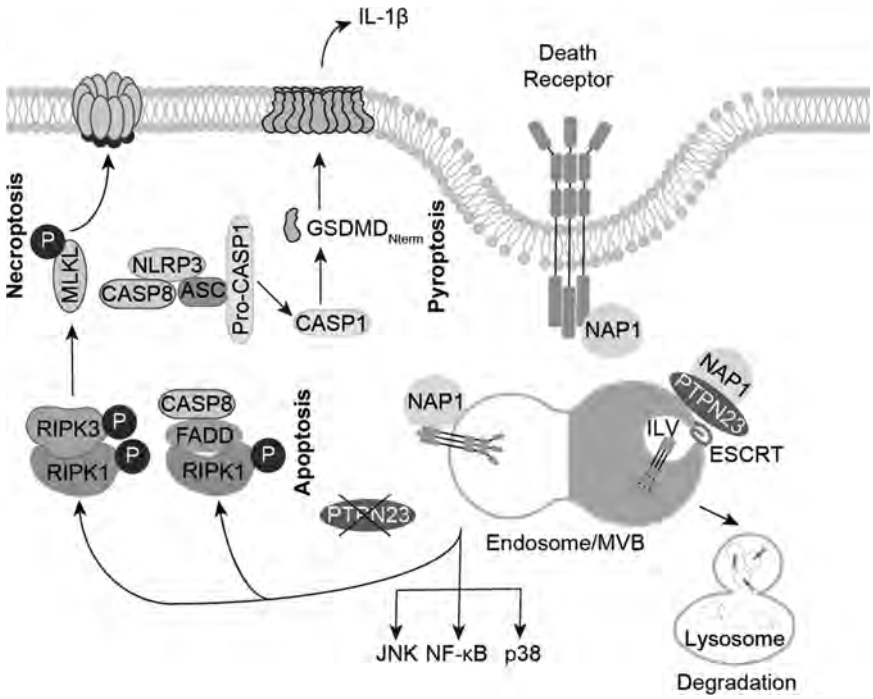



**Figure 1.** Proposed model for the role of PTPN23 in the regulation of death receptor sorting and degradation. PTPN23 collaborates with NAP1 to facilitate receptor sorting. (*Right*) Representing the normal condition, this process involves the engagement of ESCRT, leading to the lysosomal degradation of death receptors. (*Left*) Loss of PTPN23 results in the prolonged accumulation of death receptors within the endosomal compartment, triggering the activation of multiple cell death pathways.

initial polypeptide produced at the beginning of the viral replicative cycle. This prompted the initiation of clinical trials to explore further the potential benefits of famotidine treatment for COVID-19 patients, from which initial results suggested that famotidine use reduced the risk of clinical deterioration and improved patient-reported outcomes in subjects that were not hospitalized. Mechanistically, published studies, and our own work in collaboration with Leemor Joshua-Tor's laboratory, indicated that famotidine had no effect on viral proteins or viral replication when compared to known inhibitors, suggesting that the observed effect was most likely due to an influence on the host cell or organ. In response to a viral infection, host cells release interferons (IFNs) to combat infection and prepare neighboring cells. Interferons induce the synthesis of hundreds of IFN-stimulated genes (ISGs), including *ISG15*, *OAS1*, and *MX1*, which target different stages of the viral life cycle in order to protect the cell and inhibit viral replication. Initial transmission of the IFN signal is relayed via phosphorylation of tyrosine residues in proteins through the JAK-STAT pathway and is regulated by PTPs. To examine further potential mechanisms by which famotidine may be exerting its observed effects in the context of COVID, we tested its potential role modifying cellular tyrosine phosphorylation and antiviral signaling and have revealed new links between the drug and redox regulation of PTP function. We have observed a change in the cellular levels of phosphotyrosine in A549 lung epithelial cells following treatment with famotidine. This quick change in phosphorylation was due mainly to a dose-dependent increase in cellular production of $H_2O_2$. Notably, these changes in phosphotyrosine levels were able to affect cell signaling; we detected an increased short- and long-term response to IFN-α stimulation. Our results suggest that famotidine can increase the antiviral state of noninfected cells, thereby potentially improving viral resistance and focusing attention on redox regulation of PTP function in its mechanism of action. This study has been submitted for publication.

## Developing Inhibitors of the Protein Kinases PIM and DYRK

This project represents a long-standing collaboration with the laboratories of Darryl Pappin and Leemor Joshua-Tor and more recently with Dr. Yousef Al-Abed and his team at the Feinstein Institute. The foundation for this project was our purification and characterization of a small-molecule natural product that we demonstrated to have a unique structure and a unique specificity for PIM and DYRK family kinases. PIMs and DYRKs have been implicated in a wide variety of hematological and epithelial tumors; it was our expectation that inhibitors of these kinases may have broad therapeutic utility. We completed a structure–activity relationship (SAR) optimization program to yield low-nanomolar, drug-like lead compounds that were ATP-competitive and displayed both improved potency and selectivity relative to the parent molecule. We focused on compounds that inhibited DYRK preferentially, confirming specificity by profiling against a panel of 140 kinases. We focused our attention on the DYRKs because of their interesting biology, including their potential as therapeutic targets for several cancers including glioblastomas, which are aggressive brain tumors that are unresponsive to chemo- or radiotherapy.

In our study, we have identified two potent and specific small-molecule ATP-competitive inhibitors of DYRKs, FC-2 and FC-3, which inhibited glioblastoma cell proliferation, invasion, and subcutaneous tumor formation and increased survival in U87MG intracranial orthotopic xenograft models. In parallel, ablation of DYRK1A using CRISPR-Cas9 in two GBM cell lines validated its biological significance as a target by mimicking the effects of its inhibition in our phenotypic assays. Further confirmation of the specificity of these inhibitors was provided by point mutants of DYRK1A in which catalytic function was maintained, but in which affinity for the inhibitors was markedly attenuated. By expressing these mutants in cancer cells, we established that the effects of our small-molecule drug candidates were due to "on-target" inhibition of DYRK, rather than "off-target" effects. Previous studies have shown that DYRK1A stabilizes the EGFR by preventing its ubiquitination in EGFR-dependent glioblastoma; this is mediated by the sequestration of the E3 ubiquitin ligase (c-CBL) by Sprouty2 when it is phosphorylated by DYRK1A at Thr75. Consistent with these findings, we show that both pharmacological inhibition and knockout of DYRK1A down-regulates EGFR levels and modulates the downstream MAPK signaling pathway. Collectively, our data illustrate that targeting DYRK1A with these small-molecule drug candidates, alone or in combination with EGFR inhibitors, could be beneficial in treating EGFR-dependent glioblastoma. Currently, we are exploring additional cancer models, including non-small-cell lung cancer (NSCLC). Inhibition of DYRK1A in A549 cells using our lead compounds led to the down-regulation of EGFR and c-Met. In addition, combination therapy using DYRK and EGFR inhibitors was beneficial in NSCLC models when tested using colony morphology and cell proliferation assays.

## Targeting PTP1B Therapeutically

Rett syndrome (RTT) is an X-linked neurological disorder presenting with autistic features, which is caused primarily by mutations in a transcriptional regulator, methyl CpG binding protein 2 (MECP2). Some RTT mouse models display obesity and leptin resistance, with insulin resistance also noted in some RTT patients, which suggested to us that PTP1B function might be altered in RTT. Previously, we demonstrated that the *PTPN1* gene, which encodes PTP1B, is a direct target of MECP2, which interacts with the *PTPN1* promoter and suppresses promoter activity. Furthermore, disruption of MECP2 function was associated with increased levels of PTP1B in RTT models. As expected, we have shown that pharmacological inhibition of PTP1B, with multiple structurally and mechanistically distinct small-molecule inhibitors, reduced the extent of glucose intolerance and improved metabolic status in the RTT mouse model. Of particular importance, the PTP1B inhibitors ameliorated a wide array of the effects of MECP2 disruption in RTT mice, including movement, behavior, and heart function. Furthermore, we demonstrated that the elevated levels of PTP1B in RTT represent a barrier to BDNF signaling; inhibition of PTP1B led to increased tyrosine phosphorylation of TRKB in the brain, which augmented the response to BDNF. Based on these data, which validate PTP1B as a mechanism-based, therapeutic target for Rett syndrome, an Investigational New Drug (IND)

application will be filed in the first quarter of 2024 in collaboration with DepYmed, Inc. with initiation of a Phase 1 clinical study to follow shortly thereafter. Further investigations of the mechanism of action of PTP1B in Rett syndrome and the impact of PTP1B inhibitors together with other potential therapeutics, such as trofinetide (DAYBUE™), are underway. Furthermore, we are now extending these investigations to exploit PTP1B in a new approach to understanding and targeting aspects of neurodegenerative disease and inflammation. In particular, we have shown that deletion of PTP1B in the APP/PS1 mouse model of Alzheimer's disease ameliorated recognition and spatial memory deficits, as assessed by two memory-related behavioral tests. PTP1B deletion in APP/PS1 mice significantly reduced the Aβ burdens in the brain. Interestingly, PTP1B deletion changed the distribution of phosphotyrosine in APP/PS1 brains, to a more microglia-concentrated pattern. In addition, PTP1B deletion enhanced microglia recruitment around plaques in APP/PS1 mice and promoted phagocytosis in vitro, indicating that PTP1B deletion may restore microglial phagocytic capacity. These studies are ongoing and have potential to validate PTP1B as a therapeutic target for treatment of Alzheimer's disease.

## PUBLICATIONS

Bonham CA, Mandati V, Singh RK, Pappin DJ, Tonks NK. 2023. Coupling substrate trapping with proximity labeling to identify protein tyrosine phosphatase PTP1B signaling networks. *J Biol Chem* **299:** 104582.

Qian Q, Song D, Ipsaro J, Bautista C, Joshua-Tor L, Yeh J, Tonks NK. 2023. Manipulating PTPRD function with ectodomain antibodies. *Genes Dev* **37:** 743–759.

Tonks NK. 2023. Protein tyrosine phosphatases: mighty oaks from little acorns grow. *IUBMB Life* **75:** 337–352.

# MECHANISMS AND TREATMENT OF METASTASIS

**L. Trotman** A. Aziz W. Borges E. Cheng M. Doherty D. Ghosh S. Kim S. Kuang J. Rey C. Sanchez S. Sanchez M. Swamynathan

Resistance to therapy of metastatic prostate cancer (CRPC) is responsible for the deaths of some 30,000 U.S. men each year. Although there is considerable progress with improving antihormone therapy against metastatic disease, this standard-of-care approach will invariably fail patients at some point.

Our focus is twofold: to understand the mechanisms driving primary prostate cancer to switch into lethal metastatic disease, and to develop novel therapeutic approaches. We study the human genetics behind the transition from indolent to lethal metastatic prostate cancer and combine it with advanced functional modeling in mouse. By developing somatic transgene delivery into mouse prostate, we have succeeded in generating a unique, fast, and faithful mouse model for metastatic prostate cancer, termed RapidCaP. It allows us to generate any genetically mutant mouse prostate cancer with a much-accelerated time frame compared to breeding-based approaches. Now, we use RapidCaP for analysis and therapy of metastatic disease.

Independently, we aim to better understand how the PTEN tumor suppressor actually works. This has given us unique insights into how the process of endocytosis is intimately associated with tumor suppression by PTEN, allowing us to rethink the cell's anticancer system.

## Whole-Organ Imaging: Metastasis Origins and the Role of Nerves in Metastasis

S. Kim, M. Doherty, D. Ghosh, E. Cheng, C. Sanchez

The Trotman laboratory has established a novel platform for revealing cancer by 3D reconstruction of entire organs at single-cell resolution in collaboration with Dr. Pavel Osten (Taranda et al., *Cell Rep* **37:** 110027 [2021]). The approach uses serial two-photon tomography (STPT; Ragan et al., *Nat Methods* **9:** 255 [2012]), which works on fixed and agar-embedded mouse tissue that is placed under the objective of a two-photon microscope. A motor moves the tissue under the objective so that an entire several-centimeter-wide specimen can be imaged as a mosaic of individual tiled squares. After imaging of one section, a built-in vibrating blade microtome mechanically cuts off a tissue slice from the top, and then the steps of imaging and sectioning are repeated until the entire information of the tissue is collected. The microscope can record two fluorescent reporter channels for imaging. They have used this approach (originally designed for brain connectivity mapping) to solve the needle in a haystack problem of finding the origins of RapidCaP tumor initiation and to map the earliest steps of PC evolution (Taranda et al., *Cell Rep* **37:** 110027 [2021]). The method allows them to visualize single metastatic cells in an entire liver lobe in the context of the organ. This revealed how pioneer metastatic cells are organized by the liver vasculature and how rare metastatic cells are found within the brain vasculature (Taranda et al., *Cell Rep* **37:** 110027 [2021]).

Regarding their focus on primary tumor progression to metastasis, their new technology yielded several surprising results. They found that metastatic escape can happen already at 20 days after generation of the first tumor cell in the tissue. Furthermore, they observed that the physical properties of tumor cells and escaping cells are fundamentally different. Although the former show strong increase in cell volume compared with normal prostate cells, reaching between 1,000 $\mu m^3$ and 3,000 $\mu m^3$, the metastatic escapers are thinned down in volume to ~100 $\mu m^3$. They thus fit well into the stromal cell layers that surround the tumor. Now the team is studying what guides these transitions. To this end they use cell lines derived from the RapidCaP mouse at different time points. These can be divided into nonmigratory large cells and migratory thinner cells that serve as proxies for the tumor and metastatic escapers, respectively.

In addition, they employ a co-registration method using serial frozen sections. In this method, tissue samples are embedded in OCT (optimal cutting temperature) compound and sectioned at 50 microns with a cryostat. The serial sections are imaged with slide scanner (Olympus VS200) and aligned to create a comprehensive 3D reconstruction. This method complements STPT by providing additional fluorescent reporter channels and expanding the antibody palette for molecular characterization of metastatic pioneer cells.

## Lipid Regulation and Prostate Cancer

M. Swamynathan, S. Kuang, J. Rey, A. Aziz [in collaboration with I. Ojima, A. Bialkowska, and M. Kaczocha, Stony Brook University]

Prostate cancer therapy suffers from a lack of effective targets because, typically, success with blockade of the androgen receptor gives way to drug resistance and lethal disease relapse. Large-scale genome sequencing efforts have demonstrated that lethal recurrent disease most often presents with loss of the *PTEN* and *TP53* tumor suppressors. Unfortunately, the systematic testing of PTEN/PI3-kinase pathway-specific inhibitors has shown only limited results in prostate cancer trials. Thus, there are currently no U.S. Food and Drug Administration (FDA)-approved drugs targeting this axis in prostate cancer patients. In this work we propose a new target, the FABP5 lipid carrier. Fatty acid–binding proteins (FABPs) are a large family of signaling lipid carriers. They have been suggested as drivers of multiple cancer types. *FABP5* amplification and surge in expression are strongly correlated to that of the *MYC* oncogene, a known driver of advanced *PTEN*-deficient prostate cancer. Furthermore, we present a new preclinical platform to assess the efficacy and biology of inhibiting FABP5 with small molecules. We use the RapidCaP model, which is based on a *PTEN*-deficient prostate cancer cell type that is insensitive to standard-of-care therapies. We have combined analysis of primary cancer cells from RapidCaP (RCaP cells) with large-scale patient data sets to show that among the 10 FABP paralogs, FABP5 is the PC-relevant target. Next, we showed that RCaP cells are uniquely insensitive to both androgen deprivation therapy (ADT) and taxane treatment compared with a panel of human PC cell lines. Yet, they share an exquisite sensitivity to the small-molecule FABP5 inhibitor SBFI-103. We showed furthermore that SBFI-103 is well tolerated and can strongly eliminate RCaP tumor cells in vivo. This provides a preclinical platform to fight incurable PC and suggests an important role for FABP5 in *PTEN*-deficient PC.

## Endocytosis and Cancer

M. Doherty, W. Borges, D. Ghosh, E. Cheng, C. Sanchez

The transduction of signals in the PTEN/PI3-kinase (PI3K) pathway is built around a phosphoinositide (PIP) lipid messenger, phosphatidylinositol trisphosphate, PI(3,4,5)P3 or PIP3. Another, more ancient role of this family of messengers is the control of endocytosis, in which a handful of separate PIPs act like postal codes. Prominent among them is PI(3)P, which helps to ensure that endocytic vesicles, their cargo, and membranes themselves reach their correct destinations. Traditionally, the cancer and the endocytic functions of the PI3K signaling pathway have been studied by cancer and membrane biologists, respectively, with some notable but overall minimal overlap. This is because cancer rarely mutates the endocytic pathway as the process is essential. The discovery that PTEN contains an autonomous PI(3)P reader domain fused to the catalytic PI(3,4,5)P3 eraser domain has prompted us to explore the relationship between PI3K signaling and endocytosis. We have now shown that PTEN function can be enhanced by a compound that inhibits clathrin-mediated endocytosis, revealing for the first time that PTEN activity can be enhanced by small-molecule-based approaches. As tumors frequently present with haploinsufficiency for PTEN and therefore reduced PTEN activity, this compound represents the starting point for further work on the therapeutic potential of targeting endocytosis in cancer. We have elucidated how PTEN is recruited to endosomes and defined a set of protein players that are involved in this process. Now the team is asking whether the genes that encode endosomal recruitment factors of PTEN are involved in cancer.

## Genomics of Lethal Human Prostate Cancer

S. Kuang, D. Ghosh, E. Cheng, C. Sanchez [in collaboration with S. Hall, M. Vira, C. Metz, O. Yaskiv, O. Rodriguez, and M. Ryan, Northwell Health; L. Kollath and C. Morrissey, University of Washington, Seattle]

Next-generation sequencing techniques have provided the ability to incorporate cutting-edge genomic

profiling in understanding, prognosis, and treatment of various tumor types. However, heterogeneity of the prostate tissue during tumorigenesis makes it difficult to conduct exhaustive transcriptome analyses. Spontaneous genetic changes arising in PC are a crucial imprint of this variability. Therefore, one of our aims in the laboratory is to shed some light on key genomic drivers of metastatic progression utilizing genomic information obtained from nuclear DNA (nDNA) of prostate cancer cells. The project goal was to obtain an extensive copy number landscape of visceral and bone metastases from 10 patients who had consented to subject their bodies to rapid autopsy (RAP) after death from prostate cancer. Samples were obtained from University of Washington, Seattle and Northwell Health, New York. We have successfully established a pipeline for processing of frozen tissue sample for single-nucleus sequencing to determine copy number alterations (CNAs) at the single-cell level. Recurrent CNAs involving cancer genes have emerged as the primary driver of lethal metastatic PC, whereas recurrent missense mutations are infrequent. After processing all bone metastases, we analyzed matched visceral metastases (liver, lung, lymph node, etc.) from each of the 10 patients. Normal muscle sample from each patient was obtained and processed as a baseline control for CNAs.

Based on the data collected on 2,914 cells from 31 metastatic sites of 10 patients, we first answer these general questions on the metastatic landscape seen at single-cell resolution.

1. What is the CNA-based clonality of metastasis (within tumor site and between sites)?
2. Can we infer fitness of clones based on representation (within a site and/or between sites of a patient)?
3. Are there recurrent CNAs that have been missed by bulk sequencing of metastatic PC?

Given our expertise and the emergence of PTEN deletion as the most prominent feature of lethal metastatic PC, we place special emphasis on the below questions.

1. What genes are most significantly co-deleted with PTEN at the single-cell level?
2. Does loss of PTEN dominate clonality as expected from a strong driver event?

These data are complemented by our analysis of tumors from primary PC patients using the same approach. Samples from these early patients are collected through our collaboration with clinician scientists at Northwell Health. Analysis of genome-wide DNA and RNA alteration in primary and metastatic RapidCaP samples is used for cross-species prioritization of results. Based on our preliminary results, this project allows us to discover novel markers of metastasis and new drivers of the lethal disease that remained undiscovered based on bulk sequencing analysis.

## PUBLICATIONS

Aakula A, Isomursu A, Rupp C, Erickson A, Gupta N, Kauko O, Shah P, Padzik A, Pokharel YR, Kaur A, et al. 2023. PP2A methylesterase PME-1 suppresses anoikis and is associated with therapy relapse of *PTEN*-deficient prostate cancers. *Mol Oncol* **17:** 1007–1023. doi:10.1002/1878-0261.13353

Brina D, Ponzoni A, Troiani M, Calì B, Pasquini E, Attanasio G, Mosole S, Mirenda M, D'Ambrosio M, Colucci M, et al. 2023. The Akt/mTOR and MNK/eIF4E pathways rewire the prostate cancer translatome to secrete HGF, SPP1 and BGN and recruit suppressive myeloid cells. *Nat Cancer* **4:** 1102–1121. doi:10.1038/s43018-023-00594-z

Chung T, Garcia L, Swamynathan MM, Froeling FEM, Trotman LC, Tuveson DA, Lyons SK. 2023. Internally controlled and dynamic optical measures of functional tumor biology. *Anal Chem* **95:** 5661–5670. doi:10.1021/acs.analchem.2c05450

Hillowe A, Gordon C, Wang L, Rizzo RC, Trotman LC, Ojima I, Bialkowska A, Kaczocha M. 2023. Fatty acid binding protein 5 regulates docetaxel sensitivity in taxane-resistant prostate cancer cells. *PLoS One* **18:** e0292483. doi:10.1371/journal.pone.0292483

Swamynathan M, Mathew G, Aziz A, Gordon C, Hillowe A, Wang H, Jhaveri A, Kendall J, Cox H, Giarrizzo M, et al. 2023. FABP5 inhibition against *PTEN*-mutant therapy resistant prostate cancer. *Cancers (Basel)* **16:** 60. doi:10.3390/cancers16010060

# NOVEL STRATEGIES TO DIAGNOSE AND TREAT PANCREATIC CANCER

**D. Tuveson** K. Addison, H. Alakonya, L. Baker, D.P. Budagavi, G. Caligiuri, A. Deschênes, L.P. Ferguson, V. Gaeth, A. Habowski, A. Jensen, J. Kastan, D. Kightley-Sutter, S. Kim, D. King, F. Kouassi, W. Lan, E. Lentsch, S. Nadella, J. Nigri, Y. Park, J. Peluso, Y. Qin, M. Shakiba, K. Shruti, L. Surin, J. Thalappillil, H-C. (S.) Ting, C. Tsang, C. Tonelli, K. Yu

Our laboratory investigates pancreatic ductal adenocarcinoma (PDAC), the primary form of pancreatic cancer and the third leading cause of cancer-related deaths in the United States. More specifically, we aim to generate insights into the molecular underpinnings of PDAC using organoids and mouse models of PDAC, which could inform novel strategies to detect and treat this currently incurable cancer. These tools have facilitated the discovery of new therapeutic approaches and the detailed analysis of the cell types and signaling pathways for patterning the PDAC microenvironment. Through our studies, we have found that resistance to KRAS-targeted therapy exhibited potentially different mechanisms when considering oncogenic KRAS-specific inhibition versus pan-RAS inhibition. In addition, we have discovered a new Wnt-sensing cancer-associated fibroblast (CAF) substate in PDAC that may play a pivotal role in dense desmoplastic reaction in conjunction with myofibroblastic CAFs (myCAFs), identified and characterized in our laboratory.

## Oncogenic KRAS Inhibition in Pancreatic Cancer

This work was done in collaboration with P. Pérez-Mancera (The University of Liverpool, UK) and M. Singh (Revolution Medicine, California)

PDAC is characterized by the presence of activating KRAS mutations that are known to drive tumorigenesis and tumor progression. In PDAC, glycine at amino acid position 12 is predominantly mutated to either aspartic acid (D) or valine (V). KRAS proteins were previously considered to be difficult to target. However, recent advances in the development of U.S. Food and Drug Administration (FDA)-approved agents and small molecules have provided a new avenue for treating cancer patients with KRAS mutations—including those involved in PDAC. Despite these breakthroughs, resistance to KRAS inhibitors has been observed, leading to challenges in effective treatment.

The mechanisms underlying this resistance are not yet fully understood, making therapy resistance intractable. Further research is needed to unravel the complexities of resistance mechanisms and develop combination therapy strategies to overcome therapy resistance in treatment. Our laboratory previously developed genetically engineered mouse models (GEMMs) of PDAC, including the KPC model (*Kras*$^{LSL\text{-}G12D/+}$*; Trp53*$^{LSL\text{-}R172H/+}$*; Pdx1-CRE*) as well as an unpublished FPC model (*Kras*$^{Frt\text{-}LSL\text{-}G12V\text{-}Frt/+}$*; Trp53*$^{LSL\text{-}R172H/+}$*; Pdx1-CRE; Rosa26*$^{FlpOERT2/+}$), harboring a conditional endogenous Kras$^{G12D}$ or Kras$^{G12V}$ allele, respectively (Fig. 1A). The Kras$^{G12V}$ allele in FPC mice can be subsequently deleted in vivo by the application of tamoxifen (TAM), and we have found that these FPC mice stochastically develop PDAC in a manner indistinguishable from that of KPC mice. The selective deletion of a mutant KRAS$^{G12V}$ allele upon administration of TAM mimics the pharmacological mutant-specific inhibition of KRAS$^{G12V}$ in tumors, allowing us to characterize pathways that enable resistance to Kras inhibition. Over the past year, our research was centered on evaluating the effectiveness of KRAS inhibition and identifying potential resistance mechanisms through two distinct approaches: genetic deletion of mutant KRAS as a monoallelic therapy and pharmacological inhibition of KRAS using a pan-RAS inhibitor (RM-042). Among several inhibitors targeting either pan-RAS, mutant KRAS$^{G12D}$, or mutant KRAS$^{G12C}$ that are being tested in clinical trials, RM-042 is an oral small-molecule inhibitor of GTP-bound RAS proteins. It binds to cyclophilin A noncovalently and engages RAS to form a tricomplex that sterically

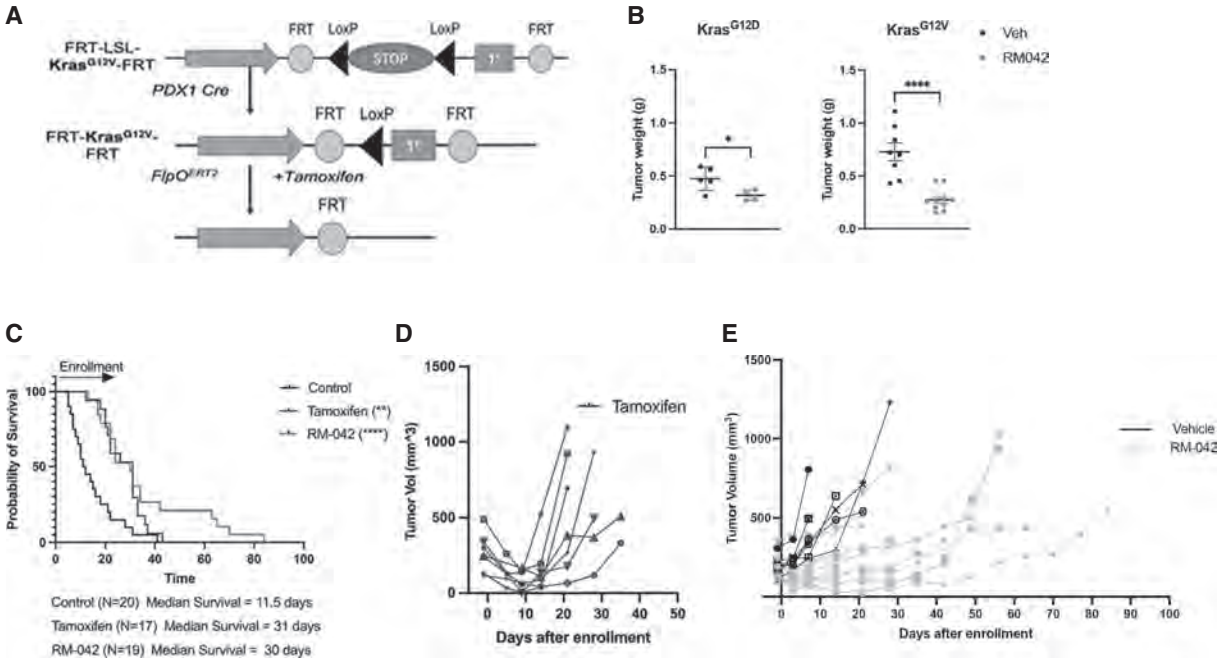



**Figure 1.** Genetic deletion of oncogenic KRAS and RM-042 treatment impairs tumor growth and extends survival in the FPC model of pancreatic ductal adenocarcinoma (PDAC). (*A*) Schematic showing alleles of the FPC model. (*B*) Tumor weight measurements of KPC ($Kras^{G12D}$) tumors that were enrolled to vehicle or RM-042. Data are mean ± SEM of five (vehicle) and four (RM-042) biological replicates (*left*). Tumor weight measurements of FPC ($Kras^{G12V}$) tumors (*right* panel) enrolled to vehicle or RM-042 (*right*). Results show mean ± SEM of eight (vehicle) and 10 (RM-042) biological replicates. (*C*) Kaplan–Meier survival analysis for FPC mice treated with RM-042 ($n = 19$, light gray) or TAM ($n = 17$, dark gray). ****, $p < 0.0001$; **, $p = 0$; 014 by log-rank (Mantel–Cox) test. (*D,E*) Tumor volume changes in the FPC mice treated with vehicle (corn oil, or vehicle-treated).

blocks RAS downstream effectors. RM-042 is a pan-RAS inhibitor that can inhibit multiple RAS proteins (HRAS, NRAS, KRAS, and mutant KRAS). First, we evaluated whether initial response and resistance occur through similar mechanisms in both approaches—genetic deletion and pharmacological inhibition of $KRAS^{G12V}$. We demonstrated that treatment with TAM for 5 days (i.e., genetic deletion of mutant $KRAS^{G12V}$) led to enrichment of extracellular matrix (i.e., increased collagen deposit contents shown by Masson's trichrome stains) and the loss of phospho-ERK (a main KRAS effector) in cytokeratin 19–expressing neoplastic cells in FPC mice. As RM-042 is a pan-RAS inhibitor, we conducted a short-term therapeutic study using both the $KRAS^{G12V}$ (FPC) and $Kras^{G12D}$ (KPC) mice treated with RM-042 for 7 days, and demonstrated that tumor weights were decreased within 7 days in both models (Fig. 1B). Similarly, treatment with RM-042 resulted in increased collagen deposits and reduced phospho-ERK level and Ki67 (proliferation marker) positivity in FPC tumors 7 days after enrollment. Taken together, our short-term preclinical studies strongly indicate that deletion of $KRAS^{G12V}$ or treatment with RM-042 each exert antitumor activity in the FPC mouse model of PDAC.

Next, we performed survival studies using the FPC model to determine the overall survival rates and to investigate KRAS inhibition–induced resistance mechanisms if present. Long-term treatment with RM-042 or TAM significantly increased median survival to a similar extent, and by about threefold compared with the control groups (Fig. 1C). Initial tumor regressions were observed upon both RM-042 and TAM treatment (Fig. 1D,E), but all mice eventually relapsed and succumbed to their disease. Interestingly, administration of TAM yielded more mice that relapsed quickly in comparison to RM-042 treatment, whereas relapse rates with RM-042 exhibited greater variability, with a minority of mice relapsing quickly

and a majority showing disease stabilization before progression. Indeed, the RM-042 cohort contained a subgroup of mice with a prolonged survival beyond 43 days, whereas none of the mice treated with TAM lived beyond 43 days. We found a complete absence of phospho-ERK in relapsed tumors treated with RM-042, whereas relapsed tumors following TAM treatment retained comparable amounts of phospho-ERK compared with the vehicle group. The deletion of $KRAS^{G12V}$ in the relapsed tumor was confirmed by polymerase chain reaction (PCR) (data not shown). This suggests alternative pathways to activate MAPK in TAM-treated mice. We also observed that relapsed tumors following RM-042 treatment predominantly contain well-differentiated epithelial cells, whereas relapsed tumors in TAM-treated mice revealed a hybrid epithelial–mesenchymal transition (EMT). The different phenotypes observed in relapsed tumors following monoallelic inhibition compared with RM-042 (pan-RAS inhibitor) could be due to the influence of the wild-type KRAS or other RAS proteins with RM-042 targets all RAS proteins. Our current research is focused on exploring the underlying mechanisms that result in these distinct phenotypes. Overall, our model offers a new approach to elucidate the multiple and unique resistance mechanisms by comparing monoallelic inhibition of mutant $KRAS^{G12V}$ with pan-RAS inhibition and to guide the development of combination therapy strategies.

## WNT Sensing as a Critical Regulator of CAF Heterogeneity and Function in PDAC

This work was done in collaboration with A. Dobin and J. Preall (CSHL) and A. Califano (Columbia University, New York)

The dense fibrotic stroma in PDAC is a major factor in progression and therapeutic resistance. This is the result of the continued deposition and remodeling of the extracellular matrix (ECM) by CAFs. Our group and others have described the functional heterogeneity of CAFs and identified three main subtypes: inflammatory CAFs (iCAFs), myCAFs, and antigen-presenting CAFs. In particular, myCAFs have been shown to be dependent on transforming growth factor-beta (TGF-β) signaling, produce large quantities of ECM components, and correlate with poor survival. Previous attempts at targeting myCAFs have yielded contradicting results, partly because of a lack of understanding of the myCAF differentiation process. Therefore, understanding the underlying mechanisms of their activation is necessary for the development of appropriate therapeutic strategies to ablate these tumor-promoting fibroblasts. To address this, we set out to study the evolution of the tumor microenvironment (TME) in response to tumor progression to identify pivotal axes of cancer cell–fibroblast cross talk.

In our collaborative study with Northwell Health, we hypothesized that nucleic acid testing (NAT) drug screening on patient-derived organoids (PDOs) can predict the most suitable chemotherapy regimen for a patient. One aim of our study was to assess the feasibility of generating PDOs from both chemotherapy-naïve and NAT-treated pancreatic cancer patients. We also explored factors that might influence PDO establishment from patients who received NAT therapy. Importantly, we compared PDO sensitivity to each NAT regimen with the patient's pathologic and clinical response to determine the correlation.

To track the evolution of fibroblasts across different stages of tumor development, we performed single-cell RNA sequencing (scRNA-seq) on normal pancreas (NP), a well-established model of caerulein-induced chronic pancreatitis (CP), early PanIN lesions (*Kras*$^{LSL\text{-}G12D/+}$;*Pdx1-Cre*, KC model), and overt PDAC (KPC model). To decipher differences in transcriptional landscape and protein activity between fibroblasts in unperturbed, inflamed, and oncogenic conditions, we applied the VIPER/ARACNe workflow that infers protein activity based on regulatory gene networks. Unsupervised clustering of fibroblasts revealed the coordinated activation of common pathways across the data sets, as well as the presence of certain tumor-specific pathways in KC and KPC mice (Fig. 2A). Whereas the majority of these tumor-specific clusters represented myCAFs and myCAF-like cells, Cluster 1 was distinct in its transcriptional and protein activity. Pseudotrajectory analysis inferred that Cluster 1 preceded myCAF and myCAF-like clusters and was only observed in KC and KPC mice (Fig. 2A,B). We then analyzed the top master regulators (MRs) differentially activated in Cluster 1 compared with myCAFs and found that Cluster 1 was characterized by the activation of pathways downstream of WNT signaling (Fig. 2C). To establish the relevance of this population of CAFs in human tumors, we analyzed

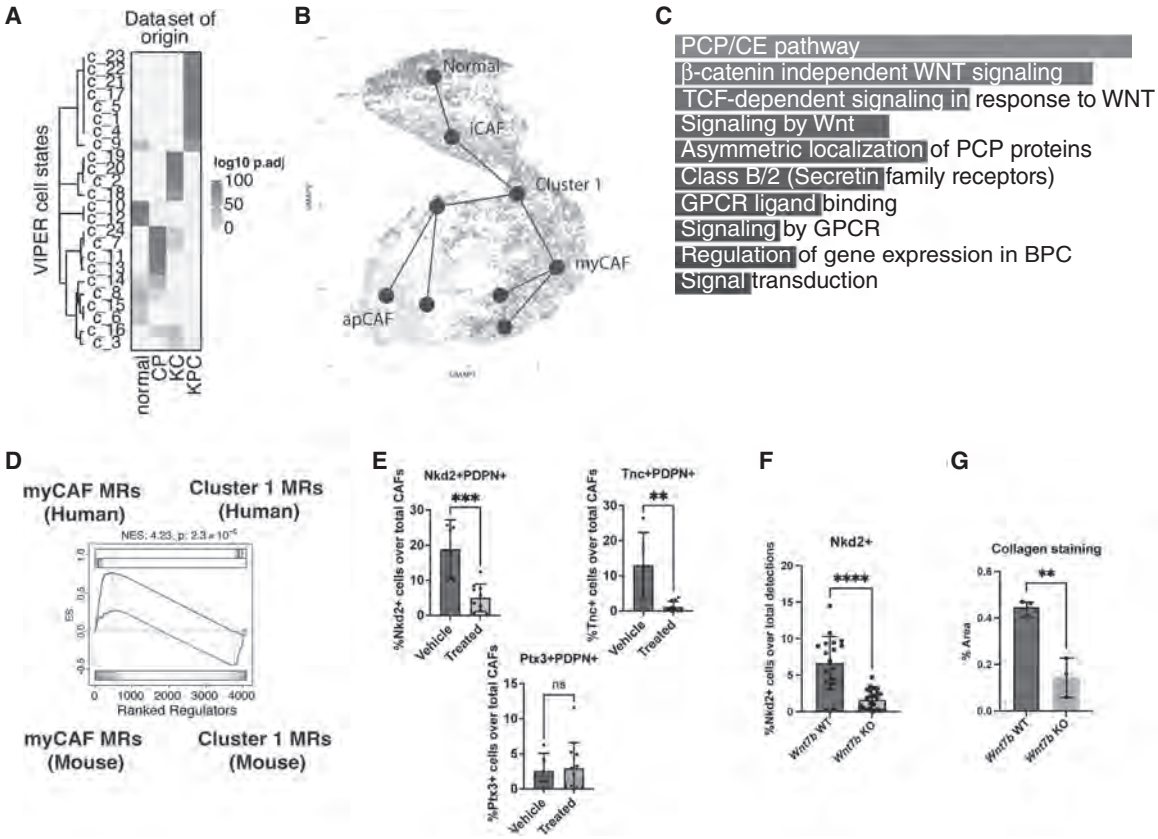



**Figure 2.** Wnt-sensing cancer-associated fibroblast (CAF) substate in pancreatic ductal adenocarcinoma (PDAC). (*A*) A heatmap depicting the enrichment of each data set in each cell state cluster identified by VIPER, showing unique VIPER cell states associated with the cancer-associated fibroblasts in KPC. A positive enrichment indicates a significant overlap by Fisher's exact test. (*B*) Pseudotrajectory plot generated by the Monocle algorithm on the VIPER cell states in normal and KPC data sets. (*C*) Pathway analysis of Cluster 1 regulatory proteins showing the enrichment of WNT-sensing pathways. (*D*) GSEA analyses showing enrichment of the top 50 most differentially activated master regulators (MRs) of mouse myCAFs and Cluster 1 CAFs in human CAFs. *Bottom* bar shows the ranked list of mouse regulators spanning myCAFs (*left*) to Cluster 1 CAFs (*right*). *Top* bars show the rank of each MR (vertical lines) in human myCAFs (blue) and Cluster 1 CAFs (red). (*E*) Quantification of % Cluster 1 CAFs, myCAFs, and iCAFs stained for *Nkd2*, *Tnc*, and *Ptx3*, respectively, in vehicle- and LGK974-treated KC mice by fluorescent in situ hybridization (FISH). (*F*) Quantification of % Cluster 1 CAFs and myCAFs in *Wnt7b* wild-type (WT) and knockout (KO) orthotopically transplanted mouse organoid model using FISH. (*G*) Quantification of Masson's trichrome stain as an indicator of the extent of collagen deposition in *Wnt7b* WT and KO orthotopically transplanted mouse organoid model.

publicly available human PDAC data sets and observed the presence of a cluster with a transcriptional profile similar to the mouse Cluster 1 and a highly conserved set of differentially activated MRs compared with myCAFs (Fig. 2D).

Although the combined role of WNT and TGF-β in fibrosis has been reported in other organs, little is known about WNT sensing and its consequences in CAFs. To investigate this, we treated pancreatic stellate cells (PSCs), the progenitors of CAFs in PDAC, in vitro with a combination of active WNT and TGF-β and found that WNT alone induced the expression of Cluster 1 gene signature and not that of myCAFs. Although TGF-β treatment resulted in up-regulation of myCAF genes, the combination of WNT and TGF-β resulted in far greater expression levels of these genes, suggesting a combined role for these two cytokines in differentiation of myCAFs. To corroborate this observation in vivo, we blocked WNT signaling in KC mice by administering the porcupine (PORCN) inhibitor LGK974 and characterizing the resulting CAF heterogeneity. Blocking WNT

activity through inhibition of PORCN resulted in a significant depletion of Cluster 1 cells marked by *Nkd2*, as well as myCAFs, marked by *Tnc* (Fig. 2E). Together, these data suggest that activation of WNT signaling induces the differentiation of Cluster 1 cells and that this together with TGF-β plays a critical role in differentiation of myCAFs.

To determine the source and type of WNT responsible for generating the Cluster 1 subset, we assessed the expression of various WNT ligands by different cell types in our KPC data set. Using fluorescent in situ hybridization (FISH) of RNA probes for four of the most highly expressed ligands, we were able to conclude that only the presence of Wnt7b-expressing cancer cells positively correlated with the presence of Cluster 1 CAFs ($Nkd2^+$) and myCAFs ($Tnc^+$). To assess whether Wnt7b is required for the differentiation of Cluster 1 CAFs, we performed genetic knockout (KO) of *Wnt7b* in a mouse tumor organoid line that normally expressed high levels of this gene. Orthotopic transplantation of *Wnt7b* KO organoids in wild-type mice led to concomitant depletion of Cluster 1 CAFs (Fig. 2F) and a significant reduction in collagen deposition (Fig. 2G). Thus, *Wnt7b* is necessary for the differentiation of Cluster 1 cells and lack of this ligand leads to diminished stromal response in PDAC.

Taken together, this work demonstrates the presence of a WNT-dependent CAF substate in PDAC that together with myCAFs generates the stromal reaction characteristic of PDAC and absent in other fibrotic settings, such as pancreatitis. Additionally, we uncover a previously unknown role of epithelial cell–derived WNT7B in orchestrating CAF heterogeneity. Blocking of WNT7B sensing in fibroblasts may thus provide a therapeutic avenue for preventing or reducing ECM deposition and generation of desmoplasia in PDAC.

## PUBLICATION

Caligiuri G, Tuveson DA. 2023. Activated fibroblasts in cancer: perspectives and challenges. *Cancer Cell* **41:** 434–449. doi: 10.1016/j.ccell.2023.02.015

# Ras AND Rho REGULATORS IN DEVELOPMENT AND DISEASE

**L. Van Aelst** F. Boato B. de la Cruz Thea C. Gonzalez B. Leasure
S. Chan S. George V.H. Le

Research in our laboratory centers on investigating the molecular and cellular mechanisms underlying both cancer and neural disorders. Particularly, we focus on elucidating the role of Ras and Rho GTPase family members in signal transduction. These proteins play crucial roles in cellular activities such as cell growth control, differentiation, and morphogenesis. Alterations in Ras and Rho functions have been implicated in cancer as well as neural disorders, including intellectual disability, autism, schizophrenia, epilepsy, and mood disorders. Our interests lie in understanding how defects in Ras- and Rho-linked proteins contribute to the development of these diseases/disorders. To this end, our laboratory continues to define the functions of select GTPases, their regulators, and their effectors, using animal models of cancer and neurodevelopmental/neurological disorders. Below we highlight key ongoing projects.

## Shedding Light on Chandelier Cell Connectivity and Contribution to Neural Disorders

Proper assembly and functioning of cortical circuits rely on the formation of specific synaptic connections between excitatory pyramidal neurons (PyNs) and different types of GABAergic interneurons. Among the various cortical interneuron types, chandelier cells (ChCs; also known as axo-axonic cells), in particular, have decisive influence over the output of excitatory PyNs because of their unique morphology and selective subcellular innervation. Specifically, ChCs possess a very distinctive, highly branched axon with multiple arrays of short, vertically oriented terminals of presynaptic boutons, called cartridges. This unique architecture enables a single ChC to connect to a large population of PyNs. Moreover, ChCs form synapses exclusively on the axon initial segment (AIS) of PyNs, the site of action potential initiation. As such, ChCs are physiologically poised to exert powerful yet precise control over PyN firing and population output. The importance of proper ChC function is further underscored by the detection of aberrant ChC cartridge/bouton number/size and axo-axonic synapse density in disease states such as schizophrenia, autism spectrum disorder, and epilepsy. Despite the importance of ChCs, still little is known about the mechanisms governing ChC structure and connectivity. To date, the only molecules implicated in neocortical ChC morphogenesis are the atypical Rac activator DOCK7, receptor tyrosine kinase ErbB4, and fibroblast growth factor 13. In particular, we found that silencing of DOCK7 in ChCs via a novel ventral medial ganglionic eminence (vGME)-directed in utero electroporation (IUE) approach decreases ChC cartridge bouton size and density and does so by modulating the activity of ErbB4. Noteworthy though, we found that ChCs depleted for DOCK7, ErbB4, or FGF13 still make contact with PyN AISs, indicating that other molecules must regulate ChC/PyN axo-axonic innervation.

To identify molecular determinants of ChC/PyN AIS innervation, we initiated an in vivo RNA interference (RNAi) screen of PyN-expressed axonal cell adhesion molecules (CAMs) and select Ephs/ephrins. In doing so, we found the L1 family member L1CAM to be essential for ChC/PyN AIS innervation. We further showed that L1CAM is required during both the establishment and the maintenance of innervation and that selective innervation of PyN AISs by ChCs requires AIS anchoring of L1CAM by the cytoskeletal ankyrin-G/βIV-spectrin complex. Thus, our findings identify PyN-expressed L1CAM as the first known CAM required for the innervation of PyN AISs by ChCs in the neocortex. More recently, we collaborated with the Rasband laboratory (Baylor College of Medicine) to identify AIS-enriched cell surface proteins using a selective proteomic proximity labeling assay with tyramide (SPPLAT). This led to the identification of the CAM CNTN1 as being highly enriched at PyN AISs, which is particularly interesting as it was previously reported to interact in cis with L1CAM in cellular assays. These findings prompted us to investigate whether PyN CNTN1 is also

involved in mediating PyN AIS innervation by ChCs. To this end, we used the CRISPR-Cas9 system and ventricular zone–directed IUE to knock out CNTN1 in neocortical PyNs. We found that CNTN1 depletion in PyNs significantly reduced the percentage of PyN AISs innervated by ChCs. Together, our data unveil a critical role for PyN-expressed CNTN1, which forms an in cis interaction with L1CAM, in mediating proper neocortical ChC/PyN AIS innervation.

In addition to the above intrinsic molecular players, we also uncovered a novel, previously unrecognized role for microglia, the brain's resident immune cells, in the regulation of PyN AIS synapse formation/maintenance by ChCs in the neocortex. Specifically, we identified a synaptogenic/growth-promoting role for microglia in regulating PyN AIS synapse formation by ChCs. We showed that a population of microglia closely interact with PyN AISs and ChC cartridges and that such tripartite interactions, which rely on the unique AIS cytoskeleton and microglial $GABA_{B1}$ receptors, are associated with increased ChC cartridge length and bouton number and AIS synaptogenesis. Conversely, microglia depletion or disease-induced aberrant microglia activation (e.g., under Alzheimer's disease–associated neuroinflammation conditions) impairs the proper development and maintenance of ChC cartridges and boutons as well as AIS synaptogenesis. These findings unveil key roles for homeostatic, AIS-associated microglia in regulating proper ChC axonal morphogenesis and synaptic connectivity in the neocortex. Importantly, given our findings that microglia are critical for proper ChC structure/function and because perturbations in both microglia and GABAergic ChCs have been reported in patients with neurological conditions associated with aberrant excitatory/inhibitory balance, our work also sheds new light on potential microglial-dependent interneuronopathies underlying brain disorders.

## Delineating the Multifunctional Role of the X-Linked Intellectual Disability Protein Oligophrenin-1

*Oligophrenin-1* (*OPHN1*), which encodes a Rho-GTPase activating protein, was the first Rho-linked gene identified for intellectual disability (ID). It was initially discovered through the analysis of a balanced translocation t(X;12) in a female patient with mild ID. Subsequent studies revealed *OPHN1* mutations in families with a syndromic form of ID, with affected individuals exhibiting mild to severe ID, behavioral problems, epileptic seizures, and developmental delay. We began the functional characterization of *OPHN1* to understand how mutations in this gene contribute to these symptoms. Our initial efforts focused on understanding the role of OPHN1 in the hippocampus, a brain structure crucial for learning and memory. These studies uncovered a multifunctional role for OPHN1 at hippocampal CA1 synapses. Specifically, we found that OPHN1 is essential for controlling activity-driven glutamatergic synapse development and that the temporal regulation of OPHN1 translation plays a key role in mGluR-dependent long-term depression (LTD), a form of plasticity associated with cognitive disorders that requires rapid new protein synthesis. These findings provided first insights into how *OPHN1* mutations could contribute to learning and memory deficits in affected individuals.

As to the behavioral problems reported in *OPHN1* patients, these include hyperactivity, emotional imbalance, and intolerance to frustration, which are often precipitated or exacerbated by stressful events. This prompted us to examine the role of OPHN1 in stress-related behavioral problems using the well-established learned helplessness procedure, a model of depressive-like behaviors. We discovered that Ophn1 deficiency in mice significantly enhanced helpless and depressive-like behavior when faced with repeated, uncontrollable stress. Remarkably, deleting Ophn1 exclusively in parvalbumin (PV) interneurons in the prelimbic medial prefrontal cortex (PL-mPFC) was sufficient to induce helplessness. This behavioral phenotype was driven by a reduced excitatory drive onto Ophn1-deficient PL-mPFC PV interneurons, leading to hyperactivity in this region. Crucially, suppressing neuronal activity or RhoA/Rho-kinase signaling in the PL-mPFC reversed the helpless behavior. These findings identified OPHN1 as a crucial regulator of adaptive behavioral responses to stress and revealed the mechanistic link between *OPHN1* genetic deficits, mPFC circuit dysfunction, and abnormalities in stress-related behaviors.

Finally, recognizing that malformations of cortical development significantly contribute to epileptic seizures and developmental delay, we recently explored the role of OPHN1 in neocortical development. By combining IUE methodology with RNAi, we discovered a critical role for OPHN1 in the migration of neocortical pyramidal (NP) neurons in the developing mouse neocortex. Furthermore, using molecular replacement

and live-cell imaging, we found that OPHN1 regulates NP neuron migration by controlling both the multipolar to bipolar (MP-BP) transition and the proper morphogenesis of the leading process (LP) and nuclear translocation through separate pathways. Specifically, we discovered that OPHN1 inhibits the RhoA/Rho-kinase pathway via its GAP domain to control MP-BP transition and, in a proline-rich domain 2 (PRD2)-dependent mechanism, regulates LP morphogenesis and nuclear translocation. Importantly, the coordinated action of both pathways is required to ensure efficient NP neuron migration in the developing neocortex. These findings may provide a mechanistic explanation for the epileptic seizures and developmental delay observed in *OPHN1* patients.

In summary, our studies have revealed new mechanistic insights into how mutations in the *OPHN1* gene could lead to the various manifestations observed in individuals affected by OPHN1.

## Molecular and Cellular Mechanisms Contributing to Tumor Formation and Progression

In addition to its role in regulating cognitive and behavioral functions, we have recently uncovered the implication of OPHN1 in glioma tumorigenesis. We found a significant upregulation of OPHN1 expression in glioblastoma (GBM) tumors compared to normal tissue. To assess its role in glioma tumorigenesis, we implemented several GBM mouse models. Using these models, we observed that OPHN1 overexpression significantly enhanced glioblastoma formation, whereas OPHN1 knockdown decreased the formation of tumors. Additionally, our investigations revealed that OPHN1 is not only expressed in GBM cells but also in surrounding microglia. This observation prompted further inquiry into the role of microglia in GBM, leading to a collaboration with the Lukey (CSHL) and Cross (Memorial Sloan Kettering Cancer Center) laboratories to investigate the exchange of metabolites between GBM cancer cells and microglia. These efforts led to the identification of a potentially immune-modulatory metabolic cycle that can facilitate GBM progression, and metabolomics approaches are now being applied to explore the mechanistic details of this process.

Motivated by our examinations on OPHN1 in stress-related disorders, we embarked on a collaboration with the Egeblad laboratory (CSHL) to investigate how chronic stress promotes lung metastasis from disseminated cancer cells. Through our utilization of animal models to study stress, we found that chronic stress induces notable modifications in the lung microenvironment. Specifically, we identified an accumulation of fibronectin, a decrease in T-cell infiltration, and an increase in neutrophil infiltration. Subsequent depletion experiments revealed the essential role of neutrophils in stress-induced metastasis. Moreover, we found that chronic stress disrupted the regular circadian variations in neutrophil numbers and marker expression and caused increased neutrophil extracellular trap (NET) formation. Glucocorticoids released during chronic stress were responsible for the gene expression changes, including in circadian-regulating genes, and for inducing neutrophils to form NETs. In mice with glucocorticoid receptor deletion in neutrophils, chronic stress failed to increase NET levels and metastasis. Furthermore, digesting NETs with DNase I prevented chronic stress from causing fibronectin accumulation and inducing lung metastasis. Together, our data show that glucocorticoids cause formation of NETs, and these are critical for establishment of a stress-induced metastasis-promoting microenvironment. NETs therefore could be markers and targets to prevent metastatic recurrence in cancer patients, many of whom will experience chronic stress due to their disease.

### PUBLICATIONS

de Ruiter Swain J, Michalopoulou E, Noch EK, Lukey MJ, Van Aelst L. 2023. Metabolic partitioning in the brain and its hijacking by glioblastoma. *Genes Dev* **37:** 681–702. doi:10.1101/gad.350693.123

Ogawa Y, Lim BC, George S, Oses-Prieto JA, Rasband JM, Eshed-Eisenbach Y, Hamdan H, Nair S, Boato F, Peles E, et al. 2023. Antibody-directed extracellular proximity biotinylation reveals that Contactin-1 regulates axo-axonic innervation of axon initial segments. *Nat Commun* **14:** 6797. doi:10.1038/s41467-023-42273-8

*In Press*

Amit M, Anastasaki C, Dantzer R, Demir IE, Deneen B, Dixon KO, Egeblad M, Gibson EM, Hervey-Jumper SL, Hondermarck H, et al. 2024. Next directions in the neuroscience of cancers arising outside the CNS. *Cancer Discov* **14:** 669–673. doi:10.1158/2159-8290.CD-23-1495

He XY, Gao Y, Ng D, Michalopoulou E, George S, Adrover JM, Sun L, Albrengues J, Daßler-Plenker J, Han X, et al. 2024. Chronic stress increases metastasis via neutrophil-mediated changes to the microenvironment. *Cancer Cell* doi:10.1016/j.ccell.2024.01.013

# NEUROSCIENCE

The **Florin Albeanu** laboratory takes advantage of expertise in behavioral analysis, computational, and systems neuroscience via synergistic interactions with other CSHL groups. Recent and current projects have been aimed at understanding the reciprocal relationship between sensation and action, the circuit-level decoding mechanisms of specific odor features, and the wiring logic of olfactory system, including the interplay between feedforward and feedback signals via parallel processing loops. They leverage closed-loop behaviors in rodents and large-scale optical imaging, optogenetic manipulations, and electrode recordings to identify brain-wide neuronal circuits that mediate egocentric predictions of sensory inputs given specific motor actions, as well as the basis set (red–green–blue analogs) of olfactory perception. By comparing across sensory modalities (olfaction, vision, audition), they also aim to distinguish between neuronal circuits that predict moment-to-moment changes in sensory input and general-purpose prediction systems that encode more abstract models of the world. Toward finding the basis set of olfactory perception, taking advantage of behavioral analysis and patterned optogenetic manipulations, they search for the algorithms and circuit-level decoding mechanisms of specific odor features. They investigate the function of neural circuits in the olfactory stream (olfactory bulb, cortex, striatum) and the inputs these receive from motor and association neocortical areas in wild-type mice and models of psychiatric disorders. They use high-throughput DNA barcoding tools for mapping neuronal connectivity to understand the wiring logic of olfactory system, including the interplay between feedforward and feedback signals and their relationship to odorant receptors.

Social animals must interact with each other to cooperate and compete. Using sounds for such interactions is common across many taxa. Humans engaged in conversation, for example, take rapid turns to go back and forth—a feat that most of us tend to perform effortlessly, but that breaks down during neuropsychiatric disorders. Our understanding of neural circuits that underlie vocal communication, especially in mammals, remains quite rudimentary. Recently, it was discovered that a neotropical rodent, Alston's singing mouse, engages in fast vocal interactions, even in laboratory settings. The **Arkarup Banerjee** laboratory, using this novel model system, seeks to pursue two complementary questions. First, how does the auditory system interact with the motor system to generate the sensorimotor loop required for vocal communication? Second, what are the neural circuit modifications that allow behavioral novelty to emerge during evolution? Various rodent species show marked differences in vocal behaviors. Genes that determine such behavioral differences (e.g., between the singing mouse and the laboratory mouse) must act via neural circuits within the brain. However, the structural and functional changes in the brain that specify the distinct vocal repertoires across related species remain unknown. Research in the laboratory combines cutting-edge systems neuroscience and comparative evolutionary analyses of neural circuitry across rodent species to bridge this knowledge gap.

The powerful influence of sensory experience on brain development has been appreciated since the 1960s. However, even today, the fundamental cellular and molecular mechanisms through which sensory input shapes developing neural circuits remain largely mysterious. The **Lucas Cheadle** laboratory recently discovered that sensory experience alters gene and protein expression in microglia, the resident immune cells of the brain. These sensory-induced changes allow microglia to interact with neighboring neurons to strengthen and maintain a subset of synaptic connections and to eliminate others. These findings raise the exciting possibility that microglia, which are predominantly associated with immune responses to injury and disease, also decode salient features of the physical world and contribute to neural responses to the environment.

The Cheadle laboratory applies a multidisciplinary approach to the visual system of the mouse to investigate the contributions of microglia to sensory experience–dependent synapse development and plasticity. They further seek to identify the molecular mechanisms through which microglia effect changes at synapses and thereby exert control over brain function. To accomplish this, the Cheadle laboratory images microglial interactions with synapses in the brains of living mice, which allows the researchers to characterize the specific features of the environment to which microglia respond. In parallel, the research team uses cutting-edge single-cell transcriptomic and genomic strategies, such as single-cell RNA sequencing, to profile the molecular changes in microglia that are elicited by distinct sensory stimuli. With these combined approaches, the Cheadle laboratory is interrogating the ways in which environmental stimuli converge on the microglial genome to shape neural circuit development and function.

The **Benjamin Cowley** laboratory identifies data-driven models of neural responses and behavior by coupling data collection with model training during closed-loop experiments. They condense these models into compact, interpretable forms, allowing them to describe the complicated computations of the brain in a clear and concise way. To understand the brain's computations, one often seeks a model to capture the step-by-step computations of neurons. For example, a model can take an image as input and output a visual neuron's response. The most predictive models in computational neuroscience typically have millions of parameters and need large amounts of training data, making it difficult to obtain and interpret such models. The Cowley research group takes a two-pronged approach to address these problems. First, they design adaptive stimulus selection techniques to efficiently train models (e.g., deep neural networks) with as little recording time as possible. They work hand in hand with experimentalists to deploy these systems in closed-loop experiments. Second, they develop machine-learning techniques to identify highly predictive models that have as few parameters as possible. They then analyze these "compact" models to determine the computations necessary and sufficient to explain a neuron's response. Their approach of identifying highly predictive, interpretable models will shed light on computations otherwise hidden by the scale and complexities of the brain.

**Hiro Furukawa**'s laboratory studies receptor molecules involved in neurotransmission. Its members mainly focus on the structure and function of NMDA (*N*-methyl-D-aspartate) receptors—ion channels that mediate excitatory transmission. Dysfunctional NMDA receptors cause neurological disorders and diseases including Alzheimer's disease, Parkinson's disease, schizophrenia, depression, and stroke-related ischemic injuries. The Furukawa laboratory is working to solve the three-dimensional structure of the very large NMDA receptor by dividing it into several domains. They seek to understand the pharmacological specificity of neurotransmitter ligands and allosteric modulators in different subtypes of NMDA receptors at the molecular level. Toward this end, they use cutting-edge techniques in X-ray crystallography to obtain crystal structures of the NMDA receptor domains and validate structure-based functional hypotheses by a combination of biophysical techniques, including electrophysiology, fluorescence analysis, isothermal titration calorimetry, and analytical centrifugation. Crystal structures of NMDA receptors serve as a blueprint for creating and improving the design of therapeutic compounds with minimal side effects for treating neurological disorders and diseases. During the last several years, the team discovered and mapped several regulatory sites in specific classes of NMDA receptors—progress that now opens the way to the development of a new potential class of drugs to modulate receptor activity.

Facial expression, which conveys complex and nuanced emotions, is the primary focus of **Helen Hou**'s laboratory. How facial expressions are produced and adapted and how these processes change in mental disorders remain unknown, largely because of the complexity of the behavior in humans and a lack of cellular-level understanding of the underlying neurobiology. In *The Expression of the Emotions in Man and Animals*, Darwin explored the connection of emotional

states to organization of movements and characterized the universal nature of facial expression and its anatomical and musculature origin, such as lifting of the eyebrow in surprise. Underlying the rich tapestry of facial expressions is a group of facial muscles directly controlled by motor neurons in the facial nucleus in the brain stem. Many of the machineries of facial expression are highly conserved in mammals, including rodents. Hou's laboratory investigates how the brain orchestrates motor control in natural and innate behaviors—specifically the neural mechanisms underlying the dynamic control of facial expression—using rodents as a model, applying electrophysiological, imaging, neuroanatomical, behavioral, and computational analyses. To solve these puzzles, they leverage the "brain stem bottleneck," evolutionarily conserved integration nodes through which all communication between forebrain and spinal cord must funnel.

**Alexei Koulakov** and colleagues are trying to determine the mathematical rules by which the brain assembles itself, with particular focus on the formation of sensory circuits such as those involved in visual perception and olfaction. The visual system of the mouse was chosen for study in part because its components, in neuroanatomical terms, are well understood. What is not known is how projections are generated that lead from the eye through the thalamus and into the visual cortex, how an individual's experience influences the configuration of the network, and what parameters for the process are set by genetic factors. Even less is known about the assembly of the neural net within the mouse olfactory system, which, in the end, enables the individual to distinguish one smell from another with astonishing specificity and to remember such distinctions over time. These are among the challenges that engage Koulakov and his team.

Understanding the link between neural circuits and behavior has been the focus of research in the **Bo Li** laboratory. They are particularly interested in studying the synaptic and circuit mechanisms underlying reward processing, attention, and learning and memory, as well as synaptic and circuit dysfunctions responsible for maladaptive behaviors that are related to major mental disorders. They integrate in vitro and in vivo electrophysiology, imaging, molecular, genetic, optogenetic, and chemogenetic techniques to probe and manipulate the function of specific neural circuits—with a focus on the fear and reward circuits—in the rodent brain, and to determine how these circuits participate in adaptive or maladaptive behavioral responses in various tasks.

**Partha Mitra** is interested in understanding intelligent machines that are products of biological evolution (particularly animal brains), with the basic hypothesis that common underlying principles may govern these "wet" intelligent machines and the "dry" intelligent machines that are transforming the present economy. Mitra initiated the idea of brain-wide mesoscale circuit mapping, and his laboratory is involved in carrying out such mapping in the mouse (http://mouse.brainarchitecture.org) and the marmoset (in collaboration with Japanese and Australian scientists at the RIKEN Brain Science Institute and Monash University).

Mitra spent 10 years as a member of the theory department at Bell Laboratories and holds a visiting professorship at Indian Institute of Technology Madras, where he is helping establish the Center for Computational Brain Research. He has an active theoretical research program in machine learning and control theory wherein he is using tools from statistical physics to analyze the performance of distributed/networked algorithms in the "thermodynamic" limit of many variables.

The **Gabrielle Pouchelon** laboratory is interested in the interplay between environmental cues and molecular programs in the assembly of neural circuits. During development, environmental factors can affect the penetrance of genetic susceptibility to neuropsychiatric disease. Indeed, sensory experience and molecular programs have been shown to synergistically regulate neuronal maturation. Neurons in the neocortex receive activity through long-range presynaptic inputs that communicate information from the environment. In particular, sensory inputs to cortical sensory areas regulate the excitatory/inhibitory balance during development. However, excitatory and

inhibitory neurons appear to be differentially affected by external cues. What is the role of input-specific developmental activity, and what are the downstream molecular effectors in distinct cell types? In addition, as development progresses, not only sensory activity but other early postnatal environmental cues or stress are conveyed to the neocortex. Do these modulators shape neural circuits—and if so, how?

To address these questions, the Pouchelon group uses a combination of approaches from physiology to mouse genetics and genomics. They are interested in parsing out the origin of the various dysfunctions in neurodevelopmental disorders. More specifically, they focus on the differential vulnerability and contribution of distinct neuron types and their inputs in autism disorder models.

**Stephen Shea's** laboratory studies the neural circuitry underlying social communication and decisions. He uses natural social communication behavior in mice as a model to understand circuits and processes that are evolutionarily conserved and therefore shared broadly across species, likely contributing to disorders such as autism. Shea and colleagues have examined how emotion and arousal enable mice, via their olfactory systems, to store memories of other individuals and of related social signals. The team has exploited the intimate relationship between memory and emotion to effectively create memories in anesthetized mice, allowing them unprecedented access to neurobiological processes that typically only occur during behavior. The laboratory has been making a detailed analysis of the changes in neural connections that underlie odor memory. The team is particularly focused on an enigmatic cell type (granule cells [GCs]) that has long been hypothesized to be crucial for memories but has resisted direct study. They have developed methods for recording, giving them the first glimpse of the dynamics of these cells while the animal is learning an odor. The results show unexpectedly complex population dynamics among the GCs that were independently predicted by a model of odor learning developed in Alexei Koulakov's laboratory. The two laboratories are collaborating to discern how GC population activity gets integrated by olfactory bulb output neurons and to pinpoint the synaptic circuit that underlies this form of learning. In parallel, another member of the laboratory is using imaging techniques to determine how memories are stored among broad neuronal ensembles at a different level of the system. Recently, the laboratory made a key breakthrough, developing the ability to record from GCs in awake animals and discovering that their activity is dramatically modulated by state of consciousness. Finally, the Shea laboratory completed a series of studies of a different form of social recognition: auditory recognition of pup vocalizations by their mothers. Through this research, they have shown that a mouse model of Rett syndrome exhibits deficits in communication and learning not unlike those in human patients. Grants from the Simons and Whitehall Foundations are allowing the laboratory to extend this work by directly linking these deficits to the action of the gene *MeCP2* in the auditory cortex.

The **Jessica Tollkuhn** laboratory seeks to understand how transient events during brain development exert lasting effects on gene expression, circuit function, and, ultimately, behavior. They study how sex-specific neural circuits in rodents are established and modulated by the gonadal hormones estrogen and testosterone. The cognate receptors for these hormones are nuclear receptor transcription factors, which orchestrate modification of local chromatin environment and thus exert long-term effects on gene expression. However, the genes regulated by these receptors, as well as the specific mechanisms they use, remain poorly understood in the brain. This is in part because the extraordinary cellular heterogeneity of the brain complicates analysis of the small subpopulations of neurons that mediate sex-specific behaviors.

Having recently identified sex differences in both gene expression and chromatin in brain regions known to regulate sex-specific behaviors, the Tollkuhn laboratory is now working to understand how hormones generate these molecular sex differences during development, through the use of biochemical, genomic, and behavioral analyses. They have developed a method that permits

genome-wide analysis of histone modifications or DNA methylation in genetically defined populations of neurons. They hypothesize that these epigenetic data, combined with gene expression profiling, define the molecular signature of the critical period for sexual differentiation of the brain. Their goal is to provide a mechanistic link between the transcriptional effects of hormone signaling during development and the consequent social behaviors displayed in adulthood.

**Anthony Zador** and colleagues study how brain circuitry gives rise to complex behavior. Work in the laboratory is focused on two main areas. First, they ask how the cortex processes sound, how that processing is modulated by attention, and how it is disrupted in neuropsychiatric disorders such as autism. Recently, the laboratory found that when a rat makes a decision about a sound, the information needed to make the decision is passed to a particular subset of neurons in the auditory cortex whose axons project to a structure called the striatum. In the second major line of work in the Zador laboratory, they are developing new methods for determining the complete wiring instructions of the mouse brain at single-neuron resolution, which they term the "connectome." In contrast to previous methods, which make use of microscopy, these methods exploit high-throughput DNA sequencing. Because the costs of DNA sequencing are plummeting so rapidly, these methods have the potential to yield the complete wiring diagram of an entire brain for just thousands of dollars.

# UNDERSTANDING HOW THE BRAIN RELATES ACTION TO PERCEPTION

**F. Albeanu** W. Bast M. Dussauze
D. Cowan P. Gupta
Z. Dawood D. Hernández Trejo

The broad scope of our group is understanding the algorithms the brain uses to relate actions to perception by (1) discovering the logic of the odor space and the neural representations underlying olfactory perception, and (2) determining how the brain generates sensorimotor predictions and computes error signals. In recent projects, we contributed to understanding (a) the wiring logic of olfactory system, including the interplay between feedforward and feedback signals, (b) the circuit-level decoding mechanisms of specific odor features in changing environments, and (c) the reciprocal relationship between sensation and action.

## Sensorimotor Prediction Error Signals in the Olfactory Cortex

Sensation and action operate in a closed loop: Movements shape sensory input and sensory inputs guide motor commands. Through experience, the brain learns the reciprocal relationship between inputs and movements to build internal models that predict the sensory consequences of upcoming actions (sensorimotor predictions). In rodents, olfaction is intrinsically linked to motor action through sniffing and body movements. However, to date, most studies have probed olfactory processing during passive odor sampling, and the effect of movements on olfactory representations has been rarely analyzed.

We hypothesized that, in closed-loop olfaction, mice predict the sensory consequences of their actions (the next most probable odor input). Movement-related olfactory expectations get compared with current odor input within the olfactory cortex to represent olfactomotor prediction errors. To test these hypotheses, we developed a novel behavioral task (Smellocator; Fig. 1), in which head-fixed mice learn to steer the left–right location of an odor source by controlling a light lever with their forepaws. In this manner, we (1) link a precise motor action to well-defined sensory expectations (odor location), and (2) subsequently violate the learned expectations via online feedback perturbations in expert mice.

Strikingly, we find that expert mice readily counter brief sensorimotor perturbations by making precise corrective movements that provide a readout of their individually learned sensorimotor predictions. Concurrent single unit recordings from the olfactory cortex showed that olfactomotor expectations reshape odor-driven responses of individual neurons, with transient perturbations often evoking strong responses (Fig. 1). One challenge in assessing whether these responses represent error signals is that mice also change their sniffing patterns when faced with unexpected perturbations. To untangle potential sensorimotor mismatch responses from sniff-related modulation, we built a linear model aimed at predicting the responses of individual neurons as a function of both sniffing and odor stimulus state (identity, location). Although the model explained well neuronal spiking under closed-loop conditions, responses during periods of sensorimotor mismatch could not simply be explained as changes in sniffing.

Our results indicate that the olfactory cortex computes sensorimotor prediction errors by integrating odor information with movement-related predictions, presumably relayed via top-down feedback. Using cell type analysis and flexible activity manipulations, we further aim to identify the circuit elements that facilitate the comparison of olfactory inputs with predictions.

*Significance:* We devised an innovative framework to understand the neural substrates of sensorimotor predictions and errors. Mismatch between expectations and sensory experience may underlie nervous system dysfunctions such as schizophrenia (overprediction) and autism (underprediction). Progress in understanding the mechanisms of sensorimotor predictions is, thus, central to gaining novel insight for clinical intervention.

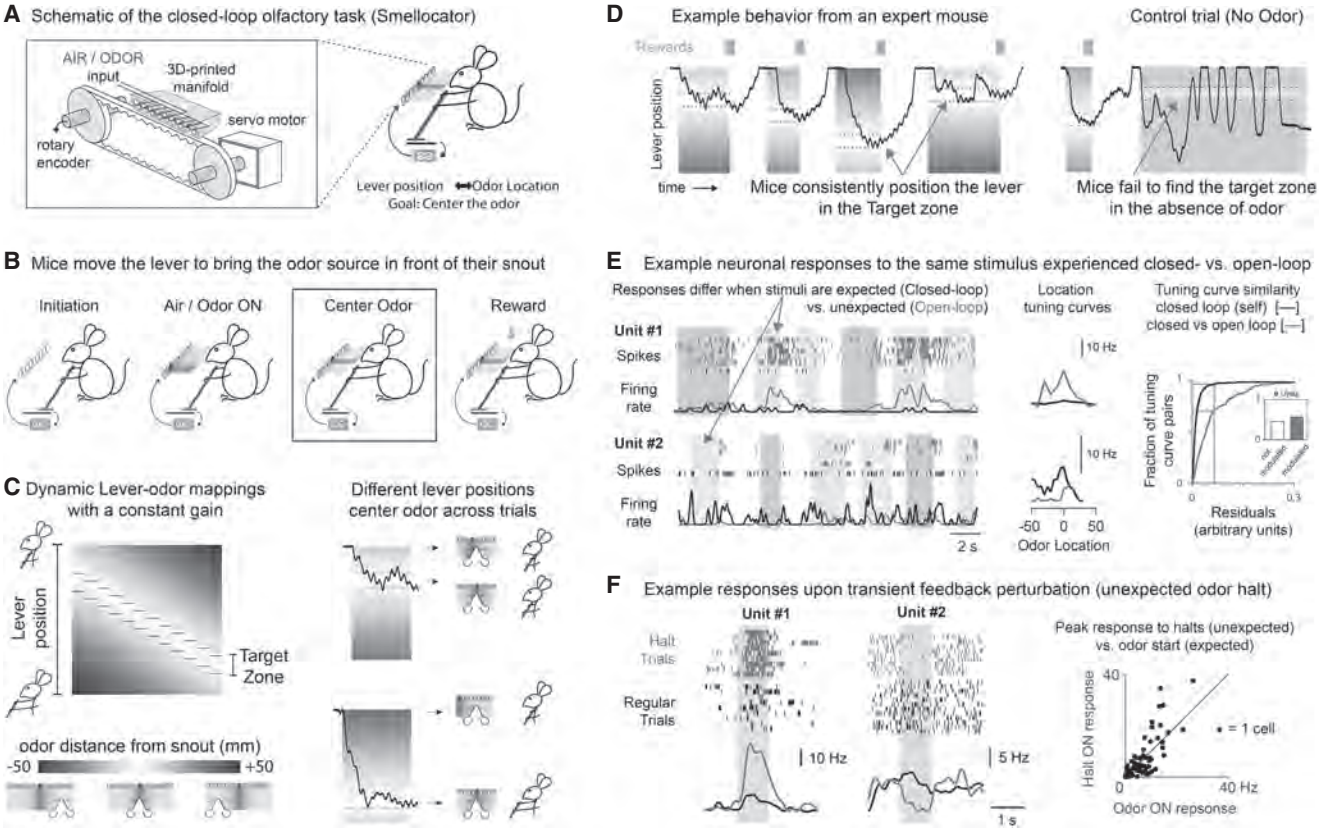



**Figure 1.** A closed-loop behavior to understand the neuronal substrates of sensorimotor predictions. (*A*) Setup: 1D lever movement is mapped to left–right odor displacement (±50 mm range). (*B*) Smellocator task: Mice initiate trials by pulling the lever to activate airflow and place the odor source at the start location. Mice then steer the odor and center it, by bringing the lever to the target zone: lever positions that place odor <2 mm from the snout. Maintaining target hold (~400 msec) triggers reward and terminates the trial. (*C*) (*Left*) Dynamic lever-odor mappings across trials, with a constant gain: The one-to-one correspondence of lever position to odor source location changes across trials as the odor source is initialized at different locations at Trial Start when the lever is held at its max position. However, the gain is kept constant across all trials—unit displacement of the lever always results in the same relative displacement of the odor source. Different shades indicate distance from snout. (*Right*) Example trials with different initial odor source locations and corresponding lever trajectories (black) in time. The gradient represents the trial-specific mapping of lever position to odor location. Schematics show odor location at trial start and in target zone. (*D*) Example session showing reliable centering of odor in each trial despite varying target zones. Catch trials (gray) with clean air result in random, untargeted movements. (*E*) Responses (*left*) and location tuning curves (*middle*) of two example olfactory cortex neurons to the same odor stimulus time series during a closed loop (black) and five passive replays (gray). The odor stimuli during open loop are a replay of the odor time series generated during closed loop. Boxes indicate Trial On periods; different shades indicate different odors. (*Right*) Summary plot showing differences between "closed-loop" and "open-loop" location tuning curves of 61 co-recorded olfactory cortex neurons in one mouse. (*F*) (*Left*) Responses of two example olfactory cortex neurons upon transient perturbation of closed-loop olfactory feedback—sudden halting of the odor source in place for one second, irrespective of the animal's action. Responses are aligned to perturbation start. Black and gray traces show time-matched response in control trials (no perturbation) and in odor halt trials, respectively. (*Right*) Summary plot comparing responses of 71 co-recorded neurons in one mouse to sudden odor halts (unexpected stimuli) versus closed-loop odor exposures (expected). Halt responses of many neurons are larger than their typical odor ON responses.

## Understanding the Neural Circuit Logic of Olfactory Perception

The representations of sensory objects that are formed by the nervous system reflect properties of these objects relevant for each organism's fitness. The principles underlying these representations have been challenging to discover in sensory modalities lacking convenient reference dimensions of space and time, such as olfaction. *In such cases, it is tempting to assume that sensory object spaces have no structure at all*. This suggestion is the leading hypothesis for the nature of olfactory sensory representations, which have been under intense debate lately. Recent machine-learning approaches enabled progress in predicting the perceptual quality (smell) of odorants, given their molecular structure. However, despite these advances, the nature of information processing and the neural network logic of representation of odorants remain elusive.

Using functional imaging and sequencing of barcoded neuron mapping strategies, we aim to relate odorant receptors to glomeruli and their odor responses in the mouse olfactory bulb as well as to the logic of downstream processing neural circuits. In mammals, odorants are sensed by the olfactory sensory neurons (OSNs) of the olfactory epithelium. Each OSN expresses one member of the odorant receptor (OR) gene family, ∼1,100 in mouse. OSNs expressing the same OR gene converge to discrete spherical structures on the surface of the olfactory bulb (OB), called glomeruli. The spatiotemporal pattern of glomerular activation is thought to represent information about odorants. A comprehensive convergence map of OSNs expressing different ORs could help discover the algorithms that relate the odor space to the underlying neural circuits and to olfactory perception. Detection of OR gene expression in situ has been successfully demonstrated recently. Although these studies provided impactful lookup tables for the location of specific subsets' ORs on the bulb surface, several technical constraints in spatial resolution or throughput have limited our ability to accurately associate a glomerulus with one or several ORs.

Jointly with the Koulakov and Zador groups at CSHL and the Znamenskiy group at the Crick Institute, we used a high-throughput in situ sequencing method (BARseq) to detect a large fraction of ORs at subglomerular resolution in 20-μm-thick olfactory bulb slices. We designed in situ sequencing padlocks (five to 10 per OR to increase detection) aimed to specifically target all OR transcripts (Fig. 2). We performed BARseq on three brains and detected ∼400 ORs that form tight spatial clusters within the glomerular layer of individual bulb slices. The number of OR transcripts detected in situ was significantly correlated with OR expression levels in bulk olfactory bulb RNA-seq data. In addition, using sequencing of barcoded neurons, we found that mitral cells in the olfactory bulb project in a graded manner to different locations along the olfactory (piriform) cortex anterior–posterior (A-P) axis, and further these piriform cortex loci project to extra-piriform brain regions that the same mitral cells send collaterals to, completing triadic circuit motifs (Fig. 3). As such, the olfactory cortex architecture is structured and thus need not rely on algorithms that assume random connectivity. We suggest an alternative model in which, akin to other senses, odor stimuli are processed along parallel, spatially segregated, and functionally distinct streams that process the perception of odor identity, valence, and actions associated with olfactory spatial navigation (Fig. 3). Our findings reveal a novel organizing wiring principle that enables parallel computations and further cross-referencing, because olfactory information reaches a given target brain region via matched direct and indirect pathways.

*Significance:* By further combining barcoded rabies tracing of inputs, in situ sequencing of viral barcodes, and in vivo calcium imaging of neuronal activity, we will investigate the fine-scale organization of synaptic connectivity in the mouse olfactory system with unprecedented throughput. We aim to map the brain-wide inputs onto individual cells in higher-order olfactory cortical and subcortical processing centers. We will further determine the identity of parent input glomeruli and corresponding ORs of bulb neurons converging on single downstream neurons. Thus, we will bridge the gap of scales of circuit description by yielding long-range brain connectomes of thousands of individual neurons in the same animal. These connection patterns will be linked to functional responses to large odor panels (∼500) and compared between animals. Our strategy has the potential to overturn the leading model of random connectivity in olfaction, and reveal the neural algorithms underlying olfactory processing (Chen et al., *Cell* **185:** 4117 [2022]).

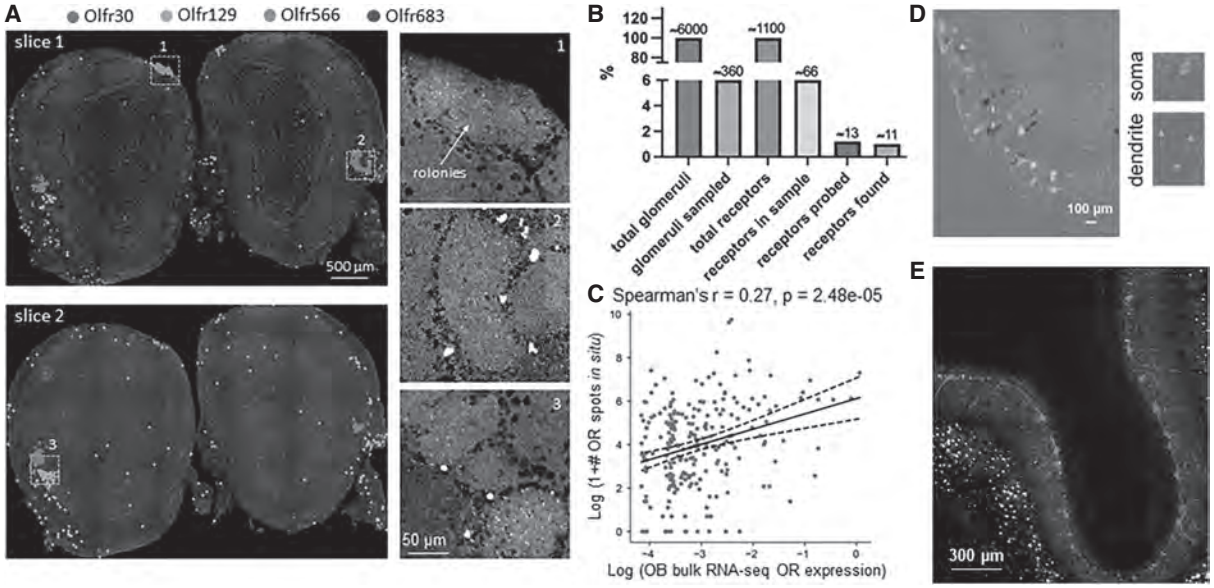



**Figure 2.** Detection of odorant receptor (OR) mRNA via in situ sequencing. (*A*) Here, we overlay shade-coded dots to show signal from 4 ORs, each detected at the same coordinates across two consecutive slices. *Insets 1*, *2*, and *3* show the raw signal of transcripts corresponding to a particular OR, confined to glomerular borders. (*B*) Across the 12 slices processed, we estimated ~360 glomeruli (~6% of the total 3,000/bulb hemisphere). Hence, we reasoned that we should have the same fraction of receptors (~66 out of ~1,100). Having designed padlock probes for ~20% of the ORs, we expected to find 13 in our slices, and, based on a conservative count, found 11 within glomeruli. (*C*) The number of rolling circle colonies (rolonies) detected in situ correlates with ORs' bulk gene expression. (*D*) In situ sequencing image of OB infected with barcoded Sindbis virus. Shades indicate matching sequences in the glomeruli, dendrites, and mitral cell bodies. (*E*) Example OB with mitral cells (gray) infected with rabies virus injected in the piriform cortex (PCx).

## A Glomerular Hierarchy for Olfactory Discriminations Aids the Search for Perceptually Similar Stimuli

Can we predict the perceptual similarity of two odorants by knowing which ORs they activate? This seemingly simple question remains unsolved as difficulties in controlling stimuli at the level of receptor types preclude disentangling the contribution of individual odorant receptors in shaping olfactory perception. To overcome this limitation, we exploited the anatomical clustering of ORs to individual glomeruli and identified them in transgenic mice using multiphoton and widefield imaging. After determining their responses to 123 monomolecular odorants, we created synthetic olfactory stimuli by optogenetically activating combinations of glomeruli with subglomerular resolution. To determine the perceptual distances between glomerular sets, we asked mice to report differences in stimulus identity and quantified the contribution of each glomerulus in shaping stimulus perception. Our psychophysical model revealed a striking glomerular perceptual hierarchy: Some glomeruli were up to six times more potent than others in shaping a reference percept (Fig. 4). We further investigated whether this hierarchy is rooted in the glomerular (ORs) odor response spectra. We found a robust correlation between the perceptual weight of each glomerulus and the average similarity of its odor responses to those of other glomeruli in the pattern.

Alternatively stated, the more a glomerulus odor response spectrum resembles those of other glomeruli in the activity pattern, the lower its perceptual importance. Thus, the perceptual relevance of a glomerulus appears to be an emerging property of the pattern of co-activated glomeruli rather than solely an intrinsic feature of its OR identity. Furthermore, using an unsupervised method for identifying latent factors in glomerular odor response patterns, we generated testable behavioral response predictions for arbitrary patterns of optically addressable glomeruli. We are systematically designing and testing glomerular activity

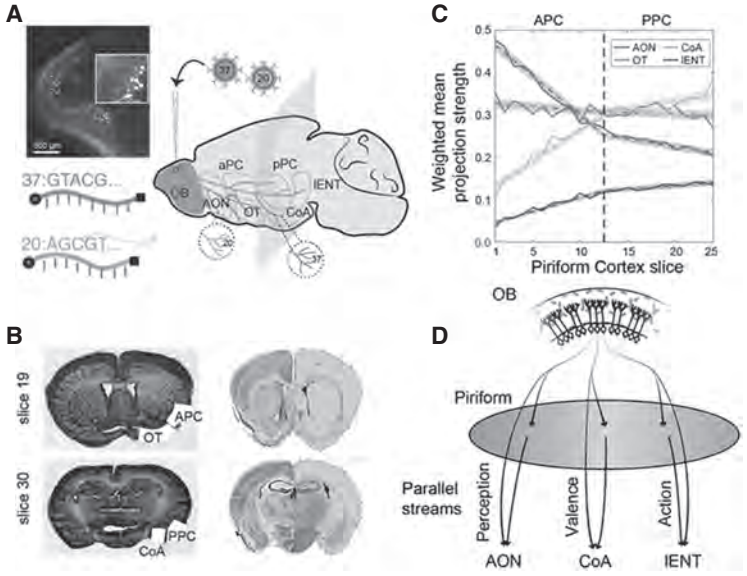



**Figure 3.** MAPseq and BARseq projection mapping of 5,000 individual olfactory bulb (OB) and 30,000 piriform neurons discovers structure in olfactory connectivity. (*A*) Schematics of the MAPseq strategy, which uses RNA barcodes to label neurons and map their brain-wide projections (*inset*: infection of MTCs by Sindbis virus carrying the barcodes and enhanced green fluorescent protein [EGFP]). (*B*) Laser capture microdissection of target brain regions from Nissl-stained coronal sections registered to the Allen Brain reference atlas. (*C*) Weighted mean projection strength for OB neurons to four major extra-piriform targets as a function of location of PCx co-innervation (the conditional probability of co-innervation P(target|PCx location), solid lines). Dashed lines/shaded areas show piecewise linear fits in aPCx and pPCx with the 95% confidence interval. (*D*) Schematics of parallel olfactory processing streams structured along the A-P axis of the PCx according to triadic connectivity motifs discovered by MAPseq.

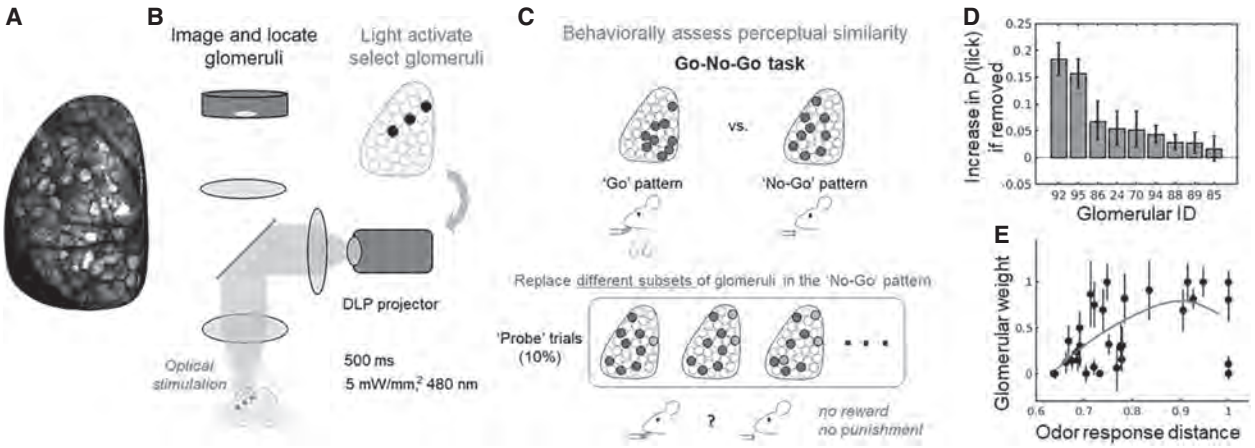



**Figure 4.** Glomerular optogenetic patterned stimulation. (*A*) Expression of opsin (ReaChr-citrine) in the olfactory sensory neurons (OSNs). Dorsal aspect of the olfactory bulb with glomeruli fluorescently labeled. (*B*) Schematics of digital light processing (DLP)-based patterned illumination rig used to trigger neuronal activity in OSNs expressing ReaChr with subglomerular spatial precision (~15 µm resolution; Avg. glomerulus diameter = ~75 µm). The same setup can be used to monitor fluorescence responses to odors of the OSNs (glomeruli) expressing a genetically encoded calcium indicator (GCaMP). (*C*) Schematics of Go-No go task and perturbations of the optogenetically induced glomerular activity patterns. (*D*) Perturbing different glomeruli in the No-go reference pattern in catch trials in expert animals results in substantially different changes in the lick probability, suggesting that different glomeruli in the pattern contribute with different weights to the emerging percept. (*E*) Relationship between glomerular perceptual weight (*D*) and the average distance (one correlation) in odor response space between the odor response spectrum of the perturbed glomerulus and the odor response spectra of other glomeruli in the pattern.

patterns of choice targeted at triggering specific perceptual differences from the reference pattern.

*Significance:* By combining behavioral analysis and circuit dissection, we aim to bridge the gaps between the biophysical features of OR activation, the structure of the perceptual space, and the underlying neural circuits in a search for the olfactory analog of the red–green–blue basis set of odorant perception.

## Corticobulbar Feedback Supports Behavioral Flexibility during Rule Reversal

Successful goal-directed behaviors rely on the ability to modify actions in accordance with changes in evolving environments. Although feedback from higher brain regions reformats visual and auditory sensory representations to enable differential evaluation of same sensory inputs depending on behavioral needs, whether this is the case for olfactory processing remains unknown. To determine whether top-down corticobulbar feedback from the piriform cortex supports behavioral flexibility, we used multiphoton imaging in mice engaged in a multimodal rule-reversal task guided by olfactory and auditory cues. Both odor and, surprisingly, the sound cues triggered corticobulbar feedback bouton responses that preceded the behavioral reporting and mirrored the reversals in stimulus reward contingency applied repeatedly within a session.

Within seconds of rule reversal, we observed changes in the identity, polarity, and kinetics of corticobulbar feedback bouton responses to the same sensory stimuli. These were correlated with changes in the behavioral performance across multiple reversals within the same session (Fig. 5). In addition, optogenetic perturbation of the cortical feedback activity locally within the olfactory bulb disrupted the behavioral performance. Thus, the corticobulbar feedback multiplexes information about stimulus identity, contingency, and behavioral outcome in a sensory modality–independent manner and is rapidly reformatted according to changes in behavioral contingencies.

*Significance:* Our work enables understanding the interplay between feedforward and cortical feedback underlying cognitive flexibility.

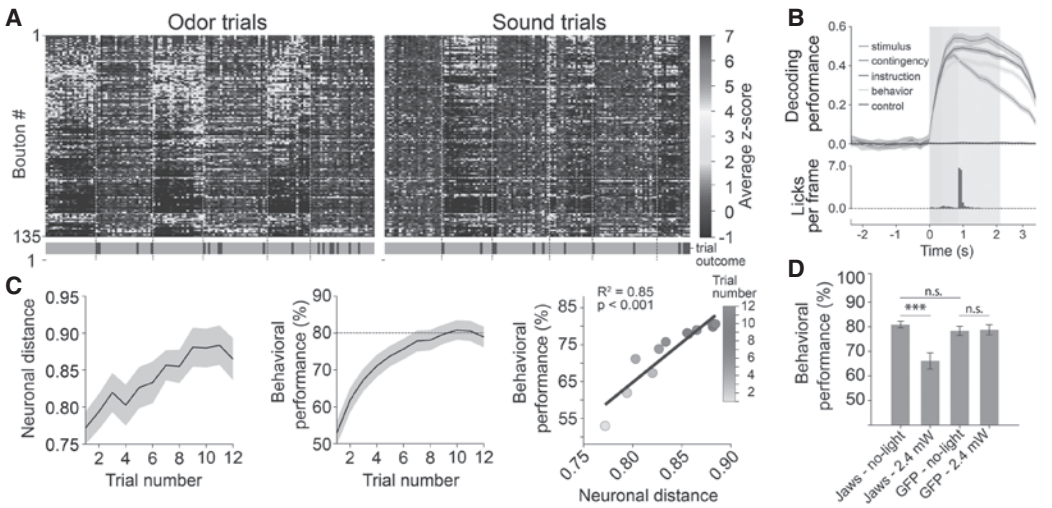



**Figure 5.** Cortical feedback supports behavioral flexibility during rule reversal. (*A*) Fluorescence maps: Average *z*-score values for all boutons from an example session. Odor trials (*left*) are parsed from sound trials (*right*). Bouton order is kept the same. Each row represents one bouton average response across trials within each block (Odor-Go; Sound-Go, Odor-Go, etc.). Correct trials (hits and correct rejections) are shown in light gray; incorrect trials (misses and false alarms) in dark gray. (*B*) Cross-validated classifiers (multilayer perceptrons) successfully decoded different task features: stimulus identity (odor vs. sound), instruction (Go/No-Go), behavioral outcome (lick/no lick), and trial contingency (hits, correct rejections, false alarms). (*C*) (*Left*) Distance between the bouton activity ensemble trajectory of the preceding block and the ensemble trajectory of each trial of the current block to the same stimulus. (*Center*) Behavioral performance after rule reversal. (*Right*) Behavioral performance versus neural distance. (*D*) Effect of bulbar optogenetic suppression of cortical feedback on behavioral performance.

## Ongoing CSHL Collaborations

*Koulakov and Zador laboratories:* Using high-throughput DNA barcoding (MAPseq, BARseq) to map odorant receptors to glomeruli and determine the logic of long-range projections and synaptic connectivity of olfactory circuits.

*Huang, Cowley, and Koulakov laboratories:* Identifying neural circuits underlying the computation of sensorimotor (olfactomotor and visuomotor) prediction error signals during closed-loop behaviors.

*Banerjee and Navlakha laboratories:* Divisive normalization, a key computation in sensory processing.

### PUBLICATION

Hernández Trejo D, Ciuparu A, Garcia da Silva P, Velasquez MC, Rebouillat Gross M, Davis MB, Chae H, Muresan RC, Albeanu DF. 2023. Fast updating feedback from piriform cortex to the olfactory bulb relays multimodal reward contingency signals during rule reversal. bioRxiv doi:10.1101/2023.09.12.557267

# FUNCTION AND EVOLUTION OF NEURAL CIRCUITS FOR VOCAL COMMUNICATION

**A. Banerjee** M. Davis E. Isko X. Zheng
C. Harpole Y. Thapa

The primary goal of the Banerjee laboratory is to investigate brain-wide neural circuits for vocal communication. We use the rich vocal behavior of the Alston's singing mice to pursue two complementary questions. First, how does the auditory system interact with the motor system to generate the fast sensorimotor loop required for vocal communication (Fig. 1A)? Second, what are the neural circuit modifications that allow behavioral novelty to emerge during evolution (Fig. 1B)? Using a comparative approach, we investigate brain-wide connectivity and neural circuitry differences between the Alston's singing mouse and other rodent species (e.g., the lab mouse) that have intermediate degrees of vocal behavior. These two lines of work combine cutting-edge neural circuit analysis of a natural behavior with comparative evolutionary analyses across species to gain insight into the function and evolution of neural circuits for vocal communication. Here are the highlights of the main scientific projects we are currently pursuing in our laboratory.

## Fast and Flexible Switching between Different Vocal Modes

C. Harpole, X. Zheng, M. Davis

Adaptive social behaviors require considerable flexibility. During conversations, we constantly switch vocal modes (e.g., laughing vs. talking) at short timescales. However, the neural mechanisms underlying such vocal flexibility remain unknown, especially in mammals. We study the Costa Rican singing mouse (*Scotinomys teguina*) to answer this question. Singing mice produce long, loud, human-audible vocalizations

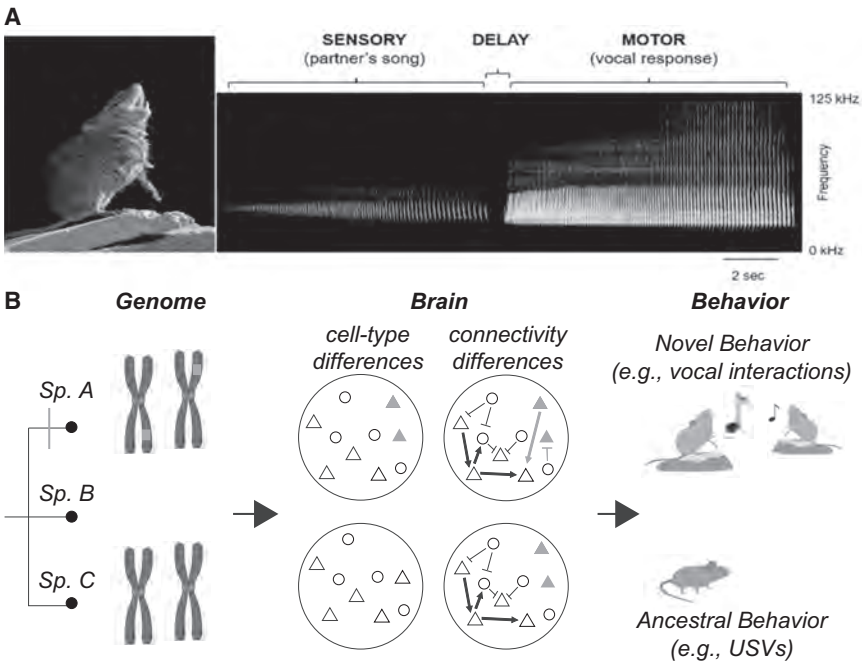



**Figure 1.** (*A*) Vocal turn-taking behavior in the singing mice—a novel mammalian model system to study neural circuits for vocal communication. (*B*) Understanding genetic, cellular, and circuit changes that lead to behavioral divergences among closely related species.

that last many seconds (songs) for long-distance communication. To understand how singing mice vocalize during social interactions in close proximity, we have developed a novel behavioral assay in which two animals interact across a perforated divider. Singing mice robustly vocalize to one another across the divider. The hardware's partial acoustic isolation allows most vocalizations to be correctly assigned. We found that singing mice also make a second mode of social affiliative vocalizations in addition to their songs: one that is quiet, higher pitched, and less stereotyped relative to songs. This quiet mode has several behavioral similarities to those common in a large majority of murine rodents, including laboratory mice: They are ultrasonic vocalizations (USVs) produced in close proximity (a few body lengths) of a conspecific and are rarely emitted when alone. Moreover, these mice are able to switch between these two vocal modes flexibly by changing instantaneous note rate and loudness. Because these two vocal modes differ in acoustic properties as well as in behavioral usage, we wondered whether the peripheral production mechanisms differ between USVs and song. By measuring the mouse's breathing using an intranasal thermistor, we found that respiratory coupling of vocalizations during USVs or songs were identical. Moreover, we found that both

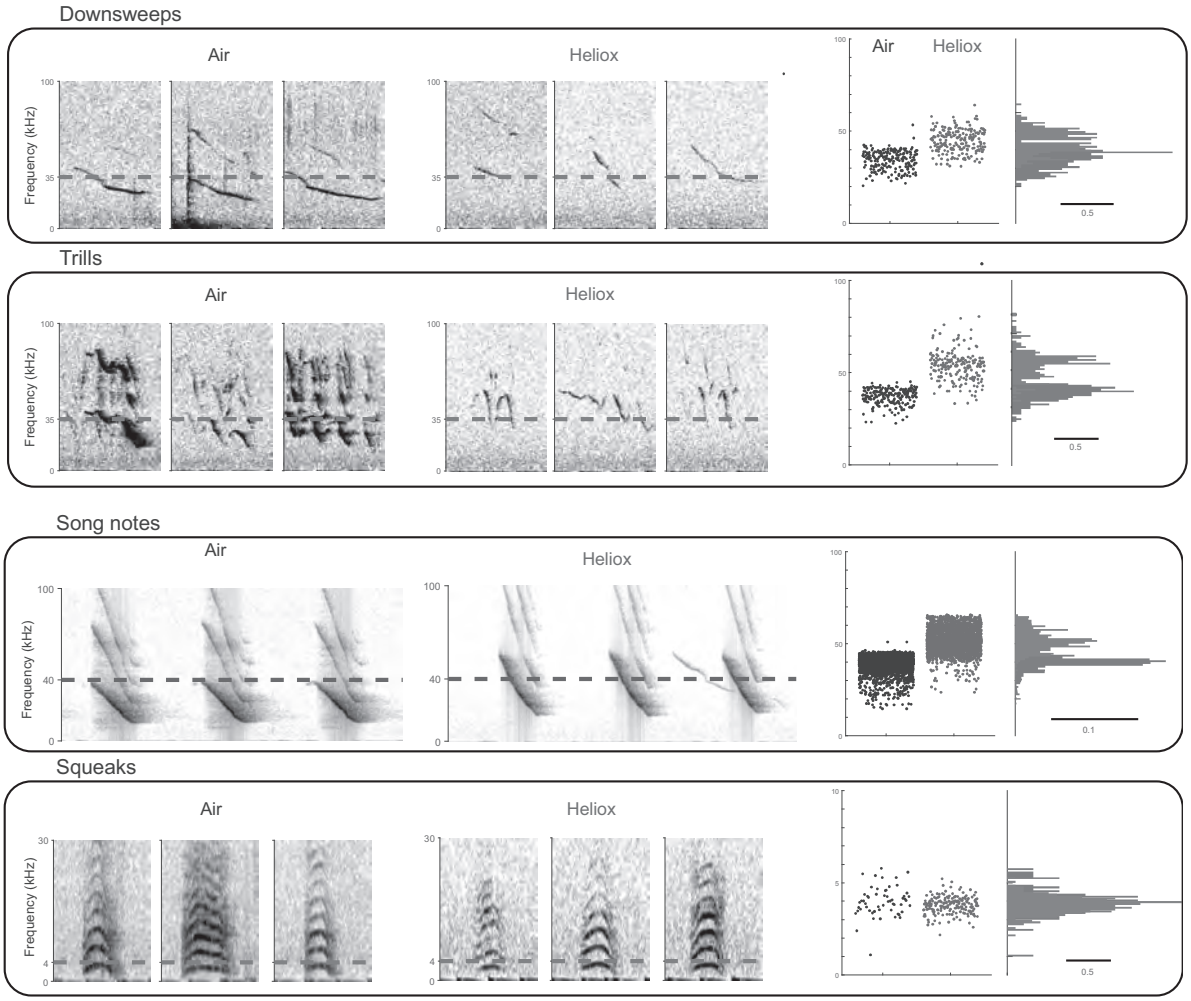



**Figure 2.** Shared peripheral mechanism of vocal production between the ancestral behavioral mode: ultrasonic vocalizations (USVs) and the behavioral novelty (songs). Reducing the density of ambient air using helium gas raises the pitch of both USV notes and song notes.

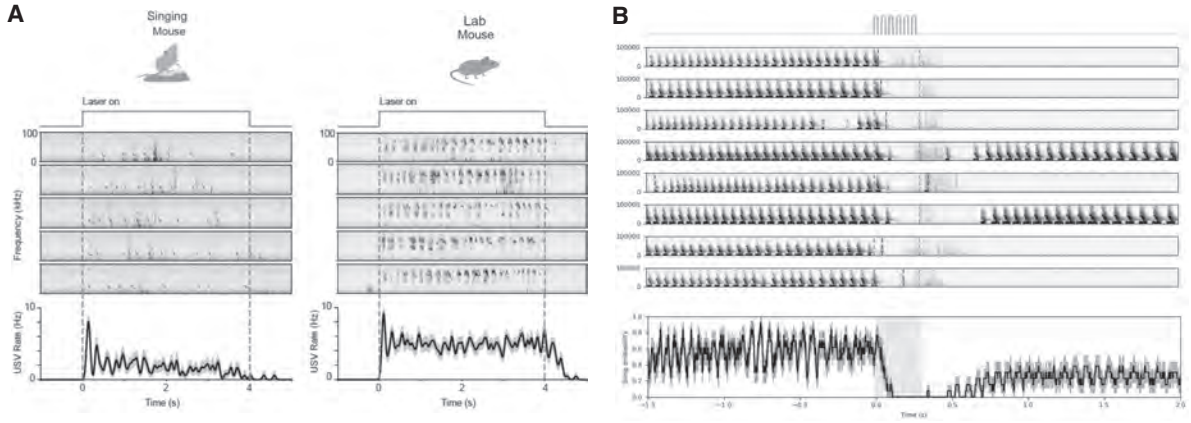



**Figure 3.** Periaqueductal gray (PAG) neurons promote ultrasonic vocalizations (USVs) while stopping/pausing songs. (*A*) Optogenetic activation of ChR2-expressing excitatory neurons in PAG with blue light generates USV bouts in both lab mice and singing mice. (*B*) Stimulating the same neurons stops/pauses ongoing songs.

USVs and songs are produced peripherally using an aerodynamic whistle mechanism by reducing the density of air using a helium–oxygen gas mixture (Fig. 2). These peripheral measurements, combined with the behavioral analyses, suggest that the mechanisms that distinguish USV and song production are centrally controlled by the brain. We are currently testing the sensory cues that elicit these two vocal modes by systematic perturbation of visual, auditory, and olfactory streams of sensory inputs.

## Neural Circuits for Mammalian Vocalization

X. Zheng, C. Harpole, M. Davis

In this project, we aim to understand how mammalian brains control diverse vocalizations in dynamic environments. This project focuses on the neural circuits for context-dependent vocalizations in singing mice. Xiaoyue (Mike) Zheng and postdoc Cliff Harpole developed a behavioral assay in which two mice interact across a perforated divider. They are able to assign a vast majority of vocalizations to their source animals. They found that vocalizations are organized into two distinct modes—calls and songs. Songs are comprised of a series of progressively longer notes that evolve predictably over six to 10 seconds. In contrast, calls are much less stereotyped and much quieter. The usage of calls suggests they may be homologous to the social USVs in other rodents and may share neural control by the periaqueductal gray (PAG). Using excitatory optogenetics and other viral tools, Mike found that the PAG is both necessary and sufficient for producing calls (Fig. 3). He then tested whether the PAG also plays any role in songs. Using similar techniques, he found that the PAG is also critical for song production, and particularly song patterning. Based on these experiments, we posit that singing mouse calls are akin to USVs and represent an ancestral behavioral mode shared across many rodents, whereas the songs represent a behavioral novelty that may be an elaboration of the ancestral mode. The PAG is critical for both modes, and it may serve as a locus for evolutionary diversification. We are gearing up to perform electrophysiological recordings in the PAG to understand the neural mechanisms underlying vocal mode switching in the singing mice.

## Comparative Connectomics and Transcriptomics in Lab Mice and Singing Mice

E. Isko [in collaboration with the Zador laboratory, CSHL]

A fundamental goal of neuroscience is to understand how neural circuits evolve to enable novel behavioral phenotypes. Because behaviors do not fossilize, our strategy is to identify neural circuit modifications—both structural and functional—among closely related species with large behavioral divergences. Despite a close evolutionary relationship, Alston's singing mice (and lab mice) exhibit divergent vocal behaviors (Fig. 4). This behavioral divergence must derive

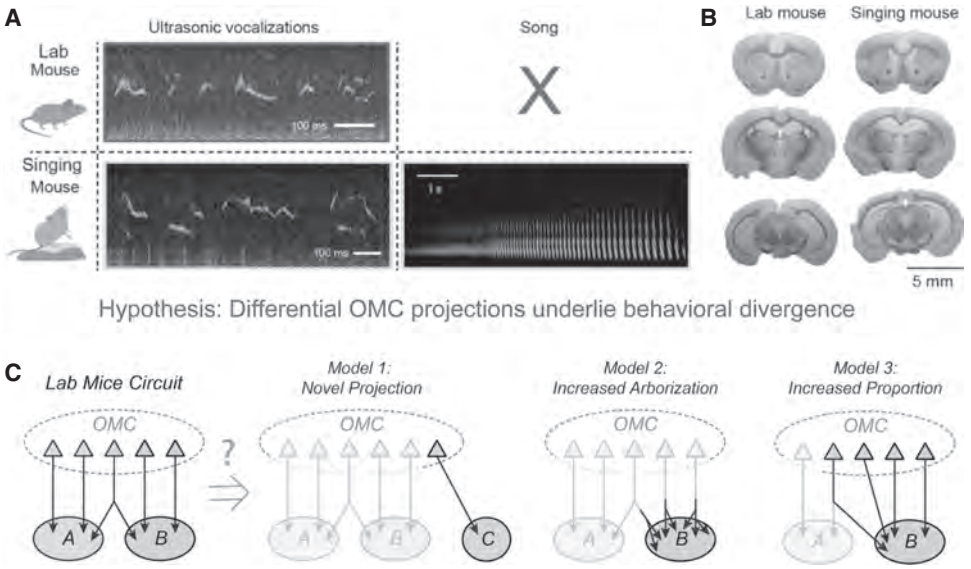



**Figure 4.** Divergent vocal behaviors between the lab mice and singing mice. (*A*) Males of both species (lab mice and singing mice) produce ultrasonic vocalizations (USVs) that last for ~200 msec. In addition to the USVs, only the singing mice produce a "song"—a series of approximately 100 notes that evolve predictably over ~10 sec. Ordinate represents frequency from 0 to 100 kHz. (*B*) Representative brain slices from a lab mouse (*left*) and a singing mouse (*right*). (*C*) Alternative hypotheses of how orofacial motor cortex (OMC) projection patterns might differ between the lab mouse and the singing mouse.

from differences in the underlying biology of neural cell types and/or neural circuits. We are collaborating with the Zador laboratory to determine differences in the neural circuits that could explain the differences in vocal behavior of lab and singing mice. We have begun our connectomic comparison in the orofacial motor cortex (OMC), a brain area involved in the singing mouse song as proven through electrical, pharmacological, and cooling experiments (Okobi et al., *Science* **363:** 983 [2019]). We are using both bulk methods (viral tracing) and single-cell methods (MAPseq, developed by the Zador laboratory) to characterize the projection patterns of neurons in the OMC of the two species (Fig. 4). To test whether there are qualitative (e.g., absent or novel) differences in projection patterns, we used viral tracing of neurons from the OMC. We identified bulk projection targets using serial two-photon tomography (STPT). We found that the OMC in each species projects to identical downstream brain areas including the contralateral cortex, striatum, thalamus, superior colliculus, PAG, and others. Having found no bulk differences between species, we next used MAPseq, a high-throughput barcoding technique, to characterize projection patterns at single-cell resolution. Compared with lab mice, we found that a larger proportion of OMC neurons in the singing mice project to the midbrain PAG and a temporal cortical area (singing mouse: $n = 5{,}114$ neurons, seven animals; lab mouse: $n = 71{,}704$ neurons, five animals). No other target region (total of 11) showed significant species-specific differences. This increased projection strength from OMC was driven by neurons with direct projections to temporal cortex showing no/few collaterals. In summary, we found evidence for expansion of existing vocal motor circuits in the singing mice compared with lab mice, which may explain their species-typical vocal behaviors. Ongoing experiments will determine the functional roles of these projections using neural circuit perturbations such as chemogenetics and optogenetics. Our future directions include using single-nucleus RNA sequencing (snRNA-seq) and BARseq2 (an in situ sequencing technique that combines gene expression and projection data, developed by the Zador laboratory) to determine differences in cell type and the spatial location of these cells in the OMC of lab and singing mice. Finding transcriptomic and neural circuitry differences between lab and singing mice will give insights into the neural substrates of vocal behavior.

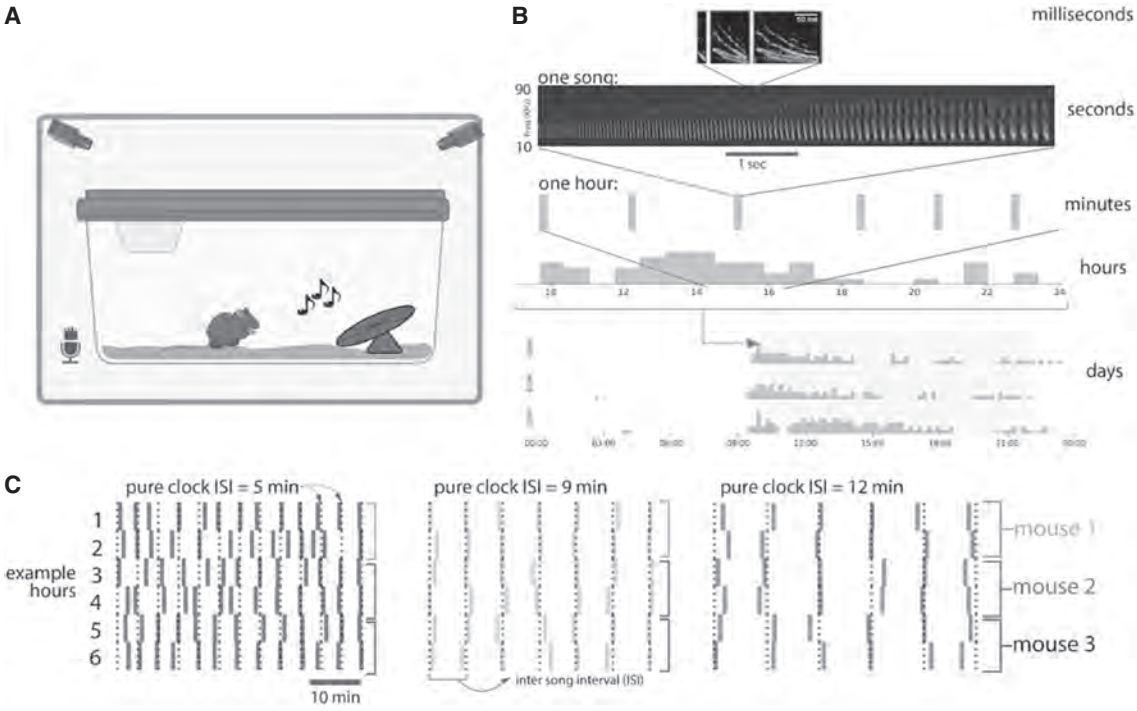



**Figure 5.** Alston's singing mice produce songs at clock-like intervals. (*A*) Experimental setup. We record song production for long periods of time (days) for, in this case, an isolated mouse. (*B*) Example vocalization spanning several timescales (i.e., milliseconds to days). First row, example notes in milliseconds; second row, example songs (each ~7 sec); third row, song histogram over a day (each row). (*C*) One hour bin with time points adhering to a clock (dashed lines) and real song times (color ticks). Clock-like song behavior was observed for short (*left*), medium (*center*), and long (*right*) intersong intervals (ISIs).

## Computational and Neural Insights into the Temporal Patterning of Songs in Singing Mice

Y. Thapa [in collaboration with the Cowley laboratory, CSHL]

Communication is the backbone of any social interaction, and a primary way animals communicate is through vocalization. From the chirps of songbirds to conversational human speech, vocalization requires temporally precise regulation. Here, we study a unique form of temporally regulated vocalization: Alston's singing mice produce dozens of human-audible songs throughout the day. A singing mouse modulates its song length and rate during social interaction—a hallmark of communication—but in isolation will also produce songs throughout the day. The temporal patterning of songs in isolation, how it changes during social interactions, and the neural mechanisms of this behavior remain unclear. We hypothesize that one of the ways the singing mice pattern their songs is by spacing them out equally in time. Based on a month-long recording of mice in solo conditions (Fig. 5A,B), we observed hours of evenly spaced songs resembling a noisy clock (Fig. 5C). We propose that these behavioral dynamics can be described using a clock-like drift-diffusion model. We also aim to extend these models to capture song interactions between two mice. Finally, we propose experiments that aim to determine the neural signature of song behavior using pharmacological perturbation, functional imaging, and manipulations; our first investigation will be dopaminergic pathways. Our research will uncover the computational and neural mechanisms underpinning an internally generated rhythmic behavior and how it changes in the context of social interaction.

### PUBLICATION

*In Press*

Banerjee A, Chen F, Druckmann S, Long MA. 2024. Temporal scaling of motor cortical dynamics reveals hierarchical control of vocal production. *Nat Neurosci* **27:** 527–535. doi:10.1038/s41593-023-01556-5

# NEUROIMMUNOLOGICAL MECHANISMS OF CIRCUIT CONSTRUCTION IN BRAIN DEVELOPMENT AND DISEASE

**L. Cheadle** A. Ademola, A. Arshad, V. Bagan, D. DiMartino, A. Ferro, J.A. Kahng, C. Kang, Q. Lin, J. Park, I. Sanchez Martin, S.X. Tang, D. Vita, A.M. Xavier, C. Zhang

Human brain development is a complex, protracted process spanning from early gestation to the third decade of life. Neural circuits are first assembled in utero and then extensively refined during postnatal brain development in response to sensory experience (Wiesel and Hubel, *J Neurophysiol* **26:** 978 [1963]). Sensory-dependent (SD) refinement shapes mature neuronal connectivity by driving the strengthening and maintenance of a subset of immature synaptic connections and the elimination of synapses that fail to strengthen (Katz and Shatz, *Science* **274:** 1133 [1996]; Hooks and Chen, *Neuron* **106:** 21 [2020]). Consistent with the importance of SD refinement for healthy brain development, deficits in this process contribute to disorders of cognitive function that arise before the brain reaches maturity, such as autism, epilepsy, and schizophrenia (Feinberg, *J Psychiatr Res* **17:** 319 [1982]; LeBlanc and Fagiolini, *Neural Plast* **2011:** 921680 [2011]; Vezzani et al., *Nat Rev Neurol* 7: 31 [2011]). Intriguingly, these disorders can be triggered or aggravated by systemic inflammation, and individuals with mutations in genes encoding regulators of immunity are at significantly heightened risk of neurodevelopmental dysfunction (Shuid et al., *Int J Environ Res Public Health* **18:** 18062817 [2021]; Schizophrenia Working Group of the Psychiatric Genomics, *Nature* **511:** 421 [2014]). These findings suggest that immune cells can influence SD circuit refinement in the healthy brain, and that the obstruction or exacerbation of neuroimmune signaling likely contributes to neurodevelopmental pathology (Michel et al., *Dev Neurobiol* **72:** 1277 [2012]; Han et al., *Nat Rev Neurol* **17:** 564 [2021]). However, the molecular mechanisms through which experience sculpts developing circuits, and the ways in which immunological factors contribute to this process, remain to be clearly defined (Fig. 1).

Despite mounting evidence that the immune system exerts a powerful influence over brain development, the fundamental mechanisms through which immune cells shape neurological function represent a key gap in knowledge with wide-ranging implications for human health. Systematically dissecting the neuroimmune pathways underlying SD circuit refinement in the brain is not only essential for understanding brain development at a deep mechanistic level, it could also lead to the identification of new diagnostic biomarkers and drug targets for treating disorders that are exacerbated by inflammatory processes during periods of SD refinement. The overarching objective of our research program is to define the precise cellular and molecular mechanisms through which the nervous and immune systems interact to organize sensory-dependent features of brain development and function. We focus on two major aspects of this question in parallel: (1) How do microglia, the resident immune cells of the brain, promote the refinement of neural circuits in response to sensory experience? And (2) how do changes in immunological state in the periphery, such as prenatal systemic immune activation, impair circuit wiring in the context of disease? We generate new viral and genetic tools and apply state-of-the-art experimental approaches, ranging from in vivo multiphoton microscopy to single-cell genomics, to the mouse visual system and to human brain tissue to address these questions. The specific ongoing and future research directions that my laboratory will pursue over the next five years are outlined below.

## DIRECTION 1. CHARACTERIZE THE MECHANISMS THROUGH WHICH MICROGLIA DRIVE THE SENSORY-DEPENDENT REFINEMENT OF DEVELOPING BRAIN CIRCUITS IN VIVO

Although SD circuit refinement is an essential component of brain development, this process has been predominantly studied in fixed tissue or in acute brain slices. Thus, how synapses are dynamically strengthened, remodeled, or

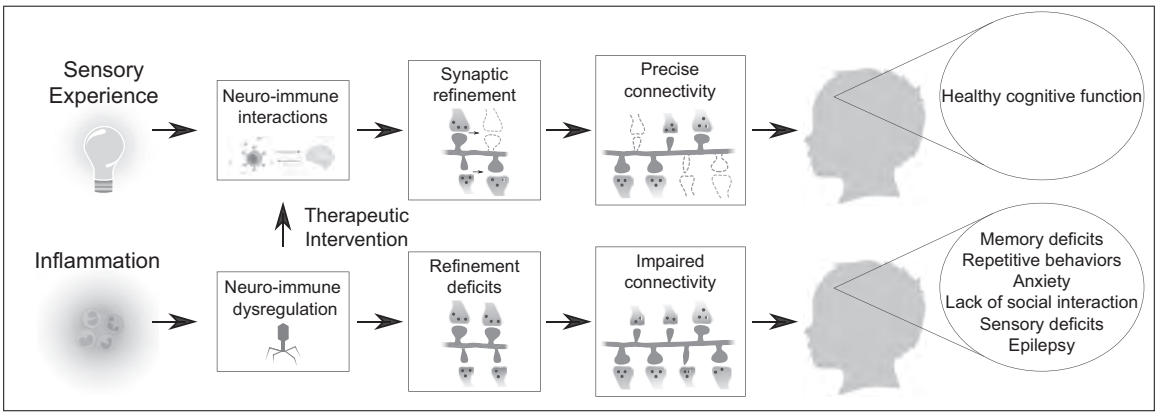



**Figure 1.** Schematic displaying the link between experience and brain wiring in the absence or presence of inflammation. Neuro-immune interactions are crucial for establishing synaptic connectivity in the healthy brain. When these interactions are disrupted (e.g., because of inflammation), circuit wiring deficits lead to neurodevelopmental disorders associated with cognitive impairments such as autism and schizophrenia. Identifying the mechanisms linking neuro-immune interactions to synaptic refinement could lead to the identification of novel therapeutic strategies for treating these conditions.

eliminated in the living brain, particularly through microglia–neuron interactions that are coordinated by sensory experience, remains to a large extent unclear. Given our discovery that microglia are engaged by experience to refine developing circuits through molecular signaling interactions with neurons, a primary goal of my research program is to characterize the intercellular signaling pathways through which microglia control SD refinement in vivo.

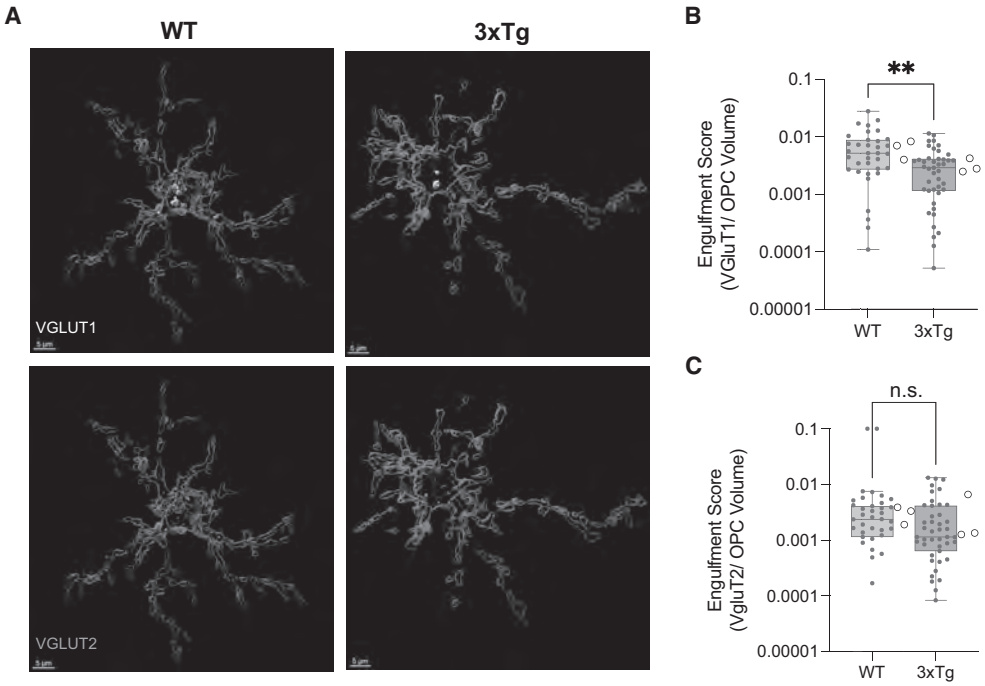



**Figure 2.** Intrahippocampal synapses are selectively spared from engulfment by oligodendrocyte precursor cells (OPCs) in a mouse model of Alzheimer's disease (AD). (*A*) Example reconstructions of OPCs containing VGLUT1$^{+}$ inputs (*top*) or VGLUT2$^{+}$ inputs (*bottom*). (*B*) Quantification of the engulfment of VGLUT1$^{+}$ synapses by OPCs in 3×TG and wild-type (wt) mice at 6 months of age. Student's *t* test, $^{**}p < 0.01$. (*C*) Quantification of the engulfment of VGLUT$^{2+}$ synapses by OPCs at 6 months. Student's *t* test, $p > 0.05$.

## Characterize the Spatiotemporal Dynamics of Microglia–Synapse Interactions in Real Time

A. Ferro

A major barrier to uncovering the mechanisms through which microglia disassemble synapses is a lack of studies characterizing the spatiotemporal dynamics of microglia–synapse interactions in real time. By developing viral and genetic tools to monitor synaptic activity, microglial motility, and physical contacts between microglia and synapses simultaneously in the visual cortex of awake mice using two-photon microscopy, we are characterizing the structural mechanisms through which microglia eliminate synapses. We also utilize electrophysiology and calcium imaging to determine the consequences of disrupting microglia–synapse interactions for circuit connectivity and function. We expect these studies to shed light on the dynamic mechanisms through which microglia interact with synapses to refine circuits in response to experience.

## Dissect the Downstream Mechanisms Through Which Microglia-Neuron Signaling Eliminates Synapses

A. Ferro, A. Arshad

Although a number of pathways are known to regulate communication between microglia and neurons, TWEAK-Fn14 signaling is unique as a mechanism linking microglia to SD refinement. Thus, the TWEAK-Fn14 pathway is appealing as a molecular handle to derive insights into how experience shapes circuits by inducing microglia–neuron cross talk. Over the past year, we discovered that Fn14 is important for shaping circuit connectivity not only during visual system development, but in the adult hippocampus as well. We showed that pyramidal (PYR) neurons in hippocampal CA1 induce *Fn14* expression in response to neuronal activity, and that mice lacking Fn14 exhibited heightened activity both at baseline and in response to epileptic seizures. Notably, Fn14 constrained neuronal activity most robustly at the daily transition between light and dark, and mice lacking Fn14 exhibited deficits in circadian rhythms, sleep–wake states, and cued and spatial memory. In mice lacking Fn14, microglia contacted and disassembled fewer synapses than in wt (wild-type) littermates. We are currently preparing a manuscript describing these findings for peer review and will continue to disentangle the mechanisms linking microglia specifically to Fn14-dependent function in the adult hippocampus over the next year.

## Define New Roles for Microglia–Oligodendrocyte Precursor Cell (OPC) Interactions in Circuit Refinement in Health and Disease

A. Ferro, S.X. Tang, C. Zhang

In the course of our studies on the roles of microglia in sensory-dependent refinement, we unexpectedly discovered a novel role for OPCs in this process. Specifically, we found that these unique yet poorly understood glial cells can engulf synapses through phagocytosis in response to sensory experience in the developing visual cortex of the mouse. Intriguingly, the ability of OPCs to engulf synapses is promoted by microglia, as depleting microglia from the brain dampened synapse engulfment by OPCs (Auguste et al., *Nat Neurosci* **25:** 1273 [2022]; Cheadle 2023; Buchanan et al. 2023). To better understand how microglia–OPC cross talk shapes developing circuits, we are applying two-photon microscopy to interrogate structural interactions between (1) microglia, (2) OPCs, and (3) synaptic inputs in awake mice in vivo. We are also applying electrophysiology to better understand direct interactions between neurons and OPCs at the functional level, with a focus on the synaptic junctions that connect these cell types. Finally, given evidence that synapse loss during aging is a major driver of dementia in neurodegenerative conditions such as Alzheimer's disease and Huntington's disease, we are testing the hypothesis that alterations in the engulfment activity of OPCs may contribute to the development of these disorders. Consistent with this possibility, we recently discovered that OPCs down-regulate their engulfment of a subset of excitatory synapses in the hippocampi of the 3×-TG mouse model of Alzheimer's disease at 6 months of age, prior to the appearance of inflammation in the brain (Fig. 2). We are following up on this exciting finding by determining how this decreased engulfment of synapses by OPCs impacts synaptic maintenance and cognitive function in the Alzheimer's disease model in vivo. Our ongoing studies will be

facilitated by robust protocols for studying synapse engulfment by OPCs that are described in a manuscript recently accepted for publication (Kahng et al. 2023).

### Screen for Novel Molecular Mediators That Promote OPC-Driven Synaptic Refinement

A.M. Xavier, Q. Lin, C. Kang, J. Kahng, A. Ferro, A. Ademola, D. Vita

The molecular mechanisms through which OPCs engulf synapses in response to sensory experience remain completely unknown. We are applying a three-pronged approach to address this gap in knowledge. First, we have identified four candidate mediators of engulfment by OPCs based on known roles for these factors in phagocytosis/circuit refinement: the complement component C1qa, the complement-like factor C1ql1, the phagocytic receptor Lrp1, and the "don't-eat-me" signal Sirp1α. We are conditionally deleting each of these factors from OPCs specifically and assessing engulfment ex vivo. Second, we are performing single-cell transcriptomics on OPCs following a sensory stimulation paradigm to uncover transcripts that are up-regulated in these cells by visual experience, an environmental cue that we previously showed to increase the engulfment of synapses by OPCs. Finally, we are applying two unbiased proteomic screening approaches to identify proteins at the interface of synapse–OPC interactions in vivo (i.e., the synapse–OPC proteome). Both approaches involve the unbiased identification of potential mediators of OPC-dependent synapse engulfment based on the selective labeling of neuronal proteins. Uncovering which neuronal proteins are present within (i.e., engulfed by) OPCs is expected to uncover the neuronal molecules that OPCs engage to remove synapses in vivo.

## DIRECTION 2. DISSECT THE MECHANISMS THROUGH WHICH SYSTEMIC INFLAMMATION TRIGGERS AND/OR EXACERBATES DISORDERS OF BRAIN DEVELOPMENT

Elucidating the contributions of microglia to the establishment of synaptic connectivity in the brain is crucial for understanding one avenue through which immune cells facilitate neurological function. However, another major unknown is how systemic inflammation beginning outside of the brain gives rise to the neurobiological deficits that underlie brain pathology. In a more recently emerging direction in our laboratory, we are moving beyond interactions between microglia and neurons to characterize the mechanisms through which the brain detects systemic inflammation and to define the molecular links between inflammation and neurodevelopmental dysfunction. The overarching goal of this direction is to establish an inventory of inflammatory signals that mediate communication between peripheral immune cells and the brain and to determine how these signals impinge upon neurons and nonneuronal cells to disrupt SD phases of circuit refinement.

### Interrogate Neuro-Immune Interactions in Mouse Models of Systemic Inflammation

I. Sanchez Martin, D. DiMartino

One of the earliest observations to suggest that impairments in neuro-immune communication contribute to neurodevelopmental disorders was the finding that maternal immune activation (MIA) during pregnancy (e.g., viral infection of a pregnant mother) leads to a significantly heightened risk of autism and/or schizophrenia in the offspring (Shuid et al., *Int J Environ Res Public Health* **18:** 18062187 [2021]). This phenomenon can be modeled in mice by exposing a pregnant dam to the viral mimetic poly(I:C), which activates the maternal immune system and triggers behavioral, synaptic, and molecular changes in the brains of the offspring that are reminiscent of core features of neurodevelopmental disorders (Malkova et al., *Brain Behav Immun* **26:** 607 [2012]; Solek et al., *Dev Dyn* **247:** 588 [2018]). Utilizing MIA as a model to understand the contributions of inflammation to impairments in brain development, we unexpectedly discovered that, mere hours after injecting the pregnant dam with poly(I:C), large-scale teratogenic effects appear in a subset of the fetal litter. Remarkably, only a subset of male embryos was impacted in this way. We are currently following up to better understand why all females and some males are resistant to this phenotype, whereas others are susceptible. We have homed in on the placenta as a potential locus of this phenomenon, identifying key transcriptional changes in the placentas of MIA-affected males compared to healthy males. In ongoing studies, we are testing the roles of macrophages (including microglia) in mediating the effects of inflammation on fetal brain development in vivo.

## Characterize Cell Type–Specific Transcriptional and Synaptic Changes in the Epileptic Brain

C. Kang, A.M. Xavier, Q. Lin

Although the majority of the work described so far has taken advantage of the mouse as a tractable model for studying circuit development, it is a priority for us to identify mechanisms that are relevant to human disease. Thus, in addition to studying mice, we are expanding our focus to encompass studies of the human brain. In particular, we aim to derive insights into the transcriptomic and circuit deficits that appear in brain samples from patients with the inflammatory pathological disorder epilepsy. Epilepsy is the most common neurological condition, affecting as many as 1 in 26 people at some point during their life span (Hesdorffer et al., *Neurology* **76:** 23 [2011]). Furthermore, ~30% of people with autism also have epilepsy, suggesting that these disorders may share fundamental disease mechanisms (Clarke et al., *Epilepsia* **46:** 1970 [2005]). Because a hallmark of epilepsy is inflammation, we reason that identifying the changes that occur in epileptiform tissue compared to relatively healthy brain tissue from the same subjects may lead to the identification of new neuro-immune signaling mechanisms underlying neurological disease. In a collaboration with neurosurgeons at one of the largest healthcare networks in the state of New York, Northwell Health, we obtain freshly resected human cortical tissue from epileptic patients undergoing surgery. We have collected samples from five patients to date. We expect that applying single-cell transcriptomics and histological analysis to these samples will allow us to identify molecular pathways underlying epilepsy, thereby shedding light on the mechanistic basis of neurodevelopmental dysfunction.

### PUBLICATIONS

Buchanan J, da Costa NM, Cheadle L. 2023. Emerging roles of oligodendrocyte precursor cells in neural circuit development and remodeling. *Trends Neurosci* **46:** 628–639. doi:10.1016/j.tins.2023.05.007

Cheadle L. 2023. Neuron–glia communication through bona fide synapses. *Nat Rev Neurosci* **24:** 395. doi:10.1038/s41583-023-00702-z

Kahng JA, Xavier AM, Ferro A, Auguste YSS, Cheadle L. 2023. Integrated high-confidence and high-throughput approaches for quantifying synapse engulfment by oligodendrocyte precursor cells. bioRxiv doi:10.1101/2023.08.24.554663

# UNDERSTANDING THE BRAIN'S STEP-BY-STEP COMPUTATIONS

**B. Cowley** R. Fayyazi R. Göndür X. Lei Y. Li J. Scott J. Skaza Y. Thapa P. Vafidis F. Vercuysse

The main goal of our research group is to understand the step-by-step computations of the brain. For example, how does a visual scene captured by the retina transform into neural signals in the brain forming our perception? Our aim is to identify the mathematical operations of these transformations by building machine-learning models that predict neural responses and behavior from input stimuli such as images.

We take the following two-stage approach. First, we train deep neural network (DNN) models to accurately predict neural responses or behavior. Because recording time is limited—meaning we have a limited amount of data to train the model—we develop closed-loop methods that adaptively choose stimuli to train DNN models with as little training data as possible. The idea is that if, for example, a visual neuron prefers images of bananas, we should not waste time showing images of tomatoes. The resulting trained DNN model is highly predictive but often large, made up of tens of millions of parameters.

Second, after obtaining this highly predictive but large DNN model, we employ machine-learning techniques to prune and distill the model in order to identify a "compact" model with the least number of parameters that still retains good prediction. We find that these compact models are often on the order of 1,000 times smaller than the original, large deep model. We then peer inside the compact models to explain their inner workings, which helps to shed light on the computations of real neurons. Overall, our approach of designing closed-loop methods and identifying compact models will yield the next generation of highly predictive, explainable models for computational neuroscience.

We apply this approach in different ways for a number of model organisms and systems: macaque monkey visual system, fruit fly visual system, and Alton's singing mice.

## Compact Deep Neural Network Models of Visual Cortex

B. Cowley [in collaboration with P. Stan and M. Smith, Carnegie Mellon University; J. Pillow, Princeton University]

DNN models are large, highly expressive models that, with enough data, can fit almost any mapping. For example, DNNs can perform object recognition—after being trained on the ImageNet data set of one million image-object pairs (Krizhevsky et al., *Comm ACM* **60:** 84 [2017]). This motivates using DNNs to fit desired mappings in neuroscience, such as predicting the response of a higher-order visual cortical neuron to a natural image. We focus on predicting responses in macaque visual area V4 whose neurons are known to prefer edges, curves, textures, objects, and colors. Over many recording sessions, we train a deep ensemble model to predict V4 responses from any input image (Fig. 1A). The reason we use an ensemble model is for active learning: For the next recording session, we choose images that maximize the disagreement across the ensemble (i.e., take images with large variance of predicted responses across the ensemble). In this way, we choose images for which the model is most uncertain. We deployed this closed-loop framework in a real experiment and found that using images chosen by active learning substantially increased prediction performance versus that of randomly chosen natural images (Fig. 1B).

In modeling, we face the trade-off between a highly accurate yet complex model and an interpretable yet less accurate model. Our deep ensemble model (Fig. 1A) achieves high accuracy but comprises approximately 60 million parameters in total—too large to understand the inner workings of its approximately 150,000 filters. We asked a simple question: How small of a model can we obtain that still yields high prediction performance? We used the machine-learning techniques of knowledge distillation (Hinton

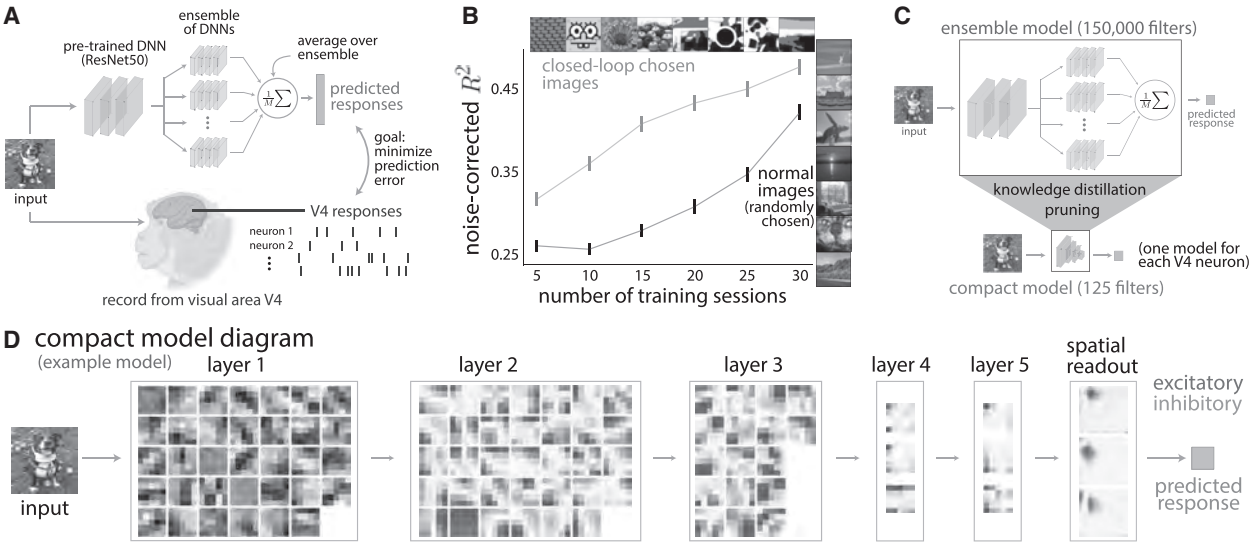



**Figure 1.** Closed-loop method to efficiently train a deep neural network (DNN) model. (*A*) We model the mapping between natural images and visual cortical responses from macaque V4 with a deep ensemble model. (*B*) Training our ensemble model with images chosen in a closed-loop manner between recording sessions (closed-loop-chosen images) leads to better prediction on held-out images versus training with randomly chosen images (normal images). (*C*) Our framework to condense our ensemble model into a compact model. We found that a compact model is typically ~1,000 times smaller than the ensemble model yet only slightly less predictive. (*D*) Diagram of an example identified compact model. Each heatmap depicts the weights for each convolutional filter; all filters are shown.

et al., arXiv 1503 [2015]) and pruning (He et al., *IEEE Computer Vision* 1389 [2017]) to obtain a "compact" model of V4 responses (Fig. 1C). These compact models were ~1,000 times smaller than the deep ensemble model (approximately 125 filters on average) yet only slightly less predictive of held-out V4 responses. In fact, a compact model is small enough to display all of their filter parameters in one diagram (Fig. 1D). We have examined the inner workings of these compact models from which we now have testable predictions about how V4 neurons form their selectivity, such as dot detection. Our work suggests that the DNN models we use in computational neuroscience are needlessly large. We have developed a modeling framework in which we sacrifice a small amount of prediction to obtain substantially smaller models.

## Mapping Model Units to Visual Neurons Reveals Population Code for Social Behavior

B. Cowley [in collaboration with A. Calhoun, N. Rangarajan, E. Ireland, J. Pillow, and M. Murthy, Princeton University; M. Turner, Stanford University]

In systems neuroscience, we are not limited to simply observing neural responses and behavior—we have a large set of genetic tools to perturb the brain in order to identify causal links between neurons and behavior. For example, we can genetically silence different visual projection neuron types in the fruit fly and observe how its behavior changes (Wu et al., *eLife*, 2016). However, there are many possible neuron types to silence, and some behavioral changes may be subtle. To address this, we built a deep neural network to model the visuomotor transformation of a fruit fly (Fig. 2A). To mimic the "visual bottleneck" of the fruit fly's visual system—comprising approximately 45 optic glomeruli, where a single neuron type innervates one glomerulus—we placed a bottleneck in our model comprising the same number of variables (or "channels") as that of the neuron types we silenced. To account for the behavior of silenced flies, we developed knockout training that "knocks out" or sets to 0 the output of the model's bottleneck variable that corresponds to a particular neuron type (Fig. 2B). In other words, we perturb the model in a manner similar to how we perturbed the fruit fly. This model, once trained, goes beyond simply predicting neural responses and behavior; it fits a one-to-one mapping between its internal bottleneck variables and real neurons. Our model is the first large-scale hypothesis of

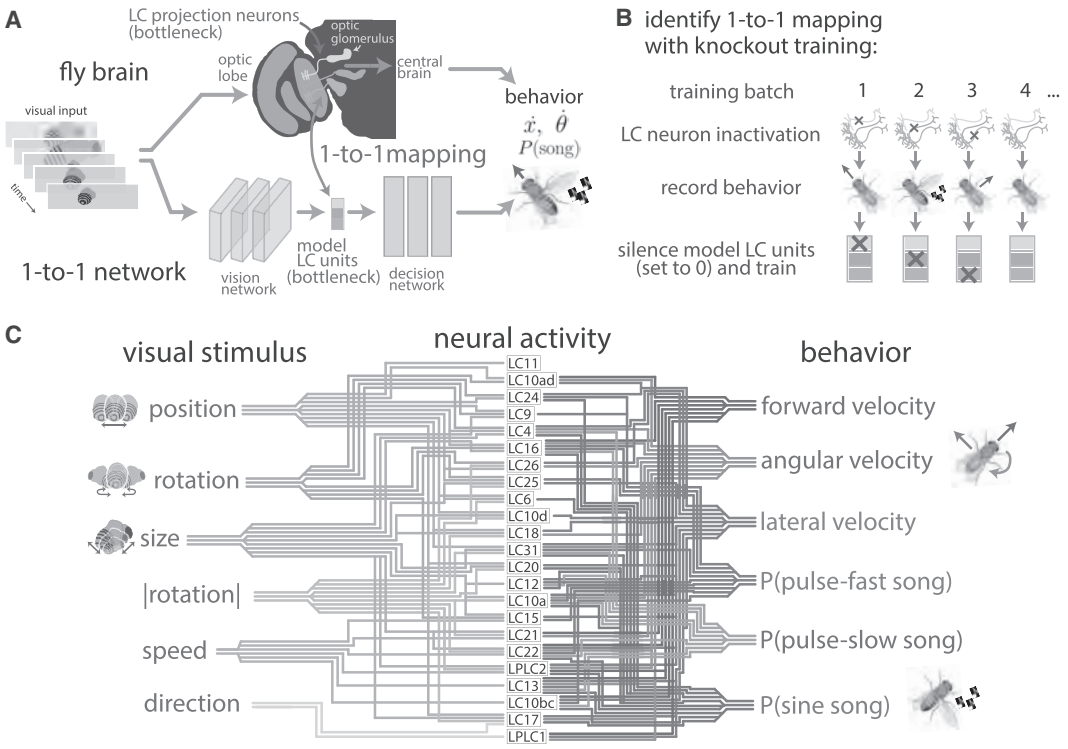



**Figure 2.** Modeling the sensorimotor transformation of the fruit fly visual system. (*A*) We model the mapping between a fruit fly's visual input to behavior with a deep neural network. Our network forms a bottleneck in the form of model units equal to the number of optic glomeruli. We seek a one-to-one mapping between model units and optic glomeruli. (*B*) Illustrative diagram of knockout training to identify the one-to-one mapping. For each batch, we consider a session for which one neuron type was genetically silenced (LC neuron inactivation) and its resulting behavior (record behavior). We train the model on this behavior, silencing the corresponding model unit (i.e., setting its value to 0). For control sessions in which no flies were perturbed (batch 4), no model units are silenced. Thus, we perturb the model units in a way similar to how we perturbed the neuron types in the fly. (*C*) Our model's predictions of what visual features are encoded by each visual projection (LC) neuron type (*left*) and which LC neuron types contribute to which behaviors (*right*). (LC) Lobular columnar.

how the visual neuron types in the fly's visual system work together to encode stimuli and contribute to behavior (Fig. 2C). This is a powerful modeling framework to aggregate different neural perturbations into one unified model.

## Computational and Neural Insights into the Temporal Patterning of Songs in Singing Mice

Y. Thapa [in collaboration with A. Banerjee, CSHL]

Communication is the backbone of any social interaction, and a primary way animals communicate is through vocalization. From the chirps of songbirds to conversational human speech, vocalization requires temporally precise regulation. Here, we study a unique form of temporally regulated vocalization: Alton's singing mice produce dozens of human-audible songs throughout the day. A singing mouse modulates its song length and rate during social interaction—a hallmark of communication—but in isolation will also produce songs throughout the day. The temporal patterning of songs in isolation, how it changes during social interactions, and the neural mechanisms of this behavior remain unclear. We hypothesize that one of the ways the singing mice pattern their songs is by spacing them out equally in time. Based on a month-long recording of mice in solo conditions (Fig. 3A,B), we observed hours of evenly spaced songs resembling a noisy clock (Fig. 3C). We propose that these behavioral dynamics can be described using a clock-like drift-diffusion model. We also aim to extend these models to capture song interactions between two

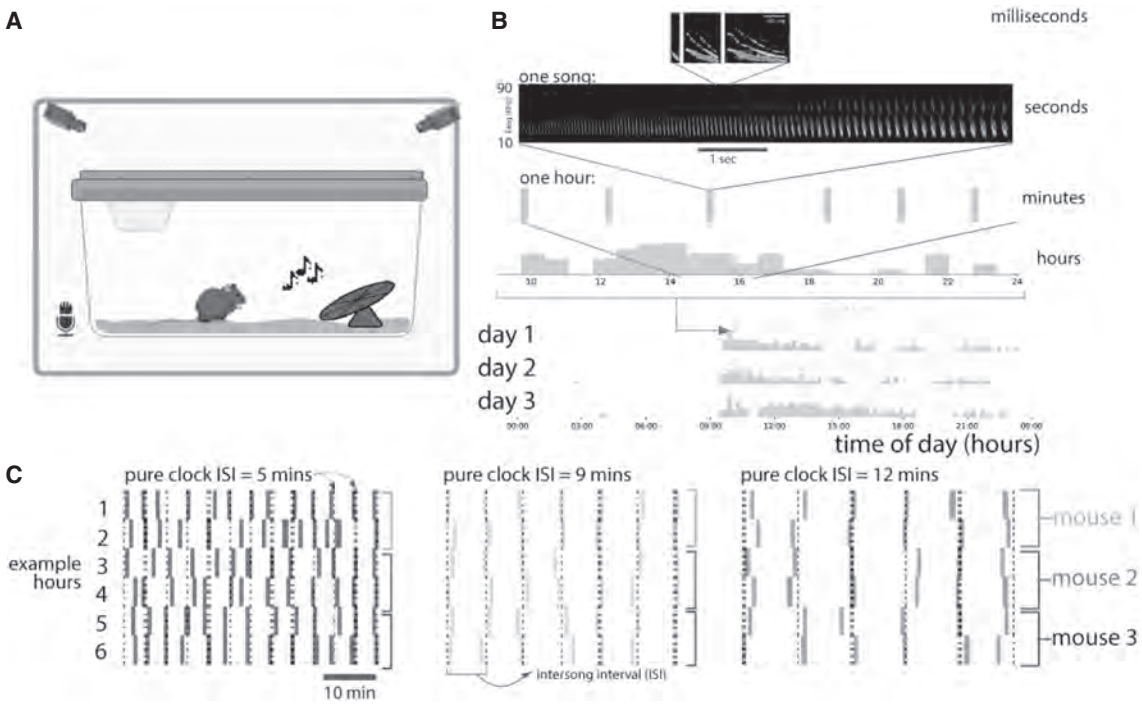



**Figure 3.** Alton's singing mice produce song at clock-like intervals. (*A*) Experimental setup. We record song production for long periods of time (days) for, in this case, an isolated mouse. (*B*) Example vocalization spanning several timescales (i.e., milliseconds to days). (*First row*) Example notes in milliseconds, (*second row*) example songs (each ~7 sec), (*third row*) song histogram over a day (each row). (*C*) One-hour bin with time points adhering to a clock (dashed lines) and real song times (ticks). Clock-like song behavior was observed for short, medium, and long intersong intervals (ISIs).

mice. Finally, we propose experiments that aim to determine the neural signature of song behavior using pharmacological perturbation, functional imaging, and manipulations; our first investigation will focus on dopaminergic pathways. Our research will uncover the computational and neural mechanisms underpinning an internally generated rhythmic behavior and how it changes in the context of social interaction.

## Modeling Responses of the Visual Projection Neurons in the Fruit Fly Visual System

J. Skaza

The visual projection neurons in the fruit fly visual system form a bottleneck between the optic lobe and central brain. How these neurons bundle visual information to guide behavior is an important piece in understanding visuomotor transformations during natural behavior such as flight and courtship. Here, we design and compare DNN models to predict the responses of lobula columnar (LC) neurons in *Drosophila melanogaster* (Fig. 4A). We consider a diverse range of network architectures: classic convolutional DNNs, DNNs whose weights are constrained by the synaptic connections found in the connectome (Dorkenwald et al., *Nat Methods* **19:** 119 [2022]), and task-driven DNN models optimized for optic flow (Lappalainen et al., bioRxiv doi:10.1101/2023.03.11.532232 [2023]) or natural courtship behavior (Cowley et al. 2024). To train our DNN models, we combined data sets from previous studies of recorded LC responses to both artificial and natural stimuli. Although still in the early stages, we find that data-driven convolutional neural network (CNN) models are able to accurately predict LC responses (Fig. 4B); we suspect the other models will perform just as well. This leads us to our main conclusion: The stimuli used to probe LC function are impoverished, and we cannot distinguish between DNN models with vastly different architectures. We overcome this by synthesizing a large set of new stimuli optimized to drive different predictions across models. Still, we analyze the best-performing DNN model to

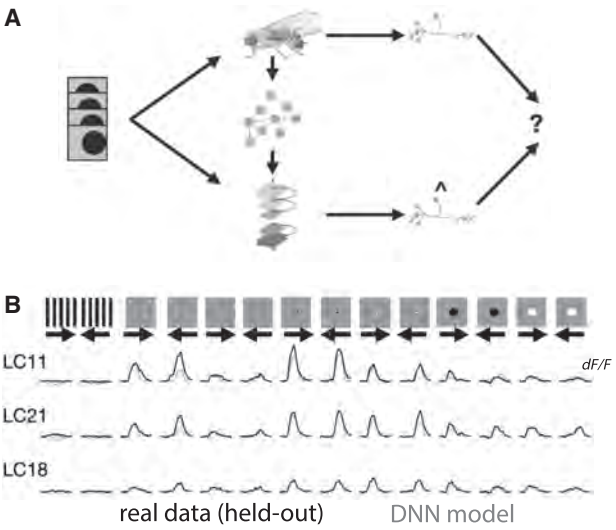



**Figure 4.** Predicting responses in visual projection (lobular columnar [LC]) neurons in the fruit fly. (*A*) Visual overview of our modeling setup. Experimentalists present dynamic stimuli while recording responses in LC neurons. We train deep neural network (DNN) models, including those inspired by the *Drosophila*'s connectome, using the same set of stimuli to obtain in silico predictions of in vivo LC activity. (*B*) Model predictions (gray traces) versus real responses (black traces) for three example LC neurons recorded with calcium imaging (Turner et al., *eLife* **11:** e82587 [2022]).

better understand its inner workings, shedding light on the true computations underlying fruit fly vision. This work advances our understanding of LC function and gives a call to action for collecting more LC responses to a large number of diverse stimuli.

## ImageBeagle: Large Image Data Set with Efficient Search for Closed-Loop Methods

R. Göndür

To understand the computations of visual neurons, we need to present images that drive neural responses. However, visual neurons, especially in higher-order visual areas V4 and IT, tend to be selective for specific visual features (e.g., an arrangement of dots, high spatial frequency textures, human faces, etc.), so we must consider a large number of candidate images in hopes that some of them are the preferred stimuli. Exacerbating this problem, closed-loop methods (combining data collection and model training) (Cowley et al. 2023) need to quickly search this candidate pool to identify the most promising image. For example, it is not uncommon to pass one million images through a DNN model to find a preferred stimulus—and yet, one million images are not enough to truly capture the preferred stimuli (e.g., how many images of bananas would be in one million randomly chosen images?).

To overcome these challenges, we developed a large image data set, ImageBeagle, to be used in neuroscience and machine-learning tasks. The goal is to collect 100 million images from diverse sources (e.g., ImageNet, YouTube, AI-generated, artificial, etc.). Importantly, we developed an efficient algorithm (Beagle) to navigate through these images to find an image that maximizes some value (Fig. 5A). Beagle starts with a randomly chosen image and moves to the nearest neighbor with the largest value, akin to a discrete version of gradient ascent. In a simulation to find the reddest image (out of 150,000 candidate images), the Beagle algorithm quickly finds one of the top red images after 400 evaluations; randomly choosing images does not outperform Beagle even after 3,000 evaluations. This means that in closed-loop experiments, we can find more appropriate images while reducing the amount of computation time. We now plan to run large-scale simulations on our full data set of 100 million images. We also plan to collect a large stimulus data set for video clips (termed VideoBeagle). This approach is a key component of the next generation of closed-loop methods.

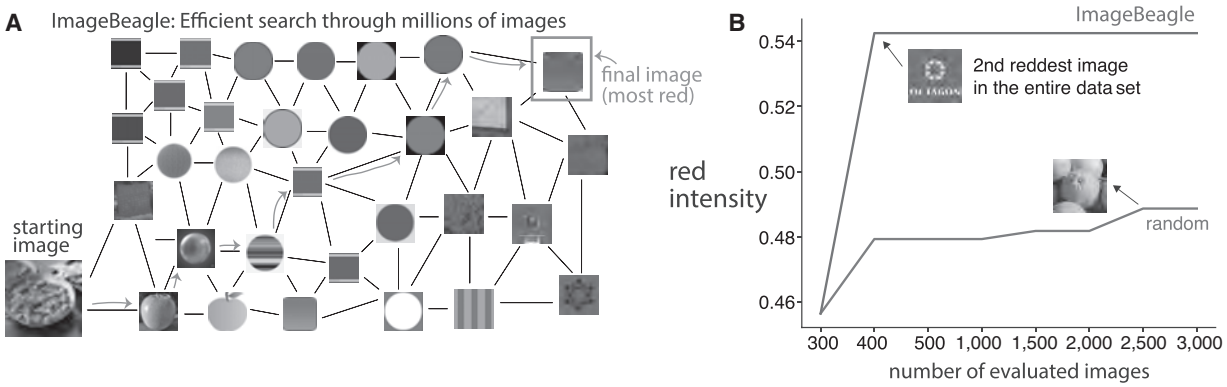



**Figure 5.** Efficiently searching through a large image data set (100 million images) with the ImageBeagle algorithm. (*A*) Consider a large number of images as a graph in which nearby images are similar (here similarity is defined by embeddings from a deep neural network). For a given query image (starting image), ImageBeagle navigates by assessing each image's nearest neighbors and moving to the one that maximizes the objective (in this case, to find the reddest image—*top rightmost* image). This is a greedy, local search akin to gradient descent. (*B*) Simulation on 150,000 images. The ImageBeagle algorithm (*top* trace) almost immediately identifies one of the reddest images after 400 evaluations. Evaluating randomly chosen images fails to find a top red image after 3,000 evaluations. Thus, ImageBeagle greatly reduced the amount of computation time needed to search for an image for any given objective. Grayscale versions of this image will not show redness; readers may rely on the red intensity scale (the proportion of red intensity versus blue and green intensities for RGB channels of the image).

## Mapping DNN Features to Visual Cortical Neurons with Factorized Linear Mappings and Active Learning

F. Vercuysse

In recent years, DNNs have demonstrated remarkable effectiveness in modeling and predicting the processing of visual information streams. Establishing a linear relationship between the visual features extracted by DNNs and neural responses proves to be a valuable method for assessing the similarity in the processing of visual information between DNNs and the visual system. However, when the number of extracted features surpasses the available training data, regression models tend to overfit on the training set. This becomes a challenge when the number of training samples is limited, typically on the order of thousands, as seen in experimental data collections. Recent work has proposed a factorized linear mapping to greatly reduce the number of parameters (and,

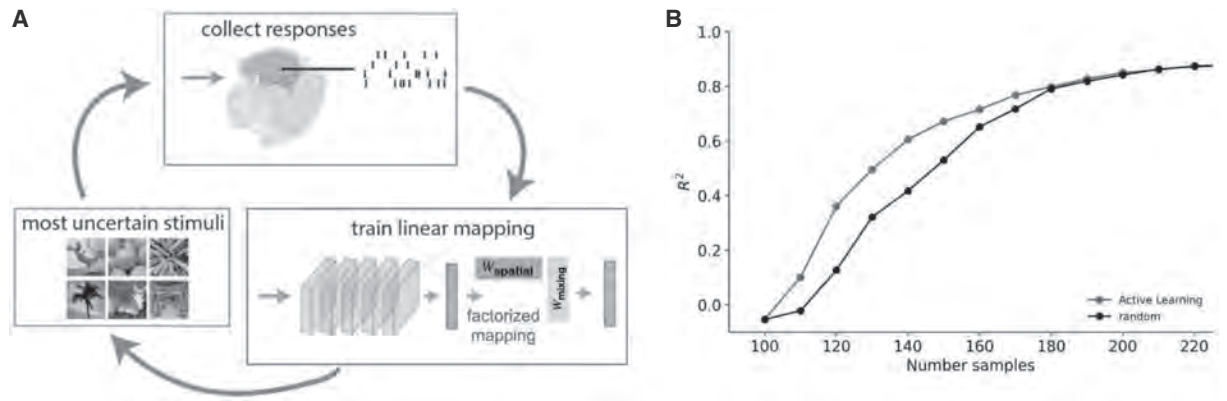



**Figure 6.** Training a factorized linear mapping to predict visual cortical responses. (*A*) Applying active learning to a factorized mapping enables the selection of the most informative stimuli to train our model. (*B*) Choosing data points guided by active-learning metrics results in a reduction in the number of samples required to achieve equivalent test performance.

thus, the chances of overfitting to a small amount of training data) (Klindt et al., *NeurIPS* **30** [2017]). This model assumes each visual neuron has the same spatial receptive field for any input channel. We extend this model to allow for multiple factors (i.e., different spatial receptive fields for different input channels). In addition, we are developing a closed-loop method to efficiently fit the parameters of any factorized linear mapping (Fig. 6A). In simulations, using active learning (i.e., a closed-loop method) outperforms randomly choosing images (Fig. 6B). This work is in early stages; we are now evaluating this model on real data as well as testing other active learning algorithms. The overall goal is to fit a mapping between DNN features and visual cortical responses with as little recording time as possible; we can then use this mapping to probe other functions of visual cortex (spatial attention, noise correlations, etc.).

## PUBLICATIONS

Cowley BR, Stan PL, Pillow JW, Smith MA. 2023. Compact deep neural network models of visual cortex. bioRxiv doi: 10.1101/2023.11.22.568315

*In Press*

Cowley BR, Calhoun AJ, Rangarajan N, Ireland E, Turner MH, Pillow JW, Murthy M. 2024. Mapping model units to visual neurons reveals population code for social behaviour. *Nature* doi:10.1038/s41586-024-07451-8

# DECIPHERING STRUCTURES AND FUNCTIONS OF NEURONAL RECEPTORS AND CHANNELS

**H. Furukawa** E. Chou M. Epstein H. Kang S. Kleeman K. Michalski R. Polfer N. Simorowski J. Syrjanen D. Thomas

Precise regulation of neurotransmission is crucial in shaping both the development and function of the brain, underpinning essential processes such as learning, memory, and cognition. These mechanisms largely depend on the finely controlled transport of ions and molecules across biological membranes facilitated by a wide variety of molecular machinery. Among these, neurotransmitter-gated ion channels and large-pore channels are particularly notable. Research in the Furukawa laboratory is focused on uncovering the molecular mechanisms underlying neurotransmission and neuroplasticity, with the ultimate aim of developing reagents with therapeutic potential for a range of neurological disorders, including schizophrenia, depression, stroke, and Alzheimer's disease.

Our investigations involve a detailed exploration of the structural and functional properties of cell surface receptors and ion channels that regulate intracellular calcium signaling in response to voltage changes and neurotransmitter stimuli, thus modulating the strength of neurotransmission. The Furukawa laboratory is currently engaged in two major research endeavors. The first is dedicated to elucidating the determinants of subtype specificity in various *N*-methyl-D-aspartate receptors (NMDARs). The second centers on the structural and functional characterization of a large-pore channel, the calcium homeostasis modulator (CALHM) channel. These projects are driven by the recognition that aberrant activation of these channels is implicated in neurodegeneration, neuroinflammation, and disturbances in neuronal function.

To achieve these objectives, the Furukawa laboratory employs an array of cutting-edge techniques, including X-ray crystallography and single-particle electron cryomicroscopy (cryo-EM), alongside a broad spectrum of biochemical and biophysical methodologies, such as electrophysiology. These diverse approaches enable us to investigate the molecular underpinnings of neurotransmission and present new avenues for therapeutic intervention in debilitating neurological diseases.

## Structure of Human CALHM1 and Functional Modulation by Lipids

Although our group published the first structure of CALHM1 derived from chicken ortholog, there has not been any structure of human CALHM1. This is due to challenges in the production of CALHM1 proteins. We overcame this technical difficulty, determined the cryo-EM structures of human CALHM1 with the carboxy-terminal domain truncation (hCALHM1Δct), and compared them with the chicken CALHM1 structure we previously revealed. The hCALHM1 proteins were expressed in HEK293 cells ($GnTI^-$ cells), purified, and reconstituted into a lipid nanodisc. By implementing single-particle cryo-EM, we were able to resolve the octameric hCALHM1Δct at 3.67 Å resolution. Our cryo-EM study revealed that chCALHM1Δct and hCALHM1Δct have a similar protein fold and octameric assembly pattern where the carboxy-terminal helix (CTH) fine-tunes the oligomeric state in each (Fig. 1). The structure of hCALHM1Δct also allowed us to pinpoint locations of the Alzheimer's risk factor mutations.

Inspection of the hCALHM1Δct and chCALHM1Δct structures revealed a hydrophobic pocket ~15 Å in length (parallel to the transmembrane domains [TMDs]) and ~10 Å wide, located between TMD3 and TMD4 of one subunit and TMD2 of the neighboring subunit and filled with a phospholipid-like density (Fig. 2A). We next estimated the binding stability of a phospholipid, 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC), and cholesterol by coarse-grained molecular dynamics (CG-MD) simulations on hCALHM1Δct embedded into a membrane bilayer consisting of a 10:3 ratio of POPC

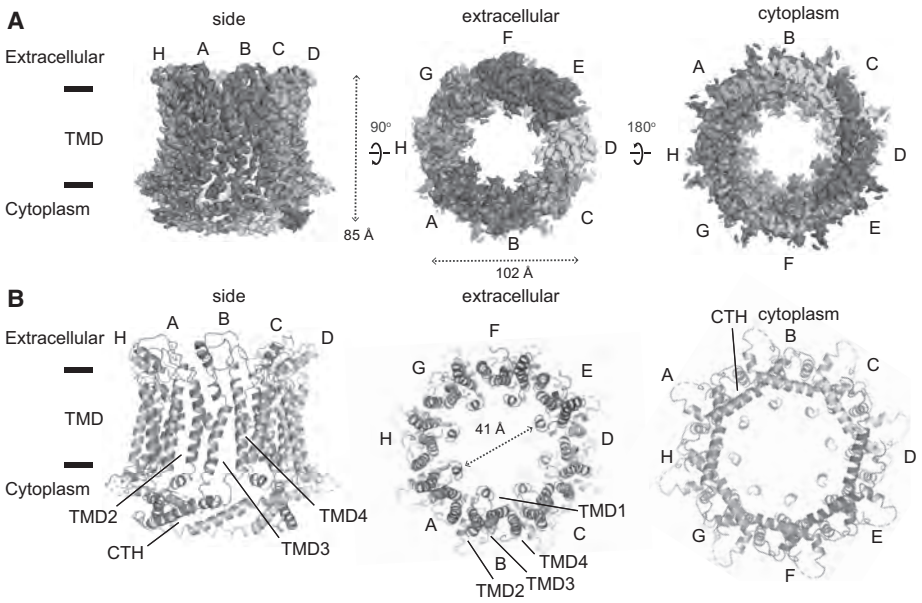



**Figure 1.** The cryo-EM structure of human CALHM1. Cryo-EM density (*A*) and molecular models (*B*) of human CALHM1 viewed from the side of the membrane, the extracellular region, and the cytoplasm. The pore distance indicated by the double-headed arrow is measured between the Gln33 Cα positions of chains A and E.

to cholesterol (Fig. 2B). Computation of the potentials of mean force (PMF) in the CG representation revealed that the energy minimum of POPC corresponds to ~–6.4 kcal/mol, with a clearly defined energy well (Fig. 2B). In contrast, the cholesterol PMF is comparatively flatter, with an energy minimum of ~–1.9 kcal/mol (Fig. 2B), indicating that the binding of a POPC is energetically more favorable than that of cholesterol in this cavity. Thus, both the cryo-EM and MD simulations support phospholipid as the preferred physiological lipid in the hydrophobic pocket.

To understand the functional role of the lipid-filled hydrophobic pocket, we generated tryptophan point mutations of the pocket-forming residues and assessed their functional effects by measuring voltage-gated currents by whole-cell patch-clamp electrophysiology (Fig. 3). Specifically, we measured and compared current density (*pA*/*pF*) at +100 mV membrane potential (jumping from –60 mV), where we observed sufficiently large amplitudes for comparative analyses. Among the mutants, the functional effects of Val192Trp and Ile109Trp stood out (Fig. 3). First, the Val192Trp mutation at the deep, narrow region of the pocket eliminated current density. Second, incorporation of the tryptophan mutation at the "entrance" of the hydrophobic pocket (Ile109Trp) exhibited an approximately fivefold increase in current density compared with the wild-type hCALHM1Δct. The other mutants showed no statistically significant changes compared with the wild type (Fig. 3). Overall, the mutations above alter both channel functions and cell surface expressions, indicating the hydrophobic pocket to be a crucial motif for controlling the CALHM1 activities. The impact of mutations on cell surface expression and current density varies depending on their location. Specifically, Ala199Trp causes an increase in surface expression, whereas Leu67Trp, Val112Trp, Ala116Trp, and Thr196Trp result in decreased expression at different levels, and Val192Trp shows no detectable expression. Regarding current density, Ile109Trp leads to an increase, whereas Leu67Trp, Val112Trp, Ala116Trp, and Thr196Trp exhibit similar levels, and Val192Trp shows no current at all. Comparison of current density with cell surface expression suggests that the Ile109Trp and Ala199Trp mutations result in up-regulation and down-regulation of channel activities, respectively (e.g., open probability or conductance level). In contrast, Val192Trp is structurally disruptive, as our cryo-EM structure suggests it could cause steric clashes with Leu120, Trp189, Leu193, and Val63. Furthermore, the Leu67Trp, Val112Trp, Ala116Trp, and Thr196Trp mutants may form up-regulated channels.

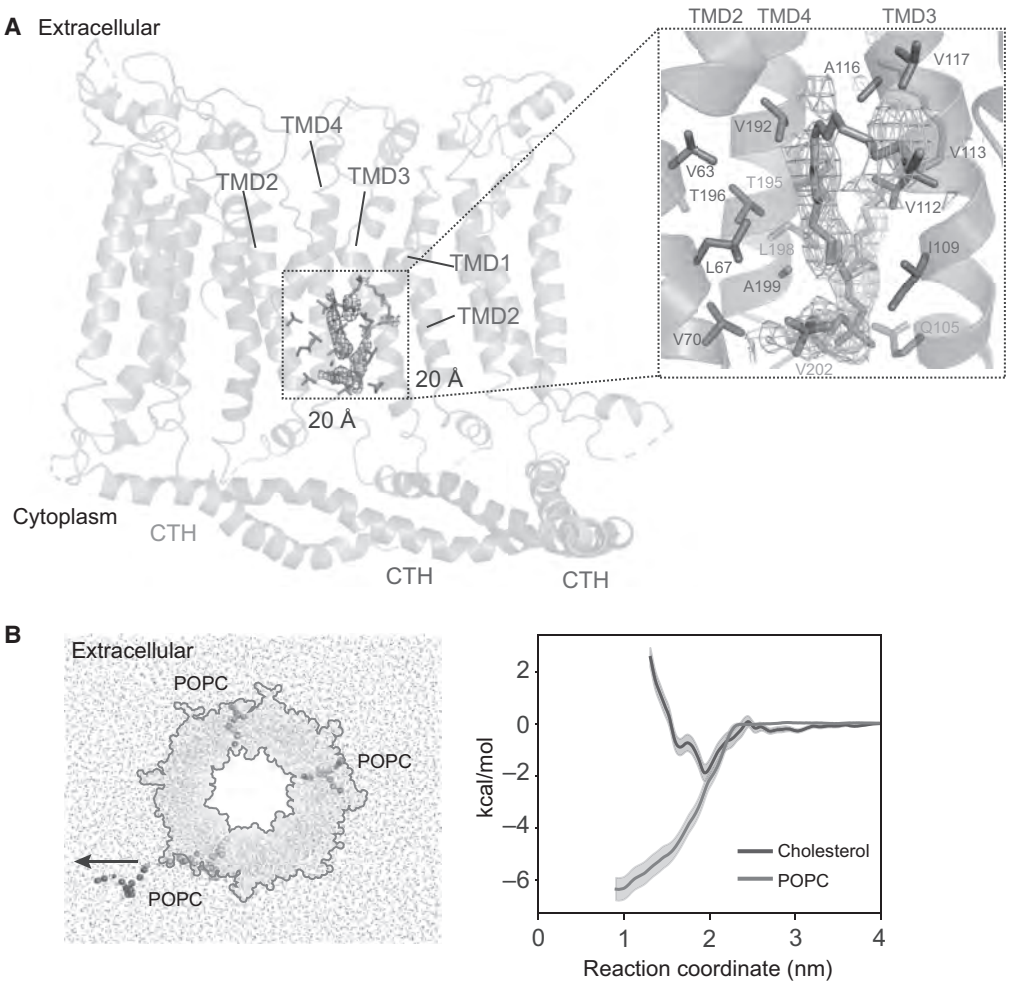



**Figure 2.** The conserved hydrophobic pocket preferentially binds to phospholipid. (*A*) A cross section of human CALHM1 shows the location of the hydrophobic pocket filled with a phospholipid. Residues that form the hydrophobic pocket are shown in stick representation. There are eight hydrophobic pockets per CALHM1 octamer. The *inset* shows a zoomed-in view of the hydrophobic pocket viewed from the pore. Residues are numbered. (*B*) Coarse-grained PMF calculations suggest that binding of POPC into this pocket is thermodynamically favored over that of cholesterol with energy minimums of approximately –6.4 kcal/mol and –1.9 kcal/mol, respectively (*right*). The arrow indicates the direction of steering for which the reaction coordinate was generated (*left*). Error bands are 1 standard deviation generated from 200 rounds of bootstrap analysis.

## Channel Blockade of CALHM1 by Ruthenium Red

The single-particle cryo-EM on hCALHM1$_{I109W}$Δct in the presence of RuR showed oblong density consistent with the size of RuR in the central pore surrounded by the amino-terminal helixes [NTHs] (Fig. 4A). The resolution of the cryo-EM map is lower around the NTHs and pore region compared with the TMDs, possibly suggesting a number of subtly different binding modes within this binding pocket because the pore diameter is large. Nevertheless, the cryo-EM density is of sufficient quality for the placement of RuR molecule, although an atomic-resolution pose cannot be determined from these data. This density was not observed in the cryo-EM structure of hCALHM1$_{I109W}$Δct in the absence of RuR, supporting the view that it represents RuR. The structures of hCALHM1$_{I109W}$Δct in the presence and absence of RuR are similar in TMD2-4 and CTH (RMSD = 0.541 Å; over 197 Cα positions), indicating that the binding of RuR only affects the structure of the central pore comprising TMD1 and NTH. Moreover,

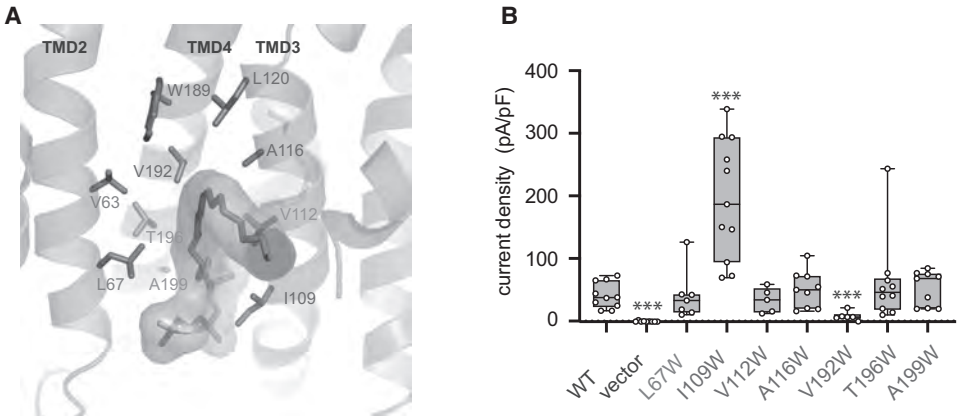



**Figure 3.** The conserved hydrophobic pocket is a key locus for structural integrity and channel functions. (*A*) Positions of residues (sticks) within the hydrophobic pocket, analyzed by site-directed mutagenesis. (*B*) Current density *pA/pF* at +100 mV for each point mutant. Each data point represents a measurement on a different cell (wild-type hCALHM1Δct, $n = 10$; Ile109Trp hCALHM1Δct, $n = 11$; Ala116Trp hCALHM1Δct, $n = 9$; Ala199Trp hCALHM1Δct, $n = 9$; Thr196Trp hCALHM1Δct, $n = 10$; Val192Trp hCALHM1Δct, $n = 6$; vector, $n = 9$; Val112Trp hCALHM1Δct, $n = 5$; Leu67Trp hCALHM1Δct, $n = 7$). Boxes represent the median, 25th, and 75th percentile values, and the whiskers represent the minimum and maximum values. *** denotes $p < 0.001$ versus wild-type. An unpaired two-tailed *t*-test with Welch's correction was used to analyze data. *P* values are as follows: hCALHM1Δct vs. vector, $p = 0.0001$; hCALHM1Δct vs. Leu67Trp hCALHM1Δct, $p = 0.9925$; hCALHM1Δct vs. Ile109Trp hCALHM1Δct, $p = 0.003$; hCALHM1Δct vs. Val112Trp hCALHM1Δct, $p = 0.4962$; hCALHM1Δct vs. Ala116Trp hCALHM1Δct, $p = 0.4395$; hCALHM1Δct vs. Val192Trp hCALHM1Δct, $p = 0.0006$; hCALHM1Δct vs. Thr196Trp hCALHM1Δct, $p = 0.3810$; hCALHM1Δct vs. Ala199Trp hCALHM1Δct, $p = 0.3253$. WTΔct is used to refer to hCALHM1Δct in the graph.

the structure of $hCALHM1_{I109W}$Δct in RuR-free conditions does not clearly resolve the NTHs, indicating a higher degree of conformational flexibility of the NTHs in the absence of RuR. We therefore suggest that the TMD1-Ile109Trp interactions, lipid-Phe19 interactions, and RuR-NTH interactions all contribute toward NTH conformational stabilization.

The RuR molecule is positioned to plug the pore through associations with the NTHs (Fig. 4A). To validate this structural observation, we mutated the pore-lining residues proximal to RuR to arginine (Gln10Arg, Gln13Arg, and Gln16Arg) and assessed RuR-mediated channel blockade by patch-clamp electrophysiology (Fig. 4B,C). We reasoned that arginine residues that face the pore of the channel, being positively charged and bulkier than glutamine, would disfavor the binding of the positively charged RuR and decrease the extent of channel blockade. Indeed, the Gln10Arg, Gln13Arg, and Gln16Arg mutants all decreased the extent of channel blockade to different extents (Fig. 4D). Thus, the experiments above validated our structural observation that RuR binds and plugs the central pore formed by the NTHs.

Overall, we revealed the structure of human CALHM1 and allosteric coupling between the lipid binding site and channel pore. We also showed RuR binding to hCALHM1 at the channel pore. CALHM2 is reported to bind RuR at the top of each TMD1 helix next to residue Phe39. However, the validity of the proposed binding site remains inconclusive because the similar cryo-EM density assigned as RuR is present in the CALHM2 gap junction structure without RuR as well (Chou et al. 2024; Michalski et al. 2024). Furthermore, whether CALHM2 forms a channel or not is under debate at this point. Nevertheless, the RuR binding to hCALHM1 is similar to that in TRPV6 in that RuR binds and plugs the TMD pore. The RuR binding in CALHM1 is not as defined as in TRPV6 because of the large diameter of the CALHM1 pore. It may be possible that the presence of RuR in the CALHM1 pore creates a highly charged electrostatic barrier that prevents the flux of ions.

## NMDAR Splice Variant—Potential Identification of Spermine Binding Site

It is well established that the GluN2B subunit of the NMDAR binds spermine, leading to its up-regulation.

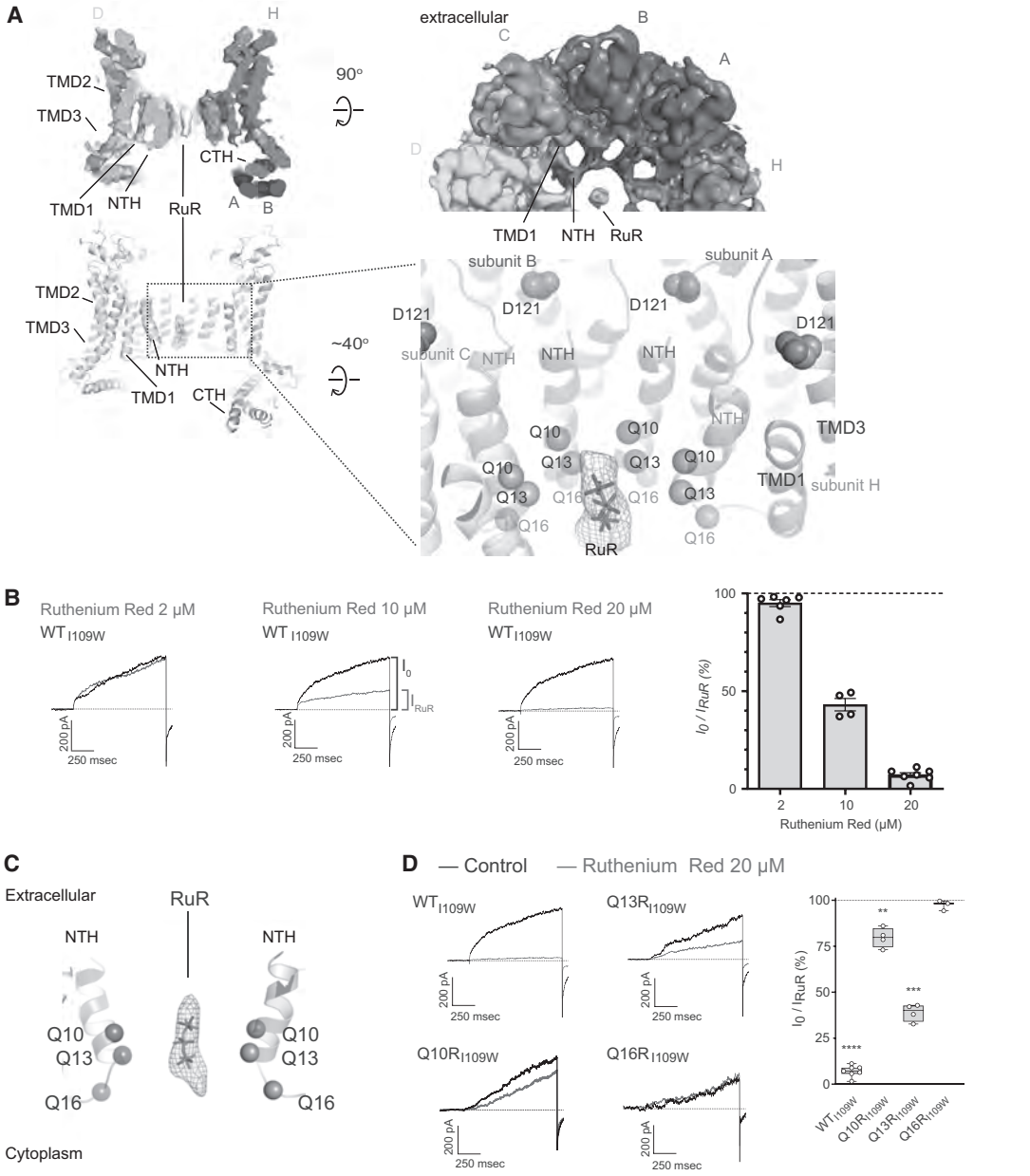



**Figure 4.** Ruthenium red binding site in the central pore. (*A*) A cross section of the cryo-EM map (*top*) and model (*bottom*) of hCALHM1$_{I109W}$Δct without imposing symmetry (c1). The RuR density is shown as white surface (*top*) and gray mesh (*bottom*). The Cα atoms of the pore-lining residues, Gln10, Gln13, and Gln16, are shown as spheres. Asp121 is in the vicinity of NTH. (*B*) Representative current traces and the concentration-response of the hCALHM1$_{I109W}$Δct channel blockade by RuR at +60 mV ($I_{RuR}/I_0$) displayed as a bar chart. Data are represented as individual points; the bars show the mean and the whiskers indicate the standard error of the mean ($n$ = 6, 4, and 7 cells for 2, 10, and 20 µM RuR application, respectively). Definitions of currents without RuR ($I_0$) and with RuR ($I_{RuR}$) are indicated in the *middle* trace. WTΔct$_{I109W}$ is used to refer to hCALHM1$_{I109W}$Δct in the chart and traces. (*C*) A cartoon of the central pore, highlighting the positions of Glu10, Glu13, and Glu16 (spheres) around the RuR binding site (cryo-EM density for RuR shown as mesh; RuR model in stick representation). (*D*) Representative current traces in the absence (black) and presence of 20 µM RuR at +60 mV for the Glu10Arg, Glu13Arg, and Glu16Arg point mutants. The graph shows the extent of the channel blockade ($I_{RuR}/I_0$) calculated from the recordings. Boxes represent the median, 25th, and 75th percentile values, and the whiskers represent the minimum and maximum values ($n$ = 7, 4, 4, and 3 cells for hCALHM1$_{I109W}$Δct, Glu10Arg, Glu13Arg, and Glu16Arg, respectively). ****, ***, and ** denote $p < 0.0001$, $p < 0.001$, and $p < 0.01$, respectively, versus basal conditions (absence of RuR) for each construct studied (two-tailed paired $t$-test: $p < 0.0001$ for wild type, $p = 0.0046$ for Glu10Arg, $p = 0.0001$ for Glu13Arg, and $p = 0.2369$ for Glu16Arg). WTΔct$_{I109W}$ is used to refer to hCALHM1$_{I109W}$Δct in the chart and traces.

The primary questions we aim to address are (1) what is the binding specificity of spermine to GluN2B and (2) why is this effect observed specifically in the GluN1a splice variant but not in GluN1b?

We first developed a cryo-EM protocol to investigate these questions, enabling us to resolve the GluN1a-2B NMDAR complex at an unprecedented resolution of 2.9 Å. This high-resolution structure was intended to facilitate detailed structural studies of spermine binding. However, despite our efforts, we have not been able to capture any discernible electron density corresponding to spermine. We hypothesize that this may be due to spermine's low binding affinity, making it difficult to detect in structural snapshots. In parallel with further experimental approaches, such as exploring different NMDAR constructs to capture spermine binding, we also employed molecular dynamics (MD) simulations to gain insight into this interaction. Specifically, we performed coarse-grained MD simulations ($3 \times 15$ μsec) on intact NMDARs, which provided preliminary insights into the potential spermine binding site, suggesting it resides within the GluN2B amino-terminal domain (ATD). This preliminary data was essential in guiding subsequent steps.

Currently, we are conducting all-atom MD simulations on the GluN1a-GluN2B ATD complex to refine our understanding of the spermine binding site. These simulations aim to map the binding interactions with greater precision and monitor the conformational changes that may be induced upon spermine binding. This detailed analysis will help elucidate how spermine stabilizes specific structural states of the receptor and how it may affect receptor function. Additionally, we are extending our MD simulations to include the GluN1b-2B ATD complex. By comparing the dynamics and binding profiles of the GluN1a- and GluN1b-containing receptors, we hope to uncover the molecular basis for spermine's selective effect on the GluN1a splice variant. This comparative approach will provide crucial insights into the splice variant–specific regulation of NMDARs by spermine.

## Structural Biology of GluN3 Subunits

The GluN1-3 (A or B) NMDAR represents an exception among NMDAR subtypes, distinguished by its unique ion channel activation mechanism, which relies solely on glycine, without the involvement of glutamate. Like the GluN1-2 NMDARs, these receptors conduct cations ($Na^+/K^+$), rendering GluN1-3 NMDARs excitatory glycine receptor channels. This starkly contrasts glycine's typical role as an inhibitory neurotransmitter via pentameric glycine receptor channels in the spinal cord and brainstem. In GluN1-3 NMDARs, glycine binding to GluN3 promotes channel opening, whereas binding to GluN1 induces rapid desensitization. Notably, in heterologously expressed GluN1-3A NMDARs in cell culture, this desensitization can be attenuated by competitive antagonists targeting GluN1, such as CGP-78608, whereas glycine remains bound to the GluN3A subunits. A similar current mediated by GluN1-3 NMDARs can be observed upon application of CGP-78608 and glycine in hippocampal neurons, indicating the presence of GluN1-3 NMDARs in the central nervous system. GluN1-3 NMDARs have been shown to play crucial roles in physiological processes, including the regulation of fear-related memories within cortical and amygdala circuits in adult mice and aversive states in the medial habenula, underscoring their role as mediators of excitatory glycine signaling. Furthermore, the GluN3 subunit has been proposed to form triheteromeric NMDARs with GluN1 and GluN2 subunits, which glycine and glutamate activate. These GluN3-containing NMDARs exhibit distinct properties, such as low $Ca^{2+}$ permeability and reduced sensitivity to channel blockers like $Mg^{2+}$, memantine, and MK-801, in contrast to GluN2-containing NMDARs.

We have successfully established recombinant expression and purification protocols for GluN1a-3A NMDARs, utilizing our innovative expression system, EarlyBac. Over the past year, significant efforts have been directed toward optimizing sample preparation techniques. This has included the introduction of disulfide cross-linking within the transmembrane domains to enhance structural stability, applying chemical cross-linking strategies to stabilize protein conformations, and the introduction of reconstitution receptors into amphipols to preserve their functional integrity during cryo-EM analysis. These methodological refinements have enabled us to collect markedly improved data sets in the presence of the competitive antagonist CNQX. We extensively evaluated various software tools to identify those most suited for high-resolution 3D reconstructions. After testing multiple platforms, we determined that cryoSPARC provided the most effective results, yielding a global 3D reconstruction at a resolution of 4.5 Å. Notably,

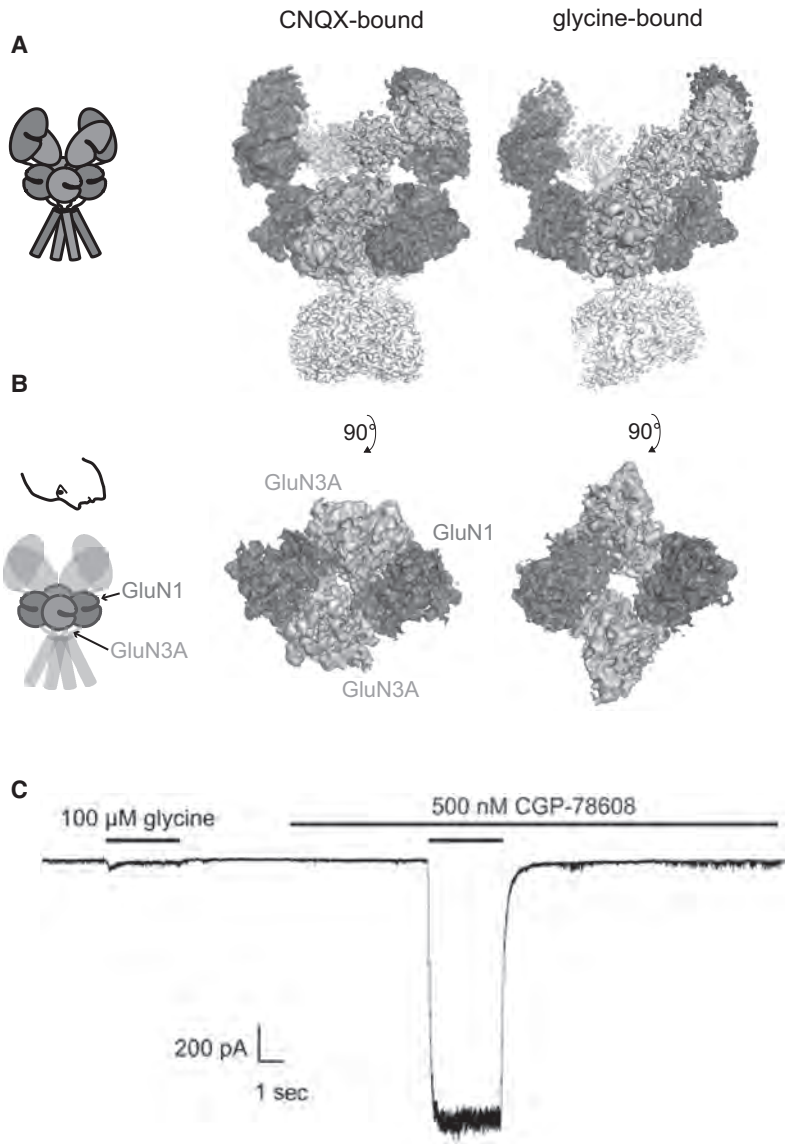



**Figure 5.** Preliminary structural analysis of GluN1-3A NMDARs. (*A*,*B*) The cryo-EM structure of GluN1a-3A NMDAR in complex with an antagonist, CNQX (*left*) and glycine (*right*) viewed from the side (*A*) and top (*B*) of the LBD layer. (*C*) Patch-clamp recording of GluN1a-3A NMDAR. Full activation is observed only in the presence of CGP-78608.

we achieved a local resolution of 3.5 Å in the ligand-binding domains (LBDs), marking the highest resolution obtained for this receptor to date. The increase in resolution has greatly facilitated more accurate model building, as illustrated in Figure 5, and represents a significant milestone in our structural characterization efforts.

Building on this advance, we intend to apply our optimized approach to resolve the receptor's structures in the agonists' presence, enabling us to capture both the active and desensitized states. This will provide invaluable insights into the dynamic conformational changes that underlie NMDAR function and modulation, offering potential avenues for therapeutic intervention.

## PUBLICATIONS

Harris LD, Regan MC, Myers SJ, Nocilla KA, Akins NS, Tahirovic YA, Wilson LJ, Dingledine R, Furukawa H, Traynelis SF, Liotta DC. 2023. Novel GluN2B-selective NMDA receptor negative allosteric modulator possesses intrinsic analgesic

properties and enhances analgesia of morphine in a rodent tail flick pain model. *ACS Chem Neurosci* **14:** 2c00779. doi:10.1021/acschemneuro.2c00779

Syrjänen JL, Epstein M, Gómez R, Furukawa H. 2023. Structure of human CALHM1 reveals key locations for channel regulation and blockade by ruthenium red. *Nat Commun* **14:** 3821. doi:10.1038/s41467-023039388-3

*In Press*

Chou TH, Epstein M, Fritzemeier RG, Akins NS, Paladugu S, Ullman EZ, Liotta DC, Traynelis SF, Furukawa H. 2024. Molecular mechanism of ligand gating and opening of NMDA receptor. *Nature* **632:** 209–217. doi:10.1038/s41586-024-07742-0

Jain A, Nakahata Y, Pancani T, Watabe T, Rusina P, South K, Adachi K, Yan L, Simorowski N, Furukawa H, Yasuda R. 2024. Dendritic, delayed, stochastic CaMKII activation in behavioural time scale plasticity. *Nature* **635:** 151–159. doi:10.1038/s41586-024-08021-8

Michalski K, Furukawa H. 2024. Structure and function of GluN1-3A NMDA receptor excitatory glycine receptor channel. *Sci Adv* **10:** eadl5952. doi:10.1126/sciadv.adl5952

Michalski K, Abdulla T, Kleeman S, Schmidl L, Gómez R, Simorowski N, Vallese F, Prüss H, Heckmann M, Geis C, Furukawa H. 2024. Structural and functional mechanisms of anti-NMDAR autoimmune encephalitis. *Nat Struct Mol Biol* doi:10.1038/s41594-024-01386-4

# NEURAL CONTROL OF MOVEMENT, PHYSIOLOGY, AND BEHAVIOR

**H. Hou** Z. Ahmad E. Davis D. Naglič J. Rodriguez X. Zhang
K. Daruwalla A. Frankel I. Nozal Martin B. Song Y. Zhang

We study how the brain orchestrates motor control in natural behaviors—specifically the neural mechanisms underlying the dynamic control of facial movement and expression—using rodents as a model. Behaviors from grimacing to chewing are a powerful reflection of our internal states in health and disease. This complex behavior is nonetheless controlled by a tractable motor circuit: A small defined cluster of motor neurons integrates brain-wide inputs, bypasses the spinal cord, and directly orchestrates muscles of facial movement. Thus, we are pursuing the idea that by understanding the nuance and control of facial movement, we may gain access to the principles of neural integration and control of natural behavior. Since the laboratory's inception, we have taken a bottom-up approach to gain insight into the brain through behavior and the effectors of behaviors (i.e., muscles and their corresponding motor neurons) and leverage electrophysiology, imaging, neuroanatomy, behavior, computational analysis, and machine-learning methods. Knowledge gained will advance our ability to decipher internal states based on facial movements in rodents, and our approach will help concretize the study of neural control of complex behavior.

## A New System to Sensitively Detect and Analyze 3D Whole-Face Movement in Mice

K. Daruwalla, I. Nozal Martin, A. Frankel, D. Naglič

Studying how coordinated movement of individual facial regions gives rise to multifunctional whole-face movements can provide unique insights into internal physiological processes. Work to date suggests we can infer pain, distress, and sensory input based on subtle facial movement patterns in humans as well as rodents. Being able to precisely and sensitively track facial dynamics has the potential to expand our understanding of how animals experience and respond to various interventions. Mice, with discernible facial responses and evolutionarily conserved mammalian facial movement control circuits, provide an ideal model to unravel the link between facial movement and internal physiological states in mammals. However, existing frameworks lack the spatial or temporal resolution to track motion of the entire mouse face, because of its small and conical form factor. We introduce Cheese3D, a computer vision system that first captures high-speed 3D motion of the entire mouse face (including ears, eyes, whisker pad, and jaw, while covering both sides of the face) using a calibrated six-camera array (Fig. 1). The interpretable framework extracts dynamics of anatomically meaningful 3D facial features in absolute world units at submillimeter precision (Daruwalla et al. 2023). In contrast to existing methods that focus on static facial images, motion of a subset of facial features, or aggregates of orofacial behavior optimized to predict cortical neural activities, Cheese3D is specifically designed to capture and represent whole-face movement while maintaining spatial and physical interpretability. Recording the motion of both individual facial regions and their spatial and temporal relationship to the whole face could be meaningful, because the building blocks of facial movement (i.e., compartments of facial musculature and the brainstem nuclei that directly control them) are highly topographically arranged. The multicamera array setup facilitates reliable, markerless identification of facial keypoints in 3D space, counteracting occlusion and distortion found in single-camera setups, and these precise spatial locations relative to other facial regions are preserved in the 3D geometric features.

## Precisely Measuring Subtle Facial Movements

A. Frankel, K. Daruwalla, I. Nozal Martin, Z. Ahmad, E. Davis

Compared with other overt body movements such as locomotion and reaching, in which limbs and appendages undergo large translations and rotations relative to their size, informative facial movements

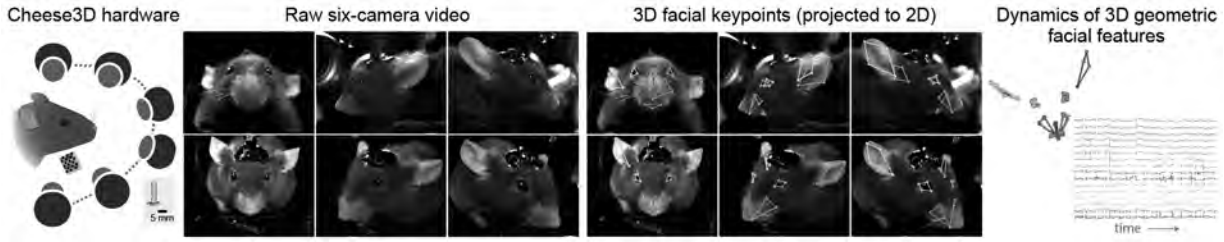



**Figure 1.** Sensitively detecting and quantifying 3D whole-face movement in mice using Cheese3D. The framework consists of a robust and calibrated multicamera system and associated analysis pipeline that extracts dynamics of anatomically meaningful 3D facial features in absolute world units at submillimeter precision.

are often small, localized, and fast. This poses unique technical challenges to sensitively track subtle movements while covering the entire face. A known issue for pose estimation methods based in convolutional neural networks is keypoint jitter, whereby local fluctuations in keypoint tracking are unrelated to genuine movement. Keypoint jitters are often mitigated using low-pass filters, but this can attenuate dynamics and reduce temporal resolution, and such filters are therefore inadequate for detecting whole-face motions that are both subtle and rapid. A key benefit of 3D multiview calibration compared with single 2D uncalibrated views is that view redundancy reduces the amplitude of keypoint jitter. As a proof-of-principle test that Cheese3D is able to capture subtle and rapid facial movements with physiological significance, we designed experiments to monitor mice emerging from ketamine-induced anesthesia. Small localized facial movements, including whisker deflections, appear during anesthesia and signal early stages of recovery. We explicitly quantified keypoint jitter in our setup in a control experiment using motionless periods and measured the reduction in jitter between 2D keypoints and 3D keypoints projected onto 2D views. These data demonstrate that Cheese3D can be used to detect small movements associated with physiologically significant events in mice.

## Uncovering Underlying Physiology from External Facial Motion

I. Nozal Martin, K. Daruwalla, A. Frankel, D. Naglič

We examined whether Cheese3D can be used to study facial movements associated with physiological processes that are not otherwise externally visible, in addition to detecting movements that are fast and subtle. Visualizing the anesthesia data over the entire period (1–2 h) revealed gradual changes in ears and eyes that are stereotyped across mice. This suggests that certain facial features can be used as a "stopwatch" to track time since anesthesia induction. To test this hypothesis, we fit a single model across all mice to predict the time elapsed since anesthesia induction (intraperitoneal injection of ketamine and xylazine) using only the initial and current value of unfiltered facial features and discovered that combining motions from different facial regions provides a useful visual indicator to track time elapsed in anesthesia. We also tested Cheese3D with facial movements that are vigorous in amplitude: Chewing in rodents is difficult to characterize externally as teeth, along with food that has entered the mouth, cannot be seen. Our data indicate that Cheese3D detects facial movements during rodent food consumption consistent with known characteristics of food placement, tooth anatomy, and muscle engagement.

Overall, our analysis revealed informative synchronous facial movement patterns that could be used to infer unseen (internal anatomy and physiological functions) from seen (external synchronized facial motion). We anticipate the method will enable important discoveries across fields in biology and medicine by allowing for noninvasive readout of moment-to-moment changes in body states in mice. The potential applications of high-resolution, whole-face kinematics data made possible by Cheese3D are vast and are likely to inspire a new era of quantitative studies linking facial movements to changes in internal states brought on by disease, drug exposure, neural processes, or other physiological functions we would otherwise have limited access to based on external observations.

## Dissecting the Neuromuscular Basis of Facial Control

D. Naglič, B. Song, X. Zhang, J. Rodriguez, Y. Zhang

The laboratory mouse, with an ever-growing list of discernible facial responses and evolutionarily conserved mammalian facial movement control circuits, can serve as an ideal model to unravel the link between facial muscle control and motor circuits that integrate cortical and subcortical inputs. The motor machinery required to generate facial movements is a group of muscles directly controlled by clusters of motor neurons and premotor networks in the brainstem, bypassing the spinal cord. Thus, it is possible to unequivocally link the activity of motor neurons and their local circuitry to the motion of facial muscle targets. However, establishing a mouse model of facial motor control has been impeded by the lack of (1) a thorough behavioral description of whole-face sequences and synchrony across different behaviors, (2) a detailed atlas and biomechanical model of mouse facial musculature that form the building blocks of facial movements, and (3) a wiring diagram of the functional connectivity with motor circuitry. Toward these goals, we have set up the technical foundation in the laboratory to (1) define the parameter space of mouse facial movement dynamics in response to specific stimuli and test the hypothesis that temporal dynamics of mouse facial expression can distinguish among sensory inputs and certain changes in internal states; (2) establish an atlas of mouse facial musculature and their functional connectivity, and we expect to provide a map of mouse facial muscles that form the basis of anatomical and biomechanical models; and (3) determine how brainstem circuitry interfaces with brain-wide input to drive precisely timed facial dynamics and test the hypothesis that local brainstem connectivity is crucial to bias facial movements.

### PUBLICATION

Daruwalla K, Nozal Martin I, Frankel A, Naglič D, Ahmad Z, Hou XH. 2023. A 3D whole-face movement analysis system to uncover underlying physiology in mice. bioRxiv doi:10.1101/2024.05.07.593051

# THEORETICAL NEUROSCIENCE AND ARTIFICIAL INTELLIGENCE

**A.A. Koulakov** B. Baserdem F. Pashakhanloo S. Shuvaev
C. Mascart K. Samoilova K. Tran

Our laboratory works on theories of neural computation. Our overall strategy is to use methods developed in mathematics, physics, machine learning, computer science, and statistics to build experimentally testable models of neural networks and their function. In most cases, we base our theories on what is known about biological systems; however, given that the principles of brain function remain unclear, in many cases we resort to building computational theories. This means that we formulate the problems solved by the brain in a mathematically rigorous fashion and hypothesize how an engineer would solve the problem, given biological and experimental constraints. We then use these solutions to form experimentally testable predictions. Testing these predictions in collaboration with our experimental colleagues helps us refute or refine our theories. For example, we are interested in understanding how connectivity is established in the brain. We have proposed several theories that may determine the rules of making connections between neurons based on a limited set of instructions contained in the genome. These theories address several levels of organization, including computational, biological, engineering, evolutionary, etc. Our theories may explain the differences between connectivities in normal and abnormal brain circuits. We are also interested in understanding the principles of perceptual invariance (i.e., how can sensory systems represent objects in the environment despite substantial variations in intensity and background). Visual percepts, for example, retain basic features, such as perceived shape and color composition, despite variable luminance, spectral composition, scale, and position of the stimuli. Although we study the question of perceptual invariance in application to well-defined problems, we believe that the principles we will uncover may generalize across sensory modalities. Finally, we are pursuing the question of how modern theories of machine learning and artificial intelligence can apply to brain function. Although reinforcement learning, deep learning, long short-term memory networks, etc., are successful in solving a variety of artificial intelligence problems, their mapping onto brain circuits remains unclear. We attempt to bring these systems closer to satisfying the constraints imposed by biology. We hope that the convergence of theoretical constructs and their biological underpinning will help us learn more about brain function.

## A Normative Theory of Social Conflict

S. Shuvaev, A.A. Koulakov [in collaboration with G. Enikolopov, SUNY Stony Brook]

Social conflict is a survival mechanism yielding both normal and pathological behaviors. To understand its underlying principles, we collected behavioral and whole-brain neural data from mice advancing through stages of social conflict. We modeled the animals' interactions as a normal-form game using Bayesian inference to account for the partial observability of animals' strengths. We find that our behavioral and neural data are consistent with the first-level Theory of Mind (1-ToM) model in which mice form "primary" beliefs about the strengths of all mice involved and "secondary" beliefs that estimate the beliefs of their opponents. Our model identifies the brain regions that carry the information about these beliefs and offers a framework for studies of social behaviors in partially observable settings (Fig. 1).

## Stochastic Gradient Descent–Induced Drift of Representation in a Two-Layer Neural Network

F. Pashakhanloo, A.A. Koulakov

Representational drift refers to changes over time in neural activation accompanied by a stable task performance. Despite being observed in the brain and in artificial networks, the mechanisms of drift and its implications are not fully understood. Motivated by recent experimental findings of stimulus-dependent

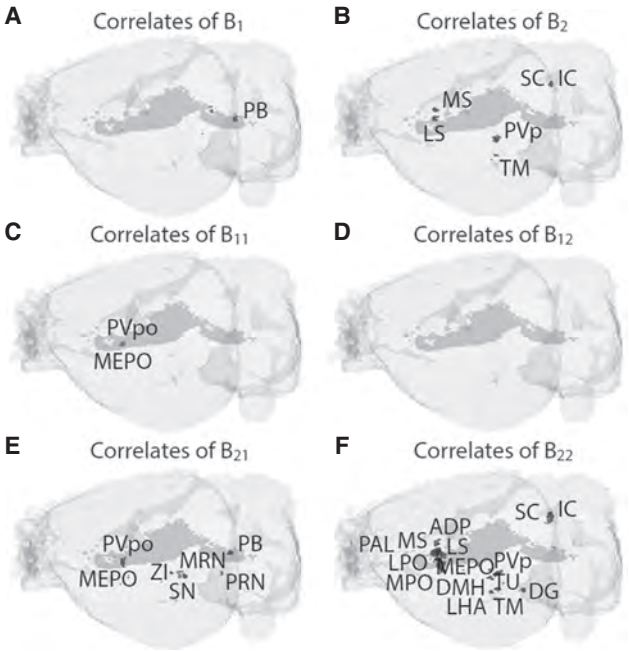



**Figure 1.** Representations of different levels of theories of mind in the mouse brain. Correlates of the reconstructed Bayesian beliefs in the c-Fos activity in the brain. (*A,B*) Zeroth-level theory of mind; (*C–F*) first-level theory of mind. The names of the brain regions with significant representations are marked.

drift in the piriform cortex, we use theory and simulations to study this phenomenon in a two-layer linear feedforward network. Specifically, in a continual online learning scenario, we study the drift induced by the noise inherent in the stochastic gradient descent (SGD). By decomposing the learning dynamics into the normal and tangent spaces of the minimum-loss manifold, we show the former corresponds to a finite variance fluctuation, whereas the latter could be considered as an effective diffusion process on the manifold. We analytically compute the fluctuation and the diffusion coefficients for the stimuli representations in the hidden layer as functions of network parameters and input distribution. Further, consistent with experiments, we show that the drift rate is slower for a more frequently presented stimulus. Overall, our analysis yields a theoretical framework for better understanding of the drift phenomenon in biological and artificial neural networks.

## PUBLICATIONS

Pashakhanloo F, Koulakov AA. 2023. Stochastic gradient descent–induced drift of representation in a two-layer neural network. *Proc 40th Intl Conf Machine Learning* **202:** 27401–27419.

Zador A, Escola S, Richards B, Öveczky B, Bengio Y, Boahen K, Botvinick M, Chklovskii D, Churchland A, Clopath C, et al. 2023. Catalyzing next-generation artificial intelligence through NeuroAI. *Nat Commun* **14:** 1597. doi:10.1038/s41467-023-37180-x

*In Press*

Shuvaev S, Amelchenko E, Smagin D, Kudryavtseva N, Enikolopov G, Koulakov AA. 2024. A normative theory of social conflict. Article 1975, pp. 45581–45601. In *Proceedings of the 37th International Conference on Neural Information Processing Systems* (NIPS '23). Curran Associates, Red Hook, NY.

# THE FUNCTION AND PLASTICITY OF NEURAL CIRCUITS IN THE BRAIN IN MOTIVATED BEHAVIORS RELATED TO PSYCHIATRIC AND METABOLIC DISORDERS

**B. Li** D. Chen M. Liu D. van de Lisdonk
W. Guan L. Ramirez Sanchez Y. Wang
L. Kimoto Q. Sun H. Zhang

The focus of research in our laboratory has been to understand the link between neural circuits and behaviors. We are particularly interested in studying the synaptic and circuit mechanisms underlying motivated behaviors and brain–body interactions, as well as synaptic and circuit dysfunctions that may underlie the pathophysiology of mental disorders and metabolic disorders, including anxiety disorders, depression, autism, drug addiction, obesity, and cancer-associated cachexia. We integrate in vitro and in vivo electrophysiology, imaging, molecular, genetic, optogenetic, and chemogenetic methodologies to probe and manipulate the function of specific neural circuits in the rodent brain and to determine their roles in adaptive or maladaptive behaviors and metabolic processes. We have made the following major progress in the past year.

## Area Postrema Neurons Mediate Interleukin-6 Function in Cancer Cachexia

Interleukin-6 (IL-6) has long been considered a key player in cancer cachexia. It is believed that sustained elevation of IL-6 production during cancer progression causes brain dysfunctions that ultimately result in cachexia. However, how peripheral IL-6 influences the brain remains poorly understood. In a recent study (Sun et al. 2024), we showed that neurons in the area postrema (AP), a circumventricular structure in the hindbrain, are a critical mediator of IL-6 function in cancer cachexia in male mice. We find that circulating IL-6 can rapidly enter the AP and activate neurons in the AP and its associated network. Peripheral tumor, known to increase circulating IL-6, leads to elevated IL-6 in the AP and causes potentiated excitatory synaptic transmission onto AP neurons and AP network hyperactivity. Remarkably, neutralization of IL-6 in the brain of tumor-bearing mice with an anti-IL-6 antibody attenuates cachexia and the hyperactivity in the AP network and markedly prolongs life span. Furthermore, suppression of *Il6ra*, the gene encoding the IL-6 receptor, specifically in AP neurons with CRISPR-dCas9 interference, achieves similar effects. Silencing Gfral-expressing AP neurons also attenuates cancer cachectic phenotypes and AP network hyperactivity. Our study identifies a central mechanism underlying the function of peripheral IL-6, which may serve as a target for treating cancer cachexia.

## The Central Amygdala Encodes Nutritional Properties and Controls Weight Gain

The perception of food-associated sensory input such as sight, smell, and taste can not only elicit consummatory behaviors but also trigger physiological responses in peripheral organs to prepare for and coordinate the digestion of foods and metabolization of nutrients. Dysregulation of these behavioral and physiological processes may cause metabolic disorders, including obesity. Although much is known about how neural circuits in the brain regulate feeding behaviors in both nonobese and obese conditions, how the brain orchestrates peripheral physiological processes in such conditions remains less understood. In a recent study (Boyle et al. 2024), we showed that the activity of protein kinase C, delta (Prkcd) neurons in the central amygdala encodes the nutritional properties of foods and exhibits obesity-associated adaptations that promote the buildup of body fat through a metabolic program. We found that these Prkcd neurons respond robustly to highly caloric foods and can accurately distinguish between them. Exposure to a high-fat diet increases the excitability, baseline activity, and food response of these neurons. Furthermore, suppressing these neurons prevents weight gain, whereas

increasing their excitability exacerbates obesity on the high-fat diet. These effects are independent of feeding behavior. Rather, they can be accounted for by profound and largely opposing changes in adipose tissue metabolic activity. Our results reveal a previously unknown link between the central amygdala and peripheral metabolism and suggest that the Prkcd neurons facilitate lipid biosynthesis and fat storage when there is a surplus of energy intake, thereby promoting weight gain.

## PUBLICATIONS

*In Press*

Boyle S, Liu M, Subhash S, Furlan A, Sharma R, van de Lisdonk D, Nunes Violante S, Yang T, Hwang G-R, He M, et al. 2024. The central amygdala encodes nutritional properties and controls weight gain. *Nature* (in press).

Sun Q, van de Lisdonk D, Ferrer M, Gegenhuber B, Wu M, Park Y, Tuveson DA, Tollkuhn J, Janowitz T, Li B. 2024. Area postrema neurons mediate interleukin-6 function in cancer-associated cachexia. *Nat Commun* **15:** 4682. doi:10.1038/s41467-024-48971-1

# THE STUDY OF INTELLIGENT MACHINES

**P.P. Mitra** S. Banerjee B. Harman C. Mezias J. Pane
P. Flannery L. Kimoto T. Miskic M. Richman
A. Gonzalez X. Li J. O'Rourke S. Savoia

## BICCN (BRAIN INITIATIVE CELL CENSUS NETWORK) CONSORTIA

BICCN Morphology and Wiring Diagram Manuscript Working Group [in collaboration with H. Dong and D. Tward, UCLA; G. Ascoli, GMU; H. Peng, Allen/SEU; other working group members and collaborators, including P. Osten, formerly CSHL]

The Mitra laboratory has contributed to several BICCN collaborative efforts, including a BICCN working group publishing novel and seminal results advancing cellular neuroscience and neuroanatomy in a package arranged with the Nature Publishing Group. In addition to taking a leading role in the overall consortium and in writing the flagship manuscript, Dr. Mitra and Mitra laboratory personnel and close collaborators are involved in multiple consortium projects, including (1) a new image and volume registration pipeline, based on generative diffeomorphic mapping; (2) a next-generation mouse whole-head CCF/atlas; and (3) creating neuroanatomical and ontological translations between homologous mouse and marmoset brain structures.

### A Novel Image and Volume Registration Algorithm and Pipeline

This work was done in collaboration with D. Tward (UCLA).

Mapping molecular, anatomical, and histological data sets from 2D microscopy images into common reference spaces remains a key challenge and a major focus of research in neuroscience. Although several brain-to-atlas mapping workflows exist, they do not fully address challenges of modern high-throughput neuroimaging—including multimodal and multiscale signals, missing data, or nonreference signals—and geometric quantification of individual variation. We address these issues, in conjunction with Dr. Daniel Tward (UCLA), by implementing a generative statistical model that describes the likelihood of imaging data given a sequence of transforms of an atlas image and a framework for maximum a posteriori estimation of unknown parameters capturing the issues listed above. The key idea in our approach is to minimize the difference between synthetic image volumes and real data over these parameters. We are able to quantify technical and biological sources of geometrical variation in an unprecedented manner.

We have now applied this framework across a broad range of data sets including various combinations of in vivo and ex vivo magnetic resonance imaging (MRI), 3D STP and fMOST data sets, 2D serial histology sections, and brains processed for single-nucleus RNA sequencing (snRNA-seq) with tissue partially removed. Working with Dr. Tward, we demonstrated biological utility by quantifying cell density across microscopy image data sets containing green fluorescent protein (GFP) nuclear labels on subsets of GABAergic interneurons. We demonstrate the accuracy of our atlas mapping and density estimation procedures by comparing with previously published data. To facilitate community accessibility, our algorithm is open-source, includes a web-based framework, and has defined and documented input and output data set standards. Our work establishes a quantitative, scalable, and streamlined workflow for unifying a broad spectrum of multimodal whole-brain light-microscopic data volumes into a coordinate-based atlas framework. This work enables large-scale integration of whole-brain data sets that are essential in modern neuroscience.

### A Next-Generation, Multimodal, Whole-Head Mouse Reference Space and Atlas

This work was done in collaboration with D. Tward (UCLA), A. Koretsky (NIH), and J. Zhang (NYU).

Our next-generation, whole-head mouse reference space has several advantages over existing common

coordinate frameworks (CCFs) and atlases, best understood by summarizing our approach in constructing this novel atlas (Fig. 1A). We start by acquiring whole-head in vivo T2 and diffusion-weighted magnetic resonance (MR) images, followed by ex vivo scans of the same modalities and ex vivo computed tomography (CT) (Fig. 1A). We then decalcify the skull and section the entire cranium to produce whole-head Nissl and myelin series (Fig. 1A). We use both radiological and histological as well as in vivo and ex vivo imaging modalities for the entire head, which has not been done before. Because all imaging, including Nissl and myelin histology, are performed on the whole head, we also possess intact out-of-brain features, such as reconstructable peripheral nerves, like the trigeminal nerve (Fig. 1D). This will permit reconstruction of facial musculature and sensors as well. We use CT scans to fit bregma and the tangent plane, allowing us to impose a stereotaxic origin and orientation onto our reference space and all volumes or microscopy image stacks mapped with this reference space (Fig. 1B). We plan to release histological and up-sampled radiological volumes, reconstructed at 20 μm isotropic resolution, with the stereotaxic

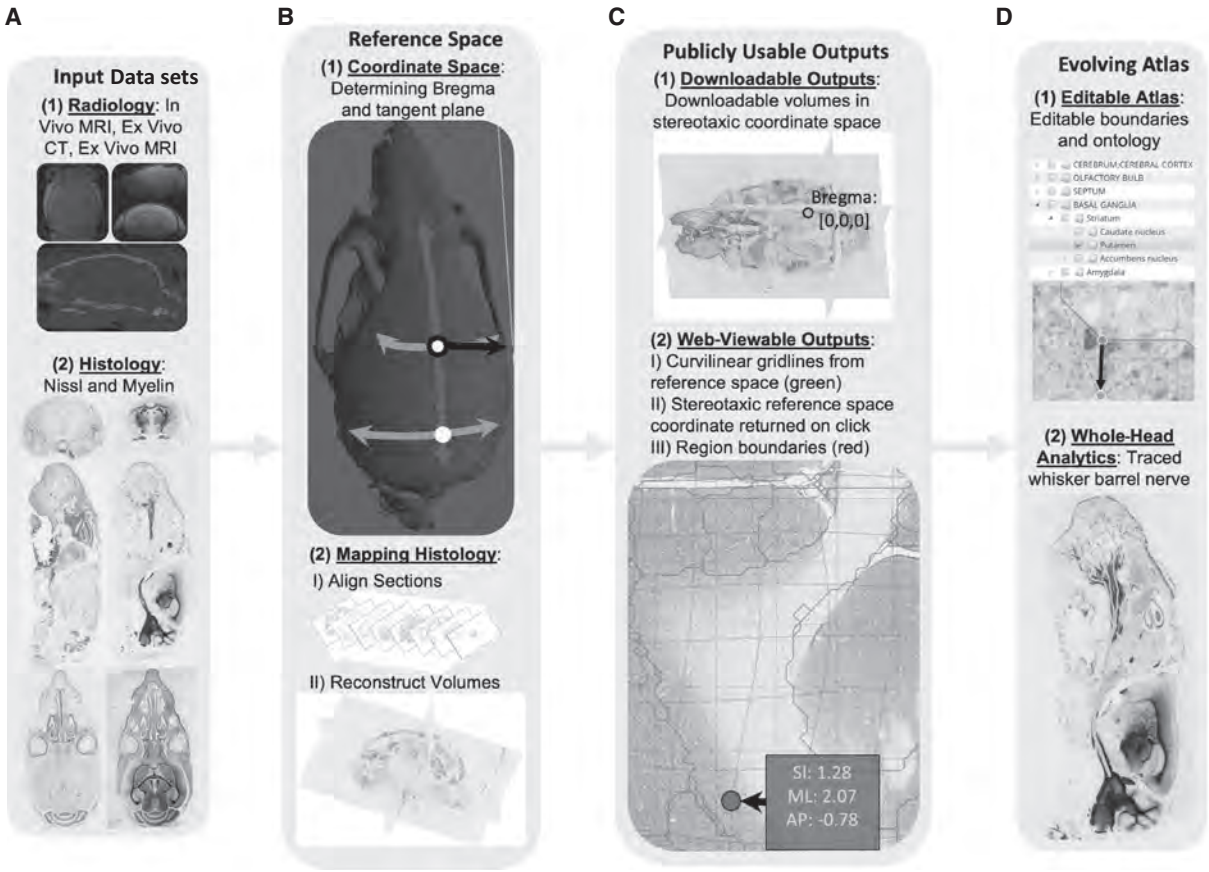



**Figure 1.** The pipeline for generating the next-generation mouse common coordinate framework (CCF) resource. (*A*) In the same animals, whole-head in vivo and ex vivo MRI, CT, and alternating Nissl and myelin histological series are acquired. (*B*) Bregma locations are fit along the sutures of the skulls in a CT scan, which is then registered with the average template volume. This reference volume, embedded within a stereotactic coordinate system, forms the basis of this atlas resource. Ex vivo MRI series and Nissl and myelin histological series from whole-brain and whole-head data sets are atlas-mapped into this reference space. (*C*) Atlas-mapped data sets in stereotactic coordinate space are released on the Brain Architecture web portal, including the averaged reference volume, aligned 2D histological sections with overlaid curvilinear coordinate grids and Allen region boundaries, and 20-μm isotropic ex vivo MRI and histological volumes reconstructed in reference space. (*D*) An editable atlas framework (EAF) is established, allowing for user-contributed ontologies to be loaded and registered and mapped segmentations to be edited, and demonstrates segmentation of biological artifacts, such as a peripheral nerve from a myelin section.

coordinate grid with bregma as origin (Fig. 1C). We will further disseminate our new atlas interactively via the Brain Architecture data portal and the associated high-resolution 2D section viewer. In addition to displaying overlaid aligned adjacent Nissl and myelin sections, we have overlays for the curvilinear gridlines derived from the 3D stereotaxic reference space and region boundaries from the initial test segmentation used, as well as the ability to query and return 3D atlas coordinates from 2D sections by mouse click (Fig. 1C). We therefore fill gaps in the technical abilities of current atlases by creating a properly multimodal reference resource embedded within a true stereotaxic coordinate space.

Conceptually, the most widely used atlases impose their particular version of region boundary segmentations onto their aligned histological and other microscopy sections and, in the case of the Allen Institute, their reconstructed volumes. These imposed image and volume segmentations themselves derive from a predefined brain region ontological tree and label set. Our new mouse brain reference space resource eschews this top-down approach in favor of a flexible, evolvable, and, eventually, user-defined set of region boundary segmentations and names. We are currently building (Fig. 1D) a platform that will allow users to edit both the hierarchical nomenclature and region boundaries. We also link with a novel online image registration platform (described in the previous section) to allow users to register their radiological volumes or 2D microscopy with the averaged T2W in vivo reference space embedded within stereotaxic coordinates (Fig. 1D). This new multimodal mouse CCF will help move the field away from top-down atlas resources toward what we are terming an evolvable atlas framework.

### Reconciling Neuroanatomical Nomenclature between Mouse and Marmoset

This work was done in collaboration with B. Huo (Broad).

To perform comparative neuroanatomical analyses of mouse and marmoset brain structures and axonal projections, it is necessary to establish a homological correspondence between the mouse and marmoset brain atlases. The rodent and primate radiations are separated by more than 30 million years of evolution, so this is a challenging task. Expert neuroanatomists and atlases disagree as to how to parcellate and name brain compartments in the mouse and marmoset. These disagreements exist both within species as well as across taxa. We found that there is often good agreement about how to draw compartment boundaries at the finest levels of segmentation of the brains (the "leaf level" in a hierarchical tree-like segmentation). The disagreements are of two sorts: how to name these leaf-level structures, and how to group the leaf-level structures into larger, hierarchical groups. We propose that it is better to not focus on the name/label hierarchies, but to establish the correspondences between the leaf-level structures across atlases with a species, and subsequently across taxa. We showcase this approach in Table 1.

## WHOLE-BRAIN, CELL TYPE–SPECIFIC PROJECTION MAPPING IN MARMOSET

Two National Institutes of Health (NIH)-funded projects, including (1) development of novel tools to probe cell-specific and circuit-specific processes in human and nonhuman primate brain [in collaboration with G. Feng, MIT; F. Krienan, Princeton]; and (2) comprehensive regional projection map of marmoset with single axon and cell type resolution [in collaboration with G. Feng, MIT; X. Wang, Broad].

### Development of Novel Tools to Probe Cell-Specific and Circuit-Specific Processes in Human and Nonhuman Primate Brain

This work was done in collaboration with G. Feng and I. Wickersham (MIT) and F. Krienan (Princeton).

Two years ago, we began a project in collaboration with Drs. Feng and Wickersham (MIT) and Dr. Krienan (Princeton) to develop enhancer-based cell type–specific viral reagents in marmoset and subsequently to obtain the first cell type–specific whole-brain projections in the marmoset. In the previous year, in addition to processing seven marmoset brains through the histological pipeline in our laboratory demonstrating successful tracer expression, we obtained our first successful cell type–specific enhancer injection data sets using Tac3 enhancer and S5E2 (parvalbumin) interneuron-specific enhancers. We have also initiated analytics on these enhancer data sets, including cell body and process detection. Additionally, we are mapping these microscopy data sets and analytic results to our novel marmoset atlas.

**Table 1.** Leaf-level structures assigned to high-level regions differently in the two atlases, compared across mouse and marmoset (Allen mouse vs. Paxinos marmoset)

| Type | Acronym | Marmoset region name | Cerebrotype | Acronym | Mouse region name | Cerebrotype |
|---|---|---|---|---|---|---|
| 1 | SNC | Substantia nigra; compact part | TEL | SNc | Substantia nigra, compact part | MID |
| 1 | SNL | Substantia nigra; lateral part | TEL | SNl | Substantia nigra, lateral part | MID |
| 1 | SNR | Substantia nigra; reticular part | TEL | SNr | Substantia nigra, reticular part | MID |
| 1 | STh | Subthalamic nucleus | TEL | STN | Subthalamic nucleus | DIEN |
| 1 | E | Ependyma & subependymal layer | TEL | SEZ | Subependymal zone | CvO |
| 1 | APT | Anterior pretectal nucleus | DIEN | APN | Anterior pretectal nucleus | MID |
| 1 | MPT | Medial pretectal area | DIEN | MPT | Medial pretectal area | MID |
| 1 | OT | Nucleus of the optic tract | DIEN | NOT | Nucleus of the optic tract | MID |
| 1 | OPT | Olivary pretectal nucleus | DIEN | OP | Olivary pretectal nucleus | MID |
| 1 | PrC | Precommissural nucleus | DIEN | PRC | Precommissural nucleus | MID |
| 2 | MCPC, PCom | Magnocellular nucleus of the posterior commissure, nucleus of the posterior commissure | DIEN, MID | NPC | Nucleus of the posterior commissure | MID |

(TEL) Telencephalon, (DIEN) diencephalon, (MID) midbrain.

## Comprehensive Regional Projection Map of Marmoset with Single Axon and Cell Type Resolution

This work was done in collaboration with G. Feng (MIT) and X. Wang (Broad).

In collaboration with Drs. Feng (MIT) and Wang (Broad) as part of the new Brain Connects program, we are establishing a pipeline to produce, analyze, and assess cell type–specific Starmap-barcoded viral tracer injection data in whole marmoset brains. This project will significantly advance nonhuman primate (NHP) anatomical knowledge. This project offers two technical leaps: (1) a large set of individually reconstructed axons spanning the entire brain, thus going beyond a simple connectivity matrix and delineating the collateral branching patterns; and (2) the ability to categorize individual neurons into specific types using the expression of a panel of approximately 1,000 genes. These individual reconstructions, obtained from systemic injections of a barcoded virus cocktail that crosses the blood–brain barrier to sparsely label a set of neurons, will be validated by a gold-standard data set employing tracer injections on a grid of injection sites. Thus far, we have completed an injection plan for both Starmap and comparison tracer injection data sets and initiated barcoded virus-infected neuron tracing in an example Starmap data set. We expect to advance this project significantly over the next year.

## POSTMORTEM HUMAN WHOLE-BRAIN HISTOLOGY

This work was done in collaboration with L. Latour (NINDS), D. Nauen (JHU), J. Zhang, D. Novikov, E. Fieremans, and T. Shepherd (NYU), M. Sivaprakasam (IIT Madras), and D. Tward (UCLA).

Last year, we reported in collaboration with David Nauen at Johns Hopkins Medical Institute and Jiangyang Zhang, Dmitry Novikov, Els Fieremans, and Timothy Shepherd at New York University that we have produced an unprecedented multimodal data set of two human hippocampi, two human amygdalas, a human brainstem, and human thalamus, as well as three hemibrains. T1, T2, and diffusion-weighted MR images were collected to reveal MR microstructure across all above-listed tissue samples. We have streamlined our tissue-processing pipeline, which has yielded improved histology results, as demonstrated with an example hemithalamus Nissl image (Fig. 2C). In the past year, we have now successfully co-registered alternating Nissl, myelin, and hematoxylin and eosin (H&E) histology data series with MRI volumes of not only the human hippocampi and amygdala, but also our hemithalamus data set (Fig. 2D). Nissl cells were segmented using machine-learning methods previously developed in this same hemithalamus data set (Fig. 2E). We have also now begun sectioning, staining (Nissl, myelin, and H&E), and imaging our

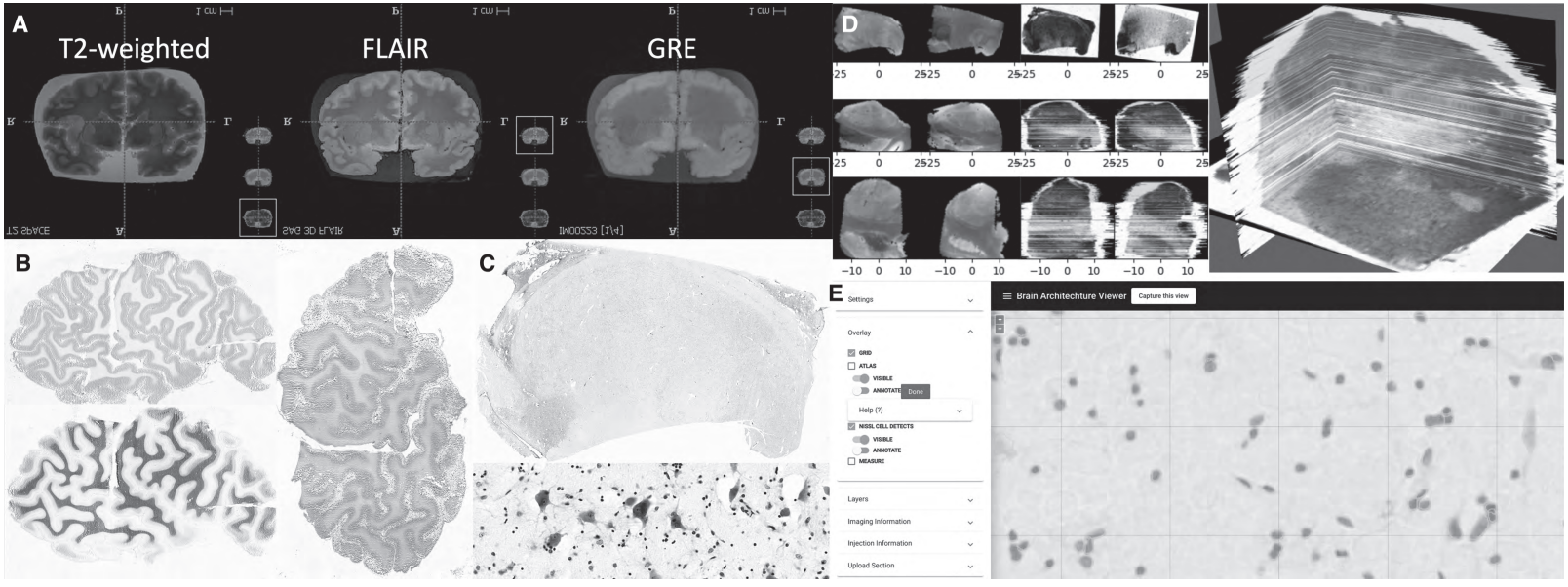



**Figure 2.** Human hemibrain and whole-brain data collection and compartment sample registration and analytics. (*A*) Whole postmortem human brain data set MRI modalities. (*B*) Hemibrain Nissl, H&E, and myelin histological data. (*C*) Nissl histology example from a human hemithalamus, including a high-resolution zoom-in. (*D*) Co-registration of histology and MRI in a human hemithalamus. (*E*) Segmented Nissl cells from a human hemithalamus displayed using the new Brain Architecture viewer.

first hemibrain and whole postmortem human brain data sets, of which we show example images (Fig. 2B). As detailed above regarding the brain compartment samples, we also obtained MRI modalities on these hemi- and whole human brain samples (Fig. 2A). We aim to complete imaging and processing of at least one hemibrain and one whole-brain data set this year; once completed, these will represent the first whole-brain multimodal MRI and histological image series collected at our currently unprecedented sampling resolution (20 microns between planes and 0.46 microns within-plane resolution).

In addition to processing and analyzing human tissue samples, we have developed automated and semi-automated tissue processing technology for large tissue samples. We have automated stainers (for Nissl, myelin, and H&E) and coverslippers for large-format glass slides for large tissue microscopy. We have also created a 3D printer bank for printing molds for human and other large animal brain tissue samples for radiological (MR and CT) imaging, tissue slabbing for further histological processing, and tissue freezing. We are working on packaging schematics, instructions, and code for recreating and using these novel tools, and we are releasing these as research products from the adult human brain mapping project. We are also creating a modified web viewer to accommodate large images such as those obtained from microscopy on human tissue samples.

## ARTIFICIAL INTELLIGENCE WITHOUT NETWORKS

This work was done jointly with C. Sire (CNRS and Université de Toulouse III) and P. Sabatier (France).

Contemporary artificial intelligence (AI) stands on two legs: large training data corpora and many-parameter artificial neural networks (ANNs). The data corpora are needed to represent the complexity and heterogeneity of the world. The role of the networks is less transparent because of the obscure dependence of the network parameters and outputs on the training data and inputs. This raises problems, ranging from technical-scientific to legal-ethical. We hypothesize that a transparent approach to machine learning is possible without using networks at all. By generalizing a parameter-free, statistically consistent data interpolation method, which we analyze theoretically in detail, we develop a network-free framework for AI incorporating generative modeling. We demonstrate this framework with examples chosen from three disciplines—ethology, control theory, and mathematics. We apply our generative Hilbert framework to the trajectories of small groups of swimming fish. The framework outperformed previously developed state-of-the-art traditional mathematical behavioral models and current ANN-based models in reproducing naturalistic behaviors. We demonstrate pure data-driven control without a system model in a canonical use case by stabilizing an inverted pendulum. Finally, we present a mathematical application by predicting zeros of the Riemann zeta function, and show results comparable to or outperforming a transformer network. We do not suggest that the proposed framework will outperform networks in all applications, as overparameterized networks can interpolate. However, our framework is theoretically sound, transparent, deterministic, and parameter-free: It does not require any compute-expensive training, does not involve optimization, has no model selection, and is easily reproduced and ported. We also propose an easily computed method of credit assignment based on this framework that could help address ethical-legal challenges raised by generative AI.

### PUBLICATIONS

Hawrylycz M, Martone ME, Ascoli GA, Bjaalie JG, Dong H-W, Ghosh SS, Gillis J, Hertzano R, Haynor DR, Hof PR, et al. 2023. A guide to the BRAIN Initiative Cell Census Network data ecosystem. *PLoS Biol* **21:** e3002133.

James RI, Verma R, Johnson LR, Manesh A, Jayakumar J, Sen M, Joseph J, Kumarasami R, Mitra PP, Sivaprakasam M. 2023. A standardized protocol for the safe retrieval of infectious postmortem human brain for studying whole-brain pathology. *Am J Forensic Med Pathol* **44:** 303–310.

Jorstad NL, Close J, Johansen N, Yanny AM, Barkan ER, Travaglini KJ, Bertagnolli D, Campos J, Casper T, Crichton K, et al. 2023. Transcriptomic cytoarchitecture reveals principles of human neocortex organization. *Science* **382:** eade9516.

Karthik S, Joseph J, Jayakumar J, Manoj R, Shetty M, Bota M, Verma R, Mitra P, Sivaprakasam M. 2023. Wide field block face imaging using deep ultraviolet induced autofluorescence of the human brain. *J Neurosci Meth* **397:** 109921.

Kumarasami R, Verma R, Pandurangan K, Ramesh JJ, Pandidurai S, Savoia S, Jayakumar J, Bota M, Mitra P, Joseph J. 2023. A technology platform for standardized cryoprotection and freezing of large-volume brain tissues for high-resolution histology. *Front Neuroanat* **17:** 1292655.

Langlieb J, Sachdev NS, Balderrama KS, Nadaf NM, Raj M, Murray E, Webber JT, Vanderburg C, Gazestani V, Tward D, et al. 2023. The molecular cytoarchitecture of the adult mouse brain. *Nature* **624:** 333–342.

Mohan H, An X, Xu XH, Kondo H, Zhao S, Matho KS, Wang B-S, Musall S, Mitra P, Huang ZJ. 2023. Cortical glutamatergic projection neuron types contribute to distinct functional subnetworks. *Nat Neurosci* **26:** 481–494.

# EXPERIENCE AND GENETIC FACTORS UNDERLYING CORTICAL NETWORK DEVELOPMENT

**G. Pouchelon** D. Dumontier D. Joseph A. Mirow D. Sanwo
D. Edelman S. Liebman A. Orozco

Multiple gene mutations have been linked to mental disorders. Nonetheless, environmental experience during development is considered as a modulator of the penetrance of susceptibility genes associated with neurodevelopmental disorders such as autism. Neurons in the neocortex receive activity through long-range presynaptic inputs that communicate information from the environment. In particular, extrinsic sensory inputs and intrinsic molecular programs have been shown to synergistically regulate neuronal maturation of sensory cortical areas. However, growing understanding of the cellular and functional diversity of the cortex unveils a much more complex picture of the balance between neural activity and genetic programs of development than previously thought. Although psychiatric symptoms are the main hallmarks of neurodevelopmental disorders, the role of experience beyond sensory activity, such as cognition and association processes, in cortical circuit development is unknown. Moreover, neuron network impairments in neurodevelopmental disorders are commonly investigated in adult models of the disease, and this overshadows primary dysfunctions developing at the onset of the disorders.

The overarching goal of our laboratory is to investigate early developmental mechanisms controlling neural network development in light of the cellular and functional diversity of the neocortex and how they are impaired in neurodevelopmental disorders. To bridge the behavioral, physiological, and molecular aspects of cortical development, we are establishing high-resolution and large-scale multifaceted approaches, currently including adeno-associated virus (AAV)-based combinatorial neuron targeting, mouse genetics, CRISPR editing, electrophysiology, functional ultrasound imaging, and transcriptomics. Our multidisciplinary approach allows for an integrated examination of circuit development and a better understanding of neural circuit formation at multiple scales. We are (1) examining early input processes underlying cortical development, (2) investigating their implications in neurodevelopmental disorders, (3) developing functional ultrasound imaging to investigate early brain patterns of activity, and (4) characterizing nonsensory, cognitive factors regulating cortical development.

## Noncanonical Activity-Dependent Development of Sensory Cortical Networks

D. Dumontier, S. Liebman, A. Mirow, D. Joseph [in collaboration with G. Fishell and B. Datta, Harvard Medical School, Boston]

Thalamocortical inputs are known to transmit sensory experience to the cortex, and they form transient connections with a population of inhibitory neurons—somatostatin (SST) neurons—during development. This transient connection has been shown to control the maturation of excitation/inhibition (E/I) balance in the cortex and therefore acts as a key trigger of cortical maturation. Using combinatorial mouse genetics and viral approaches, we found that transient thalamocortical inputs to SST neurons are controlled by a noncanonical activity-dependent mechanism that involves G-protein-coupled receptor (GPCR) signaling (Fig. 1). Disruptions in either global metabotropic engineered receptor (DREADD) signaling or the specific metabotropic glutamatergic receptor 1 (mGluR1) perturb the formation of transient connectivity. We further investigated the regulation of select candidates downstream of metabotropic signaling and their functional consequences for transient circuits and behaviors, using combinatorial CRISPR-Cas9 and viral approaches. Overall, this study provides a novel mechanism by which specific inhibitory neuron networks mature, underlining the importance of better understanding of development specific to cell types and temporal stages (Dwivedi et al. 2023).

We are building upon these findings to further investigate the downstream mechanisms of metabotropic-dependent pathways controlling thalamocortical transient connectivity. Toward this aim, we performed single-cell RNA sequencing of SST neuron

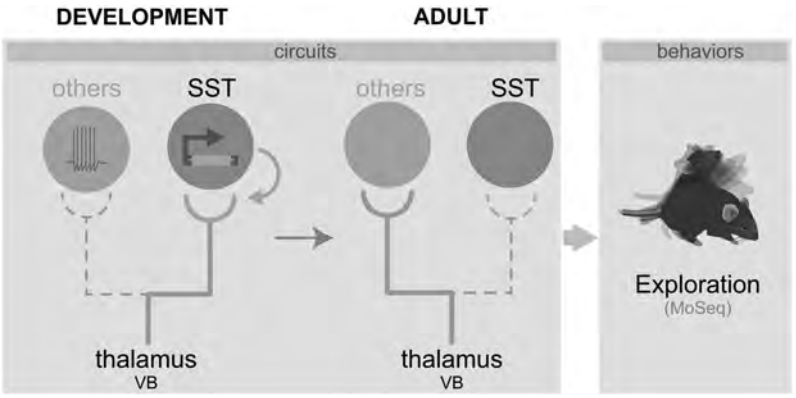



**Figure 1.** Model for the development of transient connectivity (TC). In contrast to other neuron types, to which thalamocortical connectivity develops according to Hebbian mechanisms, somatostatin (SST) inhibitory interneurons receive transient thalamocortical inputs from the ventral-basal (VB) thalamus. These TC inputs regress onto SST neurons following a non-Hebbian mechanism and involve postsynaptic metabotropic-dependent transcriptional regulation. This process is required for the maturation of exploratory behaviors in adult mice.

population upon DREADD expression. Following our current analysis, we will select differentially expressed genes upon metabotropic signaling disruption for further CRISPR-based loss-of-function characterization. Our long-term ambition is to uncover molecular candidates for modulating or rescuing transient circuit defects in neurodevelopmental disorders.

## Improving Fragile X Syndrome with the Restoration of Transient Dynamics of Early Cortical Circuits

D. Dumontier, S. Liebman [in collaboration with L. Van Aelst, CSHL]

Although early postnatal cortex dysfunctions have long been investigated, to date no early postnatal interventions have been successfully redirecting maturation trajectory and restoring adult functions. We hypothesize that disruptions in transient connectivity could be involved in the etiology of neurodevelopmental disorders. We are testing this hypothesis in a mouse model of the Fragile X syndrome (FXS), which is the most common monogenic cause of autism spectrum disorders. We found that transient connectivity from thalamocortical inputs to SST neurons is persistent in the model of FXS and that we can regenerate the dynamics of this connectivity using DREADDs. We are currently investigating whether restoring transients in thalamocortical inputs onto SST neurons further improves broader cortical maturation and sensory-associated mouse behavior. This study shows that development is not linear and early development events are essential to brain circuit formation. As our project above demonstrates that molecular mechanisms involved in early circuit plasticity are distinct from those described at adulthood, this study could shed light on temporally controlled mechanisms relevant to FXS, offering insights for the prospective development of age-specific pharmaceutical therapies.

## Establishing In Vivo Imaging to Characterize Transient and Early Activity during Development

D. Dumontier [in collaboration with S. Kaushalya, CSHL]

Thalamocortical transient connectivity is involved in the maturation of local sensory circuits. However, to date, nothing is known about their role in the development of global network activity. Studies of developmental brain activity commonly employ two-photon calcium imaging. However, such techniques require invasive procedures and anesthesia known to deeply affect brain development. In collaboration with the director of the Neuroimaging and Behavior Core at CSHL, we achieved the unprecedented noninvasive recording of postnatal brain activity in awake neonate mice, using functional ultrasound imaging (fUSI). Compared to other live-imaging techniques applicable to neonates, our fUSI approach allows for intact and simultaneous brain imaging of superficial and deep regions of awake

animals. We are currently adapting our technique to multiple postnatal stages (from postnatal day 5 to 10) covering transient events and are improving the quality and resolution of ultrasound signal. Using fUSI, we will examine brain activity patterns upon chronic and acute manipulation of transient connectivity to SST neurons using DREADDs to uncover their role in the generation of early developmental network activity and subsequent maturation trajectory. Moreover, our technique will allow us to create a longitudinal brain atlas of early patterns of spontaneous neural activity, providing regional and temporal hallmarks of healthy brain development. A preclinical research study of human newborn brain abnormalities using fUSI was recently launched (CONEXUS project). However, for this approach to be successful for diagnosis of disorders such as autism, it is essential to reveal fUSI-based spatiotemporal brain activity dynamics in normal postnatal development and models of neurodevelopmental disorders. We hope to ultimately close this knowledge gap using fUSI.

## Characterizing Nonsensory, Cognitive Experience as an Environmental Factor Controlling Cortical Network Development

S. Liebman

The primary focus of developmental studies on sensory experience contrasts with the typical efforts made to study cognitive dysfunctions in neurodevelopmental disorders. Although it is accepted that genetic programs of differentiation are refined by sensory experience to form functionally mature neural circuits, the role of other cues, such as hormonal or mental states, with respect to the developing cortex is essentially unknown. We previously found that SST neurons, which receive transient thalamocortical inputs and act as a hub for downstream network development, receive inputs from the basal forebrain at early postnatal stages (Pouchelon et al., *Cell Rep* **37:** 109993 [2021]). Neurons in the basal forebrain are mainly cholinergic and release acetylcholine (ACh), which is associated with cognition and attention in adults. Previous studies suggest a role for ACh in synaptic maturation. However, little is known about the potential role of ACh as a factor driving early cortical maturation. We hypothesize that cognition-associated cholinergic inputs control transient connectivity and cortical maturation. We are currently investigating external cues potentially activating cholinergic basal forebrain neurons in neonates and developing novel molecular and viral tools to manipulate cholinergic inputs to the cortex at birth. We believe that this study will reveal ACh as a factor associated with higher brain functions and regulating early network development.

### PUBLICATION

Dwivedi D, Dumontier D, Sherer M, Lin S, Mirow AMC, Qiu Y, Xu Q, Liebman S, Joseph D, Datta SR, et al. 2023. Metabotropic signaling within somatostatin interneurons regulates thalamocortical inputs during development. bioRxiv doi:10.1101/2023.09.21.558862

# NEURAL CIRCUITRY FOR SOCIAL COMMUNICATION

**S. Shea** H. Ansari F. Duque S. Rahman
J. Choe I. Kouba C. Rodriguez-Saltos
K. Day A. Pagliaro

The ability to recognize and communicate with potential social partners is essential for survival. This skill enables organisms to secure mates, to judiciously allocate parental resources, to be attuned to friends in distress, to cooperatively identify palatable and nourishing food, and to remain collectively vigilant to looming threats. Individual participants in these objectives must continuously monitor and integrate relevant signals to dynamically update their immediate behavioral priorities. In the world of many nonhuman animals, failure to rapidly detect and accurately balance numerous multisensory pieces of social information can have grave consequences. Humans who contend with impairments in social communication and interaction commonly experience loneliness, alienation, depression, and anxiety. My research program has therefore been designed to answer two questions: First, what are the neural mechanisms that enable an individual to detect and interpret social signals and select appropriate actions? Second, what are the specific changes to neural circuitry that lead to impaired social interaction and communication in neurodevelopmental disorders and mental illness?

## SOCIAL COMMUNICATION IN MICE

In recent years, we have addressed these questions through a multilevel analysis of a fundamentally important and spontaneously expressed mouse maternal behavior: "pup retrieval." In this behavior, altricial pups call for their mother with ultrasonic distress vocalizations (USVs) when they become separated from the litter. In response, seasoned mothers locate the pup and return it to the nest. Pup retrieval presents several experimental advantages that make it well suited for answering our questions. First, it is overt, quantifiable, and robust. Second, despite its apparent simplicity, we are discovering that the neurochemical events that underlie this behavior are highly rich and complex. Third, pup retrieval crucially depends on integration of multiple social cues, including auditory (USVs) and olfactory (the smell of pups) stimuli. Fourth, as a social communication behavior, it is useful for probing deficits of social behavior in mouse models of autism spectrum disorders (ASDs). Fifth, and perhaps most important, mothers and other females co-housed with pups ("surrogates") improve dramatically in their performance with experience. Therefore, pup retrieval is an ethologically fundamental behavior with disease relevance that is discrete and easy to measure; it also has complex determinants and is shaped by an interplay of innate behavior programs and learning.

### Maternal Behavior and Auditory Processing Are Impaired in a Mouse Model of Rett Syndrome

A. Pagliaro [in collaboration with D. Rupert, Stony Brook University]

Rett syndrome (RTT) is a neurodevelopmental disorder related to autism that includes severe impairments in speech communication. RTT is caused by loss of one working copy of a single gene, *MECP2*. We have a long-term project on the neural circuitry of the auditory cortex and how it contributes to pup retrieval in wild-type (wt) mice, as well as how it is disrupted in *Mecp2* mutant mice. We have established that expression of *Mecp2* is acutely required in the auditory cortex for accurate pup retrieval learning, and that *Mecp2* mutants ($Mecp2^{het}$) exhibit elevated expression of perineuronal nets (PNNs) and parvalbumin (PV). These are markers associated with cortical inhibitory networks and regulation of plasticity, and we also discovered that genetic and pharmacological manipulations that reversed overexpression improved behavior. We subsequently linked the dysregulation of PV and PNN expression to altered in vivo electrophysiological dynamics. Motivated by this finding, over the past few years, we

conducted a series of experiments to test the hypothesis that *Mecp2* is especially important in the neurons that express PV and PNNs ($PV^+$ interneurons). This year we published the first of two new studies regarding the specific role of PV neurons and PNNs in the pathology of *Mecp2* mutants in the *Journal of Neuroscience* (Rupert et al. 2023). We anticipate releasing the second as a preprint in 2024.

Through the first study, led by graduate student Deborah Rupert, we have found that (1) among most interneurons, *Mecp2* is essential in only $PV^+$ interneurons for normal development of retrieval behavior; (2) deletion of *Mecp2* only in $PV^+$ interneurons reproduces the neural circuit changes seen in $Mecp2^{het}$; (3) longitudinal optical recordings from $PV^+$ neurons shows widespread disinhibition in maternal wt mice. Mutant mice did not exhibit these altered inhibitory dynamics. We conclude that deleting *Mecp2* only in $PV^+$ interneurons is sufficient to recreate nearly all the consequences of deleting *Mecp2* in all cells.

In the second study, graduate student Alexa Pagliaro measured and manipulated the broad spatial and temporal patterns of activity of the $PV^+$ interneuron population during free interactions with pups. Her surprising results reveal that the population activity of $PV^+$ interneurons in the auditory cortex is dominated by slowly changing influences such as behavioral state or arousal, as opposed to responding to sensory stimuli. Substantial evidence supports the idea that plastic changes in the auditory cortex that coincide with maternal experience are triggered by decreased output from $PV^+$ interneurons ("disinhibition"). However, it remains unclear whether this disinhibition is regulated in a moment-by-moment fashion or is sustained over long periods. We compared the effects on retrieval behavior of using chemogenetic tools to interfere with the output of $PV^+$ interneurons on different timescales. Alexa discovered that only if she disinhibited the neurons continuously for several days did she see increased maternal retrieval behavior. Intermittent manipulations of the $PV^+$ interneurons do not have this effect. Long-term suppression of the neurons also degraded the expression of PNNs in the auditory cortex, another hallmark of maternal experience–induced plasticity. We conclude that $PV^+$ interneurons regulate cortical plasticity over long periods, but not through activity on the timescale of behavior or sensory information.

## Multisensory Integration in Maternal Retrieval Behavior

H. Ansari, S. Rahman, J. Choe [in collaboration with A. Nowlan, UNC–Chapel Hill]

Successful navigation of social encounters requires integration of multimodal sensory events (e.g., sound, smell, touch) with emotional states. How our emotions may shape sensory responses to social cues remains unclear. We are working to determine how social odor cues activate the emotional circuitry of the basal amygdala to influence processing of social vocalizations by the auditory cortex. In 2022, graduate student Alexandra Nowlan released a preprint of her work (Nowlan et al., bioRxiv doi:10.1101/2022.02.17.480854) suggesting that pup odor and USVs are integrated via a pathway from odor-responsive neurons in the basal amygdala (BA) to the auditory cortex (AC). First, she found that a large, scattered population of excitatory neurons in the BA project directly to the AC. Second, using optical measurement of neural activity in awake mice, she found that AC-projecting BA neurons respond to pup odors and are active during search for pups in a retrieval assay. Third, she showed that optogenetic activation of AC-projecting BA neurons elicits dramatic restructuring of AC neurons' responses to sound. Finally, she showed that the effects of activating this BA → AC circuit are magnified and switch from primarily inhibitory to primarily excitatory after maternal experience.

The origin of the odor-signaling projection in the amygdala, a structure crucial for motivated behavior, raises the possibility that it carries affective information such as salience or reward. Therefore, moving forward, our efforts are focused on understanding the behavioral significance of responses to pup and other odors in the basal amygdala. There are several important aspects of our strategy. First, we find responses in the BA → AC pathway to a broad range of scents, including predator odors, palatable food odors, and novel odors. However, the methods we have used so far do not resolve how these different responses are distributed across the neuronal population. For example, do individual neurons respond to all of these odors or only a small fraction? Second, other neurons in the BA respond to the emotional significance and valence of stimuli. We intend to compare these properties with the response characteristics of BA → AC neurons, monitoring the dynamics of responses to

innate and conditioned repetitive and aversive odors to ascertain how they encode odor meaning.

Graduate student Hoda Ansari and technicians Jane Choe and Sadia Rahman are continuing this work with three important additions. First, Jane and Sadia are recording the activity of a different pathway originating in the BA and terminating in nucleus accumbens (BA → NAc). They are finding that, despite the fact that these two populations are intermingled, they carry complementary signals during pup retrieval. Unlike BA → AC, neurons in BA → NAc are activated on contact with a pup. Second, the team is also making use of a behavioral paradigm we recently developed called "cued retrieval." This approach allows us to segment the behavior into its component events so that we can separate the neural activity associated with each. Finally, Hoda has begun using miniature head-mounted microscopes to visualize individual neurons in the BA → AC circuit.

## Reinforcement Learning Mechanisms in Social Behavior

K. Day [in collaboration with Y. Xie, Chinese Academy of Sciences]

We performed recordings from neurons that release the neurotransmitter dopamine (DA) during pup retrieval. DA is critical for motivated behavior because it both stimulates movement and helps evaluate rewards. Central to dopamine's role in processing reward is its property of signaling "reward prediction error" (RPE). What this means is that DA does not signal reward *per se*, but rather the difference between expected and encountered rewards. Therefore, when an organism receives an unexpected reward, DA neurons will fire briskly, reinforcing the action that led to that reward. As the organism continues to be rewarded for that action, the reward becomes expected, and DA neurons fire less. RPE is important because it is theoretically optimally suited as a teaching signal to update the value of a reward in reinforcement learning. We developed a novel behavioral paradigm ("cued retrieval") in which a door opens, providing a maternal female mouse with access to a chamber. The task design allows us to manipulate the expectations of the subject for whether a pup will be available in the chamber, and then we can either meet or violate those expectations. We made the following observations: First, when the female mouse is introduced to a pup and performs more than 100 retrievals, we find that the change in approach velocity (update in value) is correlated with the previous trial's DA signal. Second, when the expectation of pup availability is not met, ventral tegmental area (VTA) neurons show a below-average response, and when a pup is found unexpectedly, the VTA neurons show a larger-than-average response. Third, we used temporally precise optogenetic inhibition of VTA triggered by pup proximity to cancel the DA response to pups, and we found that the mice were far slower to learn pup retrieval. We conclude that maternal care is shaped by a dopamine RPE signal for which the pup itself is the primary reward. This year, we published a paper reporting these findings in the journal *Neuron* (Xie et al. 2023).

In June, postdoctoral fellow Katherine Day joined the laboratory to continue this project. Initially, she is measuring DA release in several targets of the VTA to determine whether the projections carry shared or divergent content. She is also performing experiments to determine what aspects of pup contact or sensory attributes of the pup make this event rewarding to the mouse.

## Sex Differences in Sensitivity to Pup Distress

This work was done in collaboration with A.C. Corona (Friedman Brain Institute, Icahn School of Medicine at Mt. Sinai).

We recently examined sex differences in parenting behavior. In laboratory mice, under certain conditions, sires will participate in parenting pups, but they do so more slowly and less reliably. To investigate the neural circuits that underlie these differences, we used whole-brain imaging in male and female mice to screen for parental behavior–specific neural activity. In addition to many of the "usual suspects," one unexpected region to emerge from this screen was the anterior cingulate cortex (ACC), which has been linked to sensitivity to distress in social partners. We made optical recordings from the ACC in freely behaving mice and found that ACC neurons exhibit sex-dependent activity during retrieval. Deactivating the ACC disrupts attentive parenting and increases parental neglect. Moreover, we identified a projection from locus coeruleus (LC)

to ACC and showed that it is active and releases noradrenaline in the ACC during parental behavior. Inactivation of only the neurons in this projection also increased parental neglect. We conclude that ACC maintains sex-dependent sensitivity to pup distress under LC modulation. We speculate that ACC is an arbiter between selfish and selfless behavioral choices in executing innate parental care behaviors. These findings are reported in a paper we published this year in *Cell Reports* (Corona et al. 2023).

## PUBLICATIONS

Corona A, Choe J, Muños-Castañeda R, Osten P, Shea SD. 2023. A circuit from the locus coeruleus to the anterior cingulate cortex modulates offspring interactions in mice. *Cell Rep* **42:** 112771. doi:10.1016/j.celrep.2023.112771

Rupert DD, Pagliaro AH, Choe J, Shea SD. 2023. Selective deletion of *Methyl CpG binding protein 2* from parvalbumin interneurons in the auditory cortex delays the onset of maternal retrieval in mice. *J Neurosci* **43:** 6745–6759. doi:10.1523/JNEUROSCI.0162-23.2023

Xie Y, Huang L, Corona A, Pagliaro AH, Shea SD. 2023. A dopaminergic reward prediction error signal shapes maternal behavior in mice. *Neuron* **111:** 557–570.e7. doi:10.1016/j.neuron.2022.11.019

# HORMONAL REGULATION OF GENE EXPRESSION IN THE BRAIN

**J. Tollkuhn** N. Adam V. Kulik A. Vouzas
K. Denney S. Sun M. Wu

Our laboratory studies hormone receptor signaling in the brain. Gonadal steroid hormones such as estrogen and testosterone are the primary drivers of sex differences in neural physiology and behavior. These hormones principally act through their cognate nuclear receptors, which bind DNA to regulate gene expression. Our overarching goal is to identify the gene programs regulated by steroid hormones in the brain and determine how these genes influence brain development, behavior, and disease risk. We anticipate that hormone-responsive genes underlie sex differences in the incidence and etiology of psychiatric and neurological diseases.

## Neuroendocrine Refinement of Sex-Variable Neural Circuits during Adolescence

S. Sun

Adolescence is a neurodevelopmental critical period marked by increases in sex hormones, resulting in emotional, cognitive, and behavioral maturation. Our work recently established that sex hormones regulate gene expression to shape neurological sex variability during perinatal development and adulthood, but how such sex variability emerges during the pubertal hormone surge remains unexplored. To address this gap in knowledge, we assayed chromatin accessibility in estrogen receptor alpha (ERα) expressing neurons with ATAC-seq before (postnatal day 28, P28), during (P35), and after (P50) puberty in three brain areas known to exhibit sex variability and play a role in sex-variable behavior: the posterior region of bed nucleus of the stria terminalis (BNSTp), the posterior medial amygdala (MeA), and the medial preoptic area (MPOA). We found in the BNSTp that there is little sex variability in chromatin accessibility at P28 and that through puberty, sex variability emerges (P35), reaching adult-like levels at P50 (Fig. 1A). Of these sex-variable regions, those more accessible in male animals were enriched for the androgen receptor (AR) binding motif, indicating that the pubertal activation of the testes reorganizes neuronal chromatin through AR activation. Surprisingly, regions more accessible in females were not enriched for hormone-response elements and instead contained motifs for other transcription factors such as Egr1 and Stat2 (Fig. 1B). These results indicate that neurological sex differences in gene regulation emerge dynamically during puberty and are primarily mediated by androgenic signaling (testosterone and dihydrotestosterone) through AR. Indeed, preliminary experiments in prepubertally orchiectomized (OrX) animals (removal of the testes) indicate that the chromatin profiles of neurons from postpubertal OrX (P50) resemble those of intact males at P28 (Fig. 1D).

Our results suggest that pubertal hormone-dependent sex-variable gene regulation underlies the emergence of sex-variable behaviors during adolescence. They also imply that hormone-dependent gene regulation tunes the electrophysiological properties of hormone-sensitive neurons during puberty in a sex-variable manner. Indeed, gene ontology analysis reveals that sex-variable sites at P50 are enriched with genes important for neurophysiological functions, such as membrane potential regulation, ion transport, and various aspects of synaptic function (Fig. 1C). Some candidate genes are provided in Fig. 1D, with their gene product's function described. We have begun to explore the neurophysiology of hormone-sensitive neurons in the BNSTp with whole-cell patch-clamp electrophysiology, to link sex hormone–dependent gene regulation with pubertal sex-variable electrophysiological refinement.

## Mapping Brain Sexual Differentiation

N. Adam

Severe mental health disorders like autism spectrum disorders, schizophrenia, and major depressive disorder manifest major sex differences in their incidence and etiology. These conditions often involve profound social impairment, yet the origins of these

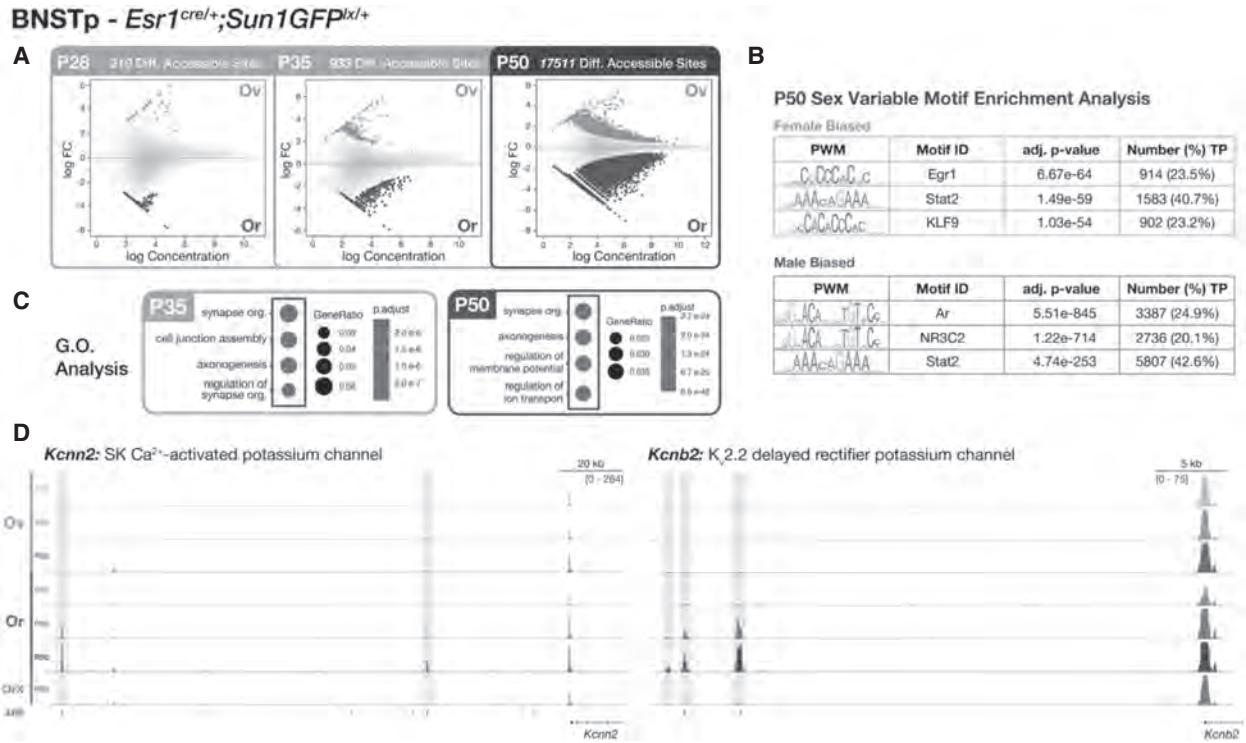



**Figure 1.** Dynamic chromatin reorganization during adolescence. (*A*) MA plots comparing sex-variable sites between animals with ovaries (Ov, female) and animals with testes (Or, male) at P28, P35, and P50 from Esr1+ neurons in the BNSTp (DiffBind, EdgeR; false discovery rate [FDR] < 0.05). (*B*) Top enriched motifs (AME) from sex-variable P50 female biased (*top*) and P50 male biased (*bottom*). (*C*) Top Gene Ontology biological process terms from sex-variable sites at P35 and P50. (*D*) Exemplar ATAC-seq tracks for two neurophysiologically relevant genes, *Kcnn2* and *Kcnb2*, from gonadally intact animals (Ov, ovaries, female; Or, testes, male) across pubertal time points and from prepubertal orchiectomized animals at P50 (OrX, orchiectomized).

social dysfunctions remain unclear, preventing effective diagnosis and treatment. Social behaviors are governed by the sexually dimorphic social behavior network (SBN), which is permanently organized by the testosterone surge occurring at birth exclusively in males. Testosterone acts through the ERα to establish sex differences in neural activity and social behaviors. Notably, the posterior bed nucleus of the stria terminalis (BNSTp) of males contains more neurons and sends stronger projections to multiple SBN regions. These major sex dimorphisms are thought to contribute to behavioral differences. However, it remains unclear how the perinatal actions of ERα create sex differences in cell number and connectivity that underlie male-typical patterns of social behavior.

Our recent work revealed the identity of a BNSTp cell type that is more abundant in males. The i1:Nfix/Moxd1 cluster is enriched in males and displays a transcriptional profile that resembles interneurons. Our aim is to leverage advanced molecular and genetic tools to characterize the role of this BNSTp$^{Nfix/Moxd1}$ population in shaping the male brain and mediating social behaviors. This includes the functional mapping of the BNSTp connections during development, with a particular focus on the Nfix/Moxd1 population. As a first step, we successfully targeted the BNSTp of newborn pups to inject cholera toxin B (CTB), a retrograde tracer labeling the projections received by all BNSTp cells. This foundational step sets the stage for identifying the specific projections sent and received by BNSTp$^{Nfix/Moxd1}$ cells. For this, we will leverage a novel Moxd1 Cre line, combined with Cre-dependent viral tracers, combined with iDISCO+ brain clearing and light-sheet microscopy to map BNSTp sexual differentiation. These methodological developments

provide a set of tools that will support our efforts to analyze at a circuit level how brain sexual dimorphism functionally translates into sex-specific behavioral outcomes.

## Defining Hormone-Mediated Aging Trajectories

V. Kulik

In addition to the ongoing work revealing the hormonal regulation of gene expression and resulting sex differentiation during critical neurodevelopmental periods like birth and puberty, the laboratory is expanding its techniques to later life time points. In humans, age-related decreases in sex hormones are linked to the onset of neurodegeneration in both males and females, although the prevalence and etiology of various neurodegenerative diseases differ substantially between the sexes. Thanks to the work published by the laboratory in Gegenhuber et al. (*Nature* **606:** 153 [2022]), we now possess the neural targets for ERα and AR. Knowing this, we can now go on to investigate whether distinct aging trajectories emerge in the brains of females and males during reproductive senescence and characterize such differences.

Our aging experiments implement a mouse model of accelerated ovarian failure (AOF) that replicates the hormonal dynamics of the human menopause transition (pre-, peri-, and postmenopause). AOF can be induced in adult mice after a 15-day treatment of a low dose of the ovotoxic chemical 4-vinylcyclohexene diepoxide (VCD). This model is particularly robust because it allows for the dissociation of the effects of hormone levels from the effects of aging. Mice do not naturally reach very low estrogen levels with age, and ovariectomy erases the perimenopause stage. VCD's effects mimic the progression from human peri- to postmenopause by resulting in decreased ovarian follicle count, increased periods of acyclicity, and eventually very low estrogen levels. Our goal is to capture the dynamics of gene expression during the critical transitions from pre- to peri- to postmenopause, which are not possible to glean from other mouse models. Ultimately, the goal is to discover potential biomarkers for hormone replacement therapies (e.g., testosterone, estradiol, estradiol and progesterone) that could potentially protect against age-related neurodegeneration.

## Evolution of the Social Brain

K. Denney

We expect that distinct expression patterns of hormone receptors underlie species differences in social behavior. Prairie voles (*Microtus ochrogaster*) are highly social, monogamous, mouse-sized rodents. Bonded prairie voles engage in prolonged physical contact with their mates (huddling) and parents of both sexes care for offspring. In addition to biparental care, both sexes of prairie vole show aggressive behavior toward novel conspecifics of either sex following pair-bond formation. Separation of bonded pairs causes symptoms consistent with partner loss in humans, including increases in stress hormones and passive coping behaviors. In contrast, meadow voles (*Microtus pennsylvanicus*), a related species, are solitary, do not form pair bonds, and only mothers care for offspring. As a first step in understanding how the brain forms social attachments, we investigated the expression patterns of androgen receptor (AR/Ar) and both estrogen receptors (ERα/*Esr1*; ERβ/*Esr2*) in the brains of mice, prairie voles, and meadow voles. We reasoned that these key drivers of innate behaviors may show unique expression patterns in prairie voles. Indeed, we find striking species differences in hormone receptor expression in the hippocampus and hypothalamus. We also observe reduced sexual dimorphism in the prairie vole brain, in line with the reduced behavioral dimorphisms in this species. These findings suggest that variation in expression of a single hormone receptor can change the context in which an innate social behavior is displayed.

### PUBLICATION

Denney KA, Wu MV, Sun SED, Moon S, Tollkuhn J. 2023. Comparative analysis of gonadal hormone receptor expression in the postnatal house mouse, meadow vole, and prairie vole brain. *Horm Behav* **158:** 105463.

# COMPUTING WITH NATURAL AND ARTIFICIAL NEURAL NETWORKS

**A. Zador** A. Benjamin, K. Daruwalla, Divyansha, G. Henry, E. Isko, S. Kerstjens, D. Maharjan, K. Matho, A. Medhat, C. Pehle, C. Soitu, L. Yuan, H. Zhan

Work in the Zador laboratory is divided into three main areas. First, we look at neural circuits underlying auditory decisions. Second, we are developing new technologies for sequencing the connectome. Finally, we are applying insights from neuroscience to artificial intelligence.

## Toward a Foundation Model of Single-Cell RNA Transcription for Neuroscience

A. Benjamin

Advances in single-cell RNA sequencing such as BARseq have revealed a dauntingly complex landscape of cellular identities in the mammalian brain. To make sense of this new scale of data, we develop computational approaches that leverage artificial intelligence to yield insights about the brain's cellular organization. Currently, we are taking a "foundation model" approach, which centers on large models trained on very large and highly diverse data sets—in essence, nearly all single-cell RNA sequencing data sets readily available in the public domain. We are working to understand what these models have learned about the structure of gene transcription, and what this might say about cell types and the organization of the brain. Already, we have shown that the brain area of cells can be better predicted from a foundation model than from standard methods based on cell types. We are working to leverage this capability to provide a tool for neuroscientists wishing to annotate the brain areas of new brains from transcription alone. Generally, AI foundation models promise to better link gene transcription to cellular function, providing a new "foundation" for studies of cellular identity in the brain.

## A Neuronal Development–Inspired Framework for Generating Networks with Learned Priors

K. Daruwalla

Deep neural networks are trained on billions of data samples, resulting in a time-consuming and computationally intensive process, whereas biological networks are able to learn within a few training examples by leveraging innate priors encoded in an organism's genome via evolution. This suggests a shared network structure across tasks that need not be relearned from scratch provided that it can be efficiently identified, stored, and retrieved. Several prior approaches explore this idea, but they either ignore or abstract away the details of evolution and gene transcription. Our work argues that the structure of the genome and how it is decoded make it well suited for generating functioning networks that perform well and adapt quickly. Specifically, genetic subprograms are executed repeatedly and compositionally during neuronal development. This iterative process allows evolution to compactly encode prior knowledge.

We seek to develop an analogous framework for artificial networks. This is achieved by leveraging denoising diffusion models (DDMs) from recent machine-learning literature. We generate exemplar networks that correctly perform a given task (i.e., a collection of "fit" individuals in evolution). Our model is trained to learn the shared structure across all exemplar networks such that we can generate novel networks that perform the original task well without any learning. Under our framework, we can train many DDMs for base networks across a variety of tasks, in which each DDM is analogous to a

genetic subprogram for the original network and task. Like gene transcription, the DDMs can be sampled iteratively and compositionally to generate a large variety of downstream networks. The final result is an approach to training artificial networks in a very sample-efficient manner. Instead of learning purely through experience, we first learn the optimal sequence of genetic subprograms, and then we train the generated networks using fewer training examples. We expect our approach to greatly reduce the energy cost of training state-of-the-art models.

## Structured Connectivity Induces Rapid Learning in Neural Networks

Divyansha

Animals are born with highly structured brain connectivity that endows them with innate behavioral capabilities. For example, a colt can stand, and a spider can hunt moments after birth. By contrast, most artificial neural networks are initialized tabula rasa, without any "innate" capabilities. Recently, there has been growing interest in exploring the capacities of artificial systems endowed with innate capabilities. Recently we proposed a method for compressing the wiring diagram of an artificial agent through a "genomic bottleneck" (Shuvaev et al. 2024). We are now extending these approaches to other architectures.

## Optimizations of the Viral System Used to Deliver Barcoded Libraries to Neurons

G. Henry

Our ability to understand the structure of neuronal circuits and the morphology of the neurons that populate such circuits was greatly accelerated by the discovery of fluorescent proteins, which enabled detailed structural analysis through the use of widely available and easy-to-deploy light microscopy methods. However, with a handful of notable exceptions, these techniques rely on bulk labeling of neurons. We have previously shown that it is possible to map the axonal projections of neurons at single-cell resolution through the use of nucleic acid–based barcodes, but an outstanding problem is how one might use a barcoding strategy to elucidate the complete structure or morphology of a neuron and more importantly that of a neuronal circuit. To achieve this goal, we have begun the process of optimizing the viral vectors that were previously developed for barcoding of axonal projections.

In the previously described MAPseq method, a strategy was developed that permits the transport of RNA barcodes to axonal terminals. We first set out to optimize this scheme by exploring alternative carriers. This involved a screen for novel carrier proteins capable of more efficiently transmitting barcodes from the soma of barcoded neurons to distant axonal termini. This effort produced two novel carrier proteins based on the presynaptic proteins VAMP2 and SNAP25. Both show a 2.5-fold enhancement in median barcode counts at various axonal target sites. We have explored a different approach that involves a targeting strategy in which the viral replicase itself is targeted to axonal terminals. Combining this strategy with previously identified replicase mutants that reduce the severity of the cytopathic effect caused by Sindbis is also being explored. In the end the goal is to enhance the barcode delivery system so that the barcode expressed by each neuron can be used to assign the cell's fluorescent fill, which in principle will allow barcode sequencing to elucidate the morphology of a neuron.

## Comparative Connectomics and Transcriptomics in Lab and Singing Mice

E. Isko [in collaboration with A. Banerjee, CSHL]

Despite a close evolutionary relationship, Alston's singing mice (*Scotinomys teguina*) and lab mice (*Mus musculus*) exhibit extremely divergent vocal behaviors. This behavioral divergence must derive from differences in the underlying biology of neural cell types and/or neural circuits. To determine these differences, we are collaborating with the Banerjee laboratory to apply techniques developed in the Zador laboratory to determine the neural circuits of vocal behavior in lab and singing mice. We have begun our connectomic comparison in the orofacial motor cortex (OMC), a brain area involved in the singing mouse song proven through electrical, pharmacological, and cooling experiments (Okobi et al., *Science* **363:** 983 [2019]). We are using both bulk methods (viral tracing) and single-cell methods (MAPseq, developed by the Zador laboratory) to characterize the projection patterns of neurons in the OMC of the two species. Our

future directions include using single-nucleus RNA-seq (snRNAseq) and BARseq2 (an in situ sequencing technique that combines gene expression and projection data developed by the Zador laboratory) to determine differences in cell type and the spatial location of these cells in the OMC of lab and singing mice. Finding transcriptomic and neural circuitry differences between lab and singing mice will give insights into the neural substrates of vocal behavior.

## Developmental Rules for Constructing Connectomes

S. Kerstjens

The brain's function depends critically on the wiring among its myriad neurons. This wiring is highly conserved within and across species, and therefore likely to be mostly genetically determined, and installed during prenatal development. However, the vertebrate connectome among billions of neurons cannot be explicitly encoded in the limited information capacity of its genome. This implies that the brain is not instantiated from a complex blueprint but rather from a relatively simple collection of developmental rules. This set of rules must solve the task of setting up a collection of address spaces and provide axons with the instructions to navigate them to their proper targets. Theoretical models for this developmental construction process are developed—both address space construction and axon navigation. These models have so far predicted and verified the existence of a global spatial "eigengene expression" profile that persists throughout embryonic development in mouse (Shuvaev et al. 2024) and that extends to zebrafish (Kerstjens et al. 2024). The ultimate aim is to explain how the developmental construction rules for the connectome endow it with an evolutionarily scalable and behaviorally functional organization.

## Cell Type–Specific Regulation of Auditory Decision-Making

D. Maharjan

Animals rely on their ability to respond appropriately to environmental stimuli in order to survive. Our research is focused on understanding how such sensory-motor associations are formed in the brain in the context of an auditory discrimination behavior. Building on previous findings, which identified the strengthening of connections from the auditory cortex to auditory striatum when animals learned to perform an auditory discrimination task, we sought to understand how this behavior is regulated by two neuronal cell types found in the striatum: Drd1a (D1 neurons) receptor-expressing and Drd2a (D2 neurons) receptor-expressing neurons. Characterized by their unique projection patterns and opposite responses to dopamine, these neurons are believed to influence behavior in divergent ways. To investigate the specific contributions of these cell types, we utilize optogenetics and fiber photometry methods in genetically modified animals performing the discrimination task. Our optogenetic experiments indicate that artificially stimulating D1 and D2 neurons exerts opposing effects on behavior. However, fiber photometry results show that only D1 neurons are active during behavior. Taken together, these findings suggest that, despite the natural antagonistic setup of D1 and D2 neurons, D1 neurons play a central role in regulating auditory discrimination behavior.

## Developmental Mechanisms Define Neuronal Diversification and Cell Type in the Mature Cerebral Cortex

K. Matho

Temporally specific transcription factors and effector genes play a key role during development in determining cortical glutamatergic pyramidal neuron (PyN) identity. Current advances in mouse genetics have resulted in an extensive set of genetic strategies based on mouse knock-in driver lines for targeting PyN cell types and their progenitors. This new framework presents an opportunity to examine the cell fate of genetically targeted progenitor pools and determine the molecular identity and connectivity of lineage- and genetically defined PyNs as they assemble into functional circuits (Matho et al., *Nature* **598:** 182 [2021]; Muñoz-Castañeda et al., *Nature* **598:** 159 [2021]). Evaluating the influence of developmental programs on cell type identity requires associating genetic strategies that define key aspects of developmental processes leading to PyN diversity with transcriptomic cell types and projection patterns systematically at cellular resolution. Recent developments from our laboratory in

sequencing of gene panels in situ and neuroanatomical projection mapping have resulted in BARseq2, a novel technique that simultaneously detects multigene expression and maps single-cell long-range projections based on RNA barcode sequencing (Sun et al., *Nat Neurosci* **24:** 873 [2021]). Further developments in the laboratory have resulted in a multigene panel encompassing all subclasses and clusters of cortical cell types, the performance of which has been compared with a comprehensive single-cell RNAseq data set from the cortex (Chen et al., bioRxiv 2022) (Fig. 1). We are working to integrate genetic strategies targeting progenitor pools and PyN subsets with BARseq2 and this new multi-gene panel to associate developmental trajectories with transcriptomic signatures and projection patterns at cellular resolution. In parallel, we have

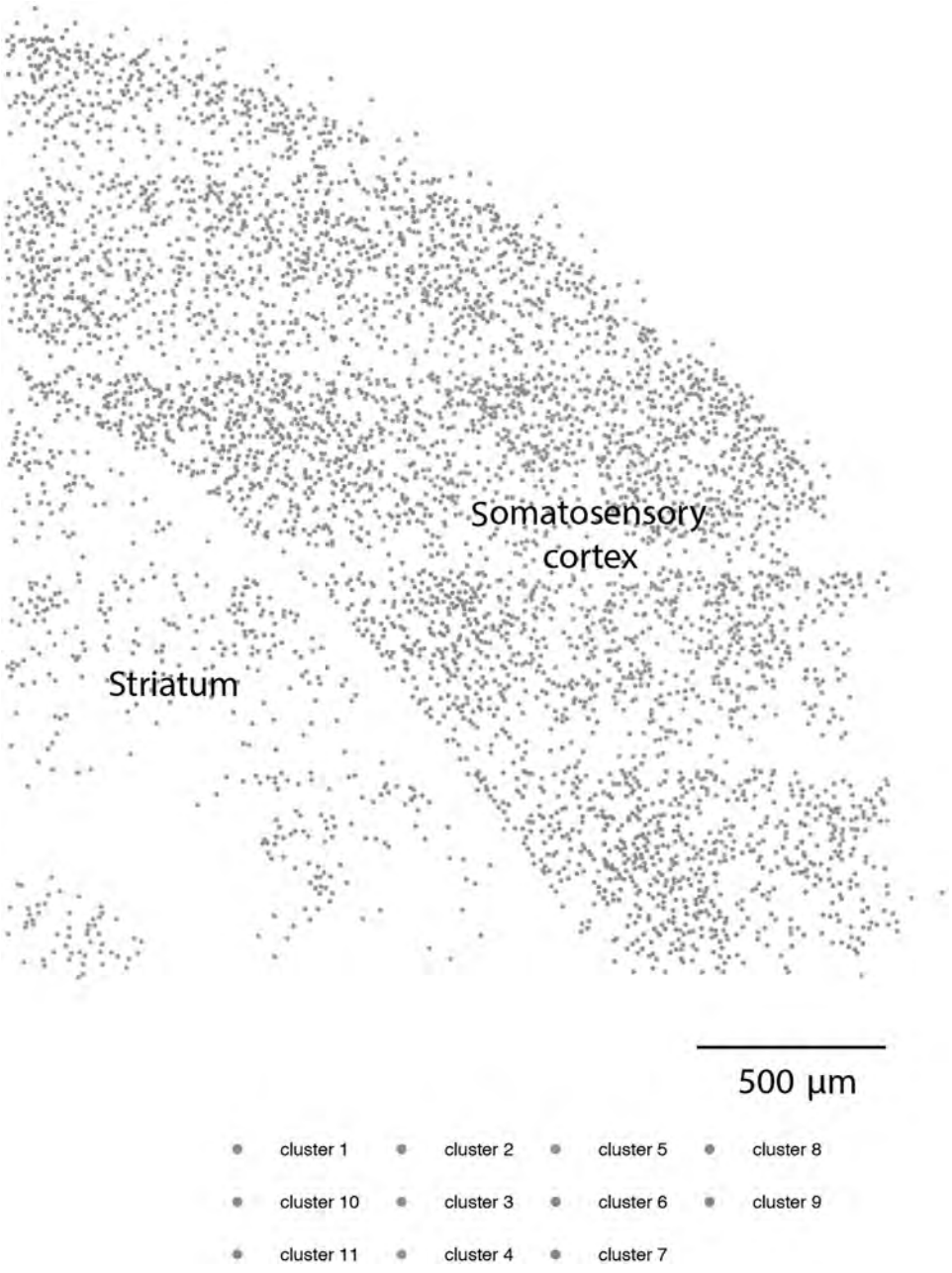



**Figure 1.** Identification of transcriptomic types based on single-cell gene expression in somatosensory cortex.

established mouse lines to investigate the link between developmental, molecular signature, and projection targets. We have validated an additional 12-gene panel to profile cortical cell types based on their developmental trajectories using gene knock-in reporter mouse lines expressing fluorescent proteins (e.g., dTomato and GFP). We have incorporated this additional 12-gene panel into a workflow of hierarchical clustering developed in the laboratory to map clusters to cell types in reference single-cell RNAseq data sets (Chen et al., bioRxiv doi:10.1101/2022.11.06.515380 [2022]). We have set up genetic strategies to report developmental mechanisms of (1) birthdating and (2) indirect versus direct neurogenesis in specific subsets of genetically specific lines (Matho et al., *Nature* **598:** 182 [2021]). These are being combined by gene panel sequencing to assess the effect of developmental mechanisms on the ultimate cell type identity in the mature cortex. Our approach has the advantage of incorporating a developmental dimension into the cell type classification of neurons in addition to the genetic expression and projection patterns achieved with BARseq2.

## Predicting Neuronal Electrical Properties from Gene Transcription

A. Medhat

Through their distinct genetic expression patterns, neurons exhibit precise control of key properties affecting their circuit role, including how they process electrical signals. Properties like a cell's spike rate in response to current injection, hyperpolarization recovery, and the action potential waveform depend on a careful balance between numerous ion channels and calcium buffers. The question is, how can we leverage foundational knowledge about single-cell gene coexpression patterns toward a new task, like predicting electrical properties, in which additional data may be limited—otherwise known as transfer learning. Toward that end, the focus has been on leveraging a BERT transformer model, pretrained on both single-cell human and mouse brain gene expressions, to predict 29 different electrical properties of mouse brain neurons, in which the training data is a PatchSeq-derived data set that maps both gene expression and electrical properties to a set of mouse brain cells. Based on early results suggesting limitations of the BERT transformer in handling gene expression data, which may be leading it to underperform traditional nonlinear supervised learning methods, we are exploring architectural adjustments to our transformer model to resolve these limitations.

## Differentiable Simulation of Spiking Neurons

C. Pehle

Neurons use spikes for communication. In the quest for artificial intelligence, spiking neural networks currently play a limited role. This is in part because a convincing demonstration of their computational advantage is missing. We have recently developed a differentiable way of simulating spiking neurons that enables gradient-based training and event-based gradient estimation. This allows us to investigate where such a computational advantage of spiking neurons can be found. Currently we are considering planning problems in connection with their combination of discrete and continuous dynamics. In a separate line of research, we have developed a differentiable way of simulating morphologically detailed neuron models. We are doing so with the aim of establishing a way to train networks of such neurons and investigate continuous learning.

## High-Throughput Integration of Neuronal Activity, Connectivity, and Gene Expression

C. Soitu

The flow of information between different regions of the brain during decision-making is a largely unknown process. This is mainly because there are no tools available to disentangle the complex interaction of myriad cell types in the brain, governed by gene expression, connectivity, spatial organization, and other properties. This work aims to investigate the anatomical substrate of interregional communication at the single-neuron level. Tools developed in our group, BARseq and MAPseq, as well as our previous work on the topic, make this question accessible for investigation. We are using two-photon imaging of activity in single neurons to identify privileged populations of neurons. Successful completion of this project would generate unprecedented data sets that bridge information at all levels—anatomical, genetic, physiological, and behavioral. Gaining knowledge of the flow of information in the brain would push the frontiers of what it is one of the grand challenges of our time—to understand how behavior arises from neural circuits.

## Mapping Neural Projections from Mouse Auditory Cortex Using Axonal BARseq

L. Yuan

The brain consists of tens of thousands of neurons interconnected through complex networks. Understanding how neurons from different brain areas communicate is fundamental to deciphering brain functions and, by extension, animal behavior. One of the challenges in mapping neural projections is that axons, the long and thin neural projections, interweave with each other and other structures within the brain. To overcome this challenge and accurately map neural projections in situ, we have developed a novel technique called axonal BARseq. This method gives unique barcodes to individual neurons. By detecting these barcodes in both the somas and axons in situ, we can determine which neuron projects to which brain area with high spatial resolution. Over the past few years, we have completed our initial proof-of-principle experiment, successfully mapping long-range projections from more than 8,000 cells in approximately a 500-µm region of the mouse primary auditory cortex. We are now prepared to scale this approach to map long-range projections from the entire mouse auditory cortex.

Last year, we made significant strides in improving our imaging throughput. Among these improvements was the customization of image acquisition codes to optimize for in situ sequencing. Additionally, we refined our current workflow, enabling us to process and image more samples simultaneously. This not only reduced potential variability between batches but also shortened the sequencing time. This year, our focus will shift to the data processing step. Given the experimental modifications, the increased volume of data, and the lessons learned from our proof-of-concept experiments, we are in the process of rewriting our data processing codes.

## MAPseq/BARseq Technique Optimization

H. Zhan

The GFP$^+$ Sindbis virus has been mainly used in MAPseq and BARseq. However, for some projects in which GFP has been used in the original samples, Sindbis expressing other fluorophores will be more useful for researchers. We made two new Sindbis viruses expressing mCherry and BFP, respectively, and made two libraries for each virus with a diversity of >20M barcodes.

We also made a large-scale Sindbis virus library with the new vamp2 carrier protein. This new virus library has ~40M diversity at a titer of $10^{10}$ infectious particles/mL. We made ~2,000 aliquots of this library and it should be enough for 8,000 regular injections.

We have been trying to optimize BARseq2 for fixed brain sections. We tried to apply the protocol of Axonal BARseq on 50-µm sections to detect endogenous genes. Although the rolony signals from endogenous genes seem weaker than those from fresh frozen tissue, there is no penetration issue and we are still trying to call bases to differentiate individual genes.

We are also responsible for running the MAPseq Core facility. We have been trying to disseminate MAPseq/BARseq techniques by helping users optimize their experimental settings for these two techniques. We meet users frequently from different institutes in different countries. With the help of our team members, we recently applied BARseq to non-Sindbis-infected barcoded brain tumor tissues and are trying to apply BARseq on barcoded rabies virus–infected human tissues.

### PUBLICATIONS

Acerbi L, Aguillon-Rodriguez V, Ahmadi M, Amjad J, Angelaki D, Arlandis J, Ashwood ZC, Banga K, Barrell H, Bayer C, et al. 2023. A modular architecture for organizing, processing and sharing neurophysiology data. *Nat Methods* **20:** 403–407. doi:10.1038/s41592-022-01742-6

Barabási DL, Bianconi G, Bullmore E, Burgess M, Chung SY, Eliassi-Rad T, George D, Kovács IA, Makse H, Nichols G, et al. 2023. Neuroscience needs network science. *J Neurosci* **43:** 5989–5995. doi:10.1523/JNEUROSCI.1014-23.2023

Findling C, Hubert F, Acerbi L, Benson B, Benson J, Birman D, Bonacchi N, Carandini M, Catarino JA, Chapuis GA, et al. 2023. bioRxiv doi:10.1101/2023.07.04.547684

Funamizu A, Marbach F, Zador AM. 2023. Stable sound decoding despite modulated sound representation in the auditory cortex. *Curr Biol* **33:** 4470–4483.e7. doi:10.1016/j.cub.2023.09.031

Yuan L, Chen X, Zhan H, Gilbert HL, Zador AM. 2023. Massive multiplexing of spatially resolved single neuron projections with axonal BARseq. bioRxiv doi:10.1101/2023.02.18.528865

Zador AM. 2023. Charles F. Stevens (1934–2022). *Nat Neurosci* **26:** 176–177. doi:10.1038/s41593-022-01241-z

Zador A, Escola S, Richards B, Ölveczky B, Bengio Y, Boahen K, Botvinick M, Chklovskii D, Churchland A, Clopath C, et al. 2023. Catalyzing next-generation artificial intelligence through NeuroAI. *Nat Commun* **14:** 1597. doi:10.1038/s41467-023-37180-x

Zhang A, Zador AM. 2023. Neurons in the primary visual cortex of freely moving rats encode both sensory and non-sensory task variables. *PLoS Biol* **21:** e3002384. doi:10.1371/journal.pbio.3002384

*In Press*

Kerstjens S, Engert F, Zador A, Douglas R. 2024. Eigengene reveals invariant global spatial patterns across mouse and fish brain development. bioRxiv doi:10.1101/2024.09.19.613507

Shuvaev S, Lachi D, Koulakov A, Zador A. 2024. Encoding innate ability through a genomic bottleneck. *Proc Natl Acad Sci* **121:** e2409160121. doi:10.1073/pnas.2409160121

# PLANT BIOLOGY

**David Jackson** and colleagues study genes and signals that regulate plant growth and architecture. They are investigating a unique way in which plant cells communicate: by transporting regulatory proteins via small channels called plasmodesmata. These channels, which direct the flow of nutrients and signals through growing tissues, are regulated during development. The team discovered a gene encoding a chaperonin, *CCT8*, that controls the transport of a transcription factor, SHOOTMERISTEMLESS (STM), between cells in the plant stem cell niche, or meristem. STM is crucial for stem cell maintenance, and studies of the *CCT8* gene indicate that movement of STM between cells is required for this function. The laboratory also continues to identify other genes that control plant architecture through effects on stem cell maintenance and identity, and their work has implications for crop yields. Recent examples include discovery of a subunit of a heterotrimeric G protein that is conserved throughout animals and plants, and their studies indicate that this gene controls stem cell proliferation. They have found that in plants, the G protein interacts with a completely different class of receptors than in animals. Their discovery helps to explain how signaling from diverse receptors is achieved in plants. This year, they also demonstrated that weak mutations in one of the receptor proteins can enhance seed production in maize, which could lead to yield increases. Separately, the laboratory has characterized system-wide networks of gene expression using "next-gen" profiling and chromatin immunoprecipitation methods that have generated many new hypotheses in developmental networks controlling inflorescence development. They are also developing a collection of maize lines that can drive expression of any reporter or experimental gene in any tissue type—tools of great interest to maize researchers that are being made available to the broader scientific community, enabling experiments never before possible in crop plants.

**Zachary Lippman**'s research team studies when, where, and how many branches, flowers, and fruits are produced on plants. All of plant development depends on small groups of stem cells at the tips of shoots known as meristems. By studying the genes that control stem cell production and maturation over space and time, within and between different developmental contexts, the Lippman laboratory is able to manipulate plant architecture and reproduction to improve crop yields. Lippman's research program integrates development, genetics, genomics, and genome editing to exploit the mechanisms that determine how plant stem cells become shoots and flowers. The laboratory takes advantage of extensive natural and mutant variation in inflorescence production and architecture in tomato and related nightshade (Solanaceae) species (e.g., potato, pepper, and groundcherry) to explore how differences in these processes explain the remarkable diversity in the architectures of these shoot systems found in nature and agriculture. Recent discoveries regarding the genes and networks underlying this diversity have led to broader questions on the significance of genomic structural variation, genetic redundancy, gene dosage, and epistasis in development, domestication, and breeding. By linking these fundamental and applied discoveries, Lippman is developing and applying innovative concepts and tools for crop improvement.

Epigenetic mechanisms of gene regulation—chemical and conformational changes to DNA and the chromatin that bundles it—have had an important impact on genome organization and inheritance and on cell fate. These mechanisms are conserved in eukaryotes and provide an additional layer of information superimposed on the genetic code. **Robert Martienssen**, a pioneer in the study of epigenetics, investigates mechanisms involved in gene regulation and stem cell fate in yeast and model plants, including *Arabidopsis* and maize. He and his colleagues have shed light on a phenomenon called position-effect variegation, caused by inactivation of a gene positioned

near densely packed chromosomal material called heterochromatin. They have discovered that small RNA molecules, arising from repeating genetic sequences, program that heterochromatin. Martienssen and colleagues have described a remarkable process by which "companion cells" to sperm in plant pollen grains provide them with instructions that protect sperm DNA from transposon damage. They found that some of these instructions, or epigenetic marks, could be inherited in the next generation. These marks, and the small RNA responsible for guiding them, can sense the number of chromosomes inherited from pollen and may allow *Arabidopsis*, a flowering plant, to produce egg cells without meiosis, an important step toward a long-time goal of plant breeding—generating clonal offspring to perpetuate hybrid vigor. The laboratory has also shown that when RNA polymerase II has transcribed a stretch of DNA, the RNA interference mechanism causes the enzyme to release its hold on the DNA and fall away. This allows the replication fork to progress smoothly and the DNA strands to be copied; histone-modifying proteins, which follow right along, establish heterochromatin. Martienssen's group also continues to work on problems related to the creation of plant-based biofuels. As part of a collaborative project to generate a high-quality full genome map of the oil palm plant, Martienssen and colleagues identified a transposon whose modification controls the yield of oil palm trees. This discovery will increase yields and should lessen the environmental burden of oil palm production, which often threatens already endangered rainforest lands.

The **Ullas Pedmale** laboratory's research goals seek to determine the mechanisms behind how a plant perceives and successfully adapts to its environment. They also aim to understand how a plant must integrate intrinsic and extrinsic cues and "decide" how best to respond to environmental cues. Understanding how plants deal with and respond to a multitude of environmental signals could help to develop crops that cope with unfavorable growth conditions without significant changes in yield.

# DEVELOPMENTAL BIOLOGY—STEM CELL SIGNALING AND CROP ARCHITECTURE

**D. Jackson** T. Clark, S. Elliot, J. Finger, M. Gleason, S.D. Iohannes, P. Lindsay, E. Morley, S. Neri, K. Pandey, N. Shanmugaraj, T. Skopelitis, K. Swentowsky, T. Tran, V. Varapparambath, X. Xu

Our research delves into the molecular pathways that control the growth and shape of plants, with a long-term goal of improving crop yields. We aim to identify genes, signals, and pathways that regulate plant architecture and development. A primary interest is in discovering the signals that carry developmental information between cells, how the signals are transmitted, and how they function. For example, we identified genes that control pluripotent stem cell signaling and developed ways to increase yield in maize by using CRISPR genome editing. We also created a "single-cell" atlas of maize ear development, providing a rich resource for research into yield traits. We continue to characterize enzymes that function in sugar metabolism but also may play a "moonlighting" role in the cell nucleus, and discovered how a signal important for stem cell fate in plants travels as an RNA message between cells.

## The Control of Meristem Size in Maize

P. Lindsay, K. Swentowsky [in collaboration with P. Boumpas, University of Heidelberg, Germany; A. Reyes and S. Xu, Stanford University; F. Xu, Shandong University, China; Y. Matsubayashi, Nagoya University]

All plant organs derive from populations of stem cells called meristems. These cells have two purposes: to divide and maintain themselves, and to give rise to daughter cells, which will differentiate into plant organs. Consequently, meristems must precisely control the size of their stem cell niche, via a network of positive and negative feedback signals. A loss of function in a negative regulator of stem cell fate results in an enlarged or fasciated meristem phenotype and a dramatic alteration in the shape of the maize ear and tassel. Maize is an excellent model system for these studies, because of a large collection of developmental mutants and a diverse genome. Our laboratory uses genetics to identify key regulators of stem cell homeostasis and meristem size. Two previously cloned mutants, *thick-tassel dwarf1* (*td1*) and *fasciated ear2* (*fea2*), encode orthologs of the *Arabidopsis thaliana* genes *CLAVATA1* and *CLAVATA2*, suggesting the well-known *CLAVATA–WUSCHEL* regulatory feedback loop is conserved in monocot crops. However, we also identified several novel stem cell regulators in maize.

We continue to study the mechanism of action of *FEA3,* which encodes a predicted leucine-rich repeat (LRR) receptor-like protein, related to *FEA2*. *FEA3* is of particular interest because it is expressed in the organizing center of the meristem and in leaf primordia, and expression of maize *WUSCHEL*, a marker for the stem cell niche organizing cells, spreads downward in *fea3* mutants, which is strikingly different from its response in the known *CLAVATA* stem cell mutants. We tested the ability of FEA3 to bind to its predicted ligand, FCP1. In collaboration with Professor Yoshikatsu Matsubayashi at Nagoya University, we found that FEA3 does not directly bind FCP1, but BARELY ANY MERISTEM1 (BAM1), an LRR receptor-like kinase that also controls meristem size, can bind FCP1. Furthermore, maize *bam1d* mutants are insensitive to the application of FCP1. *BAMs* are part of the meristem signaling pathway in *Arabidopsis*, but their function in maize has not yet been evaluated. Intriguingly, *BAM1D* is also up-regulated in the *fea3* mutant, suggesting they may act together in a pathway. In support of this, we found that BAM1D and FEA3 proteins interact, and their expression overlaps, indicating these two proteins may form a receptor–co-receptor pair. In addition to this physical interaction, *fea3;bam1d* double mutant ears are significantly smaller than either single mutant, indicating that *FEA3* and

*BAM1D* interact genetically. We are also investigating whether there is genetic redundancy with the other six *BAM* paralogous genes through analysis of higher-order *bam* mutant populations.

To confirm the interaction of FEA3 with BAM1D in vivo, and determine what other downstream signaling partners interact with FEA3, we are using proximity labeling to assess transient interactions between FEA3 and interacting proteins. In this technique, proteins of interest are tagged with TurboID, a biotin ligase that has been engineered to biotinylate nearby proteins. We optimized conditions for proximity labeling of membrane-bound proteins in maize meristem tissues and have generated affinity purification–mass spectrometry (AP-MS) data sets of FEA3 and BAM1D in both shoot apex and ear primordia tissue (Fig. 1). BAM1D peptides are enriched in the FEA3-Turbo proximity labeling data set. Furthermore, FEA3 and BAM1D interactomes have significant overlap, further supporting the idea that they are part of the same signaling pathway. Several signaling-related proteins were identified from this approach. The most promising candidates have been mutated using CRISPR-Cas9 to determine their role in meristem maintenance. We are also crossing these mutants to *fea3* and *bam1d* to ask whether they function in the same genetic pathway. These analyses begin to reveal the receptor landscape of FEA3.

Characterization of additional factors that control inflorescence architecture will improve our understanding of this process. We have identified a mutant named *fea*-Leo* that shows severe fasciation and ears and tassels that are often bifurcated (Fig. 2A). Scanning electron microscopy (SEM) of developing *fea*-Leo* ears revealed an initial invagination of the inflorescence meristem, with multiple meristem tips forming as it expands (Fig. 2B). Using quantitative trait locus sequencing (QTL-seq), *fea*-Leo* mapped to a several-megabase region near 150 Mb on Chromosome 6 (Fig. 2C). We performed RNA-seq of *fea*-Leo*

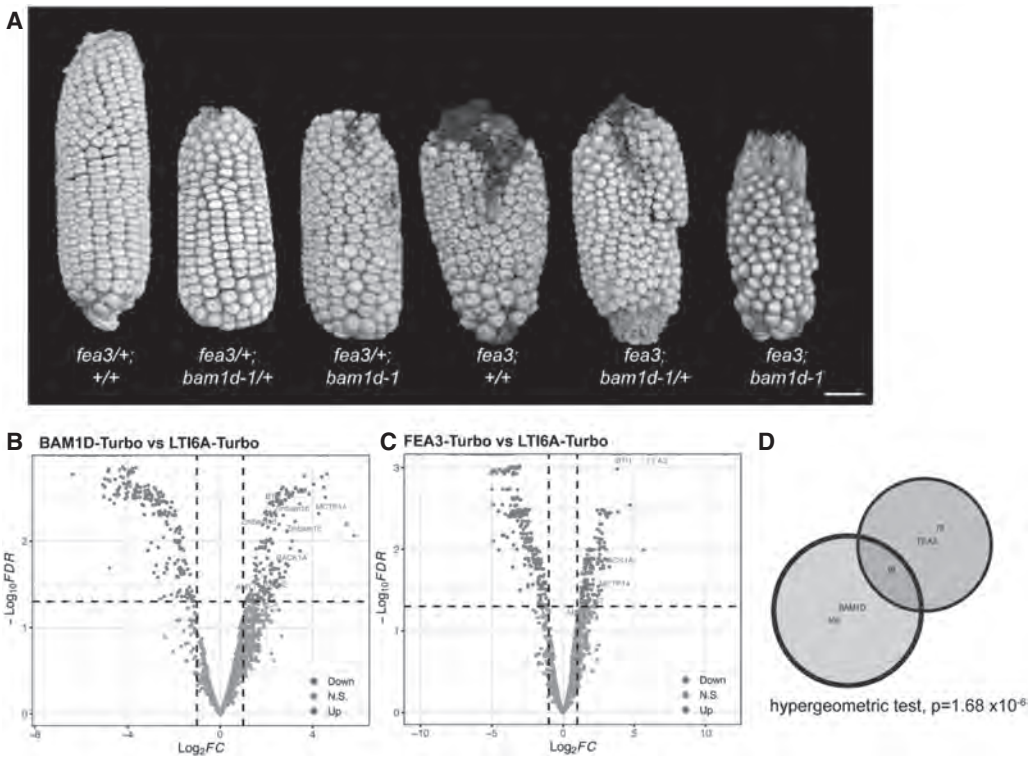



**Figure 1.** FEA3 and BAM1D control ear development. (*A*) *fea3*;*bam1d-1* double mutant ears are smaller than their single mutant counterparts. Scale bar, 2 cm. (*B*,*C*) Volcano plot showing enrichment of several signaling candidates in BAM1D-Turbo or FEA3-Turbo samples as compared with a negative control sample, LTI6A-Turbo. Candidates in the *top right* are significantly enriched relative to LTI6A-Turbo. (*D*) Significant overlap between peptides enriched in BAM1D and FEA3.

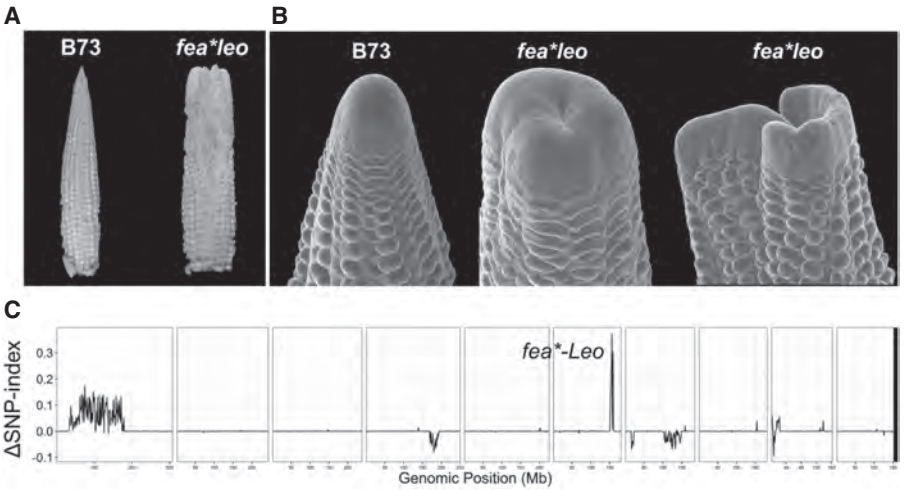



**Figure 2.** *fea*-Leo* mutants have fasciated ears and the mutation maps to chromosome 6. (*A*) Wild-type (B73) and *fea*-Leo* developing ears. (*B*) Scanning electron micrographs of B73 and *fea*-Leo* ear primordia. (*C*) Quantitative trait locus sequencing (QTL-seq) plot of *fea*-Leo* ΔSNP-index; note the strong peak on Chromosome 6.

ears to find differentially expressed candidate genes in this region, and analysis is currently underway. Using markers in the QTL region to genotype individuals segregating *fea*-Leo*, we suspect that this locus could contain structural variation, so we plan to perform Oxford Nanopore long-read sequencing to elucidate the genomic landscape surrounding *fea*-Leo*.

## Natural Variation in Inflorescence Architecture

P. Lindsay, S.D. Iohannes, S. Neri [in collaboration with M. Passalacqua, CSHL; J. Gillis, University of Toronto; R. Chen, Cardozo School of Law, New York]

Maize inflorescence architecture has been a target for extensive selection by breeders, and the maize genome is highly diverse; hence, different inbred lines vary greatly. We are investigating gene expression differences in the inflorescence meristems of the nested association mapping (NAM) lines, a panel of maize lines selected to represent the natural diversity of maize. We are particularly interested in genes that modulate inflorescence meristem size, because it correlates with kernel row number, an agronomically important trait. We performed 3′ mRNA-seq on the NAM lines, and measured their inflorescence meristems (IMs) (Fig. 3A). We correlated gene expression profiles with IM size to find candidate genes that control meristem size (Fig. 3B). We also cross-referenced our data with other NAM founder data sets to identify differentially expressed genes that are important for yield-related traits. Knockout and promoter CRISPR-Cas9 constructs for one of these genes have been transformed into maize to test their effect on meristem size and ear traits. This candidate gene has also been knocked out in *Arabidopsis*, to see whether gene function is conserved. Preliminary analyses suggest that this gene, called *EYES ABSENT*, has a general role in plant development, as mutant plants are smaller (Fig. 3C). This gene has important roles in animal development, suggesting its function is conserved in plants and animals.

Ear primordia of some NAM founders have also been imaged to ask how the positioning and number of spikelet pair meristems leads to a specific number of flower primordia and seed. We measured the positioning and number of spikelet pair meristems over time to understand how dynamics differ between NAM founders with different ear morphology (Fig. 4). We observe different dynamics depending on the line, which we are now relating to yield traits. Future work will compare promoter regions in differentially expressed genes among the NAM lines to identify *cis*-regulatory elements that control gene expression. In the future, these data will be incorporated into a predictive model to manipulate gene expression in silico and predict how gene expression changes impact meristem morphology and yield-related traits.

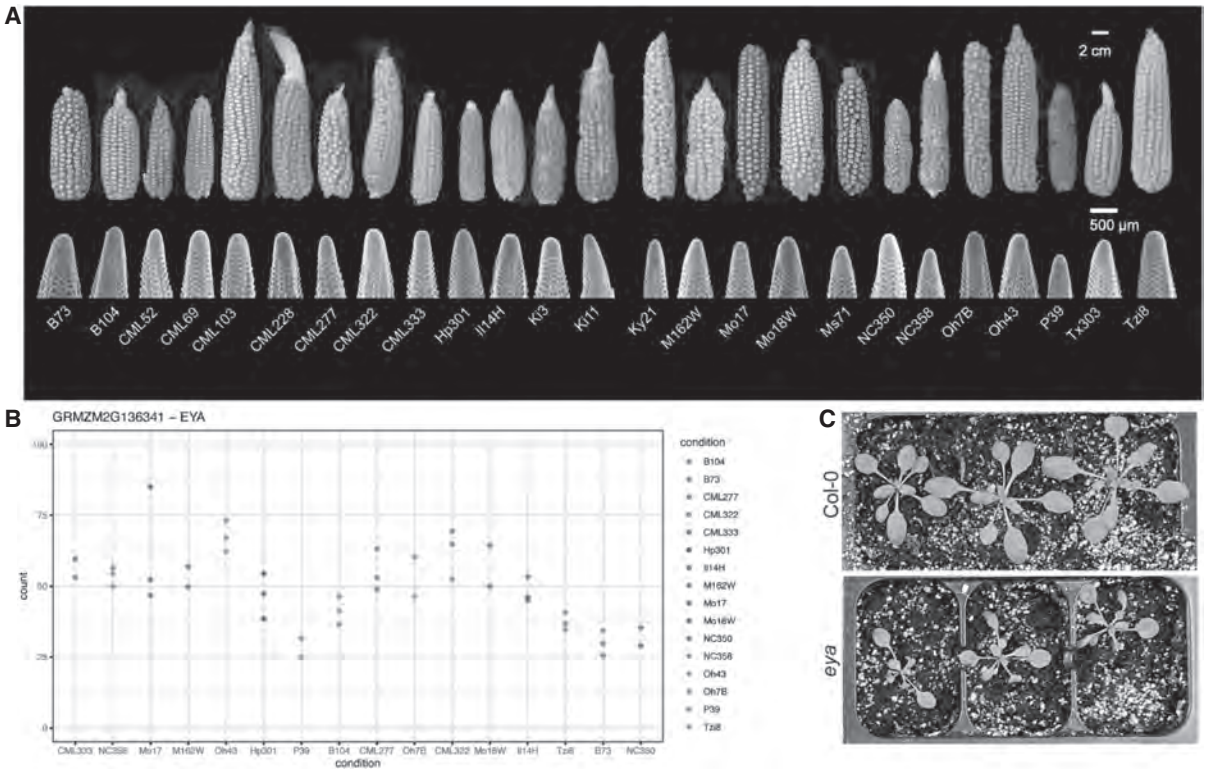



**Figure 3.** (*A*) Mature ear and immature ear primordia of nested association mapping (NAM) founders. (*B*) Gene expression levels of *EYES ABSENT*. NAM founders are sorted by meristem size. (*C*) Wild-type (Col-0) and *eya* mutant *Arabidopsis* plants; *eya* plants are smaller.

## Identification and Functional Dissection of Shared *cis*-Regulatory Elements Controlling Quantitative Trait Variation across Angiosperms

K. Swentowsky, N. Shanmugaraj [in collaboration with A. Hendelman and Z. Lippman, CSHL; M. Bartlett, University of Massachusetts Amherst; and I. Efroni, The Hebrew University of Jerusalem]

Genomic *cis*-regulatory elements (CREs) mediate transcription factor binding and gene expression. Quantitative control of gene expression affects agronomically important traits and can be modulated using CRISPR-Cas9 editing of CREs. Although editing has become routine in many species, it has been difficult to identify CREs because of the substantial size and variation of noncoding genomic sequences. We hypothesize that many genes with conserved function and expression may contain deeply conserved CREs that can be identified through comparative genomic analysis. In collaboration with Idan Efroni at The Hebrew University of Jerusalem, we have constructed a

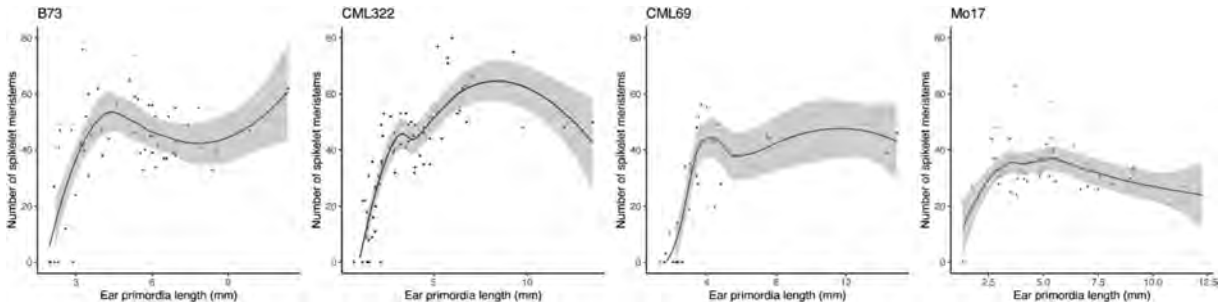



**Figure 4.** Number of spikelet meristems over time in nested association mapping (NAM) founder lines.

computational pipeline called Conservatory for identifying deeply conserved CREs. Conservatory first identifies closely related genes between species within a plant family and finds conserved CREs within the family. Then, alignments are performed between families to identify more deeply conserved CREs.

In maize, we have designed CRISPR guide RNAs (gRNAs) to target CREs identified by Conservatory in the promoter of *INDETERMINATE1* (*ID1*). *ID1* encodes a transcription factor that promotes flowering, and *id1* mutant plants are severely delayed in flowering and have reduced determinacy. We identified conserved CREs in two regions of the *ID1* promoter: one in the first ~400 bp upstream of the start codon and one ~2.5 kb upstream (Fig. 5). These regions overlap with accessible chromatin and are likely to be involved in regulation of *ID1* gene expression. Deletion of CREs important for *ID1* gene regulation should impact its expression and produce more subtle phenotypes compared with the null mutations. We obtained plants containing the *ID1* CRISPR-Cas9 construct and identified novel *ID1* promoter alleles containing deletions (Fig. 5). These alleles lack CREs that are predicted to be important for *ID1* gene expression. We are currently testing these alleles for changes in *ID1* gene expression and phenotypic consequences of these deletions.

In the 1960s, maize breeders sought to improve agricultural sustainability by breeding perennial maize plants that can grow for multiple years. They combined the *id1* null mutant with the highly branched *grassy tillers1* (*gt1*) mutant and a perennial allele from a wild species. These perennial plants had undesirable characteristics, such as severely delayed flowering and extreme branching. Once we obtain promoter-edited lines with lower *ID1* expression, we will combine them with weak *gt1* alleles generated by Madelaine Bartlett's laboratory at the University of Massachusetts Amherst and a locus for perennial regrowth that we recently mapped. We hypothesize that weak alleles of *id1* and *gt1* will allow plants to achieve perenniality without the undesirable phenotypes associated with the null alleles.

We are also targeting conserved noncoding sequences (CNSs) in RAMOSA3, which encodes a trehalose-6-phosphate phosphatase with a hypothesized moonlighting function in gene regulation. Our previous genome-wide chromatin studies suggest the presence of a distal open chromatin region (dOCR) ~56 kb upstream of RA3 (Fig. 6). This region overlaps with independent assay for transposase-accessible chromatin with sequencing (ATAC-seq), MNase-defined cistrome-occupancy analysis (MOA-seq), and chromatin immunoprecipitation sequencing (ChIP-seq) peaks, making it a target for CRE editing. We hypothesize that deletions of CNSs in this dOCR may affect RA3 expression and may lead to inflorescence branching or other subtle phenotypes that will help us to understand the transcriptional regulation of RA3 and the significance of chromatin looping with this region. To study the function of this dOCR, we used CRISPR editing to target CNSs identified from our Conservatory pipeline. We identified CNSs spanning ~5 kb of the dOCR and designed unique CRISPR guide RNAs to target them. CNSs in the proximal promoter regions both downstream and upstream will also be targeted.

We are using a similar strategy to edit promoter elements of TERMINAL EAR 1 (TE1) using Conservatory data. *te1* loss-of-function mutants have pleiotropic developmental defects, including faster leaf initiation, altered phyllotaxy, dwarf stature, and a feminized tassel, making the interpretation of gene function difficult. Thus, weak alleles targeting CNSs in the 5′ and 3′ untranslated regions and downstream of the start codon may be useful to test the potential role of TE1 in maize development

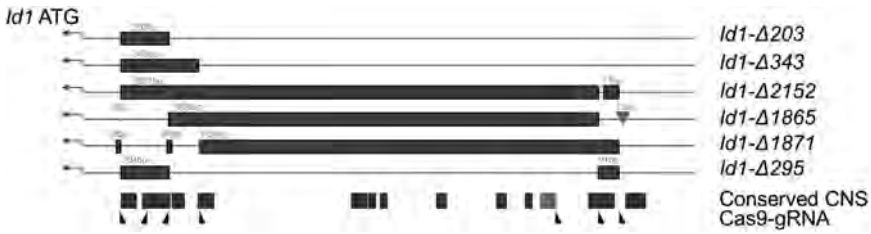



**Figure 5.** CRISPR-Cas9 mutagenesis of the *ID1* promoter. Six alleles (*y*-axis) contain deletions (black boxes) induced by Cas9-gRNAs (*bottom* arrows). Many of these overlap conserved noncoding sequences (CNSs; bottom black boxes) identified by Conservatory.

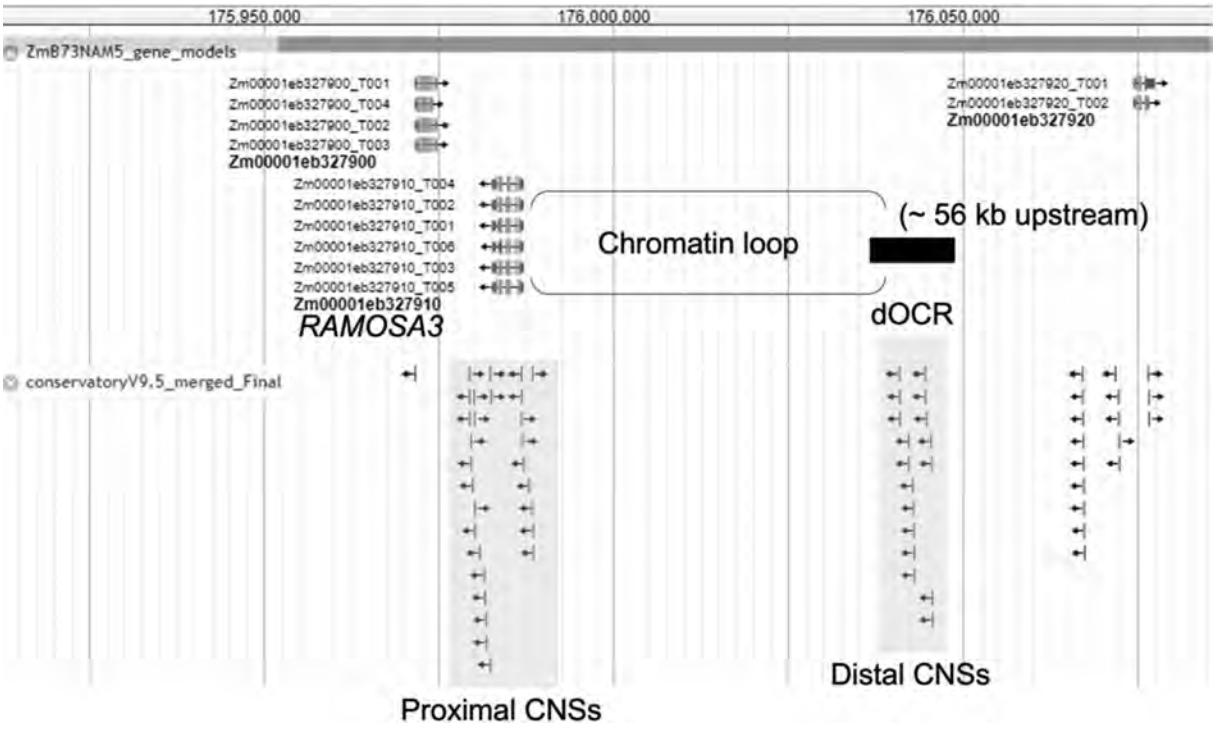



**Figure 6.** Conserved noncoding sequences (CNSs) in the proximal and distal regions of the *RAMOSA3* gene. Genome-wide chromatin studies suggest the presence of a distal open chromatin region (dOCR), ~56 kb upstream of RA3, that contains predicted CNSs.

and stress responses. Thus, studying specific functions of conserved CREs of maize meristem developmental genes that are directly related to crop improvement will pave the way for the rational editing of agronomic traits.

## Developmental Genetics and Genomics of Perenniality

K. Swentowsky

Our modern grain crops are primarily cultivated as annual plants, but breeders are interested in creating perennial crops because of their unique properties. Perennials can live for multiple years and are therefore capable of efficient and sustainable growth. Despite decades of interest in breeding perennials, we lack basic knowledge of how certain grass species can achieve perennial growth. Perenniality can be envisioned as a syndrome that occurs when several underlying traits coexist with each other. Such traits include the ability to maintain competent meristems, the regulation of source-sink dynamics, and the timing of life-history traits like flowering and senescence. For example, perennial grasses tend to delay their flowering and senescence relative to annual species in order to accommodate a perennial life cycle. With this physiological framework for how perenniality can occur in nature, we can begin to experiment and understand the genetics that underlie this phenomenon.

Maize has a close perennial relative called *Zea diploperennis* that has been the source of perennial maize breeding and genetics research. Three major QTLs for perennial regrowth, called *Regrowth1* (*Reg1*), *Reg2*, and *Reg3*, have been identified—although their underlying genes have not been characterized. We have introgressed *Reg1* and *Reg3* from *Z. diploperennis* into a maize inbred line and analyzed their effects on three traits related to perenniality: flowering time, stay-green, and tiller number. We found that relative to maize, *Reg1* and *Reg3* both delay flowering time by around a week, *Reg3* delays senescence, and neither locus affects tiller number. To understand how these QTLs affect life-history traits and identify differentially expressed genes that could underlie the QTL, we performed RNA-seq of axillary meristems and leaf tips in maize lines containing *Reg1* and *Reg3*, and our analysis is underway.

## Single-Cell Analysis of Shoot Meristems Opens a Goldmine for Functional Studies

X. Xu [in collaboration with M. Passalacqua and D. Jackson, CSHL; and J. Gillis, CSHL and University of Toronto]

Shoot meristems control plant architecture and impact crop productivity. An understanding of these structures requires insights into the component cell types, developmental domains, and the gene networks required to specify them. However, these domains are classified mainly by morphology, or insights from classical genetics, but this knowledge is limited by genetic redundancy and pleiotropy. Recently, we produced an improved single-cell RNA-seq (scRNA-seq) atlas by capturing rare stem cells in *Arabidopsis* and maize shoot meristems that were largely missed in previous studies. We identified stem cell markers and validated their expression using spatial transcriptomics. We then used cross-species analysis to discover conserved stem cell markers and cell types. Plant stem cells are maintained by a conserved CLAVATA-WUSCHEL (CLV-WUS) pathway. We thus also profiled single cells from mutants of this pathway. We found hundreds of differentially expressed genes (DEGs) in the stem cells and used multiplex CRISPR-Cas9 to knock out selected genes in a family of predicted sugar kinases. This resulted in a striking meristem termination phenotype, validating the predictive power of our single-cell atlas.

Together, this comprehensive shoot meristem single-cell atlas will open a goldmine for functional studies at a fundamentally new level and will be a valuable resource for the plant community.

## Control of Shoot Branching and Determinacy

T. Tran, X. Xu, T. Skopelitis [in collaboration with H. Furukawa, CSHL; H. Claeys, Inari Agriculture, Belgium; E. Demesa-Arevalo and P. Boumpas, Heidelberg University; K. Michalski and R. Brennan, Duke University]

Shoot branching is important for crop productivity. The *RAMOSA* (*RA*) genes impose determinacy on axillary meristem growth; consequently, ra mutants (*ra1*, *ra2* and *ra3*) have highly branched inflorescences. RA3 encodes a trehalose phosphate phosphatase, an enzyme that converts trehalose-6-phosphate (T6P) to trehalose. T6P is an important regulatory metabolite that connects sucrose levels, and thus sugar status, to plant growth and development, but its mode of action is unclear. *RA3* is expressed at the base of axillary inflorescence meristems and localizes to distinct puncta in both nuclear and cytoplasmic compartments, suggesting that its effect on development may not be simply metabolic. These data support the hypothesis that *RA* genes may serve as mediators of signals, maybe a sugar signal, originating at the boundary domain and regulating determinacy. We also found that TPP enzyme activity could be uncoupled from their mutant phenotype, suggesting that TPP proteins have an alternative or "moonlighting" function.

We also found that RA3 interacts with TPS1 and TPS12, trehalose phosphate synthase proteins that naturally lack catalytic activity. We showed that RA3 and TPS1 are expressed in the same cells, using immunolocalization. *tps1* and *tps12* mutants enhance ra3 inflorescence phenotypes (Fig. 7A), suggesting their physical interaction is biologically significant. We also found that *tps1;tps12;ra3* mutants enhanced tillering (Fig. 7B). Tillering is an important trait that was selected during domestication; selection of maize from its wild ancestor teosinte resulted in a strong suppression of tillering through selection of a gain-of-function allele of the TEOSINTE BRANCHED1 (TB1) transcription factor. Previous studies found that TPS1 and RA3 are targets of TB1, suggesting they act downstream. Our findings agree with these results and highlight an important signaling role of T6P in breaking bud dormancy.

Interestingly, in a yeast-two-hybrid experiment we found that ZmTPS1 also interacts with the two catalytically active TPSs in maize, TPS11 and 14. Next, we asked whether these interactions might affect enzymatic activity. We designed a coupled enzyme assay and found that the enzymatically inactive TPS1 protein stimulates the activity of RA3 and TPS14. We also expressed and purified RA3, TPS1, and TPS14 proteins in insect cells, and found that they could form a complex. We are optimizing conditions for cryo-electron microscopy to visualize the structure of the three-protein complex (Fig. 7C).

In addition, to further understand the function of catalytic TPSs, we mutated the *TPS11* and *TPS14* genes using CRISPR-Cas9 and found that the double mutants fail to complete embryogenesis, as in *Arabidopsis*. The *tps11;tps14* double mutants were partly rescued on the media with T6P, indicating

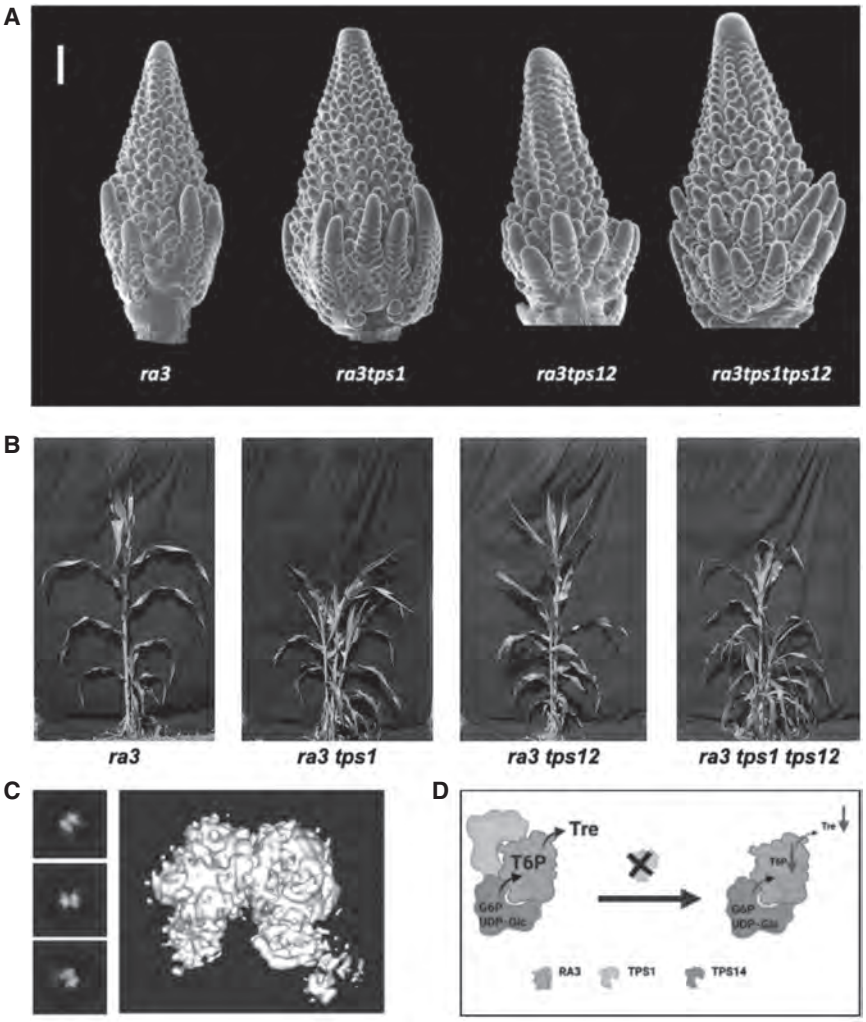



**Figure 7.** A and B *tps1* and *tps12* mutants enhance *ra3* phenotypes. (*A*) Representative immature ears of *ra3; ra3;tps1; ra3;tps12;* and *ra3;tps1;tps12* with enhanced branching compared to *ra3*. (*B*) Representative plants of *ra3; ra3;tps1; ra3;tps12;* and *ra3;tps1;tps12* with enhanced tillering compared to *ra3*. (*C*) Preliminary 2D and 3D reconstruction of the TPS–TPP complex. (*D*) Hypothesis of TPS–TPP interactions. ZmTPS11/14 and TPP (RA3) are enzymatically active. Enzymatically inactive TPS1 regulates enzymatic activity and may form a complex.

that T6P is important for maize embryo development. To further understand the functions of TPS11 and TPS14 in postembryonic development, we designed CRISPR-Cas9 constructs targeting conserved regions of the *TPS14* promoter and isolated a point mutant allele from a Chinese TILLING collection to make weaker alleles. We will phenotype the weak alleles for inflorescence branching and other vegetative phenotypes, and in combination with *tps1,11,12* and *ra3* mutations

In summary, our results suggest that plant TPPs function in a complex with both enzymatically inactive and active TPSs, and the enzymatically inactive TPSs stimulate activity of the active enzymes (Fig. 7D). Our results provide insights for the first time into the combined activity of the two major trehalose gene classes, TPSs and TPPs, in plant development.

RA3 homologs are present in *Arabidopsis*; however, we found that single mutants have no phenotype. To overcome this redundancy, we used CRISPR-Cas9 to create a complete gene family knockout of the 10-member TPP gene family. These mutants enhance branching, reminiscent of the maize phenotype. We are characterizing the TPP knockout lines by RNA-seq analysis, sugar metabolite and nutrient analysis, and meristem measurements.

## Regulation of Cell-to-Cell mRNA Trafficking by AtRRP44A in *Arabidopsis*

V. Varapparambath, M. Gleason [in collaboration with M. Kitagawa, Huazhong Agricultural University, Wuhan, China]

Transcription factors (TFs) and their mRNAs play crucial roles in cell fate determination within plant tissues. These molecules are selectively transported between neighboring plant cells via plasmodesmata (PD), which are membrane-lined channels traversing the cell wall. Despite their importance, the mechanism underlying this transport through PD has remained largely unknown. In previous research, we established a system using *Arabidopsis* seedlings to investigate intercellular transport. Specifically, we focused on a mobile homeodomain TF called KNOTTED1 (KN1). Through this system, we identified mutants with defective trafficking, which were associated with mutations in the same gene—*A. thaliana* ribosomal RNA processing protein 44A (AtRRP44A). RRP44A is a subunit of the RNA exosome complex, responsible for processing and degrading various RNAs in eukaryotes. Notably, both mutants exhibited amino acid substitutions in the conserved catalytic domain. AtRRP44A is expressed in meristem tissues, overlapping with the *Arabidopsis* KN1 homolog known as SHOOT MERISTEMLESS (STM). Intriguingly, *atrrp44a;stm* double mutants had smaller meristems, suggesting that AtRRP44A plays a critical role in regulating stem cells within the shoot meristem by controlling cell-to-cell trafficking of KN1/STM signals. Although AtRRP44A is primarily associated with RNA degradation and processing, this enzymatic activity is not required for RRP44a to transport KN1 mRNA. This indicates that its canonical function can be uncoupled from its role in KN1/STM transport. Further, we explored whether AtRRP44A directly influences KN1/STM mRNA trafficking. Using the MS2 system, we visualized KN1 mRNA movement between cells. Interestingly, this trafficking was impaired in *atrrp44a* mutants, and time-lapse imaging revealed transient targeting of KN1 mRNA to PDs. In summary, our findings suggest that AtRRP44A plays a pivotal role in regulating cell-to-cell mRNA transport, impacting stem cell maintenance, and developmental processes in *Arabidopsis*.

Next, our focus will be on understanding other interacting mRNA molecules of AtRRP44A destined for neighboring cells, especially in long-distance transport into other plant organs. If we identify long-distance mRNA targets of AtRRP44A, it will suggest a global regulation by *Atrrp44a* in mRNA trafficking. Additionally, we will explore another intriguing matter: the presence of AtRRP44A in the nucleus as well as in the plasmodesmata. Previous plasmodesmata proteome studies reported the presence of AtRRP44A in PDs. Therefore, we will investigate the uncoupling of AtRRP44A biological functions between the nucleus and PDs and explore its dual role in RNA processing and transport.

To further understand the mechanism of transport between cells, we will also investigate the dynamics of mobile mRNA particles. By fluorescently labeling the mobile mRNAs, it is possible to track their path. Using chemical inhibitors, we will attempt to disrupt motion of KN1 mRNAs. In tandem, we will use mutants defective in the proteins targeted by each chemical inhibitor to validate our results. This will uncover the role of candidate proteins in the trafficking of KN1 mRNA and will help shed light on how these mRNAs coordinate plant development.

## Mechanisms of Active Compensation between Paralogous Genes in the Maize Meristem

S.D. Iohannes, T. Skopelitis, P. Lindsay [in collaboration with D. Hernandez, CSHL; H. Claeys, Inari Agriculture, Belgium]

Evolutionary innovations are often achieved by co-opting existing molecular structures to perform new functions, a process commonly referred to as "molecular tinkering." Gene duplication is a powerful source of biological innovation, giving rise to duplicates (hereafter, paralogs) that undergo diverse fates and drive evolutionary change. Redundancy between paralogous genes is an intriguing outcome of duplicate gene evolution, and its maintenance over evolutionary time has long been considered a paradox. Genetic studies in yeast and plants have suggested that the ability of ancient redundant duplicates to compensate for dosage perturbations resulting from a loss of function depends on the reprogramming of gene expression, a phenomenon known as active compensation. Our research focuses on the maize trehalose-6-phosphate phosphatases RAMOSA3 (RA3) and TREHALOSE

PHOSPHATE PHOSPHATASE 4 (TPP4), two important meristem development regulators, as a model for studying the reprogramming of paralogs. By using a quantitative imaging approach known as RNA fluorescence in situ hybridization with hybridization chain reaction (RNA FISH-HCR), we confirmed approximately fivefold up-regulation of the TPP4 transcript in the *ra3* knockout mutant (Figs. 7A, 8A). We next performed MNase-defined cistrome-occupancy analysis (MOA-seq) to assess chromatin accessibility changes between wild-type maize ear primordia and the *ra3* mutant. Our preliminary analysis suggests reduced transcription factor binding in the *ra3* mutant, which in turn points to the removal of a repressive state as a mechanism underlying active compensation (Fig. 8B). We are investigating the hypothesis that CNSs in the *TPP4* regulatory region control active compensation by binding to factors that regulate gene expression. We have generated CRISPR-Cas9 constructs to target CNSs identified through the Conservatory computational pipeline in open chromatin regions of the promoter and distal cis-regulatory region of *TPP4*. We hypothesize that specific CNS deletions could lead to loss of active compensation and enhance the *ra3* mutant phenotype or lead to loss of compensatory drift in a wild-type background. Understanding the transcriptional mechanisms of reprogramming among duplicated genes could allow us to fine-tune traits controlled by redundant paralogs and improve the predictability of gene editing outcomes.

## Improving Plant Regeneration Using Morphogenic Factors

T. Clark, S.D. Iohannes, T. Skopelitis [in collaboration with D. Hernandez, CSHL; Z. Chen and A. Gallavotti, Waksman Institute of Microbiology, Rutgers]

Plants possess an incredible degree of developmental plasticity and ability to regenerate. Regeneration can

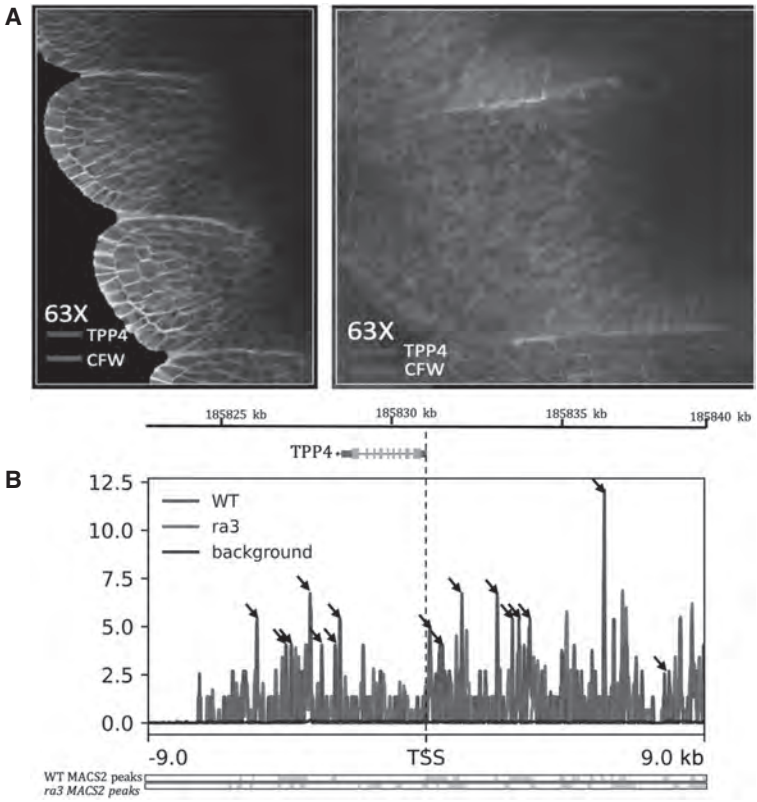



**Figure 8.** (*A*) RNA FISH-HCR micrographs of wild-type (*left*) and *ra3* mutant (*right*) ear primordia. The cell walls are labeled with Calcofluor-white dye (CFW) and the TPP4 transcript is labeled with AlexaFluor 594. (*B*) MOA-seq peaks in the TPP4 regulatory region in wild-type (WT), *ra3* mutant, and background. Reduced transcription factor binding in the mutant is indicated by arrows.

be achieved using tissue culture systems and manipulation of auxin and cytokinin hormones. However, the ability to regenerate shoots is still a bottleneck in several plant species. Ectopic overexpression of plant morphogenic genes is a promising strategy for increasing the regeneration efficiency of recalcitrant plants such as maize. Such morphogenic genes include the transcription factor BABYBOOM (BBM) and the transcription factor-cofactor complex GROWTH REGULATING FACTOR4-GRF-INTERACTING FACTOR1 (GRF-GIF), whose overexpression improves regeneration in maize while avoiding pleiotropic effects. However, the mechanisms through which these factors promote shoot regeneration are not well understood. Our project seeks to investigate the mechanisms of plant regeneration by using a transformation system overexpressing BBM and GRF-GIF, the "GGB" system. We are performing a time course RNA-seq experiment on wild-type and GGB calli (Fig. 9A,B) to assess differential expression of important stem cell developmental regulators. Moreover, we are examining auxin dynamics in GGB calli by visualizing the expression and localization of the PINFORMED1(PIN1) auxin efflux transporter using time-lapse confocal microscopy (Fig. 9D). The characterization of gene expression and hormone dynamics in the GGB system will enable us to better understand and enhance plant regeneration and will make plant transformation and agricultural biotechnology more efficient.

## Elucidating the Role of TERMINAL EAR1 (TE1) in Maize Development and Stress

N. Shanmugaraj, T. Tran, S.D. Iohannes, T. Skopelitis
[in collaboration with R. Chen, Cardozo School of Law, New York; H. Claeys, Inari, Ghent, Belgium; P. Wu, CSHL; E. Demesa Arevalo, Institute for Developmental Genetics, Heinrich Heine University Düsseldorf; M. Kitagawa and L. Liu, Huazhong Agricultural University, Wuhan, China]

RNA binding proteins (RBPs) govern a variety of important plant developmental processes by interacting with RNA targets and controlling posttranscriptional

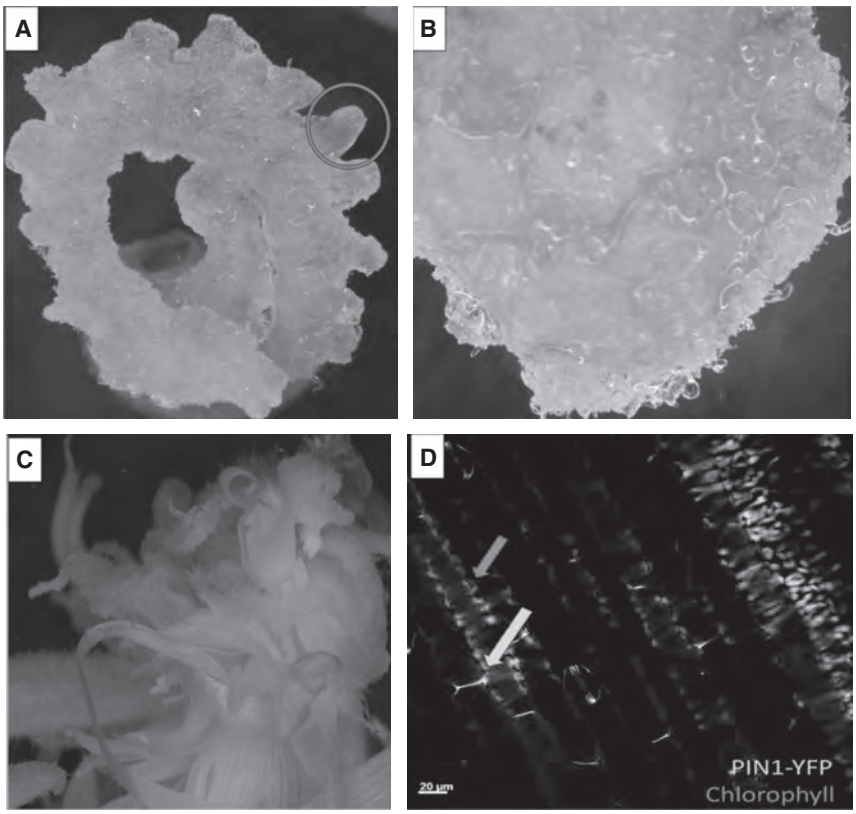



**Figure 9.** (*A*) Maize calli with developing embryos forming on GGB leaf slices (circled). (*B*) Wild-type maize calli with friable edges lacking embryos. (*C*) Developing shoots from GGB embryos. (*D*) Confocal image of PIN1-YFP in maize leaf explants.

RNA metabolism. Maize TERMINAL EAR 1 (TE1) encodes a predicted RBP containing three conserved RNA recognition motifs. Loss-of-function mutants have pleiotropic developmental defects, including faster leaf initiation, altered phyllotaxy, dwarf stature, and a feminized tassel. However, the mechanism of TE1 action is poorly understood. Previously, RNA-immunoprecipitation and sequencing (RIP-Seq) using TE1-yellow fluorescent protein (YFP)-expressing maize ear primordia revealed TE1 binding to mRNAs of multiple maize developmental genes. However, mRNA-Seq analysis of *te1* mutants did not identify a significant change in the levels of most candidate TE1-bound target mRNAs.

In contrast, our shotgun proteomics analysis revealed that the proteins encoded by candidate TE1-bound mRNAs were more abundant in *te1* mutants, suggesting TE1 may control its targets at a posttranscriptional level. However, the sensitivity was insufficient to detect proteins corresponding to many of the TE1 mRNA targets. As many of the developmental mRNAs bound by TE1 encode transcription factors, we tried to improve the sensitivity by preparing nuclear extracts for shotgun proteomics. We successfully extracted the nuclear and cytoplasmic proteins from immature tassel meristems of *te1* mutants and wild-type sibling plants and have sent the samples for proteomics analysis by the Mass Spectrometry facility at the Carnegie Institute, Stanford University. This analysis will help us reveal whether the protein levels of TE1-bound developmental transcription factors change in *te1* mutants. If we see an increase in protein levels in *te1* mutants, this will support the hypothesis that TE1 represses translation of its bound mRNAs.

We also identified TE1–YFP interacting proteins using proteomics, and Gene Ontology analysis revealed enrichment in stress granule assembly and regulation of translation categories. Consistent with this, a functional TE1–YFP fusion localized to cytoplasmic puncta. These puncta accumulate and enlarge following heat shock and colocalize with a stress granule marker. To ask whether TE1 functions in heat stress responses, we tested germination following heat treatments and found that *te1* mutants were more sensitive to heat stress. Our preliminary results suggest that TE1 binding to developmental mRNAs and sequestering them in cytoplasmic granules may be part of a mechanism to negatively regulate development. Thus, we hypothesize that TE1 may function to control the important trade-off between plant growth and stress responses. Further elucidation of TE1 functions under stress and optimization of stress treatments are ongoing.

## PUBLICATIONS

Guillotin B, Rahni R, Passalacqua M, Mohammed MA, Xu X, Raju SK, Ortiz-Ramirez C, Jackson D, Groen SC, Gillis J, Birnbaum KD. 2023. A pan-grass transcriptome reveals patterns of cellular divergence in crops. *Nature* **617:** 785–791. doi:10.1038/s41586-023-06053-0

Iohannes SD, Jackson D. 2023. Tackling redundancy: genetic mechanisms underlying paralog compensation in plants. *New Phytol* **240:** 1381–1389. doi:10.1111/nph.19267

Kitagawa M, Tran T, Jackson D. 2023. Traveling with purpose: cell-to-cell transport of plant mRNAs. *Trends Cell Biol* **34:** 48–57. doi:10.1016/j.tcb.2023.05.010

Lee JS, Lowell JL, Whitewater K, Roane TM, Miller CS, Chan AP, Sylvester AW, Jackson D, Hunter LE. 2023. Monitoring environmental microbiomes: alignment of microbiology and computational biology competencies within a culturally integrated curriculum and research framework. *Mol Ecol Resources* doi:10.1111/1755-0998.13867

Lindsay P, Swentowsky KW, Jackson D. 2023. Cultivating potential: harnessing plant stem cells for agricultural crop improvement. *Mol Plant* **17:** 50–74. doi:10.1016/j.molp.2023.12.014

Xu X, Jackson D. 2023. Single-cell analysis opens a goldmine for plant functional studies. *Curr Opin Biotechnol* **79:** 102858. doi:10.1016/j.copbio.2022.102858

Yang N, Wang Y, Luo Y, Guo X, Li Y, Yan J, Shao W, Wei W, Jackson D, Zhang Z, et al. 2023. A spatial transcriptome map of the developing maize ear. *Nat Portfolio* doi:10.21203/rs.3.rs-3037245/v1

### *In Press*

Cahn J, Regulski M, Lynn J, Ernst E, de Santis Alves C, Ramakrishnan S, Chougule K, Sharon W, Zhenyuan L, Xu X, et al. 2024. MaizeCODE reveals bi-directionally expressed enhancers that harbor molecular signatures of maize domestication. bioRxiv doi:10.1011/2024.02. 22.581585

Nie Y, Wang H, Zhang G, Ding H, Han B, Liu L, Shi J, Du J, Li X, Li X, et al. 2024. The Maize PLASTID TERMINAL OXIDASE (PTOX) locus controls the carotenoid content of kernels. *Plant J* doi:10.1111/tpj.16618

Xu X, Passalacqua M, Rice B, Demesa-Arevalo E, Kojima M, Takebayashi Y, Harris B, Sakakibara H, Gallavotti A, Gillis J, Jackson D. 2024. Large-scale single-cell profiling of stem cells uncovers redundant regulators of shoot development and yield trait variation. bioRxiv doi:10.1011/2024.03.04.583414

# EXPLORING AND EXPLOITING GENOTYPE-TO-PHENOTYPE RELATIONSHIPS IN PLANT DEVELOPMENT, EVOLUTION, AND AGRICULTURE

**Z.B. Lippman** D. Ciren, B. Fitzgerald, I. Gentile, J. He, A. Hendelman, A. Lanctot, S. Qiao, G. Robitaille, M. Santo Domingo Martinez, J. Satterlee, B. Seman, H. Shohat, U. Vrudhula, S. Zebell

Our laboratory focuses on understanding the role of genes and their variations—stemming from both natural and engineered mutations—in shaping phenotypes. The laboratory is pursuing two fundamental biological questions with applied importance. First, we explore how gene duplications and their derived paralogs evolve genetically and functionally over short and long evolutionary timescales. Second, we study how the evolution of *cis*-regulatory sequences, which control when, where, and to what level the genes they regulate are expressed, influences quantitative trait variation within and across species. To address these questions, we are leveraging the powerful and synergistic genomic and genetic resources and tools afforded by the Solanaceae family, and particularly in the *Solanum* genus. In collaboration with both internal and external partners, we have established a pan-genome of *Solanum*, which includes more than 20 major and minor crop species. Concurrently, we have implemented genome editing for numerous species within this group, allowing us to tap into an evolutionary perspective to dissect our research questions. We also leverage existing reference genomes across the plant kingdom to study the diversification of paralogs and *cis*-regulatory control over deep time, using *Arabidopsis* as a comparative model to investigate genotype-to-phenotype relationships beyond the Solanaceae. Importantly, the evolution of paralogs and *cis*-regulation are interconnected, as gene duplications allow for the accumulation of often cryptic mutations in both coding and regulatory regions, thereby promoting diversification of redundancy relationships between paralogs and paths to new functions. In this regard, the laboratory is embarking on new research to study the mechanisms underlying the emergence of new traits, or morphological innovations, using several remarkable examples in the Solanaceae that mirror similar innovations arising through convergent evolution in other plant families. Here again, both gene duplication and *cis*-regulatory change are central, and the system and tools we have developed are providing new insights into this evolutionary question, which has lacked rigorous genetic dissection. The discoveries we are making, especially on the hidden roles of cryptic variation and epistasis in phenotypic variation, are helping address the underappreciated challenge of predictability in crop engineering, especially in indigenous crops that have the potential to promote sustainability. The projects described below are presented according to the lead researcher, but all result from collaborations within our laboratory and CSHL, as well as with outside colleagues and friends.

## Conservatory: A Pan-Plant Gene-Centric Computational Approach to Reveal Conserved *cis*-Regulatory Sequences and Their Functions

A. Hendelman

Conserved noncoding sequences (CNSs) play an integral role in the genome, acting as key regulators of gene expression despite not encoding proteins themselves. Despite the known rapid evolution of sequences in *cis*-regulatory regions compared with coding regions, these sequences can maintain conservation across species at increasing evolutionary timescales. However, discovering and characterizing this conservation has been limited by biased sampling of species' genomes for sequencing, reference genome quality, and the challenge of establishing gene orthology beyond short evolutionary timescales. Given that CNSs likely harbor critical *cis*-regulatory elements (CREs) in the regulation of gene expression, providing binding

sites for transcription factors and other regulatory entities, we set out more than two years ago to use a new gene-centric approach to discover CNSs and characterize their genomic signatures and functions.

Over the past year, we have substantially enhanced our understanding and analysis of CNSs through the development of Conservatory, now upgraded to version 2.0. Conservatory is a computational tool developed in collaboration with I. Efroni (Hebrew University). This version advances our exploration into the evolutionary dynamics and regulatory functions of the most deeply conserved noncoding sequences (dCNSs) across a broad evolutionary spectrum. Our comprehensive data set now includes 314 genomes from 285 species across 72 families, covering more than 488 million years of evolutionary divergence from mosses to flowering plants.

We chose six reference families to refine our analysis, minimizing errors from genome assembly and annotations. Our improved methods led to the identification of more than 2.3 million family-specific CNSs, highlighting the role of *cis*-regulatory mutations in gene expression and speciation. Among these, approximately 55,000 dCNSs were identified, demonstrating varying levels of conservation across different plant clades and highlighting their indispensable role in gene regulation through deep evolutionary time.

We have launched a dedicated website this year (https://conservatorycns.com), providing the plant research community with access to the Conservatory analysis. This online resource allows for the download of complete Conservatory analysis results for all 314 plant genomes incorporated in the Conservatory V2.0 analysis, facilitating broader research and collaboration within the field. The conservation of these noncoding sequences across diverse taxa highlights their indispensable role in the regulatory networks that govern gene expression patterns. By investigating CNSs and dCNSs, we are gaining new insights into mechanisms that permit or restrict *cis*-regulatory change, shedding light on evolutionary adaptations and regulatory complexity that impact growth and development. We have compiled a list of candidate genes for CRISPR functional studies. Our CRISPR targeting of selected dCNSs in *Arabidopsis* aims to uncover the conservation of *cis*-regulatory functions, with ongoing characterization of the resultant mutant plants.

## Extreme Restructuring of *cis*-Regulatory Regions Controlling a Deeply Conserved Plant Stem Cell Regulator

D. Ciren

*CLV3* is a highly conserved plant stem cell regulator, and we explored how *cis*-regulatory control of this gene evolved between *Arabidopsis thaliana* and tomato over a divergence time of ~125 million years (MYs). Despite conservation of gene function and expression pattern, the *cis*-regulatory sequences of these two genes are substantially diverged. We generated more than 70 deletion alleles in the upstream and downstream regions of *CLV3* in both species using CRISPR editing techniques and compared their individual and combined effects on a shared phenotype, the carpels that make locules in fruits. Although most *cis*-regulatory function is restricted upstream of *CLV3* and additive in effect in tomato, *cis*-regulatory control of *CLV3* function in *Arabidopsis* is partitioned equally and redundantly between upstream and downstream regions. This study indicates that major reconfigurations of *cis*-regulatory sequence space is a potent evolutionary force influencing genotype-to-phenotype relationships of genes conserved over deep time. Our findings shed light on these evolutionary dynamics and also emphasize the importance of lineage-specific approaches in deciphering the spatial architecture of *cis*-regulation, which is foundational to understand how the widespread natural *cis*-regulatory variation now being revealed by pan-genomics and engineered alleles constructed by genome editing can impact phenotypes in lineage-specific ways.

In a complementary study, we explored *cis*-regulatory control of a critical flowering regulator, the anti-florigen gene *SP5G*. Wild species of tomato flower fast under short days, whereas domesticated species are largely insensitive to day length. This change allowed their geographical expansion. Again using CRISPR, we engineered 18 mutant alleles upstream of and downstream from *SP5G* in a daylength-sensitive introgression in domesticated tomato, which carries the wild species daylength-sensitive allele of *SP5G*. Mutating various CNSs, revealed by Conservatory and open chromatin analyses, generated weak, moderate, and strong changes in flowering time. Notably, in contrast to previously published work that suggested a critical downstream deletion of 52 bp caused loss of daylength sensitivity, no single allele matched the flowering time

response of domesticated tomato. Thus, it is possible that multiple mutations within CREs were required to generate phenotypic divergence in flowering time, and in the future we will explore this hypothesis with further experiments aimed at exploring interactions among CREs in the regulation of *SP5G*. We are also currently investigating the relationship between phenotype and the timing/level of gene expression, by assaying the diurnal expression of this *cis*-regulatory allelic series.

## Genomic and Functional Dissection of CLE Gene Duplication History across Angiosperms

I. Gentile

The *CLE* gene family is a well-researched group of signaling peptides in plants involved in a range of processes from development to stress response. Their rapid evolution and presence in high copy numbers present considerable challenges for computational and functional analyses. We are using a combination of genomic analysis, computational modeling, and functional genetics to uncover their evolutionary and functional relationships in angiosperms. By scanning more than 2,000 genomes from more than 900 plant species, we have de novo annotated *CLE* genes and captured the entire sequence diversity of the family in angiosperms. Our analysis revealed many previously unannotated *CLE* genes, which exposed the entire spectrum of sequence diversity to predict mutational effects through modeling (Potts model). We validated our estimations by CRISPR base-editing experiments and further demonstrated a newfound power to detect asymmetries in paralog diversification, with duplicated copies accumulating less favorable substitutions over time. This correlated with the weaker paralogs also degrading their *cis*-regulatory regions more rapidly. These strategies converged to allow us to resolve genotype-to-phenotype links of these highly redundant genes by targeting all 52 CLEs in the major *Solanum* crop tomato. Notably, in most cases, phenotypes were revealed only with the generation of higher-order mutants from *CLE* family members for which we computationally predicted putative functional relationships, thereby illustrating that a deep evolutionary–driven computational modeling approach can resolve pervasive redundancy and allow functional analysis of genotype-to-phenotype relationships at scale.

In addition to this broad approach, we are conducting a detailed analysis of paralog evolution by utilizing *CLV3* as a model gene within the *Solanum*. We uncovered substantial diversity in *CLV3* haplotypes, particularly within the lineage of the "spiny eggplants." For instance, whereas species such as *Solanum cleistogamum*—a foraged berry native to Australia known as desert raisin—maintain a single copy of this gene, other species exhibit segmental duplications, as observed in the Australian wild species *Solanum prinophyllum* (forest nightshade). By integrating CRISPR-based techniques with comparative transcriptomics, we are dissecting the functional dynamics of *CLV3* in the context of gene duplication.

## A Conserved *cis*-Regulatory Sequence of a Key Flowering Gene Encodes Antagonizing Elements that Mediate Phenotypic Robustness

A. Lanctot

Flowering is an essential developmental transition in plants and consequently is regulated by deeply conserved molecular mechanisms. The importance of flower formation for plant fitness causes canalization of many floral traits—floral morphology is remarkably consistent both within a given individual plant and within species, despite huge diversity in floral form among diverged species. This conservation is driven by the activity of two deeply conserved floral regulators, the transcription factor LFY and its cofactor, the F-box protein UFO. This project aims to query how *cis*-regulatory control of the temporal and spatial expression patterns of these two genes across a broad evolutionary range of flowering plants affects gene function and flowering in diverse plant families. Using Conservatory to identify deeply conserved noncoding sequences (CNSs) in the regulatory regions of these genes, we are targeting these regions using CRISPR in species within two evolutionarily diverged plant families, the Brassicaceae and the Solanaceae. We have found that disruption of these CNSs causes severe and distinct phenotypes in different species.

UFO promotes flower formation in *Arabidopsis*, and *ufo* mutants lack floral organs such as petals and stamens. We found that deletions of conserved *cis*-regulatory sequence in the *Arabidopsis* UFO promoter causes petal number variation, with the perturbation

of different conserved regions leading to different degrees and directions of decanalization. These results suggest that CNSs are strong predictors of *cis*-regulatory functionality and can affect their target gene's function in distinct ways. We are currently combining these *cis*-regulatory mutants with diverse *Arabidopsis* accessions to uncover potential cryptic variation in *UFO* regulators, revealed by this decanalized state. Not all species show canalized petal number; the close *Arabidopsis* relative *Cardamine* varies from zero to four petals. Using the same constructs to disrupt UFO promoter CNSs in *Cardamine*, we are assessing how UFO CNSs affect floral development in a species in which petal number is already decanalized.

In tomato and other Solanaceae species, disrupting *UFO* function causes severe flower formation defects. Targeting a dicot-level CNS in a region of open chromatin of the tomato *UFO* promoter causes severe floral defects. Some alleles affecting this CNS show proliferative branching of inflorescences, which impedes flower formation, a phenotype similar to *UFO* null mutants. Unexpectedly, other alleles showed a distinct, single flower inflorescence with extremely enlarged sepals. This phenocopies the tomato mutant *tmf*, which results from precocious activation of *UFO*, suggesting a transcriptional repressor may bind to the sequences deleted in these alleles. We are performing molecular assays to determine how *UFO* expression levels, domains, and timing are altered in these alleles and connecting these data to candidate transcriptional activator and repressor binding sites. This single deeply conserved *cis*-regulatory sequence demonstrates the information-rich nature of CNSs—as targets of distinct transcriptional regulators, CNSs may be hubs of intricate *cis*-regulatory activity, making them informative regions to study. This project is advancing our knowledge of how CNSs regulate essential developmental pathways across deep evolutionary time, while moderating disparate gene expression patterns in distinct species.

## Genetic Interactions among Cryptic Variants in Paralogs Elicit a Continuum of Phenotypic Variation through Global Epistasis

S. Zebell

Variation in gene regulatory space is the source of much of the heritable variation in phenotypes between members of the same species. Regulatory variants can vary in their phenotypic expressivity, with effects that are frequently cryptic, displaying little to no phenotype. Molecular genetics suggests that regulatory variants within the same circuit likely interact, but the phenotypic consequence of epistasis between variants in *trans*-regulatory factors and their binding motifs found in *cis* to their targets has been difficult to demonstrate because of a lack of model systems with a wide range of precise, easily quantifiable phenotypes. To address the behavior of cryptic regulatory alleles in populations, we employed branching of inflorescences in tomato as a quantitative assay reflecting modified dose-dependent relationships among developmental transcription factors. We previously characterized the MADS-box transcription factors *JOINTLESS2* (*J2*) and its ancestral paralog *ENHANCER OF JOINTLESS2* (*EJ2*) as dosage-sensitive cryptic regulators of inflorescence architecture; the inflorescences of both single mutants are normal, but varying the dosage of one factor in the background of the other mutant leads to increasingly branched inflorescences, and double null mutants display an extreme phenotype of highly proliferated inflorescences. We identified variation in *cis*-regulatory motifs upstream of *EJ2* in wild relatives of tomato that led to an inflorescence branching phenotype in the *j2* background, and validated that genome editing of this regulatory region in tomato reproduces the branching effect. We characterized two PLETHORA transcriptional factors, PLT7 and PLT3, that bind a motif in this *EJ2 cis*-regulatory region and showed that CRISPR mutants of these *PLT* genes mimic the *J2–EJ2* cryptic, dose-dependent redundancy relationship in controlling inflorescence branching.

We then investigated the interplay between variants in this gene regulatory network by generating populations segregating for multiple *ej2 cis*-regulatory alleles and null mutations in the other three network genes and quantified branching from thousands of plants from more than hundreds of genotypes. Through a global epistasis modeling approach, we were able to demonstrate that these cryptic alleles have a predominant additive component to their phenotypic contribution, with a nonlinear relationship between their latent phenotype and their actual phenotype, setting a biological threshold of overall network activity that canalizes unbranched, wild-type inflorescences. However, optimal model fitting required accounting for

synergistic epistatic interactions between members of each paralog pair, and a surprising negative network interaction between the *PLT*s and their *J2–EJ2* targets, reflecting epistatic masking of a downstream gene by an upstream regulator. Our work demonstrates that allelic variation within the same regulatory network can produce a spectrum of phenotypic outcomes through an interplay of additive, synergistic, and masking genetic interactions.

## Epistatic Contingencies among Flowering Gene Paralogs Drove Independent Domestications of *Solanum* Crops

H. Shohat

A growing population and climate change are threatening agricultural systems. Our food system relies on less than a dozen major crops. Thousands of indigenous edible plant species hold the potential to diversify and enhance our food system. Pan-genomics and genome-editing technologies are rapidly advancing the discovery and translation of knowledge of fundamental plant biology to crop breeding, offering opportunities to elevate underutilized crops. Decades of research on model crops has revealed that genetic changes in the universal florigen-based flowering system have been repeatedly selected in both natural and agricultural systems to promote adaptation and productivity. Florigens and their antiflorigen family members are protein hormones that regulate the transition to reproductive growth (flowering), hence affecting the balance between vegetative (stems and leaves) to reproductive (flowers, fruits, and seeds) growth. Engineering allelic diversity in florigen gene orthologs can rapidly improve yield traits in indigenous crops; however, predicting desired phenotypic outcomes is surprisingly challenging.

We have found that there is widespread species-specific florigen–antiflorigen genetic variation and epistatic relationships among indigenous crops in the *Solanum* family, which can hinder application of previous knowledge to new crops. Uniting our recently established *Solanum* pan-genome comprising more than 20 species with a large-scale forward genetic approach, we demonstrated that variants of an antagonistic pair of florigen genes were selected during the domestication of the indigenous African eggplant *Solanum aethiopicum* and its relative *Solanum macrocarpon*. Flowering time variations in African eggplants are driven by species-specific loss-of-function alleles of the flowering repressor *SP5G*, followed by copy number variations in the flowering-promoting gene *FTL1*. Notably, allelic variants of these paralogs were also selected to enable early flowering during the independent domestication of tomato. Despite the shared selection of a pair of flowering regulators during independent domestication events of *Solanum* crops, further genomic and genetic analyses revealed that mutational types, allelic strength, gene function, and paralog relationships are species-specific contingencies that influence breeding by both forward and reverse genetics. Our study is revealing the dynamic evolution of paralogs in shaping trait variation in *Solanum* crops and provides a new framework for leveraging current comparative genomics and functional genetic tools to improve the predictability of genotype-to-phenotype relationships for efficient and rapid domestication and breeding of indigenous crops.

## Dissecting the Genetic and Molecular Mechanisms of the Inflated Calyx Syndrome

J. He

The evolution of morphological novelties, or innovations, has long been a topic of interest among biologists. The diverse floral organ traits found in Solanaceae species have fascinated generations of scientists. One such trait, the so-called "Chinese lantern" or inflated calyx syndrome (ICS), is observed in genera such as *Physalis*, *Withania*, and *Nicandra*. In these plants, the sepals undergo dramatic growth after anthesis, forming balloon-like structures that encapsulate the fruits. However, studying the genetic and molecular mechanisms underlying these phenomena has been hindered by a lack of genomic resources and genetic tools. As a result, conjecture and questionable data have led to dubious conclusions regarding the key regulators of this developmental process. Recent follow-up studies have only brought more confusion.

To enable a more careful and thorough molecular and genetic dissection of ICS, we have established *Physalis* as a new reference system. We generated high-quality genomes and developed CRISPR in *Physalis grisea* and *Physalis pruinosa*. Interestingly, our experiments showed that loss-of-function mutations of previously proposed key regulators and other candidate

genes failed to disrupt ICS. This prompted us to start from square one and explore ICS with a systematic and unbiased, integrated phenotypic, molecular, and genetic approach.

We first used time-lapse imaging to reveal the developmental dynamics of ICS. Notably, the process of inflation begins with the expansion of the calyx base section within a critical time window of 48–72 hours, followed by a 24- to 48-hour pause and then continued progress to full inflation. Transcriptome profiling of hundreds of individual sepal whorls from flowers captured a continuous progression of calyx developmental stages. This resulted in a fine temporally resolved map of the molecular signatures of ICS. One key finding is that inflation can be detected molecularly through dozens of newly discovered marker genes before morphological inflation commences.

Following this discovery, we profiled the transcriptome of the tip, mid, and base sections of developing sepals, providing a more detailed dissection of the initiation stage of ICS. We identified differentially expressed genes (DEGs) between pre-initiation sepals and sepals committed to inflation. Notably, among the DEGs are transcription factors involved in various aspects of plant development including cell division, organ growth, and hormone response. We are currently interrogating these candidate genes as potential regulators of ICS using a pipeline of single and higher-order CRISPR targeted mutagenesis.

## Functional Analysis of Paralog *cis*-Regulatory Sequences and Links to Morphological Innovations

M. Santo Domingo Martinez

Paralogs are genes that have arisen from a recent duplication, comprising the gene body and its regulatory sequences. As time passes, mutations accumulate in both sequence contexts, modifying the redundancy relationships of paralogs and promoting functional divergence.

A pair of paralogs controlling meristem development, *CLV3* and *CLE9*, are known to have different epistatic interactions in several Solanaceae species because of mutations in their coding and regulatory sequences. One predicted critical *cis*-regulatory sequence is a 40-bp CNS shared in both genes in two species, tomato and groundcherry. By generating several CNS alleles in different tomato and groundcherry backgrounds, we have found this sequence to be important in the control of meristem size, affecting the epistatic interactions and active compensation relationship between these two genes in a background-dependent manner.

In related work, we are taking advantage of genome editing at scale in *Arabidopsis* to deepen our knowledge about the role of CNSs in paralog functional relationships. We have used the literature to identify known and candidate paralog gene pairs having epistatic relationships, and we are currently using CRISPR to mutate shared CNSs in wild-type and reciprocal paralog null mutant backgrounds. With this work, we are aiming to understand the role and importance of CNSs in paralog evolution and functionality, and how they tune their epistatic interaction dynamics.

Finally, we are linking paralogs, *cis*-regulatory evolution, and morphological innovations by investigating heteranthery, an adaptive trait in which two or more stamen morphotypes exist in the same flower, usually associated with optimizing insect pollination. Like prickles and the inflated calyx syndrome, heteranthery has evolved independently multiple times through plant history.

In *Solanum*, this trait has evolved independently several times, giving rise to different heteranthery phenotypes associated also with changes in floral symmetry. Among the diversity of heteranthery in the family, the most conspicuous cases involve species having one exceptionally large ventral anther with a dragon tongue–like shape. We are embarking on a comprehensive morphological and molecular phenotypic characterization in several pairs of sister species that are distinguished by heteranthery. Transcriptome analyses of individual early developing anthers have revealed candidate genes that control organ asymmetry in other species and organs, and we are developing genome editing to study the functions of these genes.

## Convergent Evolution of Plant Prickles Is Driven by Repeated Gene Co-Option over Deep Time

J. Satterlee

Emerging opportunities in genomics and genetics are reinvigorating efforts to address a long-standing fundamental question in evolutionary biology:

Independent emergence of the same traits (convergence) is common and often relies on shared genetic programs over the short term; however, whether this is true for convergence over long evolutionary timescales remains unknown.

We leveraged instances of independent gains and losses of a key morphological innovation in plants: prickles. Prickles are sharp projections that evolved independently at least 40 times over more than 400 million years, conferring critical fitness advantages in herbivore defense, interplant competition, climbing growth, and more. The rose is perhaps the most well-known prickly plant, although its prickles are vernacularly referred to as thorns. Prickles are the most widespread form of sharp outgrowth found in the plant kingdom. Significantly, losses of prickles have been central to domestication in multiple crops, including in the *Solanum*.

Over thousands of years, prickleless (*pl)* genotypes were selected during parallel domestication events in the "spiny" (i.e., prickled) clade of the *Solanum*. We showed that in living and herbarium samples across the *Solanum*, the same gene that activates the plant cell division–promoting hormone cytokinin was mutated in at least 16 instances of prickle loss in cultivated and wild species. Furthermore, by leveraging reference and pan-genomes across flowering plants (angiosperms), we demonstrated that homologs of this gene (*LONELY GUY: LOG*) are also mutated and associated with independent losses of prickles separated by >150 million years. Excitingly, we find this includes "thornless rose." To understand the genetic context of this co-option event, we performed coexpression analyses, which revealed that *LOG* family member duplication and expansion enabled functional diversification. Finally, we assembled genomes and developed genome editing in two new *Solanum* model systems, a common weed and an indigenous foraged berry crop from Australia, wherein we eliminate prickles without pleiotropic consequences.

Our findings highlight the role of cytokinin signaling in other major trait innovations, such as tubers and root nodules. Indeed, the genetic repeatability uncovered in our study illuminates a much wider role for cytokinin in morphological innovation. Finally, we demonstrate how this genetic knowledge can be widely used to easily eliminate prickles in food and ornamental crops.

## PUBLICATIONS

Aguirre L, Hendelman A, Hutton S, McCandlish DM, Lippman ZB. 2023. Idiosyncratic and dose-dependent epistasis drives variation in tomato fruit size. *Science* **382:** 315–320. doi:10.1126/science.adi5222

*In Press*

Ciren D, Zebell S, Lippman ZB. 2024. Extreme restructuring of *cis*-regulatory regions controlling a deeply conserved plant stem cell regulator. *PLoS Genet* **20:** e1011174. doi:10.1371/journal.pgen.1011174

Satterlee JW, Alonso D, Gramazio P, Jenike KM, He J, Arrones A, Villanueva G, Plazas M, Ramakrishnan S, Benoit M, et al. 2024. Convergent evolution of plant prickles is driven by repeated gene co-option over deep time. bioRxiv doi:10.1101/2024.02.21.581474

# EPIGENETIC MECHANISMS OF GENE REGULATION AND INHERITANCE

**R. Martienssen** S. Blau, J. Cahn, F. Campetella Mayoral, K-H. Cheng, C. de Santis Alves, E. Ernst, H-S. Kim, A.A. Lakhani, J. Lynn, C. Mateo-Elizalde, U. Ramu, M. Regulski, J. Simorowski, J. Steinberg

The genomes of plants and animals are reprogrammed before and after meiosis to reset epigenetic marks in the next generation. When reprogramming fails, marks are retained, leading to "transgenerational" epigenetic inheritance. Our laboratory studies fundamental mechanisms of epigenetic inheritance in plants and fission yeast, and we are applying our discoveries to mammalian cells as well as to novel crops such as the clonal aquatic macrophyte *Lemna* (common duckweed). We are focusing on the roles of RNA interference (RNAi), DNA methylation, and histone modification. In fission yeast and mammalian cells, we have found that RNAi is required for centromere silencing and genome stability, but is also required to resolve replication-induced transcription stress. In *Arabidopsis*, we have discovered that small RNAs in pollen are heavily modified by pseudouridine, and that this modification is shared with some PIWI-interacting RNAs (piRNAs) in the mouse, as well as some microRNAs (miRNAs) in mouse and *Caenorhabditis elegans*, and may be related to the immunogenicity of inherited small RNA. We continue to develop duckweeds for biofuel by sequencing the genomes of several species and by developing an efficient transformation system enabling substantial production of oil in *Lemna japonica*. In 2023 we said farewell to Asad Lakhani, who left for a position in industry.

## Chromatin Remodeling of Histone H3 Variants by DDM1 Underlies Epigenetic Inheritance of DNA Methylation

J. Cahn, J. Lynn, H-S. Kim, U. Ramu, R. Martienssen
[in collaboration with S-C. Lee, Abbot Pharmaceuticals, Seoul, South Korea; B. Berube, Ohalo Genetics, Santa Cruz, CA; C. LeBlanc and Y. Jacob, Yale University; J. Calarco, deceased; D. Adams, J. Ipsaro, and L. Joshua Tor, CSHL; D. Grimanelli, IRD Montpellier; P. Voigt, Babraham Institute, Cambridge, UK]

Nucleosomes block access to DNA methyltransferase unless they are remodeled by DECREASE in DNA METHYLATION 1 (DDM1$^{LSH}$/HELLS), a Snf2-like master regulator of epigenetic inheritance. We have shown that DDM1 promotes replacement of histone variant H3.3 by H3.1. In *ddm1* mutants, DNA methylation is partly restored by loss of the H3.3 chaperone HIRA, whereas the H3.1 chaperone CAF-1 becomes essential. The single-particle cryogenic electron microscopy (cryo-EM) structure at 3.2 Å of DDM1 with a variant nucleosome reveals engagement with histone H3.3 near residues required for assembly and with the unmodified H4 tail. An amino-terminal autoinhibitory domain inhibits activity, whereas a disulfide bond in the helicase domain supports activity. DDM1 colocalizes with H3.1 and H3.3 during the cell cycle and with the DNA methyltransferase MET1$^{Dnmt1}$, but is blocked by H4K16 acetylation. The male germline H3.3 variant MGH3/HTR10 is resistant to remodeling by DDM1 and acts as a placeholder nucleosome in sperm cells for epigenetic inheritance.

## Pseudouridine Guides Germline Small RNA Transport and Epigenetic Inheritance

C. de Santis Alves, F. Campetella Mayoral, R. Martienssen
[in collaboration with R. Herridge, University of Central Otago, New Zealand; J. Dolata, Adam Mickiewicz University, Poznań, Poland; F. Borges, INRA Versailles, Paris, France; F. Van Ex, Inari LLC, Ghent, Belgium; J-S. Parent, Agriculture Canada, Ottawa; T. Kouzarides, Cambridge University; A. Schorn, CSHL]

Epigenetic modifications that arise during plant and animal development, such as DNA and histone modification, are mostly reset during gamete formation, but some are inherited from the germline—including those marking imprinted genes. Small RNAs guide these epigenetic modifications, and some are also inherited by the next generation. In *C. elegans*, precursors for these inherited small RNAs have poly(UG) tails, but how inherited small RNAs are distinguished in other animals and plants is unknown. Pseudouridine

(Ψ) is the most abundant RNA modification but has not been explored in small RNAs. Here, we develop novel assays to detect Ψ in short RNA sequences, demonstrating its presence in mouse, *C. elegans*, and *Arabidopsis* miRNAs and their precursors. We also detect substantial enrichment in germline small RNAs—namely, epigenetically activated small interfering RNAs (easiRNAs) in *Arabidopsis* pollen, and PIWI-interacting piRNAs in mouse testis. In pollen, pseudouridylated easiRNAs are localized to sperm cells, and we found that *PAUSED/HEN5* (*PSD*), the plant homolog of Exportin-t, interacts genetically with Ψ and is required for transport of easiRNAs into sperm cells from the vegetative nucleus (Fig. 1). We further show that Exportin-t is required for the triploid block: chromosome dosage-dependent seed lethality that is epigenetically inherited from pollen. Thus, Ψ has a conserved role in marking inherited small RNAs in the germline. We speculate this may reduce the immunogenicity of germline small RNA on fertilization, resembling pseudouridylated mRNA vaccines and antisense oligonucleotides in this respect.

## Balancing Argonautes: The Role of AGO2 in RNAi and Transgenerational Inheritance

J. Lynn, S. Blau, R. Martienssen

The gene silencing outcome of RNAi in plants crucially relies on the interaction of small RNAs with Argonaute (AGO) proteins, forming AGO-effector complexes that target homologous RNAs. Although seminal work on plant AGOs has identified features associated with small RNA sorting, the overlap in small RNA binding preferences among closely related AGOs blurs the lines in how plants regulate ectopic RNAi. To address these questions, we leverage two model systems: *Arabidopsis thaliana*, with its extensive collection of epigenetic mutants, and the *b1* paramutation system in maize that offers unique insights into transgenerational inheritance. AGO2 is required for broad-spectrum resistance against viral infection by cleaving viral transcripts, whereas closely related AGO4-clade proteins are required for RNA-directed DNA methylation (RdDM). Our results provide support for AGO2's involvement in suppressing

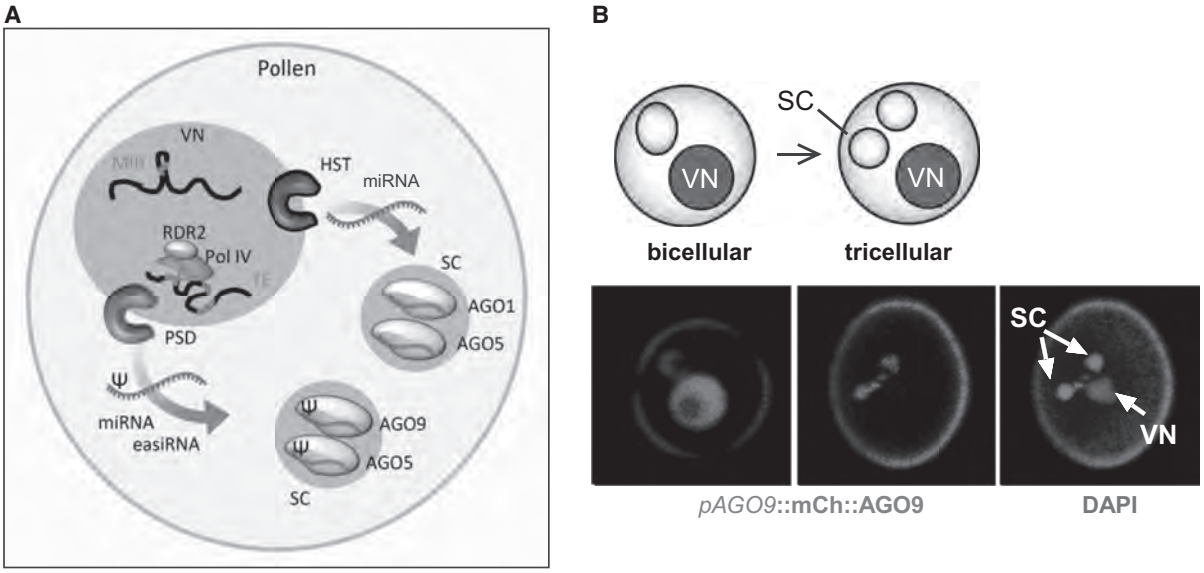



**Figure 1.** Pseudouridine and Exportin-t mediate epigenetic inheritance. (*A*) Model for transport of small RNA during pollen development. As the pollen matures, 21/22 nt easiRNAs and 24 nt siRNAs are produced by Pol IV in the vegetative nucleus (VN) and pseudouridylated along with some miRNAs. Pseudouridylated miRNAs and easiRNAs (Ψ) are exported via PSD (PAUSED, Exportin-t) into sperm cells (SC), where they are loaded onto AGO5 and AGO9, whereas unmodified miRNAs are transported via HST (HASTY, Exportin-5) and loaded onto AGO5 and AGO1. (*B*) mCherry::AGO9 fusion protein was expressed from the *AGO9* promoter (*pAGO9*) in developing pollen. DAPI was used to stain nucleic acids. AGO9 is present in the VN at the binuclear stage but is enriched in the sperm cells in mature pollen. Pseudouridylated sperm cell easiRNAs from pollen mediate dose-dependent lethality (the triploid block) by targeting maternally expressed imprinted genes in the seed.

ectopic RdDM and its contribution to the transmission of epigenetic silencing through the male germline. Plant AGOs exhibit dynamic expression, often being shuttled between cytoplasm and nucleus, and are in some cases phloem-mobile, allowing for rapid deployment of silencing throughout the plant. To investigate the functional diversity of AGOs in maize, we used a multiplexed CRISPR-Cas9 approach to generate mutations in the highly conserved catalytic PIWI domain for two different AGO genes per transgene construct. Using this approach for several AGOs, we have obtained new alleles for *Ago1a*, *Ago1c*, *Ago2a*, *Ago2b*, *Ago4a*, *Ago4b*|*Ago104*, *Ago5a*, and *Ago6*. Using this mutant collection, we will investigate the role of maize AGOs in paramutation, reproduction, and stress response.

## Biological Design of Lemnaceae Aquatic Plants for Biodiesel Production

E. Ernst, C. Mateo-Elizalde, J. Simorowski, J. Lynn, U. Ramu, R. Martienssen [in collaboration with E. Lam, Rutgers University; T. Michael, Salk Institute; J. Schwender and J. Shanklin, Brookhaven National Laboratory]

Freshwater aquatic Lemnaceae (duckweeds, watermeal) present an attractive alternative to algae as biofuel feedstocks because of their robust growth in nonsterile conditions and the relative ease of harvesting dry material. Lemnaceae species are the smallest angiosperms, producing up to 64 grams of biomass per starting gram in a week, which is far beyond that of terrestrial crops such as maize (2.3 g/g/wk). Convenient metabolic labeling makes Lemnaceae a good system for pathway modeling and engineering, and their unresponsive stomata admit high levels of atmospheric $CO_2$. Our goal is to divert substantial carbon accumulation from starch to oil metabolism in Lemnaceae. To establish the foundations of biosynthetic pathway engineering in the Lemnaceae, we produced eight new reference-quality genome assemblies of *Wolffia australiana*, *Lemna gibba*, *Lemna minor*, *Lemna turionifera*, and allotriploid hybrid *L. japonica* (*L. minor* × *L. turionifera*) clones using Oxford Nanopore long reads and Hi-C contact maps. The sequenced species retain just 18,000–24,000 protein-coding gene loci, significantly fewer than terrestrial monocots such as rice and *Brachypodium*, and comparable to *Chlamydomonas reinhardtii*. Methylome and small RNA sequencing revealed dramatic differences between the three genera consistent with known pathways of RdDM. Phylogenomic analysis of missing orthogroups and unique paralogs across the Lemnaceae, nine other monocots. and seven non-monocots uncovered variations that likely account for these and other traits, including reduced morphology, aquatic habitat, clonal reproduction, dormancy, high photosynthetic rate, and lipid production. Building on our robust genetic transformation protocol demonstrated in *L. japonica*, we completed two design–build–test–learn cycles resulting in stable clonal lines that produce high levels of triacylglycerol (TAG). Simultaneous expression of *Arabidopsis WRINKLED1*, mouse *DGAT*, and a modified sesame *OLEOSIN* in *L. japonica* accumulate TAG at up to 8.7% of dry weight (DW), an increase of 108-fold over wild type (wt), and total fatty acids reach 15% of DW. With this output, duckweeds could produce sevenfold more oil than soybean on the same growth area per year. Work is underway to model biomass synthesis, to reduce lipid futile cycling, and to further increase TAG accumulation by incorporating physiological, biochemical, genomic, and transcriptomic data.

## Deciphering the Mechanisms of Duckweed Sexual Reproduction

C. Mateo-Elizalde, E. Ernst, R. Martienssen

Duckweeds are the smallest flowering plants, although many clones do not produce flowers in nature nor in laboratory conditions. This was one of the reasons that *Lemna*, an early model for plant physiology, was left behind with the advent of plant molecular genetics. Nonetheless, duckweeds have still attracted attention for applied purposes—for example, as an alternative source of protein. Recent advances in genomic, gene silencing, and gene editing technologies have enabled duckweed molecular biology, and the number of duckweed publications has grown exponentially over the last 20 years (Mateo-Elizalde et al. 2023). Because of their predominant clonal growth habit, sexual reproduction is one of the least studied aspects of this plant family. The plant hormone salicylic acid, a floral inducer in duckweed, has been used for transcriptomic and physiological studies, along with the identification of key flowering regulatory genes such as *LgFT1* and *LaFTL1*. We have performed transcriptomic analysis of duckweed pistils, anthers, seeds, and cotyledons of

the short-day duckweed *Lemna aequinoctialis* with the aim of providing the first specific and global characterization of the duckweed transcriptome related to flowering. Our study reveals evidence of alternative strategies in gene expression control between sexual and clonal propagation, shedding light on this unexplored area of plant physiology and development in *Lemna*.

## Maizecode: DNA Regulatory Elements in Maize and Teosinte Inbreds Provide Insight into Maize Domestication

J. Cahn, M. Regulski, J. Lynn, E. Ernst, R. Martienssen [in collaboration with K. Birnbaum, New York University; M. Schatz, Johns Hopkins University; D. Micklos, W.R. McCombie, D. Ware, D. Jackson, and T. Gingeras, CSHL]

Early maize lines were domesticated from *Teosinte parviglumis (Zea mays parviglumis)*, with subsequent introgressions from neighboring *Teosinte mexicana*. Domestication traits in modern maize include increased kernel row number, loss of the hard fruit case, and dissociation from the cob upon maturity, as well as fewer tillers. Molecular approaches have identified several transcription factors involved in the development of these traits. However, these studies have also shown that a more complex regulatory network is responsible for these strong morphological differences than originally hypothesized, and our understanding of the tissue-specific regulation as well as of its variability across inbred lines is still lacking.

In this study, we investigate the transcriptional regulation that resulted from the domestication process, focusing on conservation and variability across multiple tissues and inbred lines. We generated histone modification and transcription factor ChIP-seq in parallel with transcriptomics data sets in up to five different tissues of three inbred lines that span the phenotypic diversity of maize inbreds, as well as the teosinte inbred TIL11. We developed an automatized computational pipeline to integrate these data sets as well as publicly available data. This pipeline generates metrics and outputs for both quality control and functional analyses, and it can also be applied to other species. We identified regulatory regions that emerged during the domestication process and are responsible for the tissue-specific expression of developmental genes. We show that, even though pollen grains are the most differentiated tissue on a transcriptomic level, and especially with respect to the regulation of transposable elements, ears show the least conservation, corroborating the very distinct morphological and physiological differences between maize and teosinte. This integrative analysis highlighted the role of epigenetic data in helping refine the functional annotation of the genome, notably by identifying distal "super enhancers" similar to those found in animal genomes. With this study, we hope to provide the maize community with a framework for a collaborative effort that follows the footsteps of the ENCODE project in order to better understand and potentially improve the regulatory landscape of the maize genome.

## Nuclear RNAi Mediates Transcription-Induced Replication Stress

K-H. Cheng, R. Martienssen [in collaboration with B. Roche, University of North Dakota]

Gene regulation by RNA interference is usually attributed to miRNA, but RNAi has a more ancient and fundamental role in heterochromatic silencing and genome stability. Using *Schizosaccharomyces pombe* as a model organism, which lacks miRNA, we discovered a new miRNA-independent function of Dicer on chromatin—resolving transcription-replication (T-R) conflicts. We found that *dcr1*Δ mutants become hypersensitive to genotoxic drugs in contexts in which R-loops accumulate. Using forward and reverse genetic methods, we identified multiple alleles linked to core transcriptional processes that suppress the phenotype, including the two main subunits of RNA polymerase (Rpb1, Rpb2), several general transcription factors, and the Mediator complex—particularly the subunit Med20, which was the strongest suppressor. Intriguingly, we recovered multiple independent gain-of-function alleles of Med20, whereas *med20*Δ mutant had a mild RNAi-like phenotype, indicating an unexpected tight link between Dicer and Mediator. Furthermore, we performed Precision Run-On sequencing (PRO-seq) and discovered that Dcr1 modulates transcription genome-wide, an effect that was particularly pronounced for regulating promoter proximal pausing. These results suggest that Dicer limits T-R conflicts by dynamically regulating RNA Pol II during early stages of transcription, thereby leaving way for DNA polymerase during replication.

Our ongoing work focuses on the characterization of genome instability resulting from the absence of this process in the *dcr1*Δ mutant and, in particular, the types of DNA damage seen at sites prone to R-loop formation. Given the evolutionarily conserved role of Dicer for centromeric silencing, we expect this new key function on chromatin to also be conserved in higher eukaryotes possessing miRNAs.

## PUBLICATIONS


Girish V, Lakhani AA, Scaduto CM, Thompson SL, Brown LM, Hagenson RA, Sausville EL, Mendelson BE, Lukow DA, Yuan ML, et al. 2023. Oncogene-like addiction to aneuploidy in human cancers. *Science* **381**. doi:10.1126/science.adg4521

Lee SC, Adams DW, Ipsaro JJ, Cahn J, Lynn J, Kim HS, Berube B, Major V, Calarco JP, LeBlanc C, et al. 2023. Chromatin remodeling of histone H3 variants underlies epigenetic inheritance of DNA methylation. *Cell* **186:** 4100–4116.e15. doi:10.1016/j.cell.2023.08.001

Liang Y, Yu XH, Anaokar S, Shi H, Dahl WB, Cai Y, Luo G, Chai J, Cai Y, Mollá-Morales A, et al. 2023. Engineering triacylglycerol accumulation in duckweed (*Lemna japonica*). *Plant Biotechnol* **21:** 317–330. doi:10.1111/pbi.13943

Mateo-Elizalde C, Lynn J, Ernst E, Martienssen RA. 2023. Duckweeds. *Curr Biol* **33:** R89–R91. doi:10.1016/j.cub.2022.12.036

Shi H, Ernst E, Heinzel N, McCorkle S, Rolletschek H, Borisjuk L, Ortleb S, Martienssen RA, Shanklin J, Schwender J. 2023. Mechanisms of metabolic adaptation in the duckweed *Lemna gibba*: an integrated metabolic, transcriptomic and flux analysis. *BMC Plant Biol* **23:** 458. doi:10.1186/s12870-023-04480-9

Stevenson DW, Ramakrishnan S, de Santis Alves C, Coelho LA, Kramer M, Goodwin S, Ramos OM, Eshel G, Sondervan VM, Frangos S, et al. 2023. The genome of the Wollemi pine, a critically endangered "living fossil" unchanged since the Cretaceous, reveals extensive ancient transposon activity. bioRxiv doi:10.1101/2023.08.24.554647


*In Press*


Berube B, Ernst E, Cahn J, Roche B, de Santis Alves C, Lynn J, Scheben A, Siepel A, Ross-Ibarra J, Kermicle J, Martienssen RA. 2024. *Teosinte Pollen Drive* guides maize diversification and domestication by RNAi. *Nature* **633:** 380–388. doi:10.1038/s41586-024-07788-0

Herridge RP, Dolata J, Migliori V, de Santis Alves C, Borges F, Van Ex F, Lin A, Bajczyk M, Leonardi T, Hendrick A, et al. 2024. Pseudouridine guides germline small RNA transport and epigenetic inheritance. *Nat Struct Mol Biol* doi:10.1038/s41594-024-01392-6

Shimada A, Cahn J, Ernst E, Lynn J, Grimanelli D, Henderson I, Kakutani T, Martienssen RA. 2024. Retrotransposon addiction promotes centromere function via epigenetically activated small RNAs. *Nat Plants* **10:** 1304–1316. doi:10.1038/s41477-024-01773-1

# MOLECULAR SIGNALING EVENTS UNDERLYING ENVIRONMENTAL CONTROL OF PLANT GROWTH

**U. Pedmale** Y. Hu, J. Micko, S. Palit, F. Pierdona, D. Rosado, V. Schoen, L. Taylor

All organisms experience growth during their lifetime—either by increasing in size or through cell division. This process relies on inputs from a variety of external and internal signals perceived by the organism. However, unchecked growth can lead to conditions such as cancer and other developmental defects, significantly impacting an organism's overall fitness. A central, yet unresolved, question in biology is how environmental and external signals regulate an organism's growth and development.

Both plants and animals interact with their environment, but plants primarily exhibit postembryonic growth because of their inability to move. Unlike animals, plants lack specialized organs for sensing or hearing stimuli. Nevertheless, plants are highly sensitive to their surroundings and adjust their growth in response to diverse external and internal signals. They frequently encounter variable growth conditions, including fluctuations in temperature, light quality and intensity, herbivory, pathogen attacks, water availability, and more. In response to these biotic and abiotic factors, plants adapt and thrive amid substantial environmental fluctuations.

Moreover, plants must strike a balance between mitigating potential threats and maximizing benefits, carefully allocating resources to make appropriate decisions. Despite the absence of a central nervous system, plants effectively integrate a wide range of cues and make sophisticated decisions regarding their growth. This adaptability is essential given their stationary nature. For example, when plants face challenges such as climate change or increased competition for light, they may reallocate resources to adapt, often at the cost of reduced productivity (e.g., lower yield and biomass). Such trade-offs are crucial for their survival in changing environments.

The goal of our laboratory is to uncover the mechanisms by which plants perceive and adapt to their environment. We also aim to understand how plants integrate intrinsic and extrinsic cues to determine the most appropriate responses to environmental stimuli. Gaining insight into these processes could enable the development of crops that thrive under adverse conditions without substantial reductions in yield.

Our research focuses primarily on how light environments influence plant growth and development. Light is a key environmental signal as it not only drives photosynthesis but also provides essential information about the surrounding growth conditions and seasonal changes. Plants perceive light through a diverse array of photoreceptors, including phytochromes (PHYA-E), cryptochromes (CRY1-2), phototropins (PHOT1-2), the Zeitlupe family (FKF1, LKP2, and ZTL), and UVR8.

Plants have evolved sophisticated adaptive responses to interpret and utilize light's directionality, intensity, and spectral quality. For example, under vegetational shading (Fig. 1), plants detect a reduced red-to-far-red light ratio (R:FR) because of the absorption of red light by chlorophyll and the reflection of far-red light from nearby foliage. Concurrently, there is a decrease in blue light and photosynthetically active radiation (PAR) availability.

We primarily focus on blue-light-absorbing CRYs. These molecules serve as excellent genetic and molecular tools for unraveling the complexities of growth and adaptation. However, numerous questions about the molecular function of CRYs in plants remain unanswered. Understanding the role of CRYs holds appeal not only for agriculture but also for its potential impact on human health. This aspect makes the field attractive to a wide range of funding agencies. CRYs play a pivotal role in regulating growth and development while

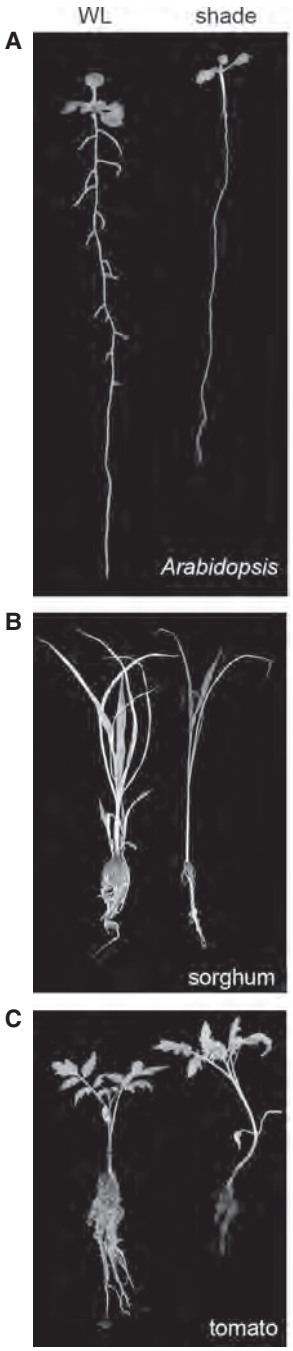



**Figure 1.** Vegetative shade perceived by the above-ground shoots restricts below-ground root growth in *Arabidopsis* (*A*), sorghum (*B*), and tomato (*C*) and decreases the overall fitness of the plants.

providing circadian entrainment to both plants and animals. In metazoans, disruptions in CRY activity have been linked to various phenomena such as cancer, altered behavior, magnetoreception, and metabolism. Therefore, comprehending the function of CRYs in plants is crucial. They are integral to plant growth and can also have a significant impact on human health.

## Cryptochromes Play a Crucial Role in Mediating the Repair of Damaged DNA

Y. Hu [in collaboration with L.N. Lindbäck, Swedish University of Agricultural Sciences]

CRYs evolved from DNA photolyases, enzymes that use blue/UV-A light energy to repair DNA damage. However, modern cryptochromes have lost the ability to directly repair DNA. In *Arabidopsis*, we observed that the *cry1cry2* mutant was more sensitive to genotoxic stress, such as UV-C radiation, compared with wild-type (WT) plants. This suggests that CRYs may positively regulate the DNA damage response, either directly or indirectly (Fig. 2). Notably, CRY2 formed punctate nuclear speckles in response to UV-C, similar to its behavior under blue light.

To further investigate how CRYs modulate the plant's response to UV radiation, we performed a time-course whole-genome transcriptome analysis in WT and *cry1cry2* double mutant plants, with and without UV-C treatment. Transcriptome analysis revealed a consistent down-regulation of the flavonoid biosynthesis pathway in the *cry1cry2* mutant, both prior to and following UV-C exposure. Flavonoids serve as natural sunscreens by absorbing harmful UV radiation and also function as scavengers of reactive oxygen species (ROS), which are typically generated during genotoxic stress. The *cry1cry2* double mutant exhibited reduced levels of flavonoids, leading to increased accumulation of cyclobutane pyrimidine dimers (CPDs), a form of DNA damage. This reduction in flavonoid biosynthesis likely underpins the heightened sensitivity of the *cry1cry2* mutant to UV-C radiation.

Additionally, our analysis revealed a delay in the activation of DNA damage–responsive genes in the *cry1cry2* mutant following UV-C exposure compared with WT plants. This finding indicates that the *cry1cry2* mutant is impaired in initiating the transcriptional response to DNA damage. Through temporal co-expression analysis, we identified the CAMTA family of transcription factors as potential regulators. *Arabidopsis* plants carrying loss-of-function CAMTA mutations exhibited UV-C sensitivity similar to that of *cry1cry2* mutants, suggesting that CAMTAs play a key role in mediating resilience to DNA damage.

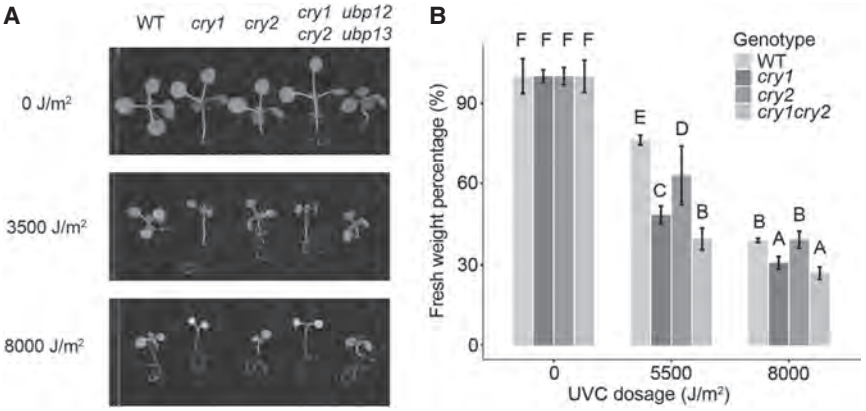



**Figure 2.** *Arabidopsis* plants lacking CRY1 and CRY2 are more sensitive to UV-C-induced DNA damage as measured by fresh weight postexposure (*A*,*B*).

## Flowering Locus VE (FVE) Controls Hypocotyl Elongation in *Arabidopsis*

V. Schoen [in collaboration with L.N. Lindbäck, Swedish University of Agricultural Sciences]

Flowering locus VE (FVE), a plant homolog of retinoblastoma-associated protein 48 (RbAp48), plays important roles in both developmental and stress-response pathways in the model plant *Arabidopsis*. FVE acts as a scaffold protein within large complexes involved in various epigenetic processes, including histone deacetylation, histone methylation, and RNA-dependent DNA methylation (RDDM). In the flowering pathway, FVE facilitates the deposition of the negative regulatory mark H3K27me3 on the *flowering locus C* (*FLC*) gene. This mark represses FLC expression, promoting the transition to reproductive growth. Conversely, mutations in FVE lead to delayed flowering as the repression of FLC is not alleviated, thereby postponing the onset of flowering.

A tandem affinity purification and mass spectrometry screen identified FVE as an interactor of CRYPTOCHROME2 (CRY2). We confirmed this interaction in *Nicotiana benthamiana* using biomolecular fluorescence complementation (BiFC) and co-immunoprecipitation assays. Imaging from the BiFC assay revealed that FVE and CRY2 interact within the nucleus. To explore the biological significance of this interaction, we analyzed flowering time and other phenotypes. Similar to *fve* mutant plants, *cry2* mutants exhibit late flowering. Intriguingly, the *fve cry2* double mutant displayed an even more delayed flowering phenotype, indicating an additive effect on flowering time.

We next examined the role of FVE and CRY2 in hypocotyl elongation, the process by which the stem grows upward during seedling germination in response to light. CRY2, a blue-light photoreceptor, interacts with PHYTOCHROME INTERACTING FACTORS (PIFs) to inhibit hypocotyl elongation. Under light conditions, CRY2 is degraded, linking light intensity to growth responses. Mutants in *cry2* exhibit elongated hypocotyls under blue light, whereas *fve* mutants show shorter hypocotyls. Interestingly, the *cry2 fve* double mutant displayed an intermediate hypocotyl phenotype, suggesting an epistatic relationship between the two genes.

This finding provides the first evidence of an epigenetic factor (FVE) influencing growth control in a context-dependent manner. Although both *fve* and *cry2* mutants exhibit late flowering, their hypocotyl elongation phenotypes are opposite: *cry2* mutants show elongated hypocotyls, whereas *fve* mutants have shortened hypocotyls. These contrasting roles highlight the nuanced and context-specific functions of the CRY2–FVE interaction in different biological pathways.

## How Does the Aboveground Shoot Control Underground Roots?

D. Rosado [in collaboration with J. Noel, Salk Institute; J. Chory, Salk Institute and HHMI; J. Dinneny, Stanford University; M. Rossi, University de São Paulo; L.N. Lindbäck, Swedish University of Agricultural Sciences; K. Schwartz, CSHL]

Roots and shoots live in different environments, yet the root knows when the shoot faces adverse

environmental conditions. During vegetational shading, many aerial organs elongate rapidly, whereas the root growth is reduced with the delay in the emergence of the lateral roots. Roots not only serve as a mechanical anchor but play a vital role in the well-being of the entire plant. Therefore, a robust and well-developed root system is required for healthy plant growth. As one can imagine, there is a negative cycle occurring during shading; shoot-perceived shade leads to reduced root growth, which in turn is unable to support the shoot—leading to unproductive plants. However, this phenomenon is an excellent model to understand growth at a systems level because of the different growth phenotypes observed in the various organs of the same plant. Additionally, the model enables us to explore the nature of the interorgan and long-distance communication, which is used to signal when a distant organ is exposed to an adverse environment. Unfortunately, and surprisingly, not much is known about the mechanisms that underlie reduced root growth seen during shading.

In this context, we performed a transcriptomic analysis of *Arabidopsis* and tomato seedlings exposed to 30 min to five days of shade (Fig. 3). In addition, we generated a time course single-cell RNA sequencing (scRNA-seq) on roots exposed to the shade and mock-shaded plants. We found that stress-induced (specifically) genes induced by defense response against pathogens were up-regulated in shaded roots, compared with those grown under nonshading conditions. To validate this observation, we monitored genes induced by defense responses against pathogens, which did confirm induction to shade. Furthermore, we discovered that MAPK signaling was activated in shade, which usually responds to abiotic and biotic stress. With this approach, we intend to generate gene-regulatory networks that will help us understand how root development is regulated under optimal and suboptimal growth conditions imposed by light. Another ongoing approach to identifying shade response regulators in roots is a forward genetic screen of mutants. The *CMT6* gene was identified in our transcriptome as a suitable marker for shade responses in roots and not in the shoot. A transgenic line harboring a reporter called GUS (β-glucuronidase) under the control of *CMT6* promoter was generated. These plants, under shade, express GUS in their roots. Ethyl methanesulfonate (EMS)-mutagenized *CMT6::GUS* lines

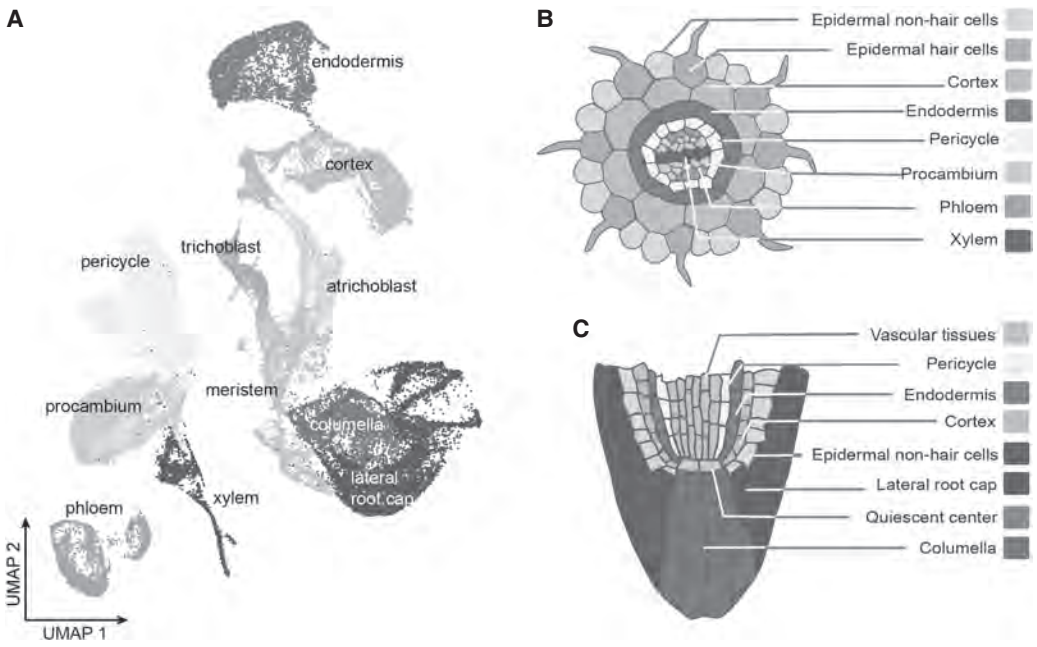



**Figure 3.** Identification of major cell types in *Arabidopsis* roots under shade and control light with single-cell RNA-seq. (*A*) UMAP dimensionality reduction plot of 57,649 root cells using scRNA-seq expression data. Each dot represents a unique cell and is shaded according to its cell type, which in turn corresponds to one or more Louvain clusters. (*B*,*C*) Cartoon representing major cell types of the *Arabidopsis* root shaded according to *A*, as seen in longitudinal (*B*) and transversal (*C*) views. Illustrations were adapted from the Plant Illustrations repository (2017).

were then generated to identify other factors using a forward genetic screen. We are currently screening the third generation of mutants that does not display GUS expression, indicating mutants. These mutants with confirmed loss of GUS expression and increased lateral roots—similar to unshaded plants—will be sequenced to find the causative loci. Next, complementation assays and functional characterization of these genes will be performed. We hope that these combined approaches will help reveal the molecular mechanisms regulating shade responses in roots and provide molecular tools to bypass the negative effect of shade avoidance in root systems.

## Decoding Light-Induced Chromatin Remodeling: The Role of Cryptochrome–ISWI Complexes in Growth and Development

S. Palit, Y. Hu

Chromatin acts as a dynamic regulator, integrating environmental signals into stable gene expression patterns that drive long-term phenotypic changes. Light, a crucial environmental cue, induces significant chromatin remodeling across species, such as histone phosphorylation in mammals during light–dark transitions and large-scale heterochromatin condensation in plants. CRYs, conserved blue-light photoreceptors, play a central role in these processes. In plants, CRYs regulate key physiological responses, including circadian rhythms, photomorphogenesis, flowering, and stress responses. In animals, they are implicated in inflammation, tumorigenesis, and other processes. Specifically, CRY2, a nuclear protein, forms liquid–liquid phase-separated (LLPS) condensates under prolonged blue-light exposure. These condensates serve as biochemical hotspots, facilitating rapid adaptation to environmental stimuli, including stress, hormonal signals, and temperature fluctuations.

In plants, CRY2 influences chromatin architecture despite lacking direct DNA-binding ability, interacting with transcriptional regulators and modulating light-induced chromatin compaction and decompaction. To explore this interaction, *Arabidopsis* expressing epitope-tagged CRY2 was used to identify CRY2-associated proteins through affinity purification and mass spectrometry. This revealed an association with imitation switch (ISWI) chromatin remodelers, including CHR11/17, RLT1/2, and ARID5, collectively termed the CRA complex. These remodelers regulate nucleosome positioning and spacing, mediating chromatin organization and gene expression. Notably, CHR11/17 share homology with mammalian SMARCA1/5, remodelers implicated in growth regulation and cancers.

The CRA complex represents a crucial link between external signals and chromatin remodeling. Although CRY2-mediated chromatin changes in plants are associated with developmental plasticity, similar mechanisms are linked to growth disorders and metastasis in humans. Understanding how CRY2 interacts with the CRA complex to mediate light-induced chromatin dynamics will provide valuable insights into the regulation of growth, development, and stress responses across organisms. *Arabidopsis* serves as an ideal model for this research because of its exceptional sensitivity to environmental cues and well-characterized genetic tools.

### PUBLICATIONS

Artz O, Ackermann A, Taylor L, Koo P, Pedmale UV. 2023. Light and temperature regulate m$^6$A-RNA modification to regulate growth in plants. bioRxiv doi:10.1101/2023.01.17.524395

Hu Y, Rosado D, Louise LN, Micko J, Pedmale UV. 2023. Cryptochromes and UBP12/13 deubiquitinases antagonistically regulate DNA damage response in *Arabidopsis*. bioRxiv doi:10.1101/2023.01.15.524001

Lindbäck LN, Ji Y, Cervela-Cardona L, Jin X, Pedmale UV, Strand A. 2023. An interplay between bZIP16, bZIP68 and GBF1 regulates nuclear photosynthetic genes during photomorphogenesis in *Arabidopsis*. *New Phytologist* **240:** 1082–1096.

# GENOMICS

Advances in high-throughput sequencing technologies have resulted in an explosive growth of multi-omics data. While presenting a tremendous opportunity for quantitative studies of numerous biological processes, crucial for both fundamental research and clinical applications, these advances also created a set of unique bioinformatics challenges for processing, integrating, and interpreting the vast amounts of data.

**Alexander Dobin** and colleagues are biological data scientists working to resolve these challenges by developing highly efficient and accurate algorithms such as STAR, the popular RNA-seq analysis software used by thousands of researchers worldwide. They are conceiving novel computational approaches to process data from emerging sequencing technologies, such as single-cell RNA-seq and long-read nanopore sequencing, with a special emphasis on detecting RNA and DNA aberrations in tumors.

Another exciting research area in the group is functional annotation of the noncoding genome via integration of multi-omics data generated by the ENCODE (Encyclopedia of DNA Elements), Roadmap Epigenomics, and GTeX consortia, essential for deciphering gene regulation mechanisms, interpreting disease-associated variants in genome-wide association studies (GWAS), and understanding epigenetic effects in cancer biology.

**Thomas Gingeras** and colleagues study where and how functional information is stored in genomes. These efforts help explain the biological and clinical effects of disease-causing gene mutations in humans and other organisms. Gingeras is a leader of the ENCODE and the mouseENCODE and modENCODE (model genome ENCODE) projects of the National Institutes of Health. His research has altered our understanding of the traditional boundaries of genes, revealing that almost the entire lengths of genomes in organisms ranging from bacteria to humans can be transcribed into RNA (pervasive transcription) and that most RNA products made by a cell are not destined to be translated into proteins (noncoding, or ncRNAs). In fact, ncRNAs are proving to be involved in a variety of other important biological functions. Some have been shown to be critical components in the pre- and posttranscriptional and translational processes as scaffolds upon which large protein complexes are assembled and as extracellular signals. The initial studies that led to these observations have been extended to cover the entire human genome. The maps of where and what types of RNA are transcribed serve as a foundation for many areas of disease diagnosis and treatment. In particular, ncRNAs are responsible for the expression of the regulatory gene *PTEN*. In endometrial cancer, dysregulation of PTEN via overexpression of a long ncRNA (HOTAIR) that binds to PTEN occurs in obese individuals. The dysregulation results in adipogenesis and inappropriate lipid metabolism, leading to endometrial cancer.

The insights of **W. Richard McCombie** and colleagues have led to the introduction and optimization of novel methods of high-throughput genome sequencing. His team has made it possible to catalog variation among individual organisms in a way that would have been unthinkable 10 years ago. They have brought online a new generation of Illumina sequencers and optimized their function to a level at which eight to 10 trillion DNA bases can be sequenced in a month. McCombie's team has been involved in international efforts culminating in genome sequences for maize, rice, and bread wheat—three of the world's most important food crops. They have also had an important role in projects to sequence the flowering plant *Arabidopsis thaliana* (the first plant genome sequenced), the fission yeast *Schizosaccharomyces pombe*, and the human genome, as well as other important genomes. McCombie's group is currently involved in several important projects to resequence

genes in patient samples that are of special interest to human health, including *DISC1* (a strong candidate gene for schizophrenia), looking for genetic variants implicated in bipolar illness and major recurrent depression. They are also looking for genes that contribute to cancer progression using whole-genome sequencing or a method called exome sequencing, which they developed with Greg Hannon to look at mutations in the regions of the genome that code for proteins.

Using multidisciplinary approaches that combine computational analysis, modeling, and prediction with experimental verification, **Doreen Ware**'s laboratory seeks a deeper understanding of the evolution of genome sequences in plants and their implications for agricultural improvement. By looking comparatively across the genomes of plants in the same lineage, they seek answers to the following questions: How are genes conserved and lost over time? What are the fates of duplicated genes? What is the impact of structural variation on phenotypic variation? Ware's team also studies gene regulation in plants, focusing on gene regulatory networks, targeting transcription factors and microRNA genes with the objective of understanding how these parts of the plant genome work together in determining spatial and temporal expression of genes. The laboratory had an important role in the project to produce a haplotype map reference genome of maize, spearheading the most comprehensive analysis of the crop yet. This has provided important information on the variation of the reference genome, as well as comparative data showing changes in the genome acquired through domestication and breeding. They have devoted special attention to examining diversity within maize, grape, and tomato, aiming to accelerate the development of strategies to introduce new germplasm that is needed to meet the demands of an increasing population and a changing environment. The laboratory also has brought fully sequenced genomes into an integrated data framework to enhance the power of their comparative studies. Ware was named principal investigator for the National Science Foundation–funded Gramene project, a comparative genomics resource for agriculturally important crops and models to support sustainable food and fuel production. Ware, as principal investigator for plants, has also helped lead an effort funded by the Department of Energy to create—out of many separate streams of biological information—a single, integrated cyber-"knowledgebase" for plants and microbial life.

# NOVEL COMPUTATIONAL ALGORITHMS AND ANALYSES OF SINGLE-CELL DATA

**A. Dobin** J. Hinds R. Utama (Bioinformatics Core) D. Yunusov

## Analyzing Transcriptomes of Cancer and Fibroblast Cells in Pancreatic Tumors

J. Hinds [in collaboration with Y. Park, G. Caligiuri, M. Shakiba, S. Nadella, C. Tonelli, S. Ting, K. Sun, and D. Tuveson, CSHL]

The complex makeup of the tumor microenvironment (TME) stands out as a defining aspect of pancreatic ductal adenocarcinoma (PDAC), leading to patients' poor treatment outcomes and establishing PDAC as the third leading cause of cancer-related mortality in the United States. A key characteristic of the PDAC stroma is the variety found within cancer-associated fibroblasts (CAFs). Although the introduction of U.S. Food and Drug Administration (FDA)-approved drugs targeting oncogenic KRAS has opened new treatment possibilities for KRAS-mutant cancers, the precise roles of KRAS in shaping the TME and its influence on tumor development are not fully understood. This uncertainty complicates the ability to foresee how inhibiting KRAS might affect mature pancreatic ductal adenocarcinoma tumors. Our research involved examining multiple single-cell and spatial transcriptomes of different mouse models, human patient samples, and organoids generated by the Tuveson laboratory.

To explore how disrupting KRAS signaling affects cancer cells, we examined data from several models: the genetically engineered mouse model designed for the permanent removal of KrasG12V (the FPC model), the drug-induced suppression of the mutant KRAS, and biochemical KRAS excision. Through the use of single-cell RNA sequencing and spatial transcriptomics applied to the FPC model before and after the excision of KrasG12V, we detected notable shifts in the transcriptomes of epithelial cells. We also noted significant shifts in the activation levels of pancreatic fibroblasts throughout the initial and advanced phases of tumor development. These activated fibroblasts are divided into three groups: inflammatory CAFs (iCAFs), myofibroblastic CAFs (myCAFs), and a recently discovered group, xCAFs, distinguished by heightened expression of Wnt pathway genes. Each group is marked by distinct indicators stemming from either inflammation or the activation of oncogenic KRAS.

The other projects we were involved in were (i) comparisons between different patient, organoid, and mouse xenograft models using single-cell transcriptomic data and (ii) investigations of how mucus production promotes classical pancreatic ductal adenocarcinoma.

## Novel Machine Learning Algorithm for Gene Quantification in Single-Cell Data

D. Yunusov

We enhanced the STARsolo framework with a novel expectation-maximization and machine learning (EM-ML) algorithm that represents a significant leap in the analysis of single-cell RNA sequencing (scRNA-seq) data. This innovation focuses on transcript-level quantification and is specifically designed to mitigate the biases associated with the distribution of scRNA-seq reads. By addressing these challenges, the enhanced algorithm not only improves the accuracy of quantification but also has profound implications for subsequent analyses and biological insights that can be derived from scRNA-seq data.

The move towards transcript-level quantification marks a departure from the more common gene-level analyses. This shift allows for a more refined view of gene expression, capturing the complexity and nuances of alternative splicing and transcript isoforms. By focusing on transcripts, researchers can gain insights into the functional implications of gene expression differences, which may be obscured when only analyzing at the gene level.

One critical advancement of the new EM-ML algorithm is its ability to address the inherent biases in

scRNA-seq read distribution. These biases can arise from various sources, such as differences in gene length, GC content, and sequencing depth. By mitigating these biases, the algorithm ensures that transcript quantification is more accurate and reflective of the true biological expression levels.

The algorithm's impact on cell clustering and differential expression analysis is particularly noteworthy. The refined quantification and bias correction allow for more precise identification of cell subtypes and the discovery of marker genes. This precision is crucial for understanding the heterogeneity within cell populations and identifying distinct cellular states and transitions. The improved resolution in detecting differential expression further enables researchers to pinpoint genes that are critical drivers of biological processes and disease states.

Another significant advantage of the new algorithm is its enhanced ability to resolve overlapping genes. In scRNA-seq data, reads mapping to regions where genes overlap can be challenging to assign correctly. The EM-ML algorithm provides a solution to this problem, allowing for better discrimination between overlapping transcripts and ensuring that reads are assigned more accurately. This improvement is particularly important for understanding the expression patterns of genes located in densely packed genomic regions.

Compared with standard EM-ML approaches, the new algorithm prefers assigning reads to known and well-annotated genes over provisionally annotated genes. This feature is particularly beneficial for ensuring that the quantification is grounded in genes with established biological functions, thereby enhancing the analysis' reliability and interpretability.

The extension of STARsolo's functionality through the introduction of a novel EM-ML algorithm represents a significant advancement in scRNA-seq analysis. By providing more accurate transcript-level quantification, addressing biases, and improving the resolution of overlapping genes, this algorithm significantly enhances our ability to dissect the complexity of gene expression in single cells. The implications for cell clustering, differential expression analysis, and the discovery of marker genes are profound, offering researchers a powerful tool for unraveling the intricate molecular landscapes of cellular heterogeneity and function.

## PUBLICATIONS

Haas BJ, Dobin A, Ghandi M, Van Arsdale A, Tickle T, Robinson JT, Gillani R, Kasif S, Regev A. 2023. Targeted *in silico* characterization of fusion transcripts in tumor and normal tissues via FusionInspector. *Cell Rep Methods* **3:** 100467. doi:10.1016/j.crmeth.2023.100467

Hazra R, Utama R, Naik P, Dobin A, Spector DL. 2023. Identification of glioblastoma stem cell-associated lncRNAs using single-cell RNA sequencing datasets. *Stem Cell Rep* **18:** 2056–2070. doi:10.1016/j.stemcr.2023.10.004

Hitz BC, Lee JW, Jolanki O, Kagda MS, Graham K, Sud P, Gabdank I, Strattan JS, Sloan CA, Dreszer T, et al. 2023. The ENCODE Uniform Analysis Pipelines. bioRxiv doi:10.1101/2023.04.04.535623

Jorstad NL, Song JHT, Exposito-Alonso D, Suresh H, Castro-Pacheco N, Krienen FM, Yanny AM, Close J, Gelfand E, Long B, et al. 2023. Comparative transcriptomics reveals human-specific cortical features. *Science* **382:** eade9516. doi:10.1126/science.ade9516

Rozowsky J, Gao J, Borsari B, Yang YT, Galeev T, Gürsoy G, Epstein CB, Xiong K, Xu J, Li T, et al. 2023. The EN-TEx resource of multi-tissue personal epigenomes & variant-impact models. *Cell* **186:** 1493–1511.e40. doi:10.1016/j.cell.2023.02.018

Sroka MW, Skopelitis D, Vermunt MW, Preall JB, El Demerdash O, de Almeida LMN, Chang K, Utama R, Gryder B, Caligiuri G, et al. 2023. Myo-differentiation reporter screen reveals NF-Y as an activator of PAX3-FOXO1 in rhabdomyosarcoma. *Proc Natl Acad Sci* **120:** e2303859120. doi:10.1073/pnas.2303859120

Suresh H, Crow M, Jorstad N, Hodge R, Lein E, Dobin A, Bakken T, Gillis J. 2023. Comparative single-cell transcriptomic analysis of primate brains highlights human-specific regulatory evolution. *Nat Ecol Evol* **7:** 1930–1943. doi:10.1038/s41559-023-02186-7

### *In Press*

Tonelli C, Yordanov GN, Hao Y, Deschênes A, Hinds J, Belleau P, Klingbeil O, Brosnan E, Doshi A, Park Y, et al. 2024. A mucus production programme promotes classical pancreatic ductal adenocarcinoma. *Gut.* doi:10.1136/gutjnl-2023-329839

# GENOME ORGANIZATION, REGULATION, AND FUNCTIONAL ROLES OF NONCODING RNAs

**T.R. Gingeras** J. Drenkow

## The ENCODE Uniform Analysis Pipelines

This work was performed in collaboration with M. Cherry (Stanford U), A. Dobin, T. Gingeras (CSHL), and the ENCODE Data Coordination Center (DCC).

Phase III of the Encyclopedia of DNA Elements (ENCODE) project continues (for more than 20 years) to focus on developing a catalog of functional elements in the human genome. Currently the database encompasses 19,000 functional genomics experiments across more than 1,000 cell lines and tissues using a wide range of experimental techniques to study the chromatin structure and the regulatory and transcriptional landscape of the *Homo sapiens* and *Mus musculus* genomes. All experimental data, metadata, and associated computational analyses created in the ENCODE project have been submitted to the Data Coordination Center (DCC) for validation, tracking, storage, and distribution to community resources and the scientific community. Over the course of the ENCODE project the DCC has engineered and distributed uniform processing pipelines in order to promote data provenance and reproducibility as well as allow interoperability between genomic resources and other consortia. All data files, reference genome versions, software versions, and parameters used by the pipelines are captured and available via the ENCODE Portal. The pipeline code, developed using Docker and Workflow Description Language (https://openwdl.org), is publicly available in GitHub, with images available on Dockerhub (https://hub.docker.com), enabling access to a diverse range of biomedical researchers. The ENCODE pipelines, maintained and used by the DCC, can be installed to run on personal computers, local high-performance computing (HPC) clusters, or in cloud computing environments via Cromwell. Access to the pipelines and data via the cloud allows small laboratories the ability to use the data or software without access to institutional compute clusters. Standardization of the computational methodologies for analysis and quality control leads to comparable results from different ENCODE collections.

## MaizeCODE Reveals Bidirectionally Expressed Enhancers That Harbor Molecular Signatures of Maize Domestication

This work was performed in collaboration with M. Schatz (JHU), M.B. Hufford (ISU), and R. Martienssen, W.R. McCombie, D. Ware, D. Jackson, and T.R. Gingeras (CSHL).

Modern maize (corn) was domesticated from *Teosinte parviglumis*, with subsequent introgressions from *Teosinte mexicana*, yielding increased kernel row number, loss of the hard fruit case, and dissociation from the cob upon maturity, as well as fewer tillers. Molecular approaches have identified several transcription factors involved in the development of these traits, yet revealed that a complex regulatory network is at play. A National Science Foundation (NSF)-funded project entitled "MaizeCODE" deploys ENCODE strategies to catalog the functional regions of genomes in the maize and teosinte genome, generating histone modification and transcription factor ChIP-seq data in parallel with transcriptomics data sets in five tissues of three inbred lines that span the phenotypic diversity of maize, as well as the teosinte inbred TIL11. The Gingeras, Martienssen, and Shatz laboratories have completed an integrated analysis of these data sets that revealed a comprehensive set of regulatory regions in each inbred, and notably of distal enhancers, which were differentiated from gene bodies by their lack of H3K4me1. Many of these distal enhancers expressed noncoding enhancer RNAs bidirectionally, reminiscent of "super enhancers" observed in the ENCODE studies of the human genome and in other animal genomes. We show that pollen grains are the

most differentiated tissue at the transcriptomic level and that they share features with endosperm that may be related to McClintock's chromosome breakage–fusion–bridge cycle. Conversely, ears have the least conservation between maize and teosinte, both in gene expression and within regulatory regions, reflecting conspicuous morphological differences selected during domestication. The identification of molecular signatures of domestication in transcriptional regulatory regions provides a framework for directed breeding strategies in maize, as well as for broader studies undertaken by others in the scientific community.

## Assessments of the Status of the Human Gene Catalog and the Appreciation of the Functional Roles for Non-Protein-Coding RNAs

This is a summary of the Amaral et al. (2023) and the Mattick et al. (2023) reviews listed below. Early conceptions of the genome treated it as a repository for genes, most of which were thought to encode a single protein-coding transcript. One goal of the Human Genome Project has been to identify every gene in the human genome since the initial draft was published in 2001. In the years since, much progress has been made in identifying protein-coding genes, currently estimated to number fewer than 20,000, with an ever-expanding number of distinct protein-coding isoforms. However, with the completion of the Telomere-to-Telomere (T2T) Consortium's T2T-CHM13 assembly, a contiguous sequence of the human genome has been absent (Nurk et al., *Science* **376:** 43 [2022]). The results of this project have provided an additional 1,956 gene predictions, 99 of which are predicted to be new annotated protein-coding genes. Hence, the exploration and characterization of the human genome and its RNA products continue. Beside the ongoing annotation of protein-coding genes, their isoforms, and their pseudogenes, the invention of high-throughput RNA sequencing and other technological breakthroughs have led to a rapid growth in the number of reported noncoding RNA (ncRNA) genes.

The various functions of ncRNAs and their many isoforms and interleaved relationships with other genes make ncRNA classification and annotation difficult. Most ncRNAs evolve more rapidly than protein-coding sequences, are cell type–specific, and regulate many aspects of cell differentiation and development and other physiological processes. Many ncRNAs associate with chromatin-modifying complexes, are transcribed from enhancers, and nucleate phase separation of nuclear condensates and domains, indicating an intimate link between long noncoding RNA (lncRNA) expression and the spatial control of gene expression during development. ncRNAs also have important roles in the cytoplasm and beyond, including in the regulation of translation, metabolism, and signaling. Often ncRNAs have a modular structure and are rich in repeats, which are increasingly being shown to be relevant to their function.

Outside of membrane-enclosed compartments, coding and ncRNAs exist in inter- and intracellular domains protected by cofactors (proteins, carbohydrates, lipids) that are specifically or nonspecifically interacting with them. RNA-binding proteins (RBPs) are the most well-studied of these cofactors. The importance of RBPs is underscored by the size of this family of genes. Depending on the evidence used to identify RBP genes, it is estimated that the size of the RBP family varies from approximately 1,550 or 8.0% (Van Nostrand et al., *Nature* **583:** 711 [2020]) to 4,000 or 20.6% (Gebauer et al., *Nat Rev Genet* **22:** 185 [2021]) of 19,393 human protein-coding genes. Because RNAs are composed of highly structured and/or intrinsically disordered regions, the diversity of RBP genes is consistent with these structural properties. Because mRNAs and lncRNAs may have both of these structural characteristics, and because binding to an RNA may involve more than one RBP, a great deal of uncertainty in identifying a complete catalog of RBPs for each RNA persists. Finally, the identity of RBPs interacting with each RNA and the binding location on the RNA remain active areas of study.

### PUBLICATIONS

Amaral P, Carbonell-Sala S, Vega FM, Faial T, Frankish A, Gingeras T, Guigo R, Harrow JL, Hatzigeorgiou AG, Johnson R, et al. 2023. The status of the human gene catalogue. *Nature* **622:** 41–47. doi:10.1038/s41586-023-06490.

Gingeras TR. 2023 Current frontiers in RNA research. *Front RNA Res* **1:** 1152146.

Hitz BC, Lee JW, Jolanki O, Kagda MS, Graham K, Sud, P, Gabdank I, Stratton JS, Sloan CA, Dreszer T, et al. 2023. The ENCODE uniform analysis pipelines. bioRxiv doi:10.1101/2023.04.04.535623

Mattick JS, Amaral PP, Carninci P, Carpenter S, Chang HY, Chen LL, Chen R, Dean C, Dinger ME, Fitzgerald KA, et al. 2023

Long non-coding RNAs: definitions, functions, challenges and recommendations. *Nat Rev Mol Cell Biol* **24:** 430–447. doi:10.1038/s41580-022-00566-8

Rozowsky J, Gao J, Borsari B, Yang YT, Galeev T, Gürsoy G, Epstein CB, Xiong K, Xu J, Li T, et al. 2023. The EN-TEx resource of multi-tissue personal epigenomes and variant-impact models. *Cell* **186:** 1493–1511.e40. doi:10.1016/j.cell.2023.02.018

*In Press*

Cahn J, Regulski M, Lynn J, Ernst E, de Santis Alves C, Ramakrishnan S, Chougule K, Wei S, Lu Z, Xu X, et al. 2024. MaizeCODE reveals bi-directionally expressed enhancers that harbor molecular signatures of maize domestication. bioRxiv doi:10.1011/2024.02.22.581585

# LEVERAGING LONG-READ SEQUENCING TECHNOLOGIES TO FACILITATE GENOMIC DISCOVERY

**W.R. McCombie** S. Downey S. Iyer S. Muller
E. Ghiban M. Kramer A. Qui
S. Goodwin S. Mavruk Eskipehlivan R. Wappel

We are continuing to work on discovering new biological information enabled by the ability to perform increasingly high-resolution analysis of genomes, transcriptomes, and epigenomes. Much of this is related to better understanding the molecular changes that contribute to cancer growth and development. We are also carrying out evolutionary studies, especially on plants, including one critically endangered species. We also continue to develop better techniques and approaches for carrying out this type of work, particularly with long-read sequencing. We published a paper on comparative bat genomics with the Siepel laboratory (Scheben et al. 2023) and other collaborators in 2023 and posted a preprint on the *Wollemia nobilis* genome analysis (Stevenson et al. 2023).

## Analysis of Early-Onset Cancer Pedigrees

W.R. McCombie, M. Kramer, S. Goodwin, R. Wappel, S. Muller, E. Ghiban [in collaboration with Z. Stadler, M. Borio, and K. Ganesh, MSKCC]

In 2023 we expanded our investigation of early-onset cancer pedigrees using Oxford Nanopore long-read germline sequencing. This collaboration with Zsofia Stadler at Memorial Sloan Kettering Cancer Center (MSKCC) focuses on early-onset cancers in which there is no immediate family history and the standard MSKCC germline panel screening did not identify any germline cancer drivers. In addition to the two colorectal cancer trios we previously described, we added a quad consisting of two brothers affected with testicular cancer and their unaffected parents. We applied our pipeline to assess variation (single-nucleotide variants [SNVs], structural variants [SVs], and methylation changes) in the probands that differed from the parents. We then prioritized rare variants that affected genes and potential regulatory regions to attempt to elucidate changes that might be functional and predispose these individuals to cancer. We identified several interesting changes in the quad, including structural variants that impact genes and methylation changes. For example, we found a homozygous deletion in both probands (heterozygous in unaffected parents) in the last exon of gene *ZNF738*, which encodes a zinc finger protein that may regulate expression.

Our findings in these families include interesting variants in genes related to cancer, and we postulate that these changes may impact known cancer pathways in a manner less understood than the typical well-defined cancer driver mutations. We continued to assess the possibility of these variants impacting gene expression using RNA-seq on tissues from one of the colorectal cancer trio probands. We performed RNA-seq on formalin-fixed paraffin-embedded (FFPE)-preserved sections of primary tumor, adjacent normal, and two adenomas. We then examined the expression of several genes in which we had noted structural variants or methylation differences. We saw expression changes in several of the genes. Using normalized read counts across three replicates of RNA-seq, we noted a change in the expression of the gene *ITIH5*, in which the proband had a homozygous deletion in the promoter region. We note decreased expression in the adenomas and tumor compared with the matched normal. *ITIH5* is a putative tumor suppressor that is inactivated in colon cancer (Kloten et al., *Epigenetics* **9:** 1290 [2014]). We also found a homozygous deletion in the proband that overlaps the last exon of *ZNF793*, a potential tumor suppressor that is hypermethylated in Barrett's esophagus (Yu et al., *Cancer Epidemiol Biomarkers Prev* **24:** 1890 [2015]). The expression signature for *ZNF793* demonstrated progressively decreasing expression from normal colon to adenoma to the lowest expression level in the tumor. We also noted the expression of *GNAS*, a known oncogene, which we found to be differentially methylated between the proband and parents, showing increased

expression in the tumor compared with the adenomas and the matched normal. Because we evaluated germline genomic variants in this trio, our collaborators are pursuing additional sample collection of saliva and/or blood from the parents and proband. This will enable us to investigate the germline effect of the differential variants in the proband compared with the parents (i.e., the proband's homozygous genotype compared with the heterozygous parents) that may contribute to the downstream changes noted in the preneoplastic and tumor tissues. These analyses will be finalized in 2024 for publication, and a grant proposal was submitted in October 2023 to expand the investigation to a larger cohort of early-onset cancer families.

## Colon Cancer Racial Disparity

W.R. McCombie, M. Kramer, S. Goodwin, S. Muller, E. Ghiban [in collaboration with A. Krasnitz, S. Beyaz, and P. Belleau, CSHL; J. Boyd, N. Chambwe, and A. Rishi, Northwell Health; L. Martello-Rooney and J. Zeng, SUNY Downstate]

Determining the molecular underpinnings of cancer health disparities presents a significant challenge. The increased incidence and poorer prognosis of colorectal cancer in African–Americans (AAs) compared with European–Americans (EAs) is only partly explained by socioeconomic inequalities or access to healthcare. To dissect the genetic factors that may contribute to this disparity, we performed deep whole-exome sequencing on tumors from a mix of self-identified AA and EA patients. Our collaborators in the Krasnitz laboratory determined each sample's genetic ancestry to confirm the ancestry. We then compared AA versus EA cohorts. We sequenced 52 AA and 51 EA tumors, as well as 40 AA polyps and 37 EA polyps. The polyps consisted of two subtypes: tubular adenomas (26 AA and 23 EA) and serrated polyps (14 AA and 14 EA). We previously reported on our pipeline to prioritize drivers and attempt to identify somatic and germline variants in these samples, which do not have matched normals. Our initial results showed a possible enrichment of putative somatic variants in *KRAS*, *CGREF1*, *CRYBG2*, and *CSMD2* in the AA cohort compared with the EA cohort. We also noted driver mutations in *TP53*, *KRAS*, and *PIK3CA* seemed to be more enriched in tumors compared with the precancerous adenomas. However, the polyps did display *APC* mutations, and in particular the serrated polyps showed multiple *BRAF* mutations.

We plan to further refine our methods to increase confidence in the identification and categorization of somatic versus germline variants and more finely assess differences between cohorts as well to delve deeper into the types of variants that differ between cohorts and their possible impact on disease. We will also further explore differences between polyps by cohort and subtype. We are preparing a publication for 2024 and submitted a grant in 2023 to increase sample size and carry out whole-genome sequencing for more in-depth analysis.

## P1000 Colon and Endometrial Projects

W.R. McCombie, M. Kramer [in collaboration with A. Krasnitz, S. Beyaz, P. Belleau, and A. Deschenes, CSHL; N. Chambwe, Northwell Health; L. Martello-Rooney, SUNY Downstate]

In collaboration with the New York Genome Center (NYGC), Laura Martello-Rooney from Downstate, Nyasha Chambwe from Northwell, and the Beyaz and Krasnitz laboratories at CSHL, we have begun analysis of two cancer cohorts for the PolyEthnic 1000 consortium. The aim of the consortium is to explore genetic factors that may contribute to cancer disparities, including disease incidence and outcome. The first is a colon cancer cohort consisting of whole-genome, short-read sequencing of 44 AA samples and 12 EA samples. Both tumor and matched normal were sequenced for each of these samples, genetic ancestry was determined, and somatic and germline variants were called by NYGC. After data quality control to exclude samples of highly admixed ancestry (<70% African or <70% European ancestry) or low quality, we employed a pipeline to prioritize somatic driver events and potentially pathogenic germline mutations in each cohort. We then compared AA versus EA samples to try to identify differentially mutated genes. Because of the small number of EA individuals, we also compared with the larger Cancer Genome Atlas Colon Adenocarcinoma Collection (TCGA-COAD) cohort (https://www.cancer.gov/tcga). Our findings show similar mutations to TCGA in known genes such as *KRAS*, *TP53*, and *PIK3CA*, with a slightly lower rate of *APC* in the African cohort compared with the TCGA cohort.

Our CSHL collaborators have performed transcriptome analyses of a subset of these tumor samples to determine consensus molecular subtypes. They have also explored HLA typing, and have determined

copy number alterations in the samples. We will begin exploring noncoding variants to investigate regulatory changes that may affect cancer progression and will integrate these findings in 2024 to try to identify ancestry-specific differences that may contribute to colon cancer onset or severity.

The second cohort is an endometrial cancer cohort consisting of tumor-normal whole-genome sequencing of 52 AA and 11 EA samples. We have further categorized these samples by subtype, including serous and carcinosarcoma, and we will assess mutations specific to each type as well as across ancestry groups. We will continue to analyze these samples and the sequenced transcriptomes in 2024.

## Living Fossils

W.R. McCombie, S. Goodwin, M. Kramer, R. Wappel, S. Muller [in collaboration with R. Martienssen and C. Alves, CSHL; G. Coruzzi, G. Eshel, M. Katari, and V. Sondervan, NYU; D. Stevenson, B. Ambrose, S. Frangos, D. Little, and S. Wilson, NYBG; M. Schatz and S. Ramakrishnan, JHU; K. Varala, Purdue; S. Kolokotronis, SUNY Downstate Medical Center]

We continued our work with the New York Plant Genome Consortium to study the genomes of "Living Fossils" plant species, which, based on the fossil record, have not substantially changed for many millions of years. We had completed Oxford Nanopore long-read sequencing of five gymnosperm genomes: *Araucaria angustifolia*, *Juniperus communis*, *Gnetum gnemon*, *Wollemia nobilis*, and *Metasequoia glyptostroboides*. In addition to being a living fossil, *W. nobilis* is critically endangered, with fewer than 100 adult trees living in the wild (in New South Wales, Australia) and a few hundred juvenile trees. We produced a highly contiguous *Wollemia* genome and performed analyses of gene families, small RNAs, and the epigenome. We found a large proportion of transposons whose activity may contribute to epigenetic diversity and explored the regulatory interplay of small RNAs and methylation patterns. Our Australian collaborators provided genotype data sampled from the very small remaining population showing very low heterozygosity. We hope our genomic analyses will aid in conservation efforts for this critically endangered species. We have continued to improve the quality of the Bionano optical mapping data for *Wollemia*. The most recent Bionano data generation yielded 361 scaffolds, with an N50 of 76 Mb (meaning half of the total assembly is in scaffolds of this length or greater). This represents half as many scaffolds and double the N50 of the previous Bionano data. Hi-C was also completed by Corteva. Together with the new Bionano data, the Hi-C will be used to improve the overall contiguity of the assembly in order to bring it to chromosomal levels. Similarly, we have carried out Bionano on *G. gnemon*. The hybrid scaffold yielded 412 scaffolds with an N50 of 20 Gb. We will continue to try to improve these metrics while we analyze the *Gnetum* genome. The initial version of the *Wollemia* manuscript was posted to bioRxiv in August 2023 (Stevenson et al. 2023) and will be updated with the new assemblies and annotations as they are generated. A *Gnetum* manuscript is in preparation, which will explore the unique morphological features of *Gnetales*, such as their ability to exist as both trees and vines, and will attempt to further classify their position in the evolution of land plants. Oxford Nanopore sequencing technology (ONT) assemblies have been completed for the three other species and are undergoing polishing to improve base quality. We completed sequencing of several of the transcriptome and methylation data sets for leaf and reproductive tissues and will prepare Bionano and Hi-C libraries to aim for chromosome-level assembly. We will then perform comparative genomic analyses of the living fossils versus their speciated sister pairs (*Wollemia* to *Araucaria*, *Metasequoia* to *Juniperus,* and *Gnetum* to the already published *Welwitschia* genome) that have evolved more recently to prepare a final manuscript. These analyses may uncover the genetic factors that enabled these species to survive despite extreme changes in climate and habitat.

## B73 T2T Assembly of Plant Samples

S. Goodwin, R. Wappel, M. Kramer [in collaboration with D. Ware, T. Gingeras, and R. Martienssen, CSHL; M. Schatz, Johns Hopkins; S. Koren and A. Philippy, NHGRI; A. Wittenberg, KeyGene]

Recent work by the McCombie laboratory and others has underscored the importance of long-read sequencing and assembly to enhance variant calling and improve the resolution of challenging repeat regions such as centromeres. In 2022, we began collaborating with researchers from Cold Spring Harbor Laboratory (Doreen Ware, Rob Martienssen, and Tom Gingeras), Johns Hopkins University (Mike Schatz), the National Human Genome Research Institute (Sergey Koren and Adam Philippy), and KeyGene (Alexander

Wittenberg) to create a high-quality long-read reference genome for maize (B73) and tomato (Heinz 1706) crucial model and crop plants.

Both tomato and maize assemblies were highly contiguous with N50s of 63.8 Mbp and 152.5 Mbp, respectively, exceeding their current reference assembly N50s of 41.7 Mbp and 47.0 Mbp. Although highly contiguous, some gaps remained in each assembly. For tomato, five of 12 chromosomes were resolved. The unresolved regions comprise a complex repeat corresponding to the 45-rDNA array on chromosome 2; an AT-rich region on chromosome 3 with a single spanning ONT read; peri-telomeric regions on chromosomes 11 and 12; and unresolved regions of heterozygosity that could not be phased on chromosomes 8, 9, 10, 11, and 12.

In the maize assembly, three chromosomes (1, 8, and 9) had four unresolved regions of heterozygosity; chromosome 6 has a complex rDNA repeat—again corresponding to the rDNA array—and coverage gaps corresponding to AT-rich regions in chromosomes 1, 2, and 4.

The methods and insight garnered through this study will be used to inform future work in T2T genome assemblies, which, in turn, will aid in understanding the complexity of highly recalcitrant repetitive regions within various plant genomes.

## Direct Sequencing of Insect Symbionts via Nanopore Adaptive Sampling

W.R. McCombie, S. Goodwin, R. Wappel, S. Mavruk Eskipehlivan, S. Muller [in collaboration with J. Bader, NIH; R. Giordano, FIU; A. Zimin, JHU, CSHL; K. Donthu, University of Illinois; P. Vieria and I. Zasada, USDA]

Insect symbionts can significantly influence host phenotype, ranging from beneficial to pathogenic effects, with many insects harboring multiple co-infections that complicate their study. However, <1% of insect species have high-quality referenced genomes available, and even fewer have their symbionts sequenced. In many cases research into these symbionts is limited as they cannot be cultured, are not present in a high enough population to be identified in lower-coverage sequencing studies, and/or are not well characterized in the literature or databases. Adaptive sampling, a feature of ONT, enables real-time selection or rejection of long-read sequences during sequencing runs. Leveraging this capability, we developed a subtractive adaptive sampling method to sequence the complete genomes of mitochondria, *Buchnera*, and its plasmids (pLeu, pTrp), as well as *Wolbachia* genomes in two aphid species, *Aphis glycines* and *Pentalonia nigronervosa*. gDNA was extracted from approximately 10 aphids and a fastq file representing the host (either *A. glycines* or *P. nigronervosa)* was provided to the MinKnow software. Reads that mapped to the host genome were rejected from the MinION flow cell in real time, whereas other reads were sequenced. This approach resulted in >100-fold enrichment of the genomes of interest over the background, and each genome was assembled into one complete contig. In the case of the plasmids, individual nanopore reads spanned the entire molecule. These reads revealed the exact number of operons present on each plasmid, a detail not available via short-read methods.

## ACME: A Method for Targeted Nanopore Sequencing

S.V. Iyer, M. Kramer, S. Goodwin, W.R. McCombie [in collaboration with A.N. Habowski, B. Yueh, D.A. Tuveson, and S. Beyaz, CSHL]

ACME (Affinity-based Cas9-mediated enrichment) is a targeted long-read sequencing method developed as part of S.V. Iyer's doctoral work over the past few years (Iyer et al., bioRxiv doi:10.1101/2022.02.03.478558 [2022]). We have since used ACME to target genes of interest in pancreatic ductal adenocarcinoma (PDAC) patient-derived organoid (PDO) lines (collaboration with Tuveson laboratory), showing its expansion to an important cancer model. Continuing our work on the PDAC PDOs from 2022, in 2023 we analyzed the ACME sequencing data generated and compared the variants called with existing short-read sequencing and expression data. We also began work using ACME to evaluate colorectal cancer (CRC) PDOs in 2023 in collaboration with the Beyaz laboratory.

### ACME Targeting of Cancer Gene Panel in PDAC PDOs

In 2022, we generated ACME data for four established PDAC PDOs lines—hT108, hT101, hT93, and hF28 (Tiriac et al., *Cancer Discov* **8:** 1112 [2018])—obtained from the Tuveson laboratory. For each organoid, two libraries were prepared (three for hT108)

and each library was run on a separate flow cell. For this work, we used the CS9109 kit on DNA we extracted using the NEB Monarch HMW DNA extraction kit. All runs were basecalled using the Guppy v6 basecaller.

In 2023, for each organoid we assessed the performance of the individual single sample runs separately and compared those results to data from "merged" runs, wherein FASTQ files from the individual runs were merged prior to mapping and variant calling. Although a merged run is akin to pooled runs that were used during benchmarking, the key difference is that for pooled runs individual sample preps were pooled together during library prep and loaded on a single flow cell, whereas for merged runs individual libraries were run on separate flow cells and the basecalled data from the multiple individual runs were merged after sequencing was complete. For each organoid, we analyzed data from the "individual runs" and "merged runs" to run comparisons.

Coverage and end-to-end target-spanning reads generated for each organoid were as expected and the performances of the individual runs were fairly similar to each other, with any differences attributed to variability typically seen between two ONT runs. As expected, the merged runs fared better than their corresponding individual runs.

#### *Variant calling from individual and merged runs*

Next, we assessed whether the better performance seen from the merged runs translated to improved variant calling. We observed a high concordance among the individual runs for each organoid, as well as between the individual runs and their corresponding merged run. On looking closely at the one to two additional SVs found by the merged runs for each organoid, we observed that those SVs were found by the corresponding individual runs as well but were filtered out because of read support of <4.

#### *Variants of interest found with ACME*

Using data from the merged runs, we were able to successfully identify known *KRAS* driver mutations—the G12D and G12V missense mutations from our ACME data (Fig. 1). The SVs detected from the merged runs were annotated using the Database of Genomic Structural Variation/Database of Genetic Variants (dbVar/DGV), the Encyclopedia of DNA Elements (ENCODE), expression quantitative trait loci (eQTL) markers, segmental duplications, and Online Mendelian Inheritance in Man (OMIM) gene and UCSC genome browser tracks to assess functional significance of the identified events. This helped us determine that the ~500-bp homozygous deletion of exon 8 we observed in the *STK11* gene in the hT101 organoid (Fig. 1B) corresponded to a highly conserved regulatory region that has been linked to somatic pancreatic cancer (Jegher et al., *N Engl J Med* **241:** 993 [1949]; Su et al., *Am J Pathol* **154:** 1835 [1999]). This deletion was unique to hT101 and was not observed in the other three organoids examined. Moreover, on looking at Illumina short-read RNA-seq expression data for all four organoids (Tiriac et al., *Cancer Discov* **8:** 1112 [2018]), we were able to detect differential expression counts in *STK11* (reduced expression) for hT101, confirming the impact of the observed deletion on gene expression in this line. Other organoids exhibited normal *STK11* expression. We also identified intronic SVs in *FGFR4*, *CDKN2A*, and *BRCA2* genes in one or more organoids, but these events did not correspond to any changes in gene expression.

### ACME Targeting of Expanded Cancer Gene Panel in CRC PDOs

In collaboration with the Beyaz laboratory, we will use ACME to compare genomic variants and methylation status between primary and metastatic tumor-derived organoids from the same patient. We extracted high-molecular-weight (HMW) DNA from four organoid lines established by the Beyaz laboratory based on primary and metastatic tumors from two patients—C19T, C19M, C22T, and C22M. For this work we also expanded the ACME panel to include guides we designed that target the entire *TP53* gene and its promoter region. We successfully used ACME to target this expanded cancer panel in one sample in 2023. In 2024 we will sequence the remaining three samples and analyze the data generated.

## 1000 Genomes Long-Read Project

S. Goodwin, R. Wappel, S. Mavruk Eskipehlivan, S. Muller, W.R. McCombie [in collaboration with the 1000 Genome ONT Sequencing Consortium; including M.C. Zody, M. Loose, M. Jain, E.E. Eichler, and D.E. Miller]

Less than half of individuals with a suspected Mendelian condition receive a clear molecular diagnosis after genetic testing. Recent work by the McCombie group

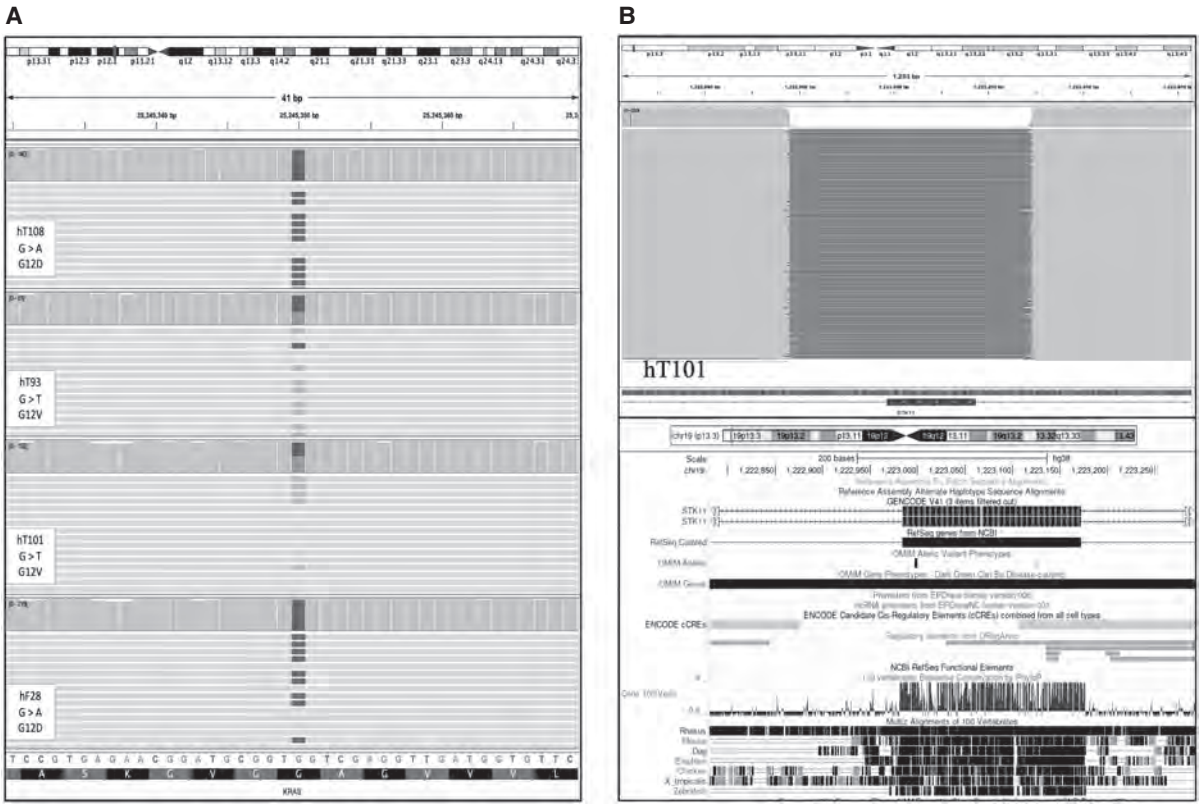



**Figure 1.** (*A*) Integrative Genomics Viewer (IGV) tracks showing reads that mapped to the *KRAS* gene across different pancreatic ductal adenocarcinoma (PDAC) organoids. Shaded central bars represent two different missense mutations (G>T, G>A) found at the same locus in all four patient-derived organoids (PDOs). (*B*) Homozygous deletion in the *STK11* gene. (*Top*) IGV track showing reads that mapped to the *STK11* gene in PDO hT101. The gap with black lines shows a homozygous deletion of exon 8 and its flanking regions in the PDAC organoid hT101. This deletion was not observed in the other organoids. (*Bottom*) UCSC Genome Browser tracks spanning Exon 8 of *STK11*. The deleted region was observed to be a highly conserved regulatory region and has been linked to somatic pancreatic cancer in the OMIM database.

and others has shown long-read sequencing (LRS) to be a promising avenue to explore for clinical applications. However, there are few control data sets for variant filtering, which poses a challenge for downstream analysis of LRS data. To tackle this issue, the McCombie laboratory has joined the 1000 Genomes Project ONT Sequencing Consortium to generate and analyze LRS data from 1000 samples included in the 1000 Genomes Project. This endeavor will uncover a broader spectrum of variation and enhance our understanding of normal human variation patterns. In the initial phase of this study the samples sequenced at CSHL and other institutions had an average depth of coverage of 37x with an average sequence read N50 (the point at which 50% of all bases sequenced are in fragments larger than the value) of 54 kbp. Utilizing multiple SV calling tools, the consortium identified an average of 24,543 high-confidence SVs per genome, including shared and private SVs with potential gene-disrupting effects and pathogenic expansions within disease-associated repeats undetected by short reads. Evaluation of methylation signatures revealed anticipated patterns at known imprinted loci, samples with skewed X-inactivation patterns, and novel differentially methylated regions. The availability of all raw sequencing data, processed data, and summary statistics offers a valuable resource for the clinical genetics community in discovering pathogenic SVs.

## ABRF Methylation Study

S. Goodwin, R. Wappel, M. Kramer, W.R. McCombie
[in collaboration with members of the ABRF DSRG]

Epigenetic modifications have been shown to have significant importance in genomic regulation, and protocols and analysis methods for interrogation of genomic methylation have evolved in recent years. To provide an overview of these methods and their utility, we began a project with the Association of Biomedical Research Facilities (ABRF) DNA Sequencing Research Group (DSRG) to assess the cross-platform performance of sequencing-based methylation profiling. The study encompasses four sample types: the well-studied NA12878 "Genome in a Bottle" human sample, the MCF7 breast cancer cell line, and two standard samples from the HCT116 DKO cell line—one fully methylated and one unmethylated. Protocols for short-read sequencing included whole-genome bisulfite conversion and whole-genome enzymatic conversion, as well as a targeted approach using the Illumina EPIC array. Both PacBio and Oxford Nanopore sequencing, which provide direct detection of modifications without conversion, were chosen for the long-read sequencing approaches. Samples were prepared for each method at three different sites, with two replicates per sample, and all libraries were then sequenced at the McDonnell Genome Institute. The Cold Spring Harbor Next Generation Sequencing Shared Resource was one of the participating sites, performing library preparation for Oxford Nanopore sequencing of the four samples (two replicates each). Sequencing was completed in 2023 for all platforms, and initial run quality control (QC) assessment showed reasonably uniform yield, read length, and quality across sites. Data analysis will commence in 2024. We will focus on 5mC methylation detection in CpG context, noting reproducibility, correlation, and concordance across platforms, as well as compare and contrast the strengths and weaknesses of each. These results will be presented at ABRF meetings to inform the community and will be prepared for publication.

### PUBLICATIONS

Akagi K, Symer DE, Mahmoud M, Jiang B, Goodwin S, Wangsa D, Li Z, Xiao W, Dunn JD, Ried T, et al. 2023. Intratumoral heterogeneity and clonal evolution induced by HPV integration. *Cancer Discov* **13:** 910–927. doi:10.1158/2159-8290.CD-22-0900

Gladman N, Goodwin S, Chougule K, McCombie WR, Ware D. 2023. Era of gapless plant genomes: innovations in sequencing and mapping technologies revolutionize genomics and breeding. *Curr Opin Biotechnol* **79:** 102886. doi:10.1016/j.copbio.2022.102886

Scheben A, Mendivil Ramos O, Kramer M, Goodwin S, Oppenheim S, Becker DJ, Schatz MC, Simmons NB, Siepel A, McCombie WR. 2023. Long-read sequencing reveals rapid evolution of immunity-and-cancer related bats. *Genome Biol Evol* **15:** evad148. doi:10.1093/gbe/evad148

Stevenson DW, Ramakrishnan S, Alves CdS, Coelho LA, Kramer M, Goodwin S, Mendivil Ramos O, Eshel G, Sondervan VM, Frangos S, et al. 2023. The genome of the *Wollemi* pine, a critically endangered "living fossil" unchanged since the Cretaceous, reveals extensive ancient transposon activity. bioRxiv doi:10.1101/2023.08.24.554647

# GENETICS OF PLANT ARCHITECTURE AND ENVIRONMENTAL RESPONSE

**D. Ware** J. Braynen, K. Chougule, A. Fahey, N. Gladman, C. Kim, V. Kumar, S. Kumari, Z. Lu, An. Olson, Au. Olson, M. Regulski, M.K. Tello-Ruiz, P. Van Buren, S. Wei, L. Zhang

The global challenges confronting agricultural security are falling into sharper relief: depleted water resources for irrigation; surging pest pressures due to longer and hotter growing seasons; degrading arable land; increasing population; and long-term geographical adjustments brought about through climate change. Overcoming these strata of obstacles necessitates nimble and reliable approaches. Predictive genetics of desirable traits in concert with rapid germplasm conversion has become the norm since high-throughput sequencing has become cost-effective and genome editing and transformation techniques continually improve. As a result, the pan-genomes and new crop genomes are available for molecular investigation and comparative genomics, thus strengthening and accelerating the output of researchers and hastening the fruits of such labor into the hands of producers.

Although the predictive genomic paradigms are still being optimized, they are showing improved reliability depending on the desired trait. Our focus is on constructing the regulatory modules that control inflorescence and root architecture as well as how plant and crop models respond to nutrient-limiting conditions in the environment. It is crucial for plant science investigators to continue the molecular dissection of pathways controlling beneficial agronomic traits like flower fertility, inflorescence architecture, root formation, and nutrient use efficiency. These research areas have noted significant improvements in crop yield, sustainability, and quality traits such as protein and metabolite content in grains.

Additionally, characterization across numerous plant species can yield a more unified systems biology model that can be effectively applied to numerous agricultural challenges. One of the primary crops we study is *Sorghum bicolor*—a C4 grass that is resilient to heat, salt, and drought stress and is consumed worldwide. Our field and computational analyses also target monocots such as rice and maize and dicots such as *Arabidopsis thaliana* and poplar to understand their gene regulatory modules and adaptability responses.

## Interplay between Transcription Factor and microRNA Gene Regulates Floral Development in the Model Plant *Arabidopsis*

An. Olson, D. Ware [in collaboration with Y. Lee, University of Science & Technology, Gunsan-si, Republic of Korea]

Floral architecture, governed by both genetic and environmental factors, profoundly influences plant development, impacting traits like seed pod formation in *Arabidopsis* and crop yield. Zinc finger homeodomain (ZF-HD) transcription factors (TFs) are found in low copies within plant genomes, compared to other TF families. Their functional significance remains a subject of exploration to dissect the intricacies of plant developmental processes. To understand the functional roles of ZF-HD TFs, mutant lines lacking the repressor function of HB31 and HB21, two paralogous ZF-HD TFs in *Arabidopsis*, were generated. The resulting mutant lines exhibited notable aberrations in plant morphology, including increased branching, reduced stature, altered floral morphology, and shortened silique length. Our investigation revealed that HB31 and HB21 act as positive regulators of genes associated with cell size determination, cell wall modification, and certain M-type MADS-box TF families. Conversely, they exert negative regulation on genes related to abiotic stress responses, vegetative-to-reproductive phase transition, and specific transcription factor families like TCP and RAV. Furthermore, analysis of noncoding microRNA gene expression profiles revealed significant down-regulation of miR164, miR822, miR396, miR2934, and miR172, alongside up-regulation of miR169, miR398, miR399, and

miR157 in the HB31 and HB21 repressor lines. Notably, phenotypic and molecular assessments elucidated the involvement of the miR396/growth-regulating factor (GRF) regulatory module, modulated by HB31 and HB21, in shaping *Arabidopsis* floral architecture. This study provides insights into the functional roles of ZF-HD TFs in orchestrating floral development in *Arabidopsis*, shedding light on the regulatory mechanisms underlying floral architecture maintenance (Lee et al. 2024).

## Characterization of Inflorescence Architecture Gene Regulatory Networks across Diverse *Sorghum* Accessions

N. Gladman, S. Kumari, A. Fahey, M. Regulski, D. Ware

Sorghum seed heads, or panicles, can present diverse morphological features between even closely related accessions. This includes varied branching number, branch lengths, seed spacing along branches, and seed size. To determine what gene regulatory networks (GRNs) impact these panicle traits, we performed an RNA-seq analysis across multiple developmental stages of developing inflorescence meristems (as quantified by meristem length) and identified conserved and divergent gene expression clusters between accessions. This includes pathways involved in general cellular growth and progression, but also well-known ontologies that have been shown to influence inflorescence characteristics, such as jasmonic acid signaling, auxin signaling, and cell polarity and organization. Additionally, we have profiled the DNA-binding sites of the sorghum orthologs of well-known maize transcription factors that influence tassel development, Tassel Sheath 4 (TSH4) and Bearded Ear 1 (BE1), and show that certain gene clusters could be under a greater influence of TSH4 regulation rather than BE1. The purpose of this pan-transcriptome is to yield improved GRNs as well as create a resource for projecting across bulk RNA-seq data sets once snRNA-seq data sets come online from our group and others in the community. See Figure 1.

## GRAS Family Transcription Factor Binding Behaviors in *Sorghum bicolor*

N. Gladman, S. Kumari, A. Fahey, M. Regulski, D. Ware

Identifying noncoding regions that control gene expression has become an essential aspect of understanding gene regulatory networks that can play a role in crop improvements such as crop manipulation, stress response, and plant evolution. TF-binding approaches such as ChIP-seq or DAP-seq can provide additional valuable insight and targets for reverse genetic approaches such as ethyl methanesulfonate (EMS)-induced or natural single-nucleotide polymorphism (SNP) variant screens or CRISPR editing techniques (e.g., promoter bashing). Three GRAS family TFs (SHR, SCL23, and SCL3) were chosen to be profiled via DAP-seq based on their ability to be produced in bacterial expression systems and as representatives of different clades within the GRAS family in the crop *Sorghum bicolor*. The binding behavior of the three GRAS TFs displays unique and shared gene targets and categories of previously characterized DNA-binding motifs as well as novel sequences that could potentially be GRAS family–specific recognition motifs. Target genes include those associated with gametogenesis, floral development, light signaling, hormone signaling, and root development. These results provide unique insight into the GRAS family of TFs and novel regulatory targets for further molecular characterization. See Figure 2.

## Characterization of a Novel Zinc Chaperone in *Arabidopsis*

L. Zhang, A. Fahey, M. Regulski, J. Braynen, D. Ware
[in collaboration with P. Cifani, CSHL; M. Pasquini, and M. Xie, Brookhaven National Laboratory; C.E. Blaby-Haas, Lawrence Berkeley National Laboratory]

Metal homeostasis has evolved to tightly modulate the availability of metals within the cell while avoiding cytotoxic interactions due to excess and protein inactivity due to micronutrient deficiency. Zinc (Zn) is prevalent because of its nontoxic nature and abundance in the environment, and it is commonly used as a protein cofactor to stabilize proteins and assist in function. In addition to being essential for transcription, translation, and regulation of protein abundance, Zn availability is critical for chloroplast biology and photosynthesis. Even though plants have evolved strategies to ensure uptake and storage, the fate of Zn ions after import into cells and before binding to Zn-dependent proteins is largely unknown. Here, we are investigating the function of two distinct subfamilies of mammalian Zn chaperones, named ZNG1 for Zn-regulated GTPase metalloprotein activator 1. AtZNG1 (AT1G26520) is

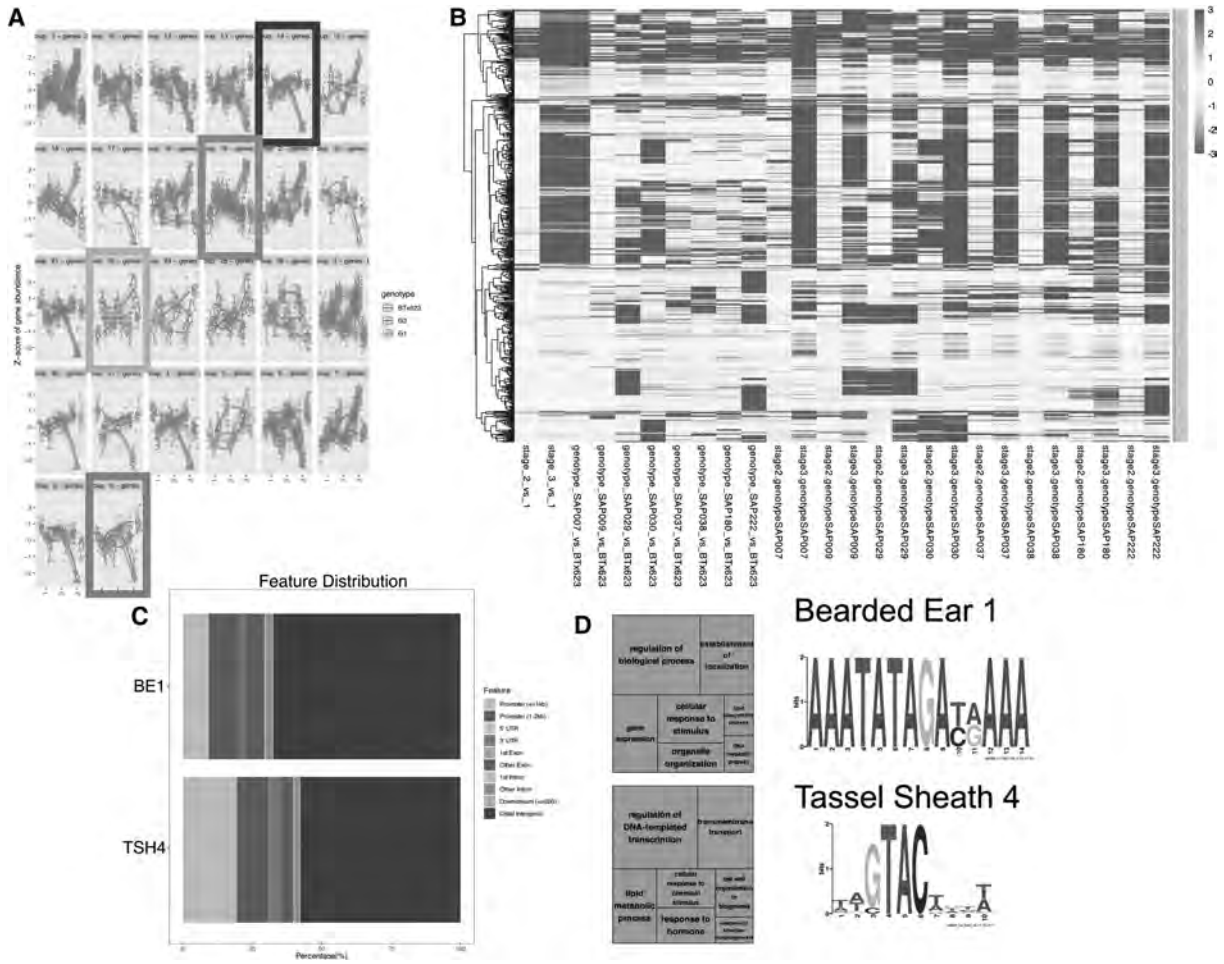



**Figure 1.** Sorghum developing meristem pan-transcriptome. (*A*) The median expression levels of highlighted clustered genes display well-known ontologies relating to inflorescence development. (*B*) Heatmap of the top 500 significantly differentially expressed genes (by adjusted *p* value) between nine sorghum accessions across three stages of inflorescence meristem development, using BTx623 stage 1 as control. (*C*) Feature distribution of DAP-seq peaks from Tassel Sheath 4 and Bearded Ear 1 in sorghum. (*D*) Gene Ontology (GO) biological process terms for genes with BE1 and TSH4 peaks in their promoter regions.

orthologous to human and fungal ZNG1 and localized to the cytosol; it interacts with methionine aminopeptidase type 1 (AtMAP1A).

The physical interaction of AtZNG1 and AtMAP1A, identified via AtZNG Y2H and BiFc assay, is further supported by structural modeling of the two protein complexes. Consistent with the hypothesized role of AtZNG1 in activating AtMAP1A, the zng1 mutant was not significantly different from Col-o in plus Zn, but in minus Zn, a growth defect was evident: zng1 has shorter primary and lateral roots, whereas the map1a mutant was severely growth inhibited. This is consistent with the role of *Arabidopsis* ZNG1 as a zinc transferase for AtMAP1A, as previously described in yeast and zebrafish. Structural modeling reveals a flexible cysteine-rich loop that, we hypothesize, enables direct transfer of zinc from AtZNG1 to AtMAP1A during GTP hydrolysis. Based on proteomics and transcriptomics, the loss of this ancient and conserved mechanism has pleiotropic consequences, impacting the expression of hundreds of genes, including those involved in photosynthesis and vesicle transport. This understanding was complemented by the functional characterization of another plant subfamily, ZNG2. There are two ZNG2 proteins in *Arabidopsis*; they are also required during zinc deficiency, but their target proteins remain to be discovered. This work was published in *Frontiers in Plant Science* (Zhang et al. 2023).

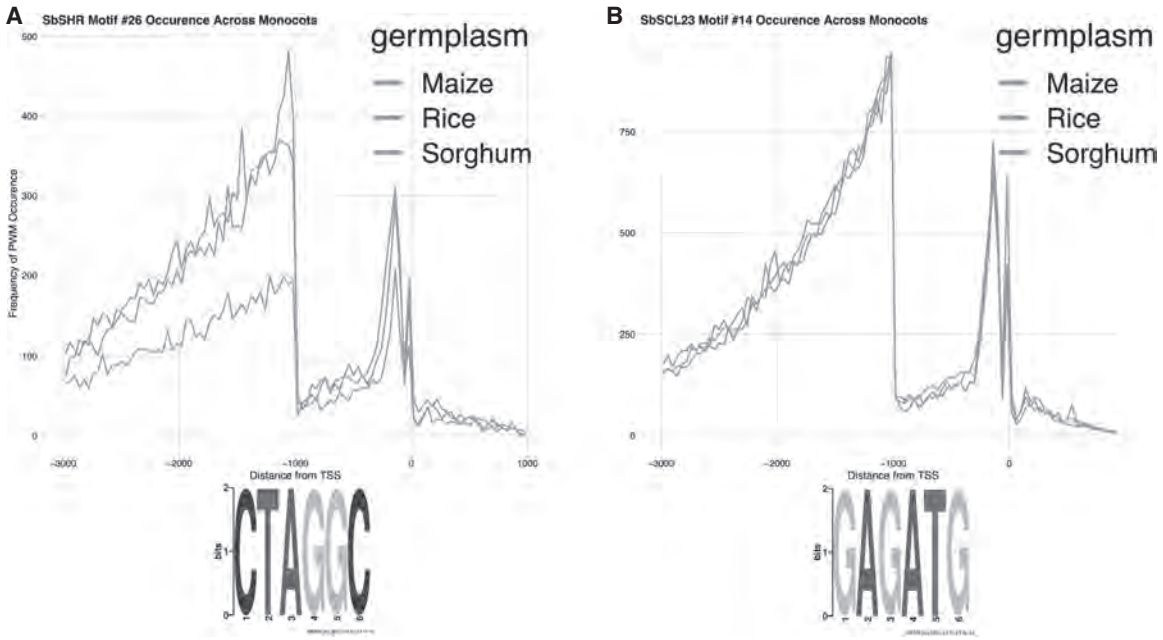



**Figure 2.** GRAS family transcription factor behavior in sorghum and other monocots. The frequency between 3,000 bp upstream of to 1,000 bp downstream from the transcriptional start site for the motif discovered for (*A*) SHR and (*B*) SLC23 in sorghum DAP-seq. The frequency projections were calculated for sorghum, rice, and maize. The frequency degradation in rice as in the case of the SHR peak (*A*) could indicate that the gene regulatory module for this motif has undergone some presence/absence variation in the gene ortholog itself, or some type of promoter region "rewiring" in rice that reduces frequency conservation between monocots.

## Functional Diversification in Heme Metabolism

L. Zhang, A. Fahey, D. Ware [in collaboration with N. Grosjean, and D. Kumaran, Brookhaven National Laboratory; C.E. Blaby-Haas, Lawrence Berkeley National Laboratory]

In this study, we combine sequence similarity, phylogenetic reconstruction, domain identification, and gene neighborhood context with genetic, biochemical, and structural characterization, revealing a large multifunctional family of enzymatic and nonenzymatic proteins involved in heme metabolism. This family, which is named homolog of HugZ (HOZ), is embedded in the "FMN-binding split barrel" superfamily, in which amino acid mutation combined with independent instances of domain fusion in prokaryotes, plants, and algae have resulted in a large group of proteins that bind free *b*-type or *c*-type heme and either catalyze its degradation or function as nonenzymatic heme sensors. Whereas in prokaryotes, these proteins are often involved in iron (Fe) assimilation, most plant and algal homologs are predicted to degrade heme in the plastid or regulate heme biosynthesis. In *A. thaliana*, which contains two HOZ subfamilies that can degrade heme (HOZ1 and HOZ2), we obtained T-DNA insertion lines for two HOZ2 paralogs, hoz2a and hoz2b. Because there is only one gene in HOZ1, we used a gene-editing approach to disrupt At*HOZ1.* Characterization of multiple loss of function mutations in the *HOZ1* and *hoz2a* causes stunted growth and delayed development, pointing to an important biological role of HOZ1-catalyzed heme turnover, whereas *hoz2b* showed no phenotype. *HOZ1* and *HOZ2a* mutants display slower growth compared with the wild type for primary root growth on MS plates and plant growth in soil. Their actions are epistatic, whereas *HOZ2b* mutants show no effects. The phenotype of the triple mutation of these three genes is similar to that of *hoz1* and *hoz2a* single mutants. See Figure 3.

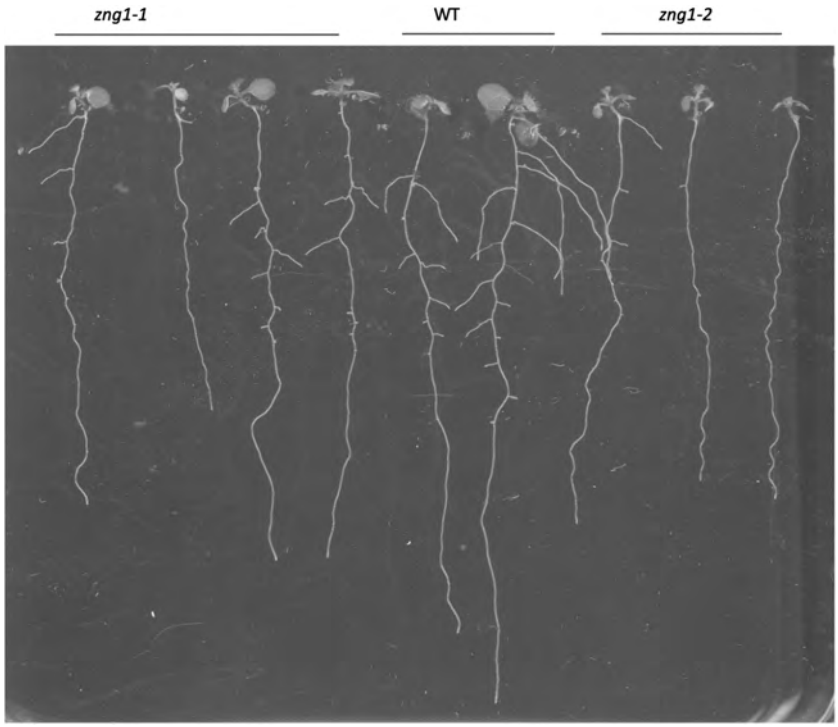



**Figure 3.** Wild type (Col-0) and mutants (*zng1-1* and *zng1-2*) were grown on MS plates without zinc for 10 days. Reduction in primary and lateral roots as well as leaf biomass observed in mutants when compared with wild type.

## Transcriptomic and Phenotypic Responses to Iron and Zinc Stress in *S. bicolor* and *Populus trichocarpa*

S. Kumari, V. Kumar, J. Braynen, D. Ware [in collaboration with T. Paape and A. Mishra, Texas A&M; D. Tedesse and M. Xie, Brookhaven National Laboratory]

Micronutrients such as Fe and Zn are required for fundamental metabolic processes including chlorophyll biosynthesis, photosynthesis, respiration, and nitrogen assimilation. Metal ion imbalances due to limiting Fe or Zn can lead to stress, impacting plant growth, biomass production, and yield. Conversely, excessive concentrations of Fe or Zn can also result in cellular toxicity. Plants depend on two primary strategies to acquire Fe/Zn: Strategy I (Reduction Strategy), adapted by non-grass-like *Arabidopsis* in which ATPase efflux transporters reduce Fe3+ to Fe2+ for uptake by roots, and Strategy II (Chelation Strategy), adapted by grasses like maize, wheat, etc., which relies on secreted phyto-siderophores (PSs) to chelate Fe3+ to become bioavailable for uptake by roots. However, rice (a grass) is known to use a combined strategy (both Strategy I and II) for root ion uptake. In the present study, *S. bicolor* BTx623 was used to quantify the response to Fe and Zn stresses (limit and excess) in root and leaf tissues over a 21-day period, with the purpose of characterizing the complex regulatory cascade of Fe and Zn homoeostasis. Using inductively coupled plasma mass spectrometry (ICP-MS), fluorometer and reflectance spectroscopy, we observed a sharp decline in the chlorophyll content, photosynthetic activity, and nitrogen assimilation in relation to the relative Fe or Zn concentration in leaf or root tissues. Time series RNA-seq analysis allowed us to identify genes with significantly up-regulated expressions (FIT, PYE, BTS, FRO, SAMs, NAS, NAAT, YSLs) in root as early stress responses. Similarly, the expression of genes involved in leaf metabolic processes (chlorophyll synthesis, photosynthesis, oxidative stress) was found to be mostly down-regulated in late stress responses. Using these data, we also built coexpression networks to model the interaction of various transporters and transcription factors that are important for Fe and Zn homoeostasis. The genes BRUTUS

(*BTS*) and FIT Binding Protein (*FBP*) were found to be involved in an ON/OFF switch that regulates the signaling cascade controlling metal ion uptake and homeostasis. Our findings also suggest that sorghum may use a combined strategy like rice to maintain ion (Fe and Zn) homeostasis during stress. This large multi-omics data set provides a valuable resource for further research into the roles of important micronutrient transport mechanisms affecting plant biomass and possible mechanisms for micronutrient allocation into tissues such as leaves or grains, where optimizing yield is essential. See Figure 4.

## Genetic Analysis of Iron Uptake and Carbon Partitioning in Four Sorghum A-Lines

J. Braynen, S. Kumari, M. Regulski, D. Ware [in collaboration with A. Bhat, D. Tedesse, and M. Xie, Brookhaven National Laboratory; T. Paape, Texas A&M]

Fe is a fundamental micronutrient essential for plant growth and development, playing a critical role in vital processes such as photosynthesis, respiration, and chlorophyll synthesis. Understanding the intricate mechanisms underlying Fe uptake and utilization in plants is crucial for enhancing crop productivity and

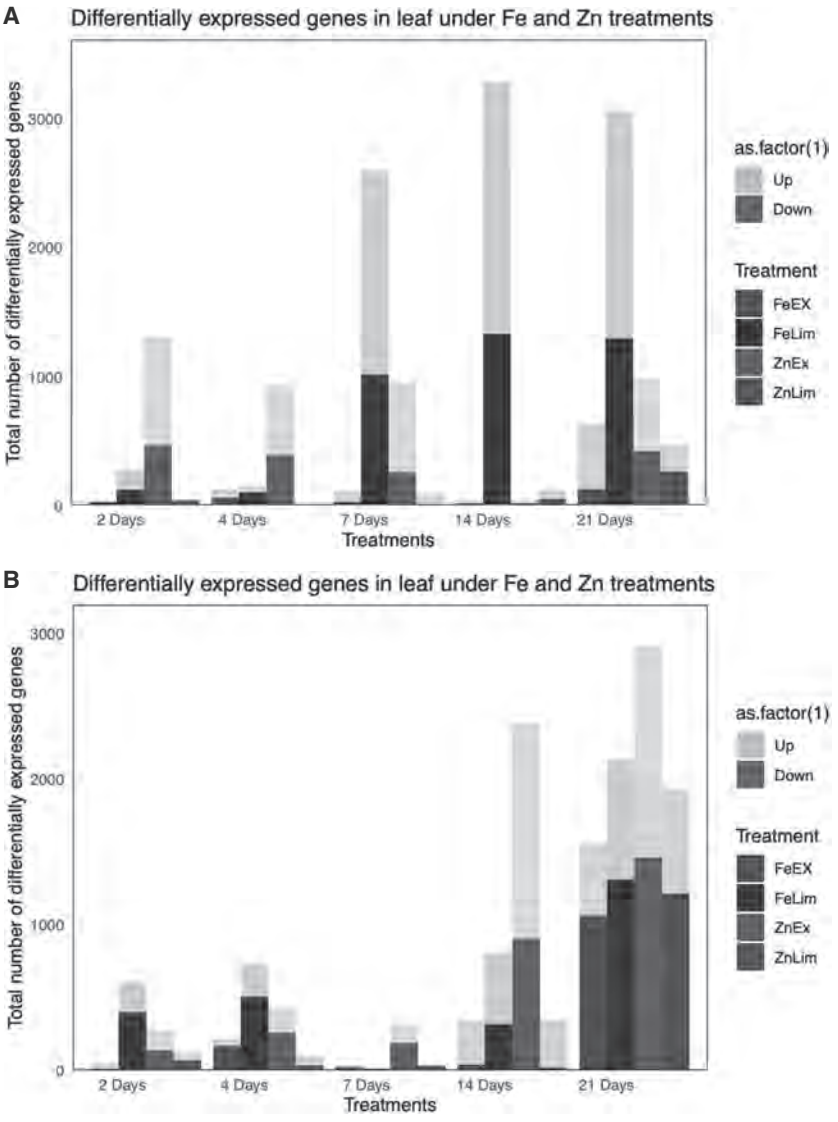



**Figure 4.** Total number of differentially expressed genes in sorghum leaf (*A*) and poplar leaf (*B*) under iron (Fe) and zinc (Zn) treatments.

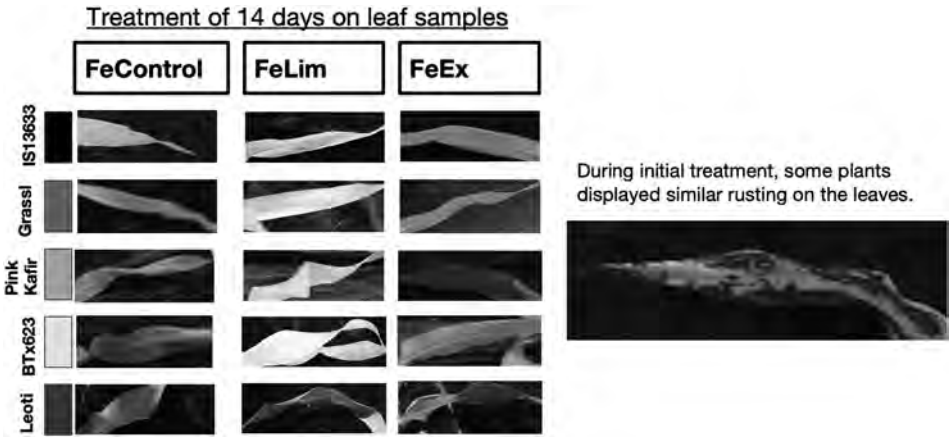



**Figure 5.** Iron availability response in sorghum CP-NAM genotypes and BTx623. Genetic response to varying Fe conditions was analyzed for four sorghum CP-NAM genotypes—Grassl, Leoti, Pink Kafir, and 1S13633—and the reference genotype BTx623. Chlorosis was observed under Fe stress at 7 and 14 days across all genotypes except for Leoti, which maintained its green coloration, indicating differential genetic responses to Fe deficiency stress among the genotypes.

resilience, particularly in environments with limited Fe availability. This annual report presents the findings of a comprehensive study aimed at deciphering the genetic response to varying Fe conditions and its impact on carbon partitioning in *S. bicolor*, a staple cereal crop in many regions worldwide. The study focused on four sorghum carbon partitioning (CP) nested association mapping (NAM) populations, utilizing a controlled hydroponic system to ensure consistent Fe availability across experimental conditions. Employing advanced molecular techniques such as RNA-seq and ICP-MS, alongside biomass measurements, enabled a thorough investigation into the genetic regulation of Fe uptake and carbon allocation in sorghum genotypes under different Fe regimes. The results of our analysis revealed intricate gene expression patterns associated with sorghum's response to varying Fe availability. Specific gene clusters were shown to play a pivotal role in the plant's ability to adapt to Fe-deficient conditions, facilitating efficient Fe uptake and utilization. Furthermore, we elucidated the mechanisms underlying the redirection of photosynthetically fixed carbon toward essential biological functions, particularly under Fe stress, highlighting the critical importance of carbon partitioning in sustaining plant growth and yield. These insights into the genetic basis of micronutrient utilization in sorghum hold significant implications for agricultural innovation and crop improvement strategies. By leveraging targeted breeding approaches, we can potentially enhance both iron-use efficiency and the optimal distribution of carbon resources within the plant, thereby addressing the challenges posed by Fe-deficient soils and improving agricultural sustainability. See Figure 5.

## Decoding Nitrogen Use Efficiency in Maize and Sorghum: Insights from Comparative Gene Regulatory Networks for Sustainable Agriculture

J. Braynen, L. Zhang, S. Kumari, V. Kumar, An. Olson, M. Regulski, D. Ware [in collaboration with C. Liseron-Monfils, National Research Council Canada; A. Gaudinier, UC Berkeley; S. Brady, UC Davis; M. Frank and B. Shen, Corteva Agriscience]

In this study, the critical role of nitrogen use efficiency (NUE) in environmental sustainability and plant nutrient uptake optimization was investigated. Through the application of yeast one-hybrid (Y1H) assays, GRNs were constructed to elucidate the regulatory mechanisms of NUE in dicot and monocot species. A significant advancement was the development of a maize GRN, incorporating 1,625 protein–DNA interactions (PDIs), and its subsequent comparison with an existing *Arabidopsis* GRN. This comparative analysis revealed conserved nitrate assimilation pathway interactions and differences in transporter gene conservation between the two species. The regulatory role of the bZIP TF family—particularly bZIP18/

bZIP30—was highlighted as a central component in the maize GRN, influencing nitrogen assimilation and carbon metabolism. To investigate temporal expression patterns, maize and sorghum were grown under controlled hydroponic conditions with varied nitrate levels to simulate nitrogen limitation and recovery. RNA-seq analysis identified species-specific and temporal variations in differentially expressed genes (DEGs), emphasizing the importance of TF families, such as bZIP and NIN-LIKE PROTEIN, in response to nitrogen stress. This study provides critical insights into the genomic regulatory frameworks essential for plant adaptability to nitrogen fluctuations, highlighting the importance of identifying candidate genes for nitrogen uptake and utilization.

## Evaluating Nitrate Tolerance in Male Sterile Sorghum Lines for Implications for Safe Forage Production

J. Braynen, S. Kumari, M. Regulski, An. Olson, D. Ware [in collaboration with W. Rooney, Texas A&M; R. Klein, USDA-ARS, Livestock, Forage, and Pasture Management Research Unit; and N. Boerman, USDA-ARS Woodward]

Sorghum plays a crucial role as a forage crop in diverse agricultural systems, with the availability of nitrogen playing a pivotal role in crop yield and plant health. To maximize crop yields and prevent nitrogen deficiency, agricultural practices often resort to the excessive use of nitrogen-based fertilizers. Such overuse, however, raises significant environmental concerns. Excessive nitrogen, beyond the absorption capacity of plants, can leach into water systems, triggering eutrophication and causing damage to aquatic life. Additionally, the surplus nitrogen in the soil may accumulate as nitrates in plant tissues, posing a potential risk of nitrate toxicity with potentially fatal consequences for ruminant animals that graze on the affected forage. Considering this, previous research done by Ralston et al. (*Agrosystems Geosci Environ* **6:** e20403 [2023]) has uncovered variable nitrate concentrations across 20 different inbred sorghum lines to determine whether there are genotypic differences in nitrate accumulation among inbred lines. In collaboration with this study, we aim to extend these findings by examining four distinct male sterile sorghum A-lines, selected for their varied nitrate concentration profiles, in addition to BTx623, which serves as the reference genome. The experimental design involved cultivating more than 50 individuals from the four selected A-lines (ATx645, ATx3408, A.11022, and A.07258bst) and BTx623 both hydroponically and in sand-based media within a greenhouse, with groups subjected to both elevated (20 mM) and standard (1 mM) ammonium nitrate conditions to simulate high and normal nitrate environments. Our findings align with previous research, confirming ample genetic variability within a sample of the maintained heterotic group of sorghum lines to breed for leaf nitrate concentrations beneath a critical toxicity threshold of 10,000 $\mu g\ g^{-1}$, while also breeding for improved overall productivity. Thus, it ensures safety for livestock consumption, while offering a path forward for producing environmentally sustainable and nontoxic forage sorghum lines.

## Engineering Homologs Provide a Fine Scale for Quantitative Traits in Polyploid

K. Chougule, D. Ware [in collaboration with J. Heo, Y.K. Lee, and S.J. Park, Wonkwang University, Iksan, Korea]

*Solanum nigrum* is a naturally occurring hexaploid black fruit nightshade belonging to the Solanaceae family. It contains beneficial metabolites, including anthocyanins, and has great potential as a human health supplement. The species is evolutionarily close to tomato (*Solanum lycopersicum*) and has small berry fruits, making it tolerant to environmental stress, similar to ancient wild tomato species. The genetic changes during tomato domestication have been extensively studied, and de novo domestication to improve and optimize agricultural traits such as fruit yield could be achieved by genetic modification of wild species in the Solanaceae family. The current study, published in *Plant Biotechnology Journal* (Lee et al. 2023), delves into the genetic modification of *S. nigrum* and its potential for agricultural improvement through the manipulation of homoeologous genes. Polyploidy, a common occurrence in flowering plants, presents challenges in genetic modification, but recent advances in genome engineering and CRISPR-Cas9 technology have enabled progress in polyploid genome analysis and gene editing. The authors sequenced and assembled the genome of *S. nigrum* and identified homologous gene sets, which exhibit similar sequence and expression profiles. The *S. nigrum* genome sequence has 8,292 scaffolds with a total length of 1.0 Gb and an N50 of 2.8 Mb compared to the reference assembly of *Solanum americanum*, which has 100 scaffolds with a total length of 2.9 Gb and an N50 of

86.1 Mb. The Ware laboratory contributed by providing genome annotation for the *S. nigrum* genome. The annotation process identified and analyzed homoeologs in *S. nigrum*, resulting in the validation of the effects of genetic dosage on quantitative traits (QTs) and the modification of a regional crop cultivar of *S. nigrum*. Using CRISPR-Cas9-mediated mutagenesis, collaborators generated various mutation combinations in homoeologous genes, resulting in quantitative phenotypic changes in hexaploid *S. nigrum*. The results suggest that engineering homoeologous genes could be useful for the agricultural improvement of polyploid crops, such as *S. nigrum*, and offer a broadly applicable path for engineering polyploid crops. Furthermore, the study demonstrated that dosage effects on homoeologous genes can lead to a broad spectrum of quantitative genetic variation, influencing traits such as inflorescence branching, fruit yield, and shoot growth patterns.

The study also explored the paralogous compensation mechanisms of CLV3-CLE9 in *S. nigrum*, revealing the potential for genetic dosage effects and paralogous compensation to impact fruit size and yield. Through the generation of multiple mutants with varying functional gene copies, the researchers observed a range of quantitative traits that could be optimized in polyploid organisms. Importantly, the study illustrated the rapid improvement in fruit yield of Boranong, an orphan cultivar of *S. nigrum*, through CRISPR-Cas9-mediated mutagenesis. The generated mutants exhibited increased inflorescence branches, flowers, and fruits, leading to enhanced fruit yield. Collectively, the results suggest that targeted gene editing and the manipulation of homoeologous genes offer promising avenues for agricultural enhancement in polyploid crops, with the potential to address challenges associated with genetic modification in such species. These findings have significant implications for the future of crop improvement and the genetic engineering of polyploid plants, offering insights into the potential for enhancing agricultural productivity and the development of novel crop varieties with improved traits. This was published in *Plant Biotechnology Journal* (Lee et al. 2023).

## PLANT GENOME STEWARDSHIP AND INFRASTRUCTURE FOR PLANT GENOMES

The Ware laboratory's endeavors in plant genomics encompass both the sequencing and annotation of complete plant genomes and the development of cyberinfrastructure to support this research. Beginning with the generation of raw sequence data by wet laboratory scientists, the process involves a collaborative effort with computational biologists and bioinformaticians to interpret the data. This interpretation includes assembling raw sequence reads into contigs, which are then organized into scaffolds to discern their position within chromosomes. Subsequent annotation identifies genes and functional elements, aided by comparisons with other genomes to infer evolutionary relationships.

Advances in sequencing technologies and data analysis have spurred the development of new software and methodologies. High-depth, low-cost RNA sequencing provides additional evidence for genome annotation, while the latest DNA sequencing enables the construction of pan-genomes that capture genetic variation across multiple genotypic backgrounds. Projects within various plant research communities, including sorghum, maize, rice, grape, and *Arabidopsis*, are sequencing diverse genotypic backgrounds to elucidate the genetic basis of phenotypic traits and enhance crop resilience.

Simultaneously, the Ware laboratory has been instrumental in developing cyberinfrastructure projects to support plant genomics research. Initially centered around the Gramene portal, which facilitated comparative genetic studies across various grass species, the platform has evolved to host data for more than 120 plant species. Leveraging advancements in long-read sequencing, genome assembly, and annotation, the laboratory has expanded its focus to include crop-specific websites for rice, maize, grapevine, and sorghum (SorghumBase).

As genomics research scales up to encompass thousands of organisms across diverse environments, the need for robust cyberinfrastructure becomes increasingly critical. Open science platforms like Gramene and SorghumBase play a crucial role in disseminating this information, adhering to FAIR principles (findable, accessible, interoperable, reusable). The Ware laboratory, along with peer databases in the AgBioData consortium, emphasizes adherence to FAIR principles and the development of standards to maintain heterogeneous data and computational resources effectively. This commitment extends to supporting the adoption of these standards within the research community, reflecting a dedication to advancing open and collaborative science in plant genomics.

## REFERENCE ASSEMBLY AND GENE STRUCTURE ANNOTATIONS

### Era of Gapless Plant Genomes: Innovations in Sequencing and Mapping Technologies Revolutionize Genomics and Breeding

N. Gladman, K. Chougule, D. Ware [in collaboration with S. Goodwin, W.R. McCombie, CSHL: E. Cooper, University of North Carolina; Z. Brenton, Carolina Seed Systems; C. Rustenholz, INRAE—University of Strasbourg]

Advances in sequencing technologies, particularly long-read sequencing, have made gapless telomere-to-telomere (T2T) genome assemblies possible for plant genomes. Gladman et al. (2023) highlights the use of hybrid technologies that combine short- and long-read single-molecule sequencing instruments to generate T2T references for all chromosomes within simple or complex genomes. Long-read sequencing technologies, such as Pacific Biosciences (PacBio) SMRT cell technology and Oxford Nanopore MinION, have enabled the sequencing of long DNA fragments, overcoming the limitations of short-read methods in assembling long, complex regions of the genome. These long-read technologies can sequence through complex regions of the genome, provide unambiguous alignments to repetitive regions, and ultimately create complete, contiguous genome assemblies. Additionally, the incorporation of long-read sequencing has resolved significant issues plaguing larger, complex genomes, such as generating nonambiguous reads and contigs that are long enough to span structural variants, repetitive regions, and transposable elements (TEs). These advancements have significantly improved the quality and accuracy of plant genome assemblies, providing researchers with a more comprehensive and reliable reference for plant genomics and molecular experimentation.

### Ten New High-Quality Genome Assemblies for Diverse Bioenergy Sorghum Genotypes

K. Chougule, Z. Lu, An. Olson, D. Ware [in collaboration with W. Voelker, K. Krishnan, L. Alexander, Jr., K. Songsomboon, C. Ponce, E. Cooper, University of North Carolina; Z. Brenton, Carolina Seed Systems; L. Boatwright, Clemson University]

Sorghum is a versatile and widely grown cereal crop valued for its efficiency, drought tolerance, and ability to grow in marginalized soils. It exhibits extensive genetic, phenotypic, morphological, and physiological diversity, making it valuable as a sustainable, fast-growing, and high-yielding bioenergy crop. Sorghum is classified into four major ideotypes—grain, sweet, cellulosic, and forage—all of which can be used in different bioenergy production methods. Understanding the genomic changes driving traits related to yield, carbon partitioning, and local adaptation is essential to fully capitalize on their potential. The study published aims to understand genomic mechanisms underlying complex traits related to yield, composition, and environmental adaptations in sorghum, a valuable bioenergy feedstock. It presents high-quality de novo genome assemblies for 10 diverse bioenergy sorghum genotypes, demonstrating similar levels of high contiguity and assembly sizes close to the expected reference genome size (Voelker et al., *Front Plant Sci* **13:** 1040909 [2022]). The assemblies contain >90% of known BTx623 genes, indicating their comprehensive coverage and accuracy. It also identified more than 24,000 large structural variants and more than 10.5 million SNPs highlighting structural variants and nonsynonymous SNPs enriched in different gene categories, emphasizing the need for long-read sequencing in crop species to identify novel variation. The research provides insights into the genetic differences between sweet and cellulosic genotypes, offering valuable resources for future mapping and trait discovery for sorghum and its diverse uses, including food, feed, and bioenergy. The gene annotation process involved building a pan-gene working set using representative pan-gene models selected from a comparative analysis of gene family trees from 18 sorghum genomes (Fig. 6). The gene structures were updated with available transcriptome evidence from BTx623 using PASA, and additional improvements to structural annotations were made using full-length sequenced cDNAs and sorghum ESTs. On average, approximately 55,000 working sets of models were generated for each sorghum line, with more than half of the protein-coding models mapping to a BTx623 reference gene. Functional domain identification was completed with InterProScan, and TRaCE was used to assign canonical transcripts based on domain coverage, protein length, and similarity to transcripts assembled by StringTie. Finally, the protein-coding annotations were imported to Ensembl core databases, verified, and validated for translation using the Ensembl API.

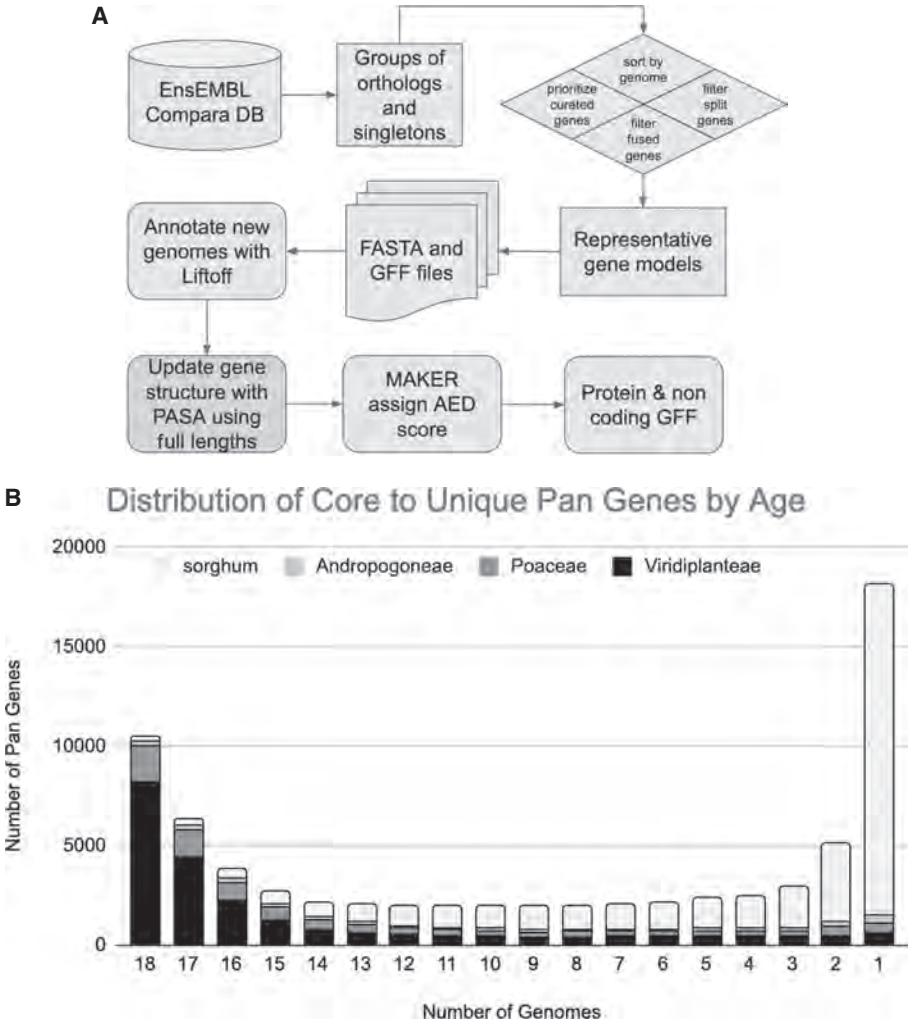



**Figure 6.** Sorghum pan gene. (*A*) Building pan gene index workflow. (*B*) Pan gene set growth by age: 72,000 pan gene sets from 18 sorghum genomes. Core genes remain stable, and pan gene set size increase is contributed by line-specific genes.

## An Improved Reference of the Grapevine Genome Reasserts the Origin of the PN40024 Highly Homozygous Genotype

C. Kim, M.K. Tello-Ruiz, D. Ware [in collaboration with A. Velt, S. Blanc, É. Duchêne, V. Dumas, P. Hugueney, M. Lahaye, and C. Rustenholz, SVQV, INRAE—University of Strasbourg; B. Frommer and D. Holtgräwe, Bielefeld University; J. Grimplet, Centro de Investigación y Tecnología Agroalimentaria de Aragón (CITA), Unidad de Hortofruticultura; J.T. Matus, D. Navarro-Payá, and L. Orduña, Universitat de València-CSIC; N. Vitulo, Università degli Studi di Verona]

The cultivated grapevine (*Vitis vinifera* ssp. vinifera) has a highly homozygous genotype known as PN40024, which serves as a reference for grapevine studies. The genome of PN40024 was generated through nine rounds of selfing, resulting in a nearly homozygous genome with several heterozygous regions. The genome assembly, PN12X.v2, was found to be fragmented and only represented the haploid state of the genome with mixed haplotypes. An improved version of the reference, PN40024.v4, was created by incorporating long genomic sequencing reads, resulting in increased continuity of the scaffolds and a reduction in *N* bases (Velt et al. 2023). The improvements made in the PN40024 genome assembly in the PN40024.v4 version include:

- Increased continuity: The PN40024.v4 assembly has a reduced number of scaffolds and a higher N50 size, indicating increased continuity and informativeness compared with PN12X.v2.
- Reduced unknown bases: The PN40024.v4 assembly has a significant reduction in the number of unknown bases (*N* bases) compared with PN12X.v2, resulting in a more complete and accurate structure.
- Enhanced gene annotation: The gene annotation in PN40024.v4 has been improved, resulting in a reliable estimation of 35,230 genes, and it outperforms the previous VCost.v3 annotation.
- Improved read mapping: The percentage of assigned reads aligned under an annotated gene was 2.4%–3% better with PN40024.v4 compared with PN12X.v2, confirming the improved quality of gene annotation and allowing identification of more differentially expressed genes.
- Reduced errors: Despite an increase in error rate due to the use of long genomic reads, the overall accuracy of the PN40024.v4 assembly remains high, with a base level quality value (QV) estimated at 36.02 and an accuracy of 99.999749801%.

The study also highlights the fact that the origin of the PN40024 genotype was determined to have resulted from nine selfings of the "Helfensteiner" cultivar instead of a single "Pinot noir." This determination was made through an analysis of the genome, which revealed that no blocks of homozygous variants could be identified, indicating that one of the two "Helfensteiner" haplotypes is always present in the PN40024 genome. This confirmation supports the conclusion that the "Helfensteiner" variety is the true parent of the first selfing, from which the PN40024 genotype was created after eight more selfings. This information was obtained through the analysis of homozygous SNP densities along the chromosomes, which allowed for the definition of haplotypic blocks and the determination of the proportion of "Pinot noir," "Schiava grossa," and common haplotypes. The Ware laboratory participated in a manual gene curation process involving the curation of 1,641 genes, with 1,579 being edited and 62 deleted. This curation was carried out using a purpose-built Apollo server, which provided a wide range of transcriptomic and genomic data for the PN40024.v4 genome. The focus was on preserving previous VCost.v3 manual curation and functional annotation efforts, particularly for genes present in the reference catalog. The resulting automated annotation, including the manually curated features, was designated as PN40024.v4.2. This version showed improved metrics, with fragmented plant core genes reduced to six and missing genes to eight, indicating 99.2% complete plant core genes. The PN40024.v4.2 gene models were considered the most updated resource for transcriptomics and functional enrichment analyses, providing an efficient resource for exploring heterozygous genetic traits. Despite these improvements, it suggests that further enhancements should focus on regions with haplotype switching and newly introduced errors due to the implementation of long genomic reads. These advances will help maintain the PN40024 genome as a gold-standard reference and contribute to eventually building a grapevine pan-genome (i.e., the entire set of genes from all grapevine varieties).

## Developing Genomic Resources to Support Crop Breeding with a Focus on Sorghum

K. Chougule, N. Gladman, V. Kumar, Z. Lu, An. Olson, M.K. Tello-Ruiz, S. Wei, D. Ware [in collaboration with G. Burrow, H. Cuevas, M. Harrison, C. Hayes, R. Klein, USDA ARS; W. Rooney, TAMU; B. Kaufman, Agriplex Inc.]

Today's plant breeding relies on advanced sequencing technologies and efficient phenotyping methods to gather extensive genetic and phenotypic data. This shift from traditional breeding to data-driven strategies has revolutionized crop improvement efforts. The Ware laboratory has partnered closely with researchers, breeders, and germplasm centers to enhance characterization and access to crop varieties for sustainable agriculture.

In the past year, the Ware laboratory focused on sorghum breeding projects, collaborating to establish a domestic sorghum community marker panel (CMP). The objective of the panel was support for standardized markers for precise genotyping, aiding in various breeding strategies such as genomic selection and diversity studies. Working with the Crop Germplasm Committee (CGC), a Sorghum CGC Genotyping Panel working group was formed to address community needs effectively. Approximately 3,500 candidate markers were identified based on existing community panels and feedback from the working group. More than 15 organizations, including universities, seed companies, and multiple USDA sites, participated by

submitting more than 25 94-well leaf tissue plates as per the sample preparation protocols provided by the genotyping service provider. After multiple rounds of genotyping tests and assay validations, approximately 2,400 markers were selected for the panel.

Standardizing genetic information is crucial for seamless data sharing and collaboration. We have engaged with organizations like the European Variation Archive (EVA) to promote standards for genetic variation representation. The Reference SNP cluster ID (rsID) serves as a unique identifier for a group of genetic variations (GVs) colocated at a specific position in the genome, facilitating standardized referencing across databases, studies, and publications. This allows GVs to be identified by rsIDs across multiple genomes, thus consolidating scattered GV knowledge, enhancing phenotype prediction, and boosting trait-based genetic marker discovery. Gramene and SorghumBase are integrating rsIDs into their databases, linking them to quantitative trait loci (QTLs), phenotype, and germplasm data.

In the last year we actively partnered with USDA Plant Genetic Conservation Unit and the Germplasm Resources Information Network (GRIN) to enhance the accessibility and curation of germplasm collections. We are genotyping approximately 600 accessions from the sorghum collections. Reference assemblies play a vital role in genomic studies and marker development.

## Gramene: Comparative Genomic Resource for Plants

K. Chougule, V. Kumar, S. Kumari, Z. Lu, An. Olson, Au. Olson, M.K. Tello-Ruiz, S. Wei, D. Ware [in collaboration with P. Jaiswal, Oregon State University; S. Dyer, G. Namati, B. Contreras-Moreira, and I. Papatheodorou, EMBL-European Bioinformatics Institute; L. Stein, Ontario Institute of Cancer Research; R. Wing, University of Arizona; N. Provart, University of Toronto]

The Gramene project provides online reference resources for plant genomes and curated pathways to aid functional genomics research in crops and model plant species. Our website (www.gramene.org) facilitates studies of gene function by combining genome and pathway annotation with experimental data and cross-species comparisons. In other words, the data and tools in Gramene enable plant researchers to use knowledge about gene function in one species to predict gene function in other species. Drawing these connections facilitates translational research in plant development and physiology that influences economically important traits—for example, grain development, flowering time, drought tolerance, and disease resistance. In 2023, the project accomplished several major milestones, culminating in our 68th data release (November 2023), which included 150 plant genomes. The Gramene team provides curation and mirroring of the Ensembl Plants project at the European Bioinformatics Institute (EMBL-EBI), and collaborates closely with the EBI's Expression Atlas project to provide manually curated, quality-controlled, and analyzed transcriptomic data, and also collaborates with the Plant Reactome pathway resource at Oregon State University (OSU) and Ontario Institute for Cancer Research (OICR). We continue to host genome and pathway annotations via the Ensembl genome browser and the Plant Reactome pathways portal.

The Gramene project is actively engaging the community through various channels including webinars, presentations, talks, posters, and demonstrations during major community events including Plant Biology, providing training and the community's feedback on our current tools and user suggestions for new functionality. In 2023, a major focus has been the continued development of pan-genome resources while closely working with research and breeding initiatives in maize, sorghum, rice, and grapevine communities. Although generating a reference assembly has progressively become easier, there are still major challenges to accurately predicting the functional features in the genomes. To this end, a major effort has been put forth to benchmark the latest sequencing platforms and to scale existing algorithms to support improved and consistent gene annotation predictions. These automated predictions are the first step, and there remains a need for human review of the models. We conducted a gene annotation jamboree with students during the summer at which the students reviewed genes in Sorghum BTx623 v3 and v5 annotation for agronomically important genes. In total the students reviewed 13,782 sorghum genes in the Gramene gene curation tool. Among them, approximately 700 genes were identified for further curation, including 258 structural adjustments made using evidence tracks within the Apollo genome annotation web-based editor.

In 2023, Gramene had two releases—R67 (August) and R68 (November)—which represented an increase of 22 genomes. With a total of 110 plant genomes, we constructed more than 156,000 protein-coding gene

family trees using the peptide encoded by the canonical transcript (i.e., a representative transcript for a given gene for each of 3,867,457 individual genes). R68 also included a total of 80 synteny maps and 279 pairwise genomic DNA alignments. The majority of species in Gramene were aligned with *Oryza sativa* Japonica (reference monocot) and *A. thaliana* (reference dicot), plus intervariety alignments for rice and wheat, for each of which there are more than 10 varieties in Gramene. Split gene predictions, a common error in gene annotations in which a gene is split into two or more predicted genes due to lack of transcript evidence joining them, were updated twice in the past year. In addition, we integrated gene functional annotations from the National Center for Biotechnology Information (NCBI)'s geneRIFs and the Rice Annotation Project's Database (rap-db)'s curated genes. Each of these curated sources provided a mapping between genes and publications, which are available from the Publications tab of the search results. A total of 13,529 gene models now have associated publications describing their function. New variation data for lettuce and cocoa were imported from the European Variation Archive (not yet available via Gramene Mart). In release 67, we introduced the Oryza CLIMtools portal that brings interactive web-based views of the environment by genome associations and RiboSNitch prediction for 658 Indica and 283 Japonica landraces. The gene search interface had three improvements: (1) integrating views of gene expression data from the Bio-Analytic Resource (BAR), (2) improving information on the closest annotated homolog by prioritizing functionally annotated genes, and (3) providing direct links to our crop-specific pangenome sites for maize, rice, grapevine, and sorghum genes. These pangenome sites provide access to additional genomes from the respective taxonomic clade for comparative analysis against the primary annotated reference genome representative.

Release #68 features 22 new reference plant genomes (including 15 rice cultivars, the IWGSC RefSeq v2.1 version of the bread wheat assembly plus two new wheat cultivars, two oat genomes, and the ash tree genome) for a total of 150 plant reference genomes. In addition, the poplar assembly was updated from version 3 to 4, and teff was replaced with a new assembly. Functional annotations were updated for Japonica rice genes. This release features 156,836 gene family trees built from 3,867,457 protein-coding genes spanning 110 of the plant genomes. Functional annotations were updated for Japonica rice genes.

Three new features in the search interface are (1) for genes in trees, the closest annotated homolog (or model species homolog if the former was not found) is provided with percent identity; (2) for rice, maize, grapevine, and sorghum genes, a link for such gene in the corresponding pangenome site (i.e., Gramene Oryza, Gramene Maize, Gramene Vitis, and SorghumBase, respectively) is also provided (RNA-seq-based gene expression levels for curated *Arabidopsis*, maize, sorghum, and soybean genes in different expression data sets are now available through embedded eFP Browsers from the BAR for Plant Biology at the University of Toronto at http://bar.utoronto.ca]) (Fig. 7); and (3) a new form enabling users to submit a gene's function was also embedded in the "Papers" tab of the search interface.

There were two releases of Gramene Oryza (https://oryza.gramene.org), R6 in January and R7 in August; one release of Gramene Maize (https://maize-pangenome.gramene.org) with eight new genomes in March; one release of SorghumBase (https://sorghumbase.org) with eight new genomes in June; and no releases of Gramene Grapevine (https://vitis.gramene.org) in 2023.

## SorghumBase: A Web-Based Portal for Sorghum Genetic Information and Community Advancement

K. Chougule, N. Gladman, V. Kumar, S. Kumari, Z. Lu, A. Olson, M.K. Tello-Ruiz, P. Van Buren, S. Wei, D. Ware [in collaboration with J. Burke, J. Chen, G. Burow, C. Hayes, Y. Emendack, and Z. Xin, USDA-ARS]

Many agriculturally important crops have dedicated web resources for their genomic and phenotypic data (MaizeGDB, SoyBase, GrainGenes, T3: Triticeae toolbox). To address this deficit for *S. bicolor*, we created SorghumBase.org, a modular web-based utility that houses genetic, genomic, pan-genomic, transcriptomic, and molecular biology data for more than a dozen different important sorghum cultivars. The currently active SorghumBase site contains data, external links, and visualizations for gene trees, syntenic regions, gene expression, molecular pathways, SNP variants, and sequence data for the primary reference genome BTx623.

In 2023, SorghumBase had one data release: R6 (June), which included the BTx623 v5.1, totaling 29 sorghum genomes and eight plant outspecies. With these 37 plant genomes, we constructed 42,926 protein-coding gene family trees with 1,233,135

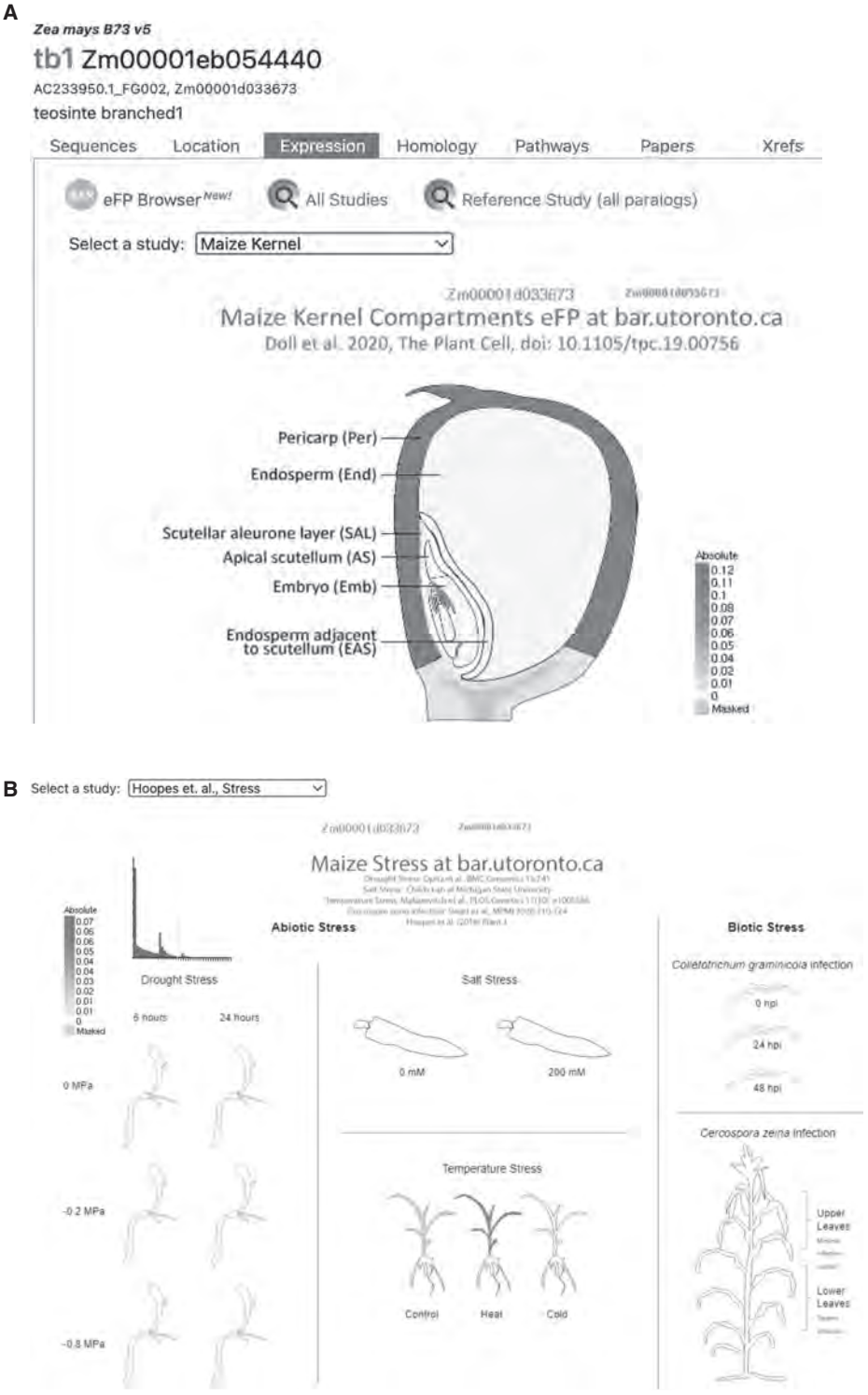



**Figure 7.** Screenshot of Maize eFP showing the visual display of the teosinte branched1 (tb1) Zm00001eb054440 gene expression levels as an example. It is based on the eFP images on the RNA-seq data of (*A*) development of maize kernel compartments and (*B*) under different abiotic (drought stress, salt stress, and temperature stress) and biotic (infection due to *Colletotrichum gramminicola* and *Crecospora zeina*) maize stress conditions. Color gradient in pictograph is based on tb1 genes' reads per kilobase per million (RPKM) expression value.

individual genes. This release also includes 8.6 million EMS-induced mutations mapped to the BTx623 v3 reference and 32.5 million SNPs in 400 Sorghum Association Panel (SAP) lines mapped to Tx2783. These additions contribute to a total of 1.6 million loss-of-function mutations. We continued to host nearly 6,000 QTLs for 223 sorghum traits and more than 150 studies from the Sorghum QTL Atlas.

We have continued to host gene expression and orthology-based pathway projections for the reference genome *S. bicolor* BTx623. We manually curated 11 sorghum bulk RNAseq studies (eight baseline and three differential gene expression), validated, configured, and submitted to further process the data at EMBL-EBI Expression Atlas. The processed baseline expression data was imported into SorghumBase. The orthology-based sorghum pathways were projected from curated Japonica rice pathways by Gramene's Plant Reactome.

In addition, we have provided a new feature called "eFP browser" under the expression tab in SorghumBase. eFP browser is developed by the Bio-Analytic Resource for Plant Biology (BAR) at the University of Toronto (http://bar.utoronto.ca) and enables researchers to explore gene expression across different plant species. The eFP browser using developmental and stress expression studies data of sorghum provides color-coded expression levels across different plant parts for any gene of interest in sorghum, thus providing a novel resource to study gene expression and function in SorghumBase.

We continued to expand our database of relevant and recent sorghum research papers by performing systematic monthly reviews of the PubMed resource. Currently, a total of 934 publications have been reviewed and indexed in the SorghumBase database so they could be matched via its integrated search interface when searching by author, keywords, PubMed ID, or any other text within the title or abstract. Publications in the database could then be linked to research highlights ($n = 81$), news items ($n = 129$), and funding sources. News items included detailed highlights of participation in conferences such as PAG and the Maize Genomics Conference. We also created an Events section on the website to list the latest webinars, conferences, meetings, etc. that are relevant to the sorghum research community. Additionally, a SorghumBase training workshop was hosted at CSHL by the Ware laboratory in December.

## PUBLICATIONS

Deng CH, Naithani S, Kumari S, Cobo-Simón I, Quezada-Rodríguez EH, Skrabisova M, Gladman N, Correll MJ, Sikiru AB, Afuwape OO, et al. 2023. Genotype and phenotype data standardization, utilization and integration in the big data era for agricultural sciences. *Database* doi:10.1093/database/baad088

Fahlgren N, Kapoor M, Yordanova G, Papatheodorou I, Waese J, Cole B, Harrison P, Ware D, Tickle T, Paten B, et al. 2023. Toward a data infrastructure for the Plant Cell Atlas. *Plant Physiol* **191:** 35–46.

Gladman N, Goodwin S, Chougule K, McCombie RW, Ware D. 2023. Era of gapless plant genomes: innovations in sequencing and mapping technologies revolutionize genomics and breeding. *Curr Opin Biotechnol* **79:** 102886.

Lee ES, Heo J, Bang WY, Chougule KM, Waminal NE, Hong NT, Kim MJ, Beak HK, Kim YJ, Priatama RA, et al. Engineering homoeologs provide a fine scale for quantitative traits in polyploid. *Plant Biotechnol J* **21:** 2458–2472.

Velt A, Frommer B, Blanc S, Holtgräwe D, Duchêne É, Dumas V, Grimplet J, Hugueney P, Kim C, Lahaye M, et al. 2023. An improved reference of the grapevine genome reasserts the origin of the PN40024 highly homozygous genotype. *G3* **13:** kjad067. doi:10.1093/g3journal/jkad067

Xin Z, Jiao Y, Burow G, Hayes C, Chen J, Burke J, Pugh NA, Ware D. 2023. Registration of 252 sequenced sorghum mutants as a community reverse genetic resource. *J Plant Regist* **17:** 599–604.

Zhang L, Braynen J, Fahey A, Chopra K, Cifani P, Tadesse D, Regulski M, Hu F, van Dam HJJ, Xie M, et al. Two related families of metal transferases, ZNG1 and ZNG2, are involved in acclimation to poor Zn nutrition in *Arabidopsis*. *Front Plant Sci* **14:** 1237722.

Zhou Y, Yu Z, Chebotarov D, Chougule K, Lu Z, Rivera LF, Kathiresan N, Al-Bader N, Mohammed N, Alsantely A, et al. Pan-genome inversion index reveals evolutionary insights into the subpopulation structure of Asian rice. *Nat Commun* **14:** 1567.

### *In Press*

George N, Fexova S, Fuentes AM, Madrigal P, Bi Y, Iqbal H, Kumbham U, Nolte NF, Zhao L, Thanki AS, et al. 2024. Expression Atlas update: insights from sequencing data at both bulk and single cell level. *Nucleic Acids Res* **52:** D107–D114.

Jiao Y, Nigam D, Barry K, Daum C, Yoshinaga Y, Lipzen A, Khan A, Parasa S-P, Wei S, Lu Z, et al. 2024. A large sequenced mutant library—valuable reverse genetic resource that covers 98% of sorghum genes. *Plant J* **117:** 1543–1557.

Lee YK, Olson An, Kim K, Ohme-Takagi M, Ware D. 2024. HB31 and HB21 regulate floral architecture through miRNA396/GRF modules in *Arabidopsis*. *Plant Biotechnol Rep* **18:** 45–55.

Zhou Y, Kathiresan N, Yu Z, Rivera LF, Yang Y, Thimma M, Manickam K, Chebotarov D, Mauleon R, Chougule K, et al. 2024. A high-performance computational workflow to accelerate GATK SNP detection across a 25-genome dataset. *BMC Biol* **22:** 13.

# QUANTITATIVE BIOLOGY

**Ivan Iossifov**'s laboratory focuses on the development of new methods and tools for genomic sequence analysis and for building and using molecular networks, and applies them to specific biomedical problems. They study the genetics of common diseases in humans using two main tools: next-generation sequencing and molecular networks representing functional relationships among genetic loci. These approaches in combination enable the kind of large-scale studies necessary for furthering our understanding of the complex etiology of disorders such as autism, bipolar disorder, and cancer.

**Justin Kinney**'s laboratory focuses on developing next-generation DNA sequencing as a tool for dissecting the biophysical mechanisms of gene regulation. As a graduate student, Kinney coinvented a widely used technique now known as the massively parallel reporter assay (MPRA). Kinney and colleagues further showed how, using ideas from information theory, such experiments could be used to infer quantitative biophysical models for how cells regulate gene expression. The Kinney laboratory continues to leverage a tightly knit combination of mathematical theory, machine learning, and experiments in order to illuminate the biophysics of gene regulation in two diverse contexts: bacterial transcriptional regulation and alternative mRNA splicing in humans. This latter context is highly relevant to understanding and treating human diseases like spinal muscular atrophy and cancer. The Kinney laboratory also develops algorithms and software for the analysis of MPRAs and other multiplex assays of variant effect (MAVEs).

The **Peter Koo** group develops methods to interpret high-performing deep-learning models to distill knowledge that they learn from big and noisy biological sequence data. Deep learning is being applied rapidly in many areas of genomics, demonstrating improved performance over previous methods on benchmark data sets. Despite the promise of deep learning, it remains unclear whether improved predictions will translate to new biological discoveries; this is due to these predictions' low interpretability, which has earned them a reputation as a black box. Understanding the reasons underlying a deep-learning model's prediction may reveal new biological insights not captured by previous methods. The Koo laboratory's goal is to elucidate biological mechanisms that underlie sequence–function relationships for gene regulation and protein (dys)function. Recently, they have teamed up with other members of the CSHL Cancer Center to investigate the sequence basis of epigenomic differences across healthy and cancer cells.

**Alexander Krasnitz** and colleagues develop mathematical and statistical tools to investigate population structure of cells constituting a malignant tumor and to reconstruct evolutionary processes leading up to that structure. These tools are designed to make optimal use of emerging molecular technologies, chief among them high-throughput genomic profiling of multiple individual cells harvested from a tumor. By analyzing these profiles, Krasnitz and colleagues derive novel molecular measures of malignancy, such as the number of aggressive clones in a tumor, the invasive capacity of each clone, and the amount of cancer-related genetic alteration sustained by clonal cells. Krasnitz and colleagues collaborate closely with clinical oncologists to explore the utility of such measures for earlier detection of cancer, more accurate patient outcome prediction and risk assessment, and better-informed choice of treatment options.

There is increasing evidence that rare and unique mutations have a significant role in the etiology of many diseases such as autism, congenital heart disease, and cancer. **Dan Levy**'s group develops algorithms to identify these mutations from large, high-throughput data sets comprising thousands of nuclear families. After earlier working with high-resolution comparative genomic

hybridization (CGH) arrays, Levy's group now uses targeted sequence data. Levy's group has developed methods for identifying de novo mutations (i.e., those seen in a child but not in his or her parents) by simultaneously genotyping the entire family; the team is currently focused on building algorithms to detect copy number variants and multiscale genomic rearrangements. Although their copy number methods are based on "read" density, there are classes of mutations that require analysis at the level of the read. Thus, they are developing algorithms to identify insertions, deletions, inversions, transpositions, and other complex events. Other projects in the Levy laboratory include analysis of single-cell RNA, phylogenetic reconstruction from sparse data sets, and disentangling haplotypes from sperm and subgenomic sequence data.

The **David McCandlish** laboratory develops computational and mathematical tools to analyze and exploit data from high-throughput functional assays. The current focus of the laboratory is on analyzing data from so-called "deep mutational scanning" experiments. These experiments simultaneously determine, for a single protein, the functional effects of thousands of mutations. By aggregating information across the proteins assayed using this technique, they seek to develop data-driven insights into basic protein biology, improved models of molecular evolution, and more accurate methods for predicting the functional effects of mutations in human genome sequences.

Critically, these data also show that the functional effects of mutations often depend on which other mutations are present in the sequence. McCandlish and colleagues are developing new techniques in statistics and machine learning to infer and interpret the complex patterns of genetic interaction observed in these experiments. Their ultimate goal is to be able to model these sequence–function relationships with sufficient accuracy to guide the construction of a new generation of designed enzymes and drugs, and to be able to predict the evolution of drug resistance phenotypes in populations of cancer cells and in rapidly evolving microbial pathogens.

The thymus generates and selects a highly variable yet specific T-cell repertoire that discriminates between healthy and nonhealthy self and dangerous nonself antigens. **Hannah Meyer**'s research group uses a systems immunology approach to dissect the mechanisms crucial to the selection processes in the thymus. Her team develops experimental techniques and combines the resulting data with innovative computational models to generate accurate and testable hypotheses about tissue-level organ physiology.

Studying thymus physiology from a qualitative and quantitative perspective will provide a more fine-grained understanding of the selection processes and their downstream consequences, such as autoimmunity, cancer immunosurveillance, and immune deficiency.

**Saket Navlakha**'s laboratory studies "algorithms in nature" (i.e., how collections of molecules, cells, and organisms process information and solve interesting computational problems critical for survival). Indeed, there are many shared goals and constraints faced by biological and engineered systems, including (1) the use of distributed networks as a backbone for information processing and communication; (2) trade-offs between optimization criteria, including efficiency, robustness, and adaptability; and (3) the need to develop low-cost, scalable solutions that conserve important metabolic or physical resources. An algorithmic perspective on biological problem-solving can lead to two ends: (1) new biological algorithms that are simple, flexible, and robust for use in computer science applications, and (2) quantitative frameworks to predict behavior, raise testable hypotheses, and guide experiments. The Navlakha laboratory has most recently focused on studying neural circuit computation and plant architecture optimization from this perspective.

Modern genomic technologies make it relatively easy to generate rich data sets describing genome sequences, RNA expression, chromatin states, and many other aspects of the storage, transmission, and expression of genetic information. **Adam Siepel**'s group focuses on a diverse collection

of research questions in this interdisciplinary area, spanning applications in cancer biology, basic molecular biology, evolutionary genetics, infectious diseases, and agriculture. For many problems in genetics today, the limiting step is no longer in data generation, but in integrating, interpreting, and understanding the available data. Addressing these challenges requires expertise both in the practical arts of data analysis and in the theoretical underpinnings of statistics, computer science, and genetics.

Over the years, the Siepel group's research has touched on topics including the identification of recombinant strains of HIV, the discovery of new human genes, the characterization of conserved regulatory elements in mammalian genomes, the identification of noncoding mutations important in cancer, and the discovery of ancient gene flow from humans to Neandertals. A general theme in their work is the development of precise mathematical models for the complex processes by which genomes evolve over time, and the use of these models, together with techniques from computer science and statistics, both to peer into the past and to address questions of practical importance for human health. They collaborate closely with experimentalists in cancer biology, transcriptional regulation, plant breeding, and many other areas.

# GENETIC VARIANTS LINKED TO AUTISM TRAITS

**I. Iossifov** Y-h. Lee B. Yamrom
S. Marks C. Yoon

The bulk of the work in our laboratory was in analyzing the large data set of whole-genome sequencing (WGS) data generated from approximately 2,400 of the Simons Simplex Collection (SSC) families and approximately 900 families from the Autism Genetic Resource Exchange (AGRE), a collection of families that have multiple children with autism. We also started the analysis of the whole-exome sequencing (WES) data from the growing Simons Foundation Powering Autism Research (SPARK) collection: Data for approximately 20,000 of the SPARK families have been released and SPARK is expected to grow to approximately 70,000 families in a couple of years. In addition, we initiated a research program to explore the potential of RNA sequencing in family collections like SSC. These data are a rich resource that we utilize in numerous projects.

Below, the abstracts of four projects we finalized in 2023 or are still working on are listed. These projects demonstrate our current efforts in studying the role of de novo noncoding variants, rare structural rearrangements, and common variants in autism's etiology. In 2023, two publications were published, demonstrating our collaborative efforts with the Tuuli Lappalainen and Neville Sanjana groups at NYGC (Einson et al. 2023; Shi et al. 2023).

## A Platform for Access and Analysis of Genetic Variants in Phenotype-Rich Family Collections

In recent years, we have witnessed impressive success in autism genetics by analyzing sequence data sets generated from deeply phenotyped family collections like SSC. However, to develop effective treatment and diagnostic strategies, an enormous amount of work must follow this early success. Future research projects will detail the effects of hundreds of genetic variants and genes at molecular, cellular, and organismic levels. Such projects will benefit from the accumulated data sets, but their large and complex structures significantly hinder their efficient use.

We addressed such difficulties by building the Genotype and Phenotype in Families (GPF) system (Chorbadjiev et al. 2024), which enabled intuitive and straightforward interaction with such data sets. The system allowed our group to integrate and analyze diverse phenotypic and genotypic data and efficiently distribute valuable resources to the broader scientific community through the GPF's intuitive web interface. We developed the GPF as an open-source project (https://github.com/iossifovlab/gpf) so different groups could use it to operate on their data. We also deployed the GPF system at the Simons Foundation (https://gpf.sfari.org), managing several large genotypic and phenotypic data sets built through the Simons Foundation Autism Research Initiative's (SFARI's) support. These data sets include the phenotypic data from the SSC, Simons Searchlight, SPARK, and AGRE and the genotypic data from the same collections generated through whole-exome, whole-genome sequencing, and high-density chip hybridization. The system currently handles genotypes derived from approximately 10,000 individuals with whole-genome and approximately 150,000 individuals with whole-exome data. The system's distributed architecture allows us to plan deployments that manage genotypes for one million whole-genome samples.

## RNAseq of SSC

Our analysis of the whole-genome data from the approximately 2,000 quad families from the SSC demonstrated convincingly that de novo variants in the noncoding regions contribute to the incidence of autism (Yoon et al., *Commun Biol* **4:** 1026 [2021]). Specifically, we observed a significantly increased rate of de novo intronic indels in affected children relative to their unaffected siblings when we restricted the rate observation to the autism genes previously implicated by whole-exome sequencing. The increase in the rate is consistent with de novo intronic indels contributing to ~5% of the autism diagnosis.

We do not observe a similar increase in the de novo intronic substitutions rate. Still, the study's size is insufficient to detect that signal given the much higher background noise rate for substitutions. Nevertheless, we expect that de novo intronic substitutions have a contribution of similar magnitude to the de novo intronic indels' contribution. Others have reported an increased de novo mutation rate in affected versus unaffected children within the intergenic space's control regions. We expect that the noncoding de novo mutation's contribution is close to 13%, perhaps only slightly less than the contribution from de novo coding mutation.

Despite the significant contribution of the noncoding variants, we have no suitable purely analytic method to distinguish the specific causal sequence variation from the many random ones. We proposed to address that through a study of the RNA. We expect that for most of the causal de novo noncoding variants, the immediate effect would be on the expression of nearby genes. We can detect such changes in expression through RNAseq data by comparing the expression of the affected gene allele to that of the unaffected allele, a method called allele-specific expression (ASE), or by identifying abnormal splicing patterns. We performed a pilot project to test this approach's feasibility using deep RNAseq profiles of the lymphoblastoid cell lines (LCLs) from 202 individuals from 48 of the SSC families. In close collaboration with Kristin Baldwin (Genetics & Development at Columbia University) and Michael Wigler, we also demonstrated that we could transform LCLs into induced pluripotent stem cells (iPSCs) and the iPSCs further into induced neurons (iNs). The iNs express nearly 90% of the autism genes identified by exome sequencing, whereas the LCLs express ~70% of these genes. Moreover, when both cell types express a gene, the allele-specific expression measures are usually similar.

The pilots' results were encouraging enough to help us convince the Simons Foundation to fund a large-scale project for generating 4,000 RNA profiles from the LCLs of all the SSC children. The project produced approximately 100 de novo noncoding variants associated with the nearby gene's perturbed transcription. We are currently trying to fund a validation effort to use CRISPR and cell transformation and determine whether these candidate variants are responsible for the observed transcriptional abnormality and whether the abnormal transcription is preserved in the relevant iN cells.

## Genome Sharing in Siblings Concordant or Discordant for Autism

We developed a method to measure the extent to which siblings, concordant and discordant for autism, share their parental genomes using genotypes derived from whole-genome data or chips. We applied the method to approximately 1,300 pairs of concordant and approximately 4,500 pairs of discordant siblings from the SSC, AGRE, and SPARK collections. Surprisingly, we observed that the fathers have an increased sharing consistent with them carrying causal determinants in at least half of the multiplex families—more than the families in which the mother has a causal variant. With lesser significance, the discordant siblings from the simplex families corroborate the observation: The antisharing tended to be stronger for the paternal genome. The more extensive sharing of paternal than maternal genomes contradicted our expectations that mothers would be the primary source of damaging variants in the high-risk multiplex families and forced us to rethink our unified hypothesis of the genetic contribution to autism.

This work is a collaboration with Michael Wigler, Kenny Ye, and Dan Levy from CSHL and Abba Krieger and Andreas Buja from the Department of Statistics at Wharton, University of Pennsylvania. Our first manuscript (Wroten et al. 2023) on the topic was published in early 2023. The manuscript describes the method, its first application, the surprising observation, and our hypotheses for what could cause it. One possible explanation is that fathers may carry strong protective alleles; another is that maternal–fetal immune incompatibility may cause autism.

## Exploring Maternal–Fetal Incompatibility

Our observation of paternal transmission in multiplex families suggests that maternal–fetal incompatibility caused by paternal alleles' expression in the fetus may contribute to autism. This immunological hypothesis is compatible with unexplained features of autism:

- Sibling risk diminishes over time.
- Increased risk in prematurely born offspring.
- Lack of multigenerational pedigrees.
- Higher prevalence in men.

Others have reported that maternal–fetal immune incompatibility contributes to the risk of neurological conditions, such as neuropsychiatric diseases and learning disabilities. In principle, the corresponding immune reaction could be monitored and, importantly, mitigated by intervention.

Last year, we initiated the exploration of this hypothesis following two separate paths. In an effort led by Tobias Janowitz, we are attempting to develop mouse models of maternal–fetal incompatibility. We will use a combination of in- and interbreeding of inbred mouse strains and immune modulation in the mothers hoping to demonstrate that such incompatibility can lead to abnormal neurobiological development and function. Second, we apply computation approaches to identify traces of maternal–fetal incompatibility using the deep whole-genome data from the SSC. We use the mothers' human leukocyte antigen (HLA) types and missense variants carried by the fathers, but not by their spouses, together with HLA-peptide-binding prediction tools to identify likely mother-immunogenic variants carried by the fathers and transmitted to autistic children or not transmitted at all. The computational approach is a collaboration between our laboratory, Alex Krasnitz, Hannah Meyer, and Tobias Janowitz.

## PUBLICATIONS

Einson J, Glinos D, Boerwinkle E, Castaldi P, Darbar D, de Andrade M, Ellinor P, Fornage M, Gabriel S, Germer S, et al. 2023. Genetic control of mRNA splicing as a potential mechanism for incomplete penetrance of rare coding variants. *Genetics* **224:** iyad115. doi:10.1093/genetics/iyad115

Shi X, Lu C, Corman A, Nikish A, Zhou Y, Platt RJ, Iossifov I, Zhang F, Pan JQ, Sanjana NE. 2023. Heterozygous deletion of the autism-associated gene *CHD8* impairs synaptic function through widespread changes in gene expression and chromatin compaction. *Am J Hum Genet* **110:** 1750–1768. doi:10.1016/j.ajhg.2023.09.004

Wroten M, Yoon S, Andrews P, Yamrom B, Ronemus M, Buja A, Krieger AM, Levy D, Ye K, Wigler M, Iossifov I. 2023. Sharing parental genomes by siblings concordant or discordant for autism. *Cell Genom* **3:** 100319. doi:10.1016/j.xgen.2023.100319

### *In Press*

Chorbadjiev, L., M. Cokol, Z. Weinstein, K. Shi, C. Fleisch, N. Dimitrov, S. Mladenov, S. Xu, J. Hall, S. Ford, et al. 2024. The Genotype and Phenotypes in Families (GPF) platform manages the large and complex data at SFARI. bioRxiv doi: 10.1011/2024.02.08.579330

# MASSIVELY PARALLEL REPORTER ASSAYS, MACHINE LEARNING, AND THE BIOPHYSICS OF GENE REGULATION

**J.B. Kinney** A. Ayaz J. Desmarais E. Seitz
Y. Dan K. Loell D. Tenenbaum

The long-term goal of our research is to illuminate the biophysical mechanisms of gene regulation through the high-throughput measurement and quantitative modeling of sequence–function relationships. Our laboratory pursues this goal using a combination of experimental, computational, and mathematical approaches. Our experimental work develops high-throughput mutagenesis assays for measuring sequence–function relationships in alternative mRNA splicing (in humans), transcriptional regulation (in bacteria), and other biological systems. Our computational work develops methods for quantitatively modeling data from high-throughput mutagenesis assays, as well as for analyzing the sequence–function relationships described by genomic AI models. Our mathematical work addresses fundamental questions about the quantitative nature of sequence–function relationships and their statistical inference from high-throughput data.

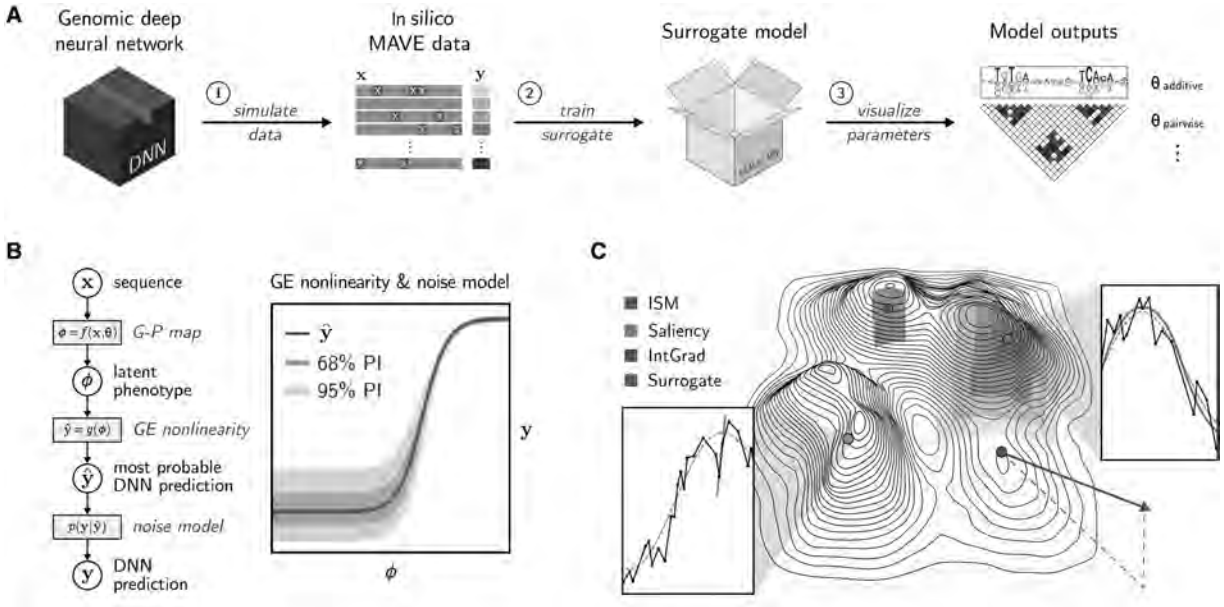



**Figure 1.** Overview of SQUID. (*A*) Schematic of the SQUID modeling framework. Analysis using SQUID comprises three main steps: (1) generate an in silico multiplexed assay of variant effects (MAVE) data set; (2) train a surrogate model on the MAVE data set; and (3) visualize parameters of the surrogate model to uncover biological mechanisms. (*B*) Structure of the latent phenotype surrogate models supported by SQUID. (G-P) Genotype–phenotype, (GE) global epistasis, (PI) prediction interval. (*C*) Schematic diagram of a deep neural network (DNN) function on a two-dimensional projection of sequence space. Each point in the plane corresponds to a unique sequence, and elevations represent DNN predictions. The large shaded region schematizes the ability of surrogate models to approximate the DNN function over an extended region of sequence space. *Insets* show an example DNN function (in 1D profile) centered about a sequence of interest with the ground-truth function (dashed line) overlaid. The *left inset* illustrates the sensitivity of saliency maps to nonsmooth local function properties. The *right inset* illustrates the ability of surrogate models to better approximate ground truth.

## Interpreting *cis*-Regulatory Mechanisms from Genomic Deep Neural Networks Using Surrogate Models

Deep neural networks (DNNs) have greatly advanced the ability to predict genome function from sequence. Interpreting genomic DNNs in terms of biological mechanisms, however, remains difficult. To address this need, Evan Seitz (a joint postdoc with Dr. Kinney and Dr. Peter Koo) developed SQUID, a genomic DNN interpretability framework based on surrogate modeling. SQUID approximates genomic DNNs in user-specified regions of sequence space using surrogate models (i.e., simpler models that are mechanistically interpretable). Importantly, SQUID removes the confounding effects that nonlinearities and heteroscedastic noise in functional genomics data can have on model interpretation. Benchmarking analysis on multiple genomic DNNs shows that SQUID, when compared with established interpretability methods, identifies motifs that are more consistent across genomic loci and yields improved single-nucleotide variant-effect predictions. SQUID also supports surrogate models that quantify epistatic interactions within and between *cis*-regulatory elements. SQUID thus advances the ability to mechanistically interpret genomic DNNs. (See Fig. 1.)

### PUBLICATION

*In Press*

Seitz EE, McCandlish DM, Kinney JB, Koo PK. 2024. Interpreting *cis*-regulatory mechanisms from genomic deep neural networks using surrogate models. *Nat Mach Intell* **6:** 701–713. doi:10.1038/s42256-024-00851-5

# TOWARD TRUSTWORTHY AND INTERPRETABLE DEEP NEURAL NETWORKS FOR REGULATORY GENOMICS

**P.K. Koo** N. Chuzhoy A. Crnjar K. Engel P. Garcia J. Kaczmarzyk T. Luo S. Muthukumar A. Nemshin A. Prakash C. Rajesh K. Rizzo A. Sarkar E. Seitz Z. Tang S. Toneyan Y. Yu J. Zhou

Our research studies the functional impact of genomic mutations through a computational lens using data-driven machine-learning solutions. We are broadly interested in applications of studying gene regulation and protein (dys)function. Our approach develops methods to interpret high-performing deep-learning models to distill knowledge that they learn from big, noisy biological data. Our goal is to elucidate biological mechanisms that underlie sequence–function relationships, with a broader aim of advancing precision medicine for complex diseases, including cancer.

## Improving Deep Learning for Regulatory Genomics with Evolution-Inspired Data Augmentations

P.K. Koo [in collaboration with N.K. Lee, CSHL]

The application of deep-learning models to predict epigenomic activity measured via high-throughput sequencing assays is increasing rapidly. Despite their early success, it is known that deep-learning models are data-hungry, often requiring large amounts of data to work well. However, in biology, there is a fundamental limitation of the data set size that is governed by the underlying biology. For instance, a transcription factor might only bind to high-affinity sites within accessible chromatin. A powerful workaround that the machine-learning community relies heavily on is data augmentation, which further increases the variations of training data. Data augmentations help to combat overfitting and improve generalization. However, there are limited strategies for data augmentation available in genomics. This issue primarily arises from the fact that functional activity can be altered in an unknown way when a perturbation is applied to DNA.

To address this issue, we introduce EvoAug, which comprises a set of new data augmentation methods for genomics inspired by evolutionary sampling strategies. EvoAug is employed as a two-stage procedure (Fig. 1). First, we pretrain a deep-learning model with evolutionary augmentations and we assume that predictions should be the same as wild type. This assumption is problematic as it is seemingly not reflective of the underlying biology. However, it guides the deep-learning model to learn more robust *cis*-regulatory features, albeit the regulatory rules may be distorted. Second, we fine-tune the deep-learning model on the original unperturbed data, which then guides the learned function back to biological reality. We show that this approach is very effective at improving generalization and interpretability across several established deep-learning models trained across the most prominent prediction tasks in regulatory genomics; these tasks span the majority of use cases for existing deep-learning models in this space.

More broadly, this work will greatly advance the field by providing a simple training method to enhance the performance of most deep-learning models in genomics. We have made EvoAug widely accessible as an application programming interface (API) built on top of PyTorch—we provide code as pip installable package and comprehensive documentation hosted on ReadTheDocs.org. We also provide several example Colab notebooks, extending the user base to those without the required GPU hardware. We anticipate that EvoAug will become a standard toolkit to train deep-learning applications in regulatory genomics.

## Facilitating Deep-Learning Analysis of Histopathology Data

J. Kaczmarzyk, P.K. Koo

Histopathology is the gold standard for diagnosing many types of cancers, and the application of deep learning in computational pathology is increasing rapidly. However, the planning of deep-learning

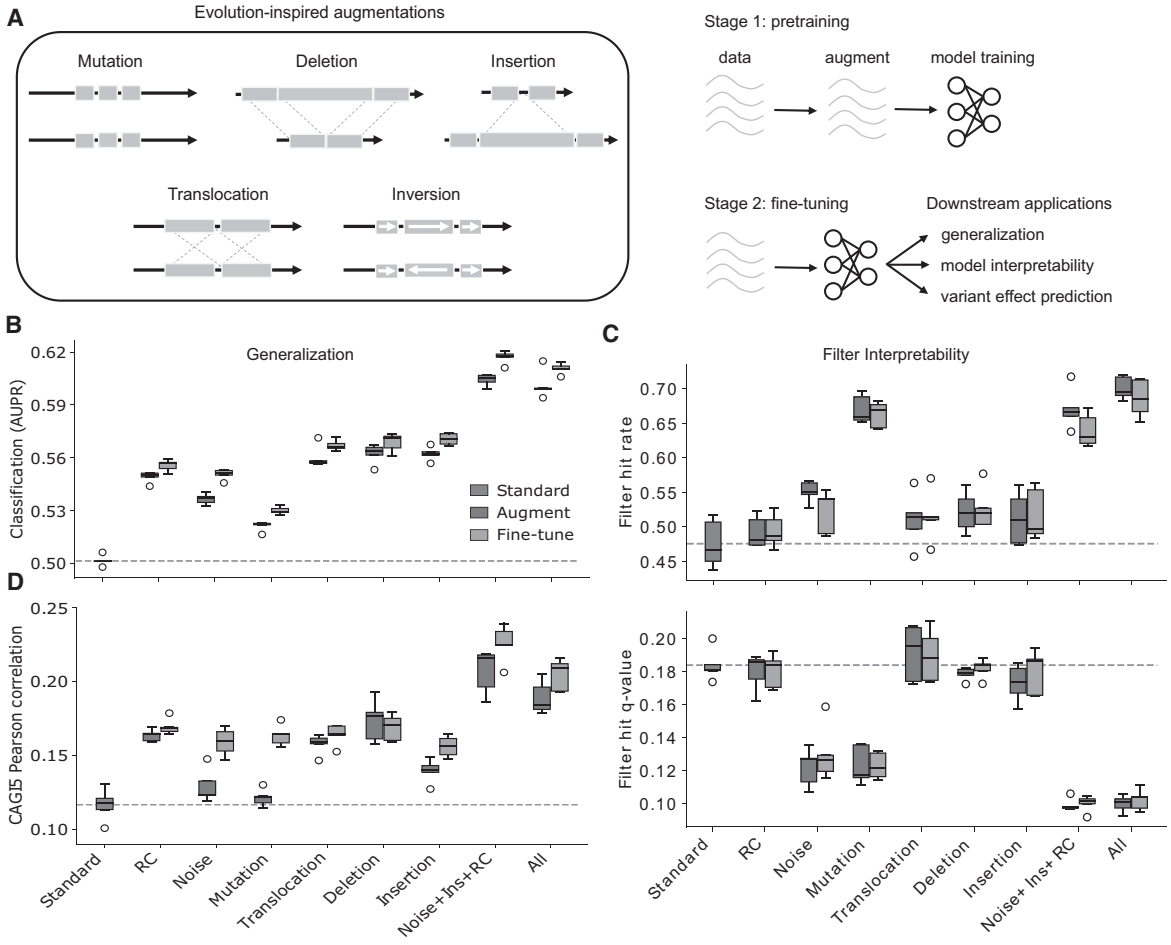



**Figure 1.** EvoAug improves performance. (*A*) Schematic of evolution-inspired data augmentations (*left*) and the two-stage training procedure (*right*). (*B*) Generalization performance (area under the precision-recall curve) for Basset models pretrained with individual and combinations of augmentations (i.e., Gaussian noise + insertion + reverse complement; all augmentations) and fine-tuned on the Basset data set. Standard represents no augmentations during training. (*C*) Comparison of the average hit rate of first-layer filters to known motifs in the JASPAR database (*top*) and the average *q*-value of the filters with matches (*bottom*). (*D*) Comparison of the average Pearson correlation between model predictions and experimental data from the CAGI5 Challenge. (*B–D*) Each box plot represents five trials with random initializations.

studies in this field is nontrivial for several reasons. First, there are hundreds of neural network architectures to choose from, and within each there are hyperparameters that can greatly affect classification performance. Second, there are diverse sources of computational pathology data for patch-based classification tasks, but these are not curated in a single location for deep-learning research. Third, there does not exist an easy-to-use toolkit to evaluate patch-based histology deep-learning models, yet this would be of great value to both deep-learning methods research groups and biomedical research groups.

To address these issues, we introduce ChampKit, which is an evaluation framework that includes hundreds of deep-learning architectures (many of which are pretrained on ImageNet), six publicly available data sets, and methods to evaluate model architectures, transfer learning, and training parameters (e.g., optimizers, regularization like early stopping) across data sets (Fig. 2). ChampKit is open-source and highly reproducible, and it enables rapid evaluation of neural networks for histopathology patch classification. Users can use ChampKit with custom patch classification data sets and custom models. Indeed, we evaluate transfer learning from a

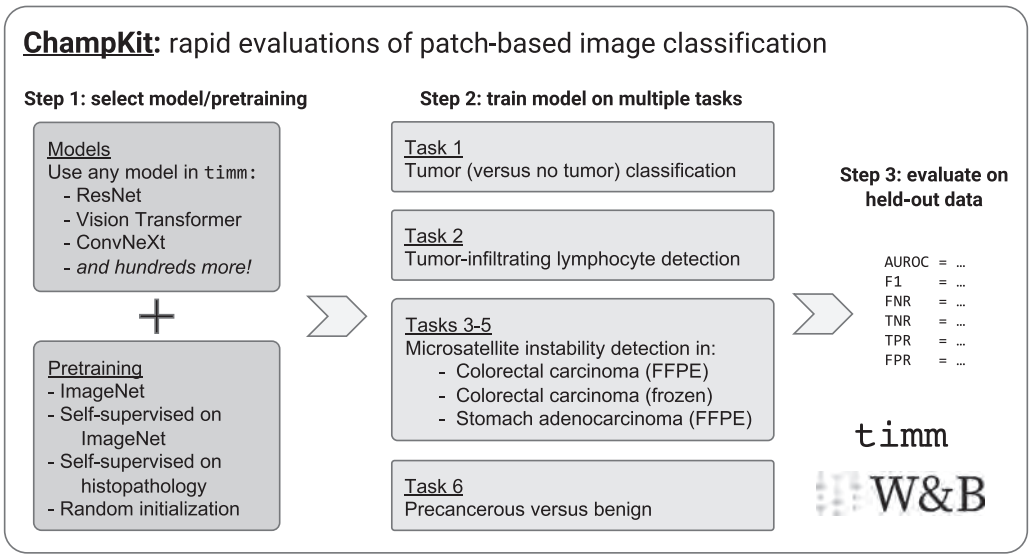



**Figure 2.** Overview of ChampKit. ChampKit enables systematic comparisons of model architectures and transfer learning across several patch-based image classification data sets. First, users select a model and pretrained weights from those available in timm model repository or a custom model with specification of pretrained weights or random initialization. Second, the models are trained on multiple tasks using identical training hyperparameters. Third, the trained models are evaluated on held-out test data for each task. Performance is tracked with Weights and Biases (W&B).

model pretrained on histopathology patches in a self-supervised fashion. We demonstrate the utility of ChampKit by evaluating three different neural network architectures on six histopathology data sets and show that ChampKit enables systematic comparisons.

More broadly, this work will advance research in histopathology by providing a platform for rapid evaluation of neural networks. Histopathology researchers can then choose the best neural networks for their data, which is difficult to do at present. We have made ChampKit widely accessible as a Python package built on top of PyTorch, timm, and Weights and Biases. We provide detailed instructions on GitHub. We anticipate that ChampKit will be a useful tool to evaluate deep-learning models for patch classification in histopathology.

## Correcting Gradient-Based Interpretations for Deep Learning in Genomics

C. Rajesh, P.K. Koo [in collaboration with A. Majdandzic, CSHL]

The application of deep-learning models to predict functional activity measured by high-throughput genomic assays is increasing rapidly. Biological insights are often gained through post hoc model interpretability analysis with attribution methods. This provides a way to understand the *cis*-regulatory patterns that drive gene regulation, and the attribution scores can also be utilized to prioritize the functional effects of noncoding mutations in human diseases. However, there remain many issues with attribution methods, many of which have been already identified within the machine-learning community, including benign overfitting or learning nonrobust representations.

Here we introduce a new noise source in attribution maps that arises specifically from the way deep neural networks (DNNs) fit DNA sequences (Fig. 3). We derive a simple correction and systematically test it comprehensively across various synthetic and in vivo data sets, covering the most prominent prediction tasks in regulatory genomics. This includes single-task binary prediction of transcription factor binding, multitask binary classification of chromatin accessibility across numerous cell types, quantitative regression of enhancer activity measured via a massively parallel reporter assay, and quantitative predictions of epigenetic profiles at base and binned resolutions. In

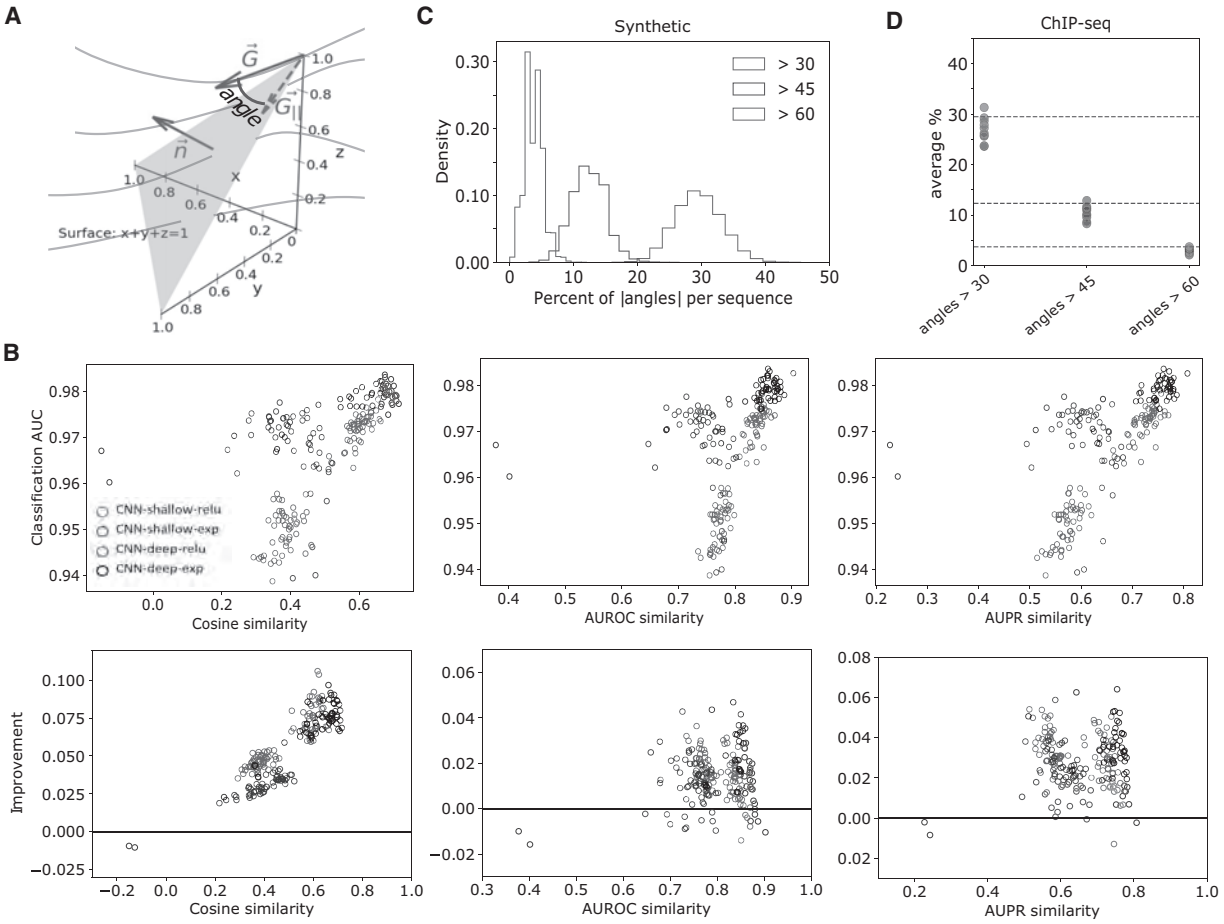



**Figure 3.** Gradient correction performance. (*A*) Toy diagram of geometric relationship between the input gradient and the simplex defined for three-dimensional categorical data. Curves represent gradient lines of a hypothetical learned function. The gray plane represents the data simplex. The vector represents the gradient pointing off of the simplex. (*B*) Performance comparison on synthetic data. (*Top row*) Scatter plot of interpretability performance measured by different similarity scores versus the classification performance (area under the curve) for saliency maps. (*Bottom row*) Interpretability improvement for saliency maps for different similarity metrics when using gradient correction. Improvement represents the change in similarity score after the gradient correction. Each point represents one of 50 trials with a different random initialization for each model. (*C*) Histogram of the percentage of positions in a sequence with a gradient angle larger than various thresholds for a deep convolutional neural network (CNN) with ReLU activations (CNN-deep-relu) trained on synthetic data. (*D*) Scatter plot of the percentage of positions in a sequence with a gradient angle larger than various thresholds for CNN-deep-relu trained on chromatin immunoprecipitation sequencing (ChIP-seq) data. Each point represents the average percentage across all test sequences for each ChIP-seq data set. For comparison, horizontal dashed lines indicate the mean value from the corresponding analysis using synthetic data in *C*.

each case, we highlight the pervasiveness of off-simplex gradient angles and demonstrate how corrected attribution maps yield improved definition of motifs and less spurious attribution noise.

More broadly, this work will greatly advance the field by providing more reliable attribution analysis from which trustworthy biological insights can be gained. Our proposed gradient correction can be implemented as a single line of code and thus can be easily extended to any genomic deep-learning application that employs attribution methods for model interpretability or variant effect predictions. Although this work is demonstrated on DNA sequence–based models, the proposed correction is general and should also extend to other categorical input data, such as protein and RNA sequences.

## PUBLICATIONS


Gao Y, He X-Y, Wu XS, Huang Y-H, Toneyan S, Ha T, Ipsaro JJ, Koo PK, Joshua-Tor L, Bailey KM, et al. 2023. ETV6 dependency in Ewing sarcoma by antagonism of EWS-FLI1-mediated enhancer activation. *Nat Cell Biol* **25:** 298–308. doi:10.1038/s41556-022-01060-1

Kaczmarzyk JR, Gupta R, Kurc TM, Abousamra S, Saltz JH, Koo PK. 2023. ChampKit: a framework for rapid evaluation of deep neural networks for patch-based histopathology classification. *Comput Methods Programs Biomed* **239:** 107691. doi:10.1016/j.cmpb.2023.107631

Koo PK, Ploenzke M, Anand P, Paul S, Majdandzic A. 2023. ResidualBind: uncovering sequence-structure preferences of RNA-binding proteins with deep neural networks. In *RNA structure prediction, Methods in molecular biology* (ed Kawaguchi RK, Iwakiri J), pp. 197–215. Springer Nature, London.

Lee NK, Toneyan S, Tang Z, Koo PK. 2023. EvoAug: improving generalization and interpretability of genomic deep neural networks with evolution-inspired data augmentations. *Genome Biol* **24:** 105. doi:10.1186/s13059-023-02941-w

Majdandzic A, Rajesh C, Koo PK. 2023. Correcting gradient-based interpretations of deep neural networks for genomics. *Genome Biol* **24:** 109. doi:10.1186/s13059-023-02956-3

Seitz EE, McCandlish DM, Kinney JB, Koo PK. 2023. Interpreting cis-regulatory mechanisms from genomic deep neural networks using surrogate models. bioRxiv doi:10.1101/2023.11.14.567120

Tang Z, Koo PK. 2023. Building foundation models for regulatory genomics requires rethinking large language models. *ICML 2023 Workshop on Computational Biology*, Honolulu, HI.

Toneyan S, Koo PK. 2023. Interpreting cis-regulatory interactions from large-scale deep neural networks for genomics. bioRxiv doi:10.1101/2023.07.03.547592

# COMPUTATIONAL GENOMICS OF CANCER

**A. Krasnitz** P. Belleau

The bulk of our research belongs to the field of computational cancer biology. Our choice of research goals within this field is dictated by (a) the potential impact of our work on clinical research and practice at present and in the foreseeable future, (b) the need to maximize the utility of emerging molecular technologies and research platforms in cancer biology, and (c) the opportunity to bring to bear quantitative techniques developed in other areas of science such as computational physics, applied mathematics, and computer science. These goals include (a) examination of intratumor genomic heterogeneity, its origin in cancer evolution, and its predictive value for aggressive and invasive potential of cancer; (b) reducing the complexity of genomic data for better interpretability while retaining their biological content; (c) derivation of clinically relevant molecular subtypes of the disease; and (d) design of predictive models for response to pharmacological interventions. As pursuit of these goals often reveals the inadequacy of the existing tools and necessitates the development of novel computational tools, toolmaking is an important component of our activity. In this report we focus on an ongoing project that is representative of our activity in 2023.

## Comprehensive Genetic Ancestry Annotation of Existing Human Molecular Data

Genetic ancestry has long been known as a major determinant of phenotype in humans, in health and disease. On a molecular level, ancestral dependence of the phenotype has been observed for a great variety of observables, including protein levels, RNA expression, and DNA methylation. Major ancestral effects on the molecular makeup of cancer have been observed for multiple cancer types. These interancestral genetic differences correlate with those firmly established in cancer epidemiology with regard to disease incidence, severity, and clinical outcome. Similarly, ancestrally determined differences in molecular phenotypes have been discovered in conditions as diverse as Alzheimer's disease, systemic lupus erythematosus, and influenza infection. Further research to uncover molecular phenotypic manifestations of genetic ancestry will require massive amounts of human molecular data alongside reliable donor ancestry annotation. A lack of genetic ancestry metadata for a vast majority of existing human molecular data represents a major obstacle to progress in this direction. Ancestry can be inferred using existing, well-established methods if the donor germline genotype is known reliably and in sufficient detail. This is the case for whole-exome or whole-genome sequences derived from cancer-free DNA, but not for multiple other sequence data types, including RNA-seq, ATAC-seq, or sequences of cytosine-converted DNA, both from cancer and cancer-free sources. Neither is this the case for tumor-derived nucleic acid sequences of cancer patients, in which the difficulty in genotype determination is exacerbated by massive, genome-wide alterations, including copy number variants and loss of heterozygosity. Donor race and/or ethnicity are often used as surrogates for genetic ancestry, but these are unknown for the vast majority of public sequence data, often are inaccurate, and always lack detail. Lack of ancestral annotation for multiple types of existing molecular data is a major obstacle to ancestry-oriented study of human disease. With our work, we seek to remove this obstacle by (a) developing a suite of reliable computational tools for inference of genetic ancestry from multiple types of nucleic-acid sequences and (b) using these methods for genetic ancestry annotation of open-access resources of human molecular data. We have developed a set of computational methods for robust and accurate inference of genetic ancestry from a variety of molecular data types. These tools rely on innovative computational synthesis of molecular data sets representing diverse ancestral backgrounds and using these synthetic data for inference parameter optimization (Belleau et al. 2023a).

One particularly important application of our inferential methodology is to cancer genomics. We designed an inference procedure having in mind a

clinical scenario with an input molecular profile of a tumor from a single patient and no matching cancer-free sequence available. The profile in question comes with its unique set of sequence properties, resulting both from the cancer-specific and germline genetics of the donor and from details of the molecular protocol. The former include somatic sequence variants due to cancer—in particular, copy number and structural variants—loss of heterozygosity, single-nucleotide variation, and microsatellite-related alterations. The latter include the target sequence and uniformity of its coverage, depth, read length, and sequencing quality. These profile-specific properties would make it impossible to confidently assess the accuracy of the inference procedure for the input profile from its benchmarking performance with the public cancer-derived data.

To overcome this difficulty, we developed a computational technique termed data synthesis, wherein the inherited genotype of the patient is supplanted in the input tumor-derived profile by one derived from an unrelated individual from a reference panel, with known ancestry. The resultant synthetic profile retains the essential properties of the tumor-derived data, both biological and protocol-specific. At the same time, the profile acquires the ancestral background of the chosen reference individual. The remainder of the reference panel is then used to infer the ancestry from the synthetic profile. This cross-validation is an integral part of our approach. We next apply established methods of ancestry inference to this synthetic profile and compare the result with the known ancestry of the reference individual. Synthesizing multiple such profiles, we are able to assess how accurate the ancestry inference is for the patient, both overall and as a function of the profile's continental-level ancestry. Furthermore, using synthetic data, we are able to optimize the inference procedure with respect to its parameters. Importantly, this assessment and optimization procedure does not require the profile in question to be part of a larger data set from a cohort of patients with a similar diagnosis. Very often in public cancer-derived data such cohorts do not provide statistically meaningful representation of non-European ancestries. Our methodology compensates for this insufficiency.

We illustrate this procedure as applied to cancer-derived molecular profiles resulting from five commonly used protocols (Fig. 1), for the purpose of inferring global genetic ancestry at the continental level of resolution. As our population reference, we used a set of more than 3,000 genotypes sampled from individuals worldwide by the International Genome Sample Resource (IGSR). These genotypes fall into five continental superpopulations: African (AFR), American (AMR), East Asian (EAS), European (EUR), and South Asian (SAS). We employed a well-established inference method wherein the genotype of the donor, as estimated from the cancer-derived profile, is projected onto the subspace spanned by D top principal directions of the IGSR genotype matrix. The ancestry of the donor is then determined as that of the majority among the nearest K IGSR neighbors of the donor in that subspace. Both D and K are inference parameters available for optimization of performance. The area under the receiver operating characteristic curve (AUROC) is used as a measure of performance.

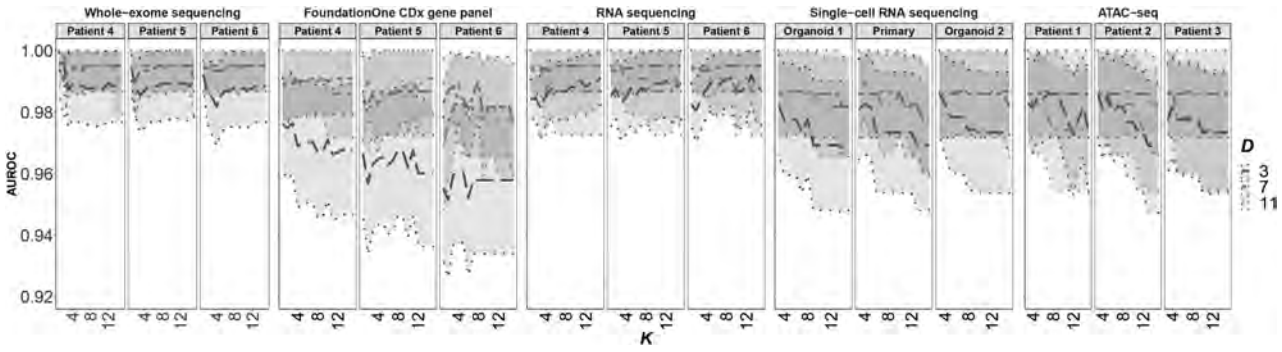



**Figure 1.** American (AMR)-specific AUROC measure of performance for global ancestry inference, as a function of the inference parameters *D* and *K*, for three pancreatic-cancer profiles (whole-exome sequencing, RNA sequencing, gene panel, in-house collection), three breast cancer profiles (ATAC-seq, The Cancer Genome Atlas), and two organoid culture profiles and a primary tissue profile of endometrial cancer (single-cell RNA-seq aggregated over 1,000–6,000 cells). Central values (dashed) and 95% confidence intervals are indicated in matching colors.

As Figure 1 shows, our method attains high performance across all five molecular platforms considered in a challenging case of donors from the AMR superpopulation, known for strong admixtures of African, native American, and European ancestries.

Alongside these advances in inferential method development, there has been progress in making our tools public and expanding their user base. Importantly, a software package embodying these novel tools and termed Robust Ancestry Inference using Data Synthesis (RAIDS) is now available from Bioconductor, a leading, peer-reviewed R language software repository for life sciences (Belleau et al. 2023b).

## PUBLICATIONS

Belleau P, Deschênes A, Chambwe N, Tuveson DA, Krasnitz A. 2023a. Genetic ancestry inference from cancer-derived molecular data across genomic and transcriptomic platforms. *Cancer Res* **83:** 49–58. doi:10.1158/0008-5472.CAN-22-0682

Belleau P, Deschênes A, Tuveson DA, Krasnitz A. 2023b. Accurate inference of genetic ancestry from cancer sequences. *Bioconductor* doi:10.18129/B9.bioc.RAIDS

# COMPUTATIONAL GENETICS

**D. Levy** S. Negi Z. Yu

Our laboratory focuses on harnessing computational methods to deepen our understanding of sporadic human diseases such as autism and various cancers. By leveraging extensive data resources, we develop innovative algorithms, protocols, and analytical tools to uncover new insights in biology and genetics.

## Ultra-Sensitive Variant Detection

Detecting genomic variants present at extremely low frequencies is crucial for applications like monitoring residual disease in cancer patients. Traditional sequencing methods are limited by error rates that prevent reliable detection of variants below a frequency of 1%. Our objective is to push these boundaries to identify variants occurring as infrequently as one in a million (0.0001%).

Together with the Wigler laboratory, we have developed the MASQ (multiplex accurate sensitive quantitation) protocol. This technique attaches unique sequence identifiers, or "varietal tags," to each original DNA molecule. By grouping and analyzing reads that share the same tag, we can distinguish true variants from sequencing errors, effectively enhancing detection sensitivity beyond standard limitations.

Tumor DNA fragments circulate in the plasma, or cell-free component of the blood, making it possible to assess tumor load noninvasively. Working with Drs. Gary Goldberg and Marina Frimer at Northwell, we applied MASQ to measure cell-free tumor DNA in patients with endometrial and ovarian cancers. We analyzed 60 samples from 46 patients, quantifying tumor load at diagnosis and, in some cases, before and after treatment. Our findings revealed a significant correlation between detected tumor DNA levels and patient survival, as well as the stage of disease. Importantly, reductions in tumor DNA levels following surgery suggested the possibility of successful treatment outcomes, which we are now evaluating further in a follow-up study.

To improve the method's utility, we are expanding MASQ to include 500 genomic loci and enhance the assay's efficiency through new laboratory protocols. Given our findings so far with MASQ, a modest improvement in efficiency and a 10- to 20-fold expansion in the target set will enable the level of precision needed for broad clinical value.

## Mutational Sequencing for Long-Range Assembly

High-quality genome assemblies often require long-read sequencing data, which are essential for studying genomes where genetic variants may be separated by large distances. However, third-generation sequencing platforms that provide long reads are costly and have higher error rates. To address this, we developed mutational sequencing (muSeq), a method that infers long-range genomic information from standard short-read sequencing data. By introducing random mutations—specifically converting cytosine to thymine—we create unique molecular signatures that enable the reconstruction of long DNA sequences from overlapping short reads.

This year, we initiated a project to measure repeat length expansions in the *FMR1* gene associated with Fragile X syndrome. The cytosine-to-thymine mutations enhance the amplification and sequencing of the challenging, long repeat regions in this gene. Preliminary results are encouraging, suggesting that muSeq could become a valuable tool for studying genetic disorders involving repeat expansions and could improve diagnostic capabilities for conditions like Fragile X syndrome.

## Precise Measurement of Microsatellite Instability

Microsatellites—short, repetitive DNA sequences—are prone to mutations, making them useful biomarkers for certain cancers. Unlike single-nucleotide variants, microsatellites have well-defined locations, allowing for the development of universal assays

applicable across different patients. Accurately measuring microsatellite lengths is challenging because of the tendency of these sequences to slip during replication, leading to errors in length determination. Building on our work with muSeq, we developed a technique called STORM (stabilizing technique of random mutagenesis) that partially mutates microsatellite sequences before amplification. This process disrupts the repeat structure, preventing slippage and enabling precise length measurements.

All cells make mistakes in microsatellite lengths during replication, but the mismatch repair (MMR) system normally corrects these errors. In some cancers, this repair system is deficient (dMMR), resulting in many uncorrected changes in microsatellite length. This hallmark of dMMR tumors is known as microsatellite instability (MSI). Knowing the MSI status of cancers is important because it informs therapeutic decisions. Current MSI detection methods are effective for colon cancer but less so for other cancer types. We applied STORM to determine the MSI status in 100 ovarian and endometrial cancer samples and found that it works exceptionally well, achieving high accuracy even when the tumor sample is mostly normal cells. Improving MSI detection could lead to better treatment selection and outcomes, particularly in those cancers not adequately served by existing tests.

We are also developing advanced algorithms for the precise alignment of complex microsatellites. By extending existing algorithms to include gap allowances in sequence alignment, we enhance the precision of analyzing these repetitive sequences. This work, coupled with advanced visualization tools, will enable detailed studies of genetic variations in microsatellites at both the population and individual levels. Because our methods allow us to read sequences that are challenging for other technologies, this research provides new insights into microsatellite variation in the human genome.

Additionally, we are developing methods for mutating adenine repeats to stabilize A-rich microsatellites, which are far more abundant in the genome than C-rich repeats. We are exploring both enzymatic and chemical approaches for inducing these mutations and are developing the analytical techniques necessary to study them. Expanding our focus to A repeats will broaden our understanding of microsatellite instability and its implications in disease.

## Integrative Single-Cell DNA and RNA Sequencing

Understanding the heterogeneity within tumors is essential for advancing cancer research and treatment. Although single-cell DNA sequencing provides insights into tumor evolution and genetic variations and single-cell RNA sequencing reveals gene expression profiles, integrating both types of data from the same cell has been a significant challenge. We have developed a high-throughput method called hybrid BAG-seq that captures both DNA and RNA from the same single nucleus. This technique involves encapsulating individual nuclei within tiny acrylamide gel beads (BAGs) containing specially designed primers that bind to nucleic acids. These primers are co-polymerized into the gel matrix, anchoring the DNA and RNA within each BAG. Through a barcoding process, each BAG is assigned a unique identifier, allowing us to track nucleic acids from individual cells during sequencing and analysis.

Using hybrid BAG-seq, we analyzed samples from five different uterine cancers, uncovering a complex landscape of genetic identities linked to gene expression profiles. We observed various relationships between genomic and transcriptomic data, including straightforward one-to-one correspondences and more complex interactions in which multiple genetic clones shared similar expression patterns or single clones displayed diverse expression profiles. These findings highlight the fact that the relationship between a cell's genetic makeup and its gene expression cannot be assumed and must be empirically determined.

In our analysis, tumor expression clusters were distinct from normal stromal cell types, yet tumor cells clustered according to their cell type of origin. Although tumor expression patterns varied between patients, the expression states of stromal cells remained largely consistent. Notably, in one patient, we identified a significant population of epithelial cells with abnormal expression profiles that, without accompanying DNA data, could not have been definitively classified as tumor cells. In another case, we detected a subset of T cells that had lost one copy of the X chromosome, indicating a clonal expansion of these cells within the tumor environment. This demonstrates the effectiveness of hybrid BAG-seq in providing a detailed view of tumor heterogeneity and the tumor microenvironment, offering new opportunities for targeted therapeutic strategies.

## PUBLICATIONS

Li S, Alexander J, Kendall J, Andrews P, Rose E, Orjuela H, Park S, Podszus C, Shanley L, Ma R, et al. 2023. High-throughput single-nucleus hybrid sequencing reveals genome-transcriptome correlations in cancer. bioRxiv doi:10.1101/2023.10.04.560973

Wroten M, Yoon S, Andrews P, Yamrom B, Ronemus M, Buja A, Krieger AM, Levy D, Ye K, Wigler M, Iossifov I. 2023. Sharing parental genomes by siblings concordant or discordant for autism. *Cell Genomics* **3:** 100319. doi:10.1016/j.xgen.2023.10031

# PREDICTING EFFECTS OF MUTATIONS FROM HIGH-THROUGHPUT DATA

**D. McCandlish** B. Gitschlag C. Martí-Gómez M. Sun

Understanding the relationship between the DNA sequence of an organism's genome (genotype) and the measurable characteristics of that organism (phenotype) is one of the fundamental goals of biology. Recent progress in high-throughput experimental techniques now enables us to simultaneously measure, in a single experiment, the cellular or molecular effects of thousands to millions of DNA sequence alterations. In the McCandlish laboratory, we are focused on developing new computational and mathematical techniques for making sense of this wealth of data. Our ultimate goals are to be able to predict the pathogenicity of mutations observed in human genome sequences, understand somatic evolution in cancer and the evolution of drug resistance in human pathogens, and help to construct highly optimized enzymes for biotechnology applications.

An important challenge in predicting the effects of mutations is that the effect of any given mutation may depend on which other mutations are already present, a phenomenon known as genetic interaction or epistasis. Our group is particularly interested in developing techniques to quantify and better understand the form and causes of these genetic interactions, with the dual goals of improving our ability to predict the effects of combinations of mutations and to understand the influence that these interactions have on the process of biological evolution.

This year we started a new collaboration with Dr. Samantha Petti from the Tufts University Department of Mathematics on a project to identify which sites and alleles in a genotype display the greatest degree of epistasis and to leverage this knowledge to make better predictions.

## Genetic Interactions and High-Throughput Data

Modern high-throughput experimental techniques are for the first time providing large-scale measurements not only for how pairs of mutations interact, but for higher-order interactions involving three or more mutations. However, how to model and understand these higher-order interactions remains an open question.

This year we have been working in three main directions. First, postdoc Carlos Martí-Gómez has been writing a software package that integrates and extends several of our previous methods for modeling and analyzing high-throughput genotype–phenotype data. Second, many of our methods are based in Gaussian processes, which provide a flexible and mathematically tractable framework for modeling high-dimensional genotype–phenotype maps. Although Gaussian processes have historically been difficult to apply to large data sets, our previous methods overcame this difficulty by using specially designed Gaussian processes that could be implemented using only sparse matrix multiplication, allowing us to infer genotype–phenotype maps for genotypic spaces containing low millions of possible genotypes. However, inference of Gaussian processes has recently seen substantial progress because of graphics processing unit (GPU) acceleration, and we have been working to leverage these advances to apply our methods at a genome-wide scale. Finally, in collaboration with Dr. Samantha Petti and former postdoc Dr. Juannan Zhou (currently an assistant professor at the University of Florida), we have been working to develop new Gaussian process methods for modeling the genotype–phenotype map. The main idea behind these methods is that there are certain key mutations that alter the effects of other mutations, and by identifying these key mutations we can develop models that make better predictions and better quantify prediction uncertainty.

Besides developing new methods, we continue to apply our existing methods by collaborating with experimentalists to analyze new genotype–phenotype map data. In collaboration with Dr. Sarel Fleishman's group at the Weizmann Institute of Science, we analyzed genetic interactions and the structure of the genotype–phenotype map in green fluorescent protein (GFP), based on a high-throughput data set mutagenizing 14 amino acid positions in the core of

the protein. The resulting analysis of the genotype–phenotype map for GFP has now been published in *Nature Communications* (Weinstein et al. 2023). In collaboration with Dr. Zachary Lippman's group here at CSHL, we analyzed how patterns of genetic interaction between mutations scale with the phenotypic strength of the individual mutations. Using CRISPR to create a set of mutations with a diversity of individual strengths of effects on tomato fruit size, we combined these mutations with mutations in two other genes previously shown to be important in regulating fruit size. Interestingly, our analysis showed that although interactions with one of these genes showed a regular and predictable pattern of epistasis consisting of a dose-dependent saturating response, interactions with the other gene showed a highly idiosyncratic pattern in which the double-mutant phenotypes would be quite different even for mutations that produce very close single-mutant phenotypes. The resulting manuscript was published in *Science* (Aguirre et al. 2023). We have also begun follow-up work by asking similar questions for another reproductive trait in tomato, inflorescence branching.

Carlos in particular has also been involved in a number of other collaborations during the past year. Working with members of our collaborator Dr. Joshua Payne's group at ETH Zürich, we have been analyzing how incorporating information on the structure of the genetic code changes our understanding of amino acid–level genotype–phenotype maps, which we submitted as a preprint this year. Carlos has also been working with members of Dr. Justin Kinney's and Dr. Adrian Krainer's groups to understand the mechanisms of action for splice-modifying drugs. Finally, we have been working with Dr. Peter Koo's and Dr. Justin Kinney's joint postdoc, Dr. Evan Seitz, to use techniques for analyzing experimental genotype–phenotype data to analyze the predictions of genomic deep neural networks and together submitted a preprint on this work.

This year we also welcomed a new postdoc to the group, Dr. Mengyi Sun. Mengyi has been working on a project in collaboration with Dr. Arlin Stoltzfus at the National Institute of Standards and Technology to integrate data between experiments that comprehensively measure the effects of all single amino acid substitutions in a protein. Asking how predictable the effects of mutations are when viewed as a function of the resident and mutant amino acids, he has found that although the individual effects are highly unpredictable, the distribution of effects appears to be drawn from a one-parameter family of probability distributions with a characteristic shape. Moreover, he has shown that further analysis of these distributions can be used to explain patterns of amino acid replacement observed during long-term molecular evolution.

## Influence of Mutational Biases on Molecular Adaptation

Evolutionary adaptation often occurs by the fixation of beneficial mutations. However, because of the specific mechanisms of DNA damage and repair at play in any given species, different beneficial mutations appear within an evolving species at different rates. Whether and to what extent these mutational biases influence the genetic basis of adaptive evolution has been a long-term subject of interest for the McCandlish laboratory, and this year we published a synthetic review of this area and its implications for the predictability of evolution (Cano et al. 2023).

In collaboration with Joshua Payne at ETH Zürich and Arlin Stoltzfus at the National Institute of Standards and Technology, postdoc Dr. Bryan Gitschlag has been working to develop further theoretical predictions for the genetic architecture of adaptation when mutational biases are present. His results show that even if mutation rates and selection coefficients are not intrinsically correlated with each other, we nonetheless expect to see a negative correlation between mutation rates and selection coefficients among fixed mutations, and that this effect also contributes to a decrease in the frequency of parallel evolution. We have also applied Bryan's theory to data from several systems, including paired mutational and fitness data from Dengue virus, and data on the frequency of different somatic amino acid changes in the tumor suppressor protein p53. These results have now been published in the journal *American Naturalist* (Gitschlag et al. 2023). Bryan has also been working on a project to compare the patterns of mutational biases in different species and to ask how closely the frequency of different types of mutations among adaptive substitutions track these biases. Finally, Bryan developed a new statistical framework for understanding changes in the frequency of selfish mitochondrial genomes in *Caenorhabditis elegans*, and submitted a preprint applying this method to data generated during his Ph.D. work.

Also in collaboration with Joshua Payne at ETH Zürich and Arlin Stoltzfus at the National Institute of Standards and Technology, we have been working to understand how frequently different classes of mutation contribute to adaptation. For instance, changes in repeat number for trinucleotide repeats occur at a high rate but produce few distinct variants, whereas complex insertions produce an almost unlimited supply of new genetic sequences but occur at a much lower rate. We have been using the contrast between single-nucleotide mutations and adjacent dinucleotide mutations in coding sequences as a model system to address these questions. Although adjacent dinucleotide mutations occur at a rate 100-fold lower than single-nucleotide mutations, because of the structure of the genetic code they access more possible amino acids, so that a majority of the time the most beneficial single amino acid substitution to a protein is accessible via a dinucleotide mutation but not a single nucleotide change. This year we developed a theory for the frequency of dinucleotide mutations, and applied this theory to data from p53 to show that competition between mutations increases the frequency of dinucleotide mutations observed in cancer patients twofold to threefold higher than would be expected based on the frequency of dinucleotide mutations among all beneficial (i.e., cancer-causing) mutations in p53.

## Mathematical Modeling of Cancer Cachexia

Cachexia is a wasting disease that is frequently the proximal cause of death in cancer and that also arises because of a wide variety of other causes including organ failure and chronic inflammation. In collaboration with members of Dr. Tobias Janowitz's laboratory, we have been working to construct simple differential equation–based models for the dynamics of cachexia and the course of disease progression during cancer cachexia. These models explain certain qualitative phenomena, such as the observation that cachectic patients often have a sudden worsening after a long period of apparent stability. For example, in our theory, this sudden worsening corresponds to the loss of stability of a homeostatic set point, at which this loss of stability can be explained mathematically using dynamical systems methods as a particular type of bifurcation known as a saddle-node bifurcation. This year we helped contribute to a review of this area, which was published in the journal *Cell* (Ferrer et al. 2023).

### PUBLICATIONS

Aguirre L, Hendelman A, Hutton SF, McCandlish DM, Lippman ZB. 2023. Idiosyncratic and dose-dependent epistasis drives variation in tomato fruit size. *Science* **382:** 315–320. doi:10.1126/science.adi5222

Cano AV, Gitschlag BL, Rozhoňová H, Stoltzfus A, McCandlish DM, Payne JL. 2023. Mutation bias and the predictability of evolution. *Philos Trans R Soc B* **378:** 20220055. doi:10.1098/rstb.2022.0055

Ferrer M, Anthony TG, Ayres JS, Biffi G, Brown JC, Caan BJ, Feliciano EMC, Coll AP, Dunne RF, Goncalves MD, et al. 2023. Cachexia: a systemic consequence of progressive, unresolved disease. *Cell* **186:** 1824–1845. doi:10.1016/j.cell.2023.03.028

Gitschlag BL, Cano AV, Payne JL. McCandlish DM, Stoltzfus A. 2023. Mutation and selection induce correlations between selection coefficients and mutation rates. *Am Naturalist* **202:** 534–557. doi:10.1086/726014

Weinstein JY, Martí-Gómez C, Lipsh-Sokolik R, Hoch SY, Liebermann D, Nevo R, Weissman H, Petrovich-Kopitman E, Margulies D, Ivankov D, et al. 2023. Designed active-site library reveals thousands of functional GFP variants. *Nat Commun* **14:** 2890. doi:10.1038/s41467-023-38099-z

# GENERATION OF DIVERSITY AND TOLERANCE DURING T-CELL DEVELOPMENT

**H.V. Meyer** A. Banerjee S. Carcy S. Chapin V. Kovaleva Y. Lin R. Prabakar M. Syed J. Torres

The thymus plays a key role in teaching T cells the ability to distinguish self from nonself. Flaws in the teaching process can lead to autoimmune diseases or immunodeficiency. The Meyer laboratory combines genomics and mathematical modeling to understand the mechanisms of healthy and pathological thymus function. We develop experimental and computational approaches to elucidate the interactions of T cells and thymic epithelial cells that drive self-tolerance and generate diversity in the immune system.

In the past year, we investigated lineage choice and effector differentiation across human T-cell lineages and developed two complementary tools to study T-cell activation. A summary of these projects and our findings is provided below.

## How T Cells Meet Their Fate

S. Carcy, J. Torres, Y. Lin [in collaboration with L. Loh and L Gapin, University of Colorado]

We set out to shed light on a crucial, yet unanswered question in human T-cell biology: How are lineage choice and effector differentiation of T cells driven during their development and peripheral function (Loh et al. 2023)? In collaboration with Laurent Gapin's laboratory at the University of Colorado, we investigated these questions, spanning conventional (CD4 and CD8 $T_{conv}$) and innate-like T cells ($T_{inn}$) including invariant nature killer T cells (iNKT), mucosal associated innate T (MAIT) cells, and γδ-T cells. Studies in mice had already delineated the developmental trajectories of $T_{inn}$, and analyses of distinct subsets of peripheral human $T_{inn}$ cells have discovered developmental stages of human Vd2-Vg9 and functional subtypes of human MAIT cells; however, a comprehensive picture spanning development and peripheral function across $T_{inn}$ and $T_{conv}$ was lacking. To address this, we assessed the range of phenotypic states $T_{conv}$ and $T_{inn}$ cells can adopt in vivo in the human thymus and blood. We uncovered that the majority of postnatal human thymic $T_{inn}$ cells exhibit a transcriptome akin to that of naïve $CD4^+$ or $CD8^+$ $T_{conv}$ cells. Only a fraction of thymic $T_{inn}$ cells show a transcriptional signature indicative of an "effector" state. Conversely, most adult blood $T_{inn}$ cells display an effector transcriptome. Although $T_{conv}$ cells exhibit a continuum of transcriptional states, spanning from naïve to central and effector memory T cells, $T_{inn}$ cells express a distinct transcriptional program shared among iNKT, MAIT, and Vd2Vg9 T cells. However, unlike the mouse, human $T_{inn}$ cells do not differentiate into functionally distinct subsets; instead, they develop an effector program with mixed type 1/type 3 effector potential. To conduct a comprehensive cross-species analysis, we constructed a murine $T_{inn}$ developmental atlas and uncovered additional species-specific distinctions, including the absence of type II $T_{inn}$ cells in humans, which implies distinct immune regulatory mechanisms across species. Finally, our study highlights differences in the pattern of CD1D expression in the thymus between the two species, which could potentially impact the maturation process of iNKT cells in humans.

## A Two-Pronged Approach to Study T-Cell Activation

V. Kovaleva, A. Banerjee, S. Chapin [in collaboration with D. Pattinson, University of Wisconsin; P. Thomas and M. Pogorelyy, St. Jude Children's Hospital; S. Navlakha, CSHL]

T cells act as potent fighters against pathogen infection and malignantly transformed cells through direct cytotoxic activity and cytokine release. The stimulation of these responses in $T_{conv}$ is initiated through the binding of the T cell's T-cell receptor (TCR) to foreign or malignant peptide fragments bound to major histocompatibility complexes (MHCs) on

antigen-presenting cells. On a systemic level, large TCR repertoire diversity enables the orchestration of antigen-specific immune responses against the vast space of possible peptides. At the same time, this diversity would have detrimental effects if not checked for potential reactivity to self. This process takes place in the thymus. However, the processes that drive selection in the thymus and peripheral effector function are identical on a molecular level. We have developed two complementary approaches to study TCR/antigen binding pairs from the large TCR repertoire and antigen space, which is crucial across many fields of biomedical research.

First, we developed copepodTCR (Kovaleva et al. 2023), an open-access tool for the design and interpretation of high-throughput experimental assays to determine TCR specificity. copepodTCR implements a combinatorial peptide pooling scheme for efficient experimental testing of T-cell responses against large overlapping peptide libraries, useful for "deorphaning" TCRs of unknown specificity. The scheme detects experimental errors and, coupled with a hierarchical Bayesian model for unbiased results interpretation, identifies the response-eliciting peptide for a TCR of interest out of hundreds of peptides tested using a simple experimental setup. Using in silico simulation, we demonstrated the applicability of our design scheme and the sensitivity of our results evaluation across varied experimental layouts and range of TCR-peptide activation signals. We experimentally validated our approach on a library of 253 overlapping peptides covering the SARS-CoV-2 spike protein using MHC class I and MHC class II restricted cell lines with known specificity. We then provide experimental guides for efficient design of larger screens covering thousands of peptides, which will be crucial for the identification of antigen-specific T cells and their targets from limited clinical material.

Although copepodTCR will aid with the design of high-throughput assays for TCR specificity detection, the space of all possible TCR/peptide:MHC pairs is too vast to experimentally test. To facilitate the selection of target peptides to test, we developed an open-access tool for predicting TCR peptide:MHC interactions, which is still a fundamental challenge (Banerjee et al. 2024). Recent methods have tackled a version of this problem—predicting TCRs that bind one given peptide—but the opposite problem—predicting peptides that one given TCR binds—remains outstanding. This latter problem is critical for predicting TCR responses in an array of applications including off-target toxicity of T-cell-based therapies and cancer neoantigen predictions. To address this, we developed a hierarchical Bayesian framework (BATMAN) that takes TCR-peptide:MHC binding data as input and can predict a TCR's binding to a new peptide. Existing methods, including those that utilize "black box" deep neural networks, perform only slightly better than random on this task, highlighting the challenge of predicting the effect of small mutations on binding activity. In contrast, BATMAN achieves AUCs >0.80 and yields interpretable, positional weights and amino acid distance functions characteristic for TCR-peptide:MHC binding properties. To train BATMAN, we compiled the largest benchmarking data set to date, containing >10,000 TCR-peptide:MHC activity data of high-confidence positive and negative samples. The latter are crucial as biases toward reporting only positive examples in databases have greatly hampered methods development. We also developed a web application with a graphical user interface that enables the community to explore this data set. This database is fully accessible, and we believe it will serve as an important benchmarking data set for future studies.

## PUBLICATIONS

Kovaleva VA, Pattinson DJ, Barton C, Chapin SR, Minervina AA, Richards KA, Sant AJ, Thomas PG, Pogorelyy MV, Meyer HV. 2023. copepodTCR: identification of antigen-specific T cell receptors with combinatorial peptide pooling. bioRxiv doi:10.1101/2023.11.28.569052

*In Press*

Banerjee A, Pattinson DJ, Wincek CL, Bunk P, Chapin SR, Navlakha S, Meyer HV. 2024. BATMAN: improved T cell receptor cross-reactivity prediction benchmarked on a comprehensive mutational scan database. bioRxiv doi:10.1101/2024.01.22.576714

Loh L, Carcy S, Krovi HS, Domenico J, Spengler A, Lin Y, Torres J, Palmer W, Norman PJ, Stone M, et al. 2024. Unraveling the phenotypic states of human innate-like T cells: comparative insights with conventional T cells and mouse models. *Cell Reports*. doi:10.1016/j.cellrep.2024.114705

# UNCOVERING ALGORITHMS IN THE NATURAL WORLD

**S. Navlakha** A. Banerjee Y. Shen
I. Bush X. Zheng

## Reducing Catastrophic Forgetting with Associative Learning—a Lesson from Fruit Flies

Catastrophic forgetting remains an outstanding challenge in continual learning. Recently, methods inspired by the brain, such as continual representation learning and memory replay, have been used to combat catastrophic forgetting. Associative learning (i.e., retaining associations between inputs and outputs, even after good representations are learned) serves an important function in the brain; however, its role in continual learning has not been carefully studied. Here, we identified a two-layer neural circuit in the fruit fly olfactory system that performs continual associative learning between odors and their associated valences. In the first layer, inputs (odors) are encoded using sparse, high-dimensional representations, which reduces memory interference by activating nonoverlapping populations of neurons for different odors. In the second layer, only the synapses between odor-activated neurons and the odor's associated output neuron are modified during learning; the rest of the weights are frozen to prevent unrelated memories from being overwritten. We prove theoretically that these two perceptron-like layers help reduce catastrophic forgetting compared to the original perceptron algorithm, under continual learning. We then show empirically on benchmark data sets that this simple and lightweight architecture outperforms other popular neurally inspired algorithms when also using a three-layer feedforward architecture. Overall, fruit flies evolved an efficient continual associative learning algorithm, and circuit mechanisms from neuroscience can be translated to improve machine computation. See Shen et al. (2023).

## Effects of Stochastic Coding on Olfactory Discrimination in Flies and Mice

Sparse coding can improve discrimination of sensory stimuli by reducing overlap between their representations. Two factors, however, can offset sparse coding's benefits: similar sensory stimuli have significant overlap, and responses vary across trials. To elucidate the effects of these two factors, we analyzed odor responses in the fly and mouse olfactory regions implicated in learning and discrimination—the mushroom body (MB) and the piriform cortex (PCx). We found that neuronal responses fall along a continuum from extremely reliable across trials to extremely variable or stochastic. Computationally, we show that the observed variability arises from noise within central circuits rather than sensory noise. We propose this coding scheme to be advantageous for coarse- and fine-odor discrimination. More reliable cells enable quick discrimination between dissimilar odors. For similar odors, however, these cells overlap and do not provide distinguishing information. By contrast, more unreliable cells are decorrelated for similar odors, providing distinguishing information, although these benefits only accrue with extended training with more trials. Overall, we have uncovered a conserved, stochastic coding scheme in vertebrates and invertebrates, and we identify a candidate mechanism, based on variability in a winner-take-all inhibitory circuit, that improves discrimination with training. See Srinivasan et al. (2023).

## BATMAN: Improved T-Cell-Receptor Cross-Reactivity Prediction Benchmarked on a Comprehensive Mutational Scan Database

Predicting T-cell-receptor (TCR) activation is challenging because of the lack of both unbiased benchmarking data sets and computational methods that are sensitive to small mutations to a peptide. To address these challenges, we curated a comprehensive database encompassing complete single amino acid mutational assays of 10,750 TCR–peptide pairs, centered around 14 immunogenic peptides against 66 TCRs. We then present an interpretable Bayesian model, called BATMAN, that can predict the set of peptides that activates a TCR. When validated on our database, BATMAN outperforms existing methods by 20% and reveals important biochemical predictors of TCR–peptide interactions. See Banerjee et al. (2024).

### PUBLICATIONS

Shen Y, Dasgupta S, Navlakha S. 2023. Reducing catastrophic forgetting with associative learning—a lesson from fruit flies. *Neural Comp* **35:** 1797–1819. doi:10.1162/neco_a_01615

Srinivasan S, Daste S, Modi M, Turner G, Fleischmann A, Navlakha S. 2023. Effects of stochastic coding on olfactory discrimination in flies and mice. *PLoS Biol* **21:** e3002206. doi:10.1371/journal.pbio.3002206

*In Press*

Banerjee A, Pattinson DJ, Wincek CL, Bunk P, Chapin SR, Navlakha S, Meyer HV. 2024. BATMAN: improved T cell receptor cross-reactivity prediction benchmarked on a comprehensive mutational scan database. bioRxiv doi:10.1101/2024.01.22.576714

# POPULATION GENETICS AND TRANSCRIPTIONAL REGULATION

**A. Siepel** B. Hassett A. Scheben
L. Liu S. Staklinski
Z. Mo A. Xue
L. Oliveira Y. Zhao

Our research focuses on two major areas: human population genetics and the evolution of transcriptional regulation in mammals. In addition, we have smaller collaborative projects on topics ranging from comparative genomics of bats (with W. Richard McCombie, CSHL) to determining the evolutionary trajectories of metastatic cancers (with Dawid Nowak, Weill Cornell Medicine). Overall, we focus on theoretical and computational research and do not generate our own data, but we often work closely with experimental collaborators. We are also broadly interested in probabilistic modeling, machine learning, and Bayesian statistics, and our research projects often involve the development of novel statistical and computational methods. In keeping with these broad interests, the research group is highly interdisciplinary, with members trained in computer science, genetics, and evolutionary biology, among other areas. The group size is stable, with one new Ph.D. student (Stephen Staklinski) joining recently.

Below, we describe recent progress in three main research areas.

## Inference of Demographic History and Natural Selection from Complete Genome Sequences

We have a long-standing interest in reconstructing the evolutionary histories of complex, structured populations from DNA sequence data. Our group developed ARGweaver, the first scalable method for genome-wide inference of the "ancestral recombination graphs," or ARGs. Building upon this foundation, recent work has focused on leveraging ARGs to detect signatures of natural selection. Our group developed a machine-learning method called selection inference using the ARG (SIA), refining predictions of selective sweeps and estimation of selection coefficients by exploiting a high-dimensional set of features extracted from inferred ARGs. This method has been used to identify selective sweeps in diverse populations, including chestnut-bellied monarchs from the Solomon Islands.

Recently, Ph.D. student Ziyi Mo extended SIA to take advantage of new "domain adaptation" methods borrowed from image processing literature to be more robust to potential mis-specification of the simulated data used for training (Mo and Siepel 2023). This is the first application of domain adaptation to applications of machine learning in population genetic inference. By implementing an established domain-adaptation technique based on a gradient reversal layer (GRL), originally introduced for image classification, we showed that the effects of simulation mis-specification can be substantially mitigated. Our analysis centered on two deep-learning methods in population genetics: our program, SIA, which deduces positive selection based on features of the ARG, and another program, ReLERNN, which estimates recombination rates from genotype matrices. With our improved domain-adaptive SIA model, dadaSIA, we obtained better estimates of selection coefficients at selected loci within the 1000 Genomes CEU population. We anticipate domain adaptation will be widely applicable in the growing use of supervised machine learning in population genetics.

## Transcriptional Regulation and Its Evolution in Primates

For several years, our research program in transcriptional regulation has focused on developing new methods for interpreting the rich nascent RNA sequencing (NRS) data generated using the powerful GRO-seq (global run-on and sequencing) protocol or its higher-resolution successor, PRO-seq (precision nuclear run-on sequencing).

Recently, Ph.D. student Lingjie Liu and postdoc Yixin Zhao developed a simple probabilistic model

that jointly describes the kinetics of transcription initiation, pause-escape, and elongation and the generation of nascent RNA sequencing read counts under steady-state conditions (Zhao et al. 2023). We then extend this initial model to allow for variability across cells in promoter-proximal pause site locations and steric hindrance of transcription initiation from paused RNA polymerases (RNAPs). In an extensive series of simulations over a broad range of parameters, we showed that this model enables accurate estimation of initiation and pause-escape rates even in the presence of collisions between RNAPs and variable elongation rates. Furthermore, we demonstrated through simulation and analysis of real data for human cell lines that pause-escape is often more strongly rate-limiting than conventional "pausing indices" would suggest. We also observed elevated occupancy at pause sites across many genes and noted that steric effects can significantly hinder initiation, resulting in a marked reduction in apparent initiation rates. Our modeling framework broadly applies to diverse nascent RNA sequencing data sets and various inference tasks.

In a related study, Ph.D. student Lingjie Liu and postdoc Yixin Zhao introduced a probabilistic model for systematically evaluating potential determinants of the local rate of transcription elongation based on NRS data (Liu et al. 2023). Derived from our unified model describing the kinetics of RNA polymerase II (Pol II) movement along the DNA template and the generation of NRS read counts at steady state, this model allows for a continuously variable elongation rate along the gene body. A generalized linear relationship with nearby genomic and epigenomic features determines the rate of elongation at each nucleotide. High-dimensional feature vectors are accommodated through a sparse-regression extension. We showed with simulations that the model allows accurate detection of associated features and accurate prediction of elongation rates. In an analysis of public PRO-seq and epigenomic data, several features such as DNA methylation, splice sites, RNA stem-loops, CTCF binding sites, and histone marks (including H3K36me3 and H4K20me1) were identified as strongly associated with reductions in local elongation rate. By contrast, low-complexity sequences and H3K79me2 marks were associated with increases in elongation rate. Our analysis revealed that cytosine nucleotides, especially when preceded by guanines and followed by adenines or thymines, are strongly linked to decreased local elongation rates. By contrast, thymines and A + T–rich k-mers are associated with increased rates. These associations are consistent across cell types and enable the model to predict features of held-out PRO-seq data. This study marks the first instance of genome-wide predictions of relative nucleotide-specific elongation rates using diverse genomic and epigenomic factors. Our predictions are accessible for the K562, $CD14^{+}$, MCF-7, and HeLa-S3 cell types through a UCSC Genome Browser track.

## Recent Collaborative Studies

In a collaborative effort led by postdoc Armin Scheben, we teamed up with W. Richard McCombie at Cold Spring Harbor Laboratory to advance our work in comparative genomics. Together, we utilized the Oxford Nanopore long-read platform to sequence the genomes of two bat species with key phylogenetic positions, the Jamaican fruit bat (*Artibeus jamaicensis*) and the Mesoamerican mustached bat (*Pteronotus mesoamericanus*), providing high-quality genome assemblies. Leveraging these resources, we conducted an extensive comparative genomic analysis encompassing a diverse selection of bats and other mammals, aiming to unveil genetic insights into the unique traits of bats, including their robust immune systems, exceptional powered flight abilities, and extended life spans (Scheben et al. 2023). The high-quality, long-read genome assemblies revealed significant changes in the interferon (IFN)-α locus, associated with immunity, in bats. This includes a notable contraction of IFN-α, leading to a shift in relative IFN-ω and IFN-α copy numbers. Our analysis reveals that three bat species have completely lost IFN-α genes. This shift toward IFN-ω expression may explain bats' increasing viral tolerance, contributing to their role as reservoirs for viruses transmissible to humans. Moreover, our analysis uncovered rapid evolution in antiviral genes activated by type I IFNs, including lineage-specific duplications of IFN-induced transmembrane genes and positive selection in IFIT2. Additionally, the study identified 33 tumor suppressors and six DNA-repair genes displaying signs of positive selection, offering potential explanations for bats' extended life spans and reduced cancer rates. The intricate immune systems of bats were shown to rely on a combination of bat-wide and lineage-specific evolution in the immune gene repertoire, demonstrating

diverse strategies. These findings provide new genomic resources for studying bats and enhance our understanding of their molecular evolution—with potential implications for human health.

In a collaborative study with Dawid Nowak at Weill Cornell Medicine, also led by postdoc Armin Scheben, we are developing computational methods for defining evolutionary spatiotemporal trajectories of metastatic cancers. This work, currently in revision at *Cancer Discovery*, aims to determine the poorly understood patterns by which primary tumors spread to or "seed" metastatic sites. Using CRISPR-Cas9 lineage tracing data from a novel somatically engineered mouse model named EvoCap (evolution in cancer of the prostate), developed by the Nowak lab, we defined patterns of metastatic seeding in prostate cancer. Our newly published R package, EvoTraceR, allows us to use phylogenetic methods to track distinct tumor lineages containing recordable barcodes amenable to cumulative edits by Cas9. We then use our publicly available pipeline to detect migration paths and seeding patterns based on a parsimonious approach that uses shared editing patterns between clonal lineages and minimizes migrations between tissues. We detected widespread intratumoral heterogeneity from the primary tumor in metastatic seeding, with only a few clonal lineages instigating most migration. The majority of metastatic seeding occurred through migrations from the prostate to bone or liver tissue. Our findings support the view of metastatic prostate cancer as a systemic disease driven by waves of aggressive clones expanding their niche, infrequently overcoming constraints that otherwise keep them confined in the primary or metastatic site. In work led by Armin Scheben and Ph.D. student Stephen Staklinski, we are now expanding our use of the developed methods to compare prostate cancer metastasis to other cancer types, including bladder cancer, while working toward a new Bayesian analysis framework that will allow us to quantify the uncertainty in migration paths and robustly infer common patterns of metastasis.

## PUBLICATIONS

Liu L, Zhao Y, Siepel A. 2023. DNA-sequence and epigenomic determinants of local rates of transcription elongation. bioRxiv doi:10.1101/2023.12.21.572932

Mo Z, Siepel A. 2023. Domain-adaptive neural networks improve supervised machine learning based on simulated population genetic data. *PLoS Genet* **19:** e1011032. doi:10.1371/journal.pgen.1011032

Scheben A, Mendivil Ramos O, Kramer M, Goodwin S, Oppenheim S, Becker DJ, Schatz MC, Simmons NB, Siepel A, McCombie WR. 2023. Long-read sequencing reveals rapid evolution of immunity- and cancer-related genes in bats. *Genome Biol Evol* **15:** evad148. doi:10.1093/gbe/evad148

Zhao Y, Liu L, Hassett R, Siepel A. 2023. Model-based characterization of the equilibrium dynamics of transcription initiation and promoter-proximal pausing in human cells. *Nucl Acids Res* **51:** e106. doi:10.1093/nar/gkad843

# COLD SPRING HARBOR LABORATORY FELLOWS

In 1986, Cold Spring Harbor Laboratory began a Fellows Program to encourage independent research by outstanding young scientists who, during their graduate studies, displayed exceptional promise of becoming leading scientists of the future. The purpose of this program is to provide an opportunity for these young scientists to work independently at the Laboratory for a period of three to five years on projects of their own choosing. Fellows are provided with a salary, research support, and technical assistance so that they can accomplish their goals free of distraction. The interactions among research groups at the Laboratory, combined with the program of courses and meetings on diverse topics in biology, contribute to a research environment that is ideal for scientific innovation by these Fellows.

The CSHL Fellows Program has been tremendously successful and has served as a paradigm for several analogous programs at other institutions, most recently a Fellows Program sponsored by the National Institutes of Health.

The success of the program is apparent from the list of distinguished alumni. Carol Greider—recipient of the 2009 Nobel Prize in Physiology or Medicine for her work on telomerase and telomere function—joined the Fellows Program in 1998. After completing her Fellow training, Carol was on the CSHL faculty for nine years, and she is currently a Distinguished Professor at the University of California, Santa Cruz. The first CSHL Fellow, Adrian Krainer (1986), is currently a Professor at the Laboratory, as are Chris Vakoc (2008), Florin Albeanu (2008), and Ivan Iossifov (2008). Former CSHL Fellows Lingbo Zhang (2013), Semir Beyaz (2017), and Hannah Meyer (2019) were recently promoted and all hold Assistant Professor positions at CSHL. Former Fellow Justin Kinney (2008) is an Associate Professor at CSHL. Scott Lowe (1995) is a Howard Hughes Medical Institute (HHMI) Investigator. After nearly 15 years on the CSHL faculty, he took on a Professorship at Memorial Sloan Kettering Cancer Center. Marja Timmermans (1998) was a member of the CSHL faculty for more than 17 years and recently accepted the Humboldt Professorship at the University of Tübingen. Eric Richards (1989) currently is the Vice President of Research and Senior Scientist at the Boyce Thompson Institute for Plant Research at Cornell University. David Barford (1991) is a Fellow of the Royal Society and Professor of Molecular Biology at the Institute of Cancer Research in London. Ueli Grossniklaus (1994) is Professor at the Institute of Plant Biology, University of Zürich, Switzerland. Térence Strick (2000) left at the end of his fellowship to become a Group Leader at the Institut Jacques Monod in Paris. Lee Henry (2000) joined HHMI's Janelia Farm in Ashburn and joined a project headed by Thomas Südhof. Ira Hall (2004) is a Professor at Yale University and the Director of the Yale Center for Genomic Health. Patrick Paddison, who had joined the Fellows Program in 2004, currently is an Associate Member at the Fred Hutchinson Cancer Research Center in Seattle, Washington.

**Corina Amor Vegas** has been a Fellow at the Laboratory since 2021. She joined us from Scott Lowe's laboratory at the Memorial Sloan Kettering Cancer Center, where she developed novel cell-based infusion strategies for manipulating cellular senescence in vivo. As a CSHL Fellow, Amor Vegas advanced her research program into the studies of organismal aging, showing how aspects of aging can be reversed by infusion of CAR-T cells that eliminate senescent cells.

# DECONSTRUCTING AGING WITH SENOLYTIC CAR T CELLS

**C. Amor Vegas** S. Chowdhury P. Garcia Baucells J. Gewolb G. Guo A. Harris G. Hwang E. Nnuji-John A. Rock

Our laboratory studies aging with the goal of improving health span and finding effective treatments for age-related diseases. A key hallmark of aging is the accumulation of senescent cells. Senescence is a stress response program whereby damaged cells stop replicating and performing their functions and instead become highly proinflammatory. In physiologic conditions, senescent cells are eliminated by the immune system. During aging, however, the increase in damaged cells plus the decreased efficiency of the immune system leads to the accumulation of senescent cells in tissues, where they generate a chronic proinflammatory milieu that drives age-related pathologies.

Inspired by the natural role of the immune system in eliminating senescent cells, we initially developed the first senolytic chimeric antigen receptor T cells (CAR T cells) (Amor et al., *Nature* **583:** 127 [2020]). At CSHL, we have worked on tailoring senolytic CAR T cells for the context of aging and age-related diseases. Thus, in our recent study we reported how low doses of anti-uPAR CAR T cells are highly effective at eliminating senescent cells in aged mice, resulting in significant improvements in whole-body metabolism. Interestingly, unlike other therapies, CAR T cells have the potential to persist and mediate their effects for multiple years. In our work we have observed how a single infusion of low-dose senolytic CAR T cells is sufficient to provide persistent activity throughout the whole lifespan of a mouse (Fig. 1).

## PUBLICATION

*In Press*

Amor C, Fernandez-Maestre I, Chowdhury S, Graham C, Nnuji-John E, Feucht J, Boyer J, Mezzadra R, Wereski MG, Levine R, et al. 2024. Senolytic CAR T cells reverse and prevent age-related metabolic dysfunction. *Nat Aging* **4:** 336–349. doi:10.1038/s43587-023-00560-5

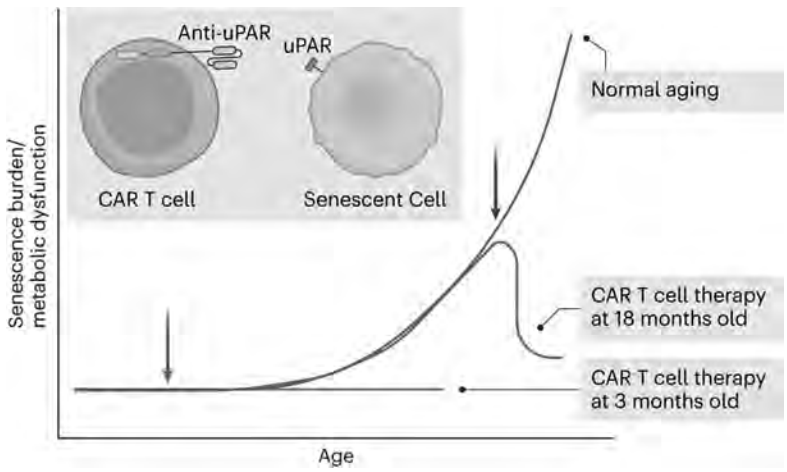



**Figure 1.** Senescent burden and metabolic dysfunction increase markedly with age. Our laboratory developed CAR T cells that target and eliminate uPAR-expressing senescent cells and recover age-related declines in metabolic function. When given at an early age, the CAR T cells persist and prevent age-related metabolic dysfunction up to 12 months later in mice. When given later in life (when metabolic dysfunction is already present), the CAR T cells eliminate urokinase-type plasminogen activator receptor (uPAR)-expressing senescent cells and recover metabolic function in aged mice. Created with Biorender.com.

# AUTHOR INDEX

# SCHOOL OF BIOLOGICAL SCIENCES

# SCHOOL OF BIOLOGICAL SCIENCES

## ADMINISTRATION

**Zachary Lippman, Ph.D.,** Professor and Director of Graduate Studies
**Monn Monn Myat, Ph.D.,** Associate Dean
**Alyson Kass-Eisler, Ph.D.,** Director of Academic Programs and Registrar
**Brianna Campmier, Ph.D.,** Assistant Director of Academic Programs
**Kimberly Creteur, M.Ed., M.S.Ed.,** Admissions and Recruitment Director
**Catherine Perez,** Administrative Coordinator of Academic Programs

## EXECUTIVE COMMITTEE

### Chair
Arkarup Banerjee

### Members
Camila dos Santos
Molly Gale Hammell (through April)
Rebecca Leshan
Zachary Lippman (ex officio)
Michael Lukey (from June)
W. Richard McCombie (from May)
David L. Spector (Director of Research)
Lloyd Trotman

### Student Representatives
Nikolas Holland, SBS
Steven Lewis, SBU

### Secretary
Monn Monn Myat

## ADMISSIONS COMMITTEE

### Chair
Adrian Krainer

### Members
Jeremy Borniger
Mikala Egeblad
David Jackson
Leemor Joshua-Tor
Zachary Lippman (ex officio)
David McCandlish
W. Richard McCombie
Stephen Shea
Jessica Tollkuhn
Linda Van Aelst

### Secretary
Kimberly Creteur

## QUALIFYING EXAM COMMITTEE

### Chair
Linda Van Aelst

### Members
W. Richard McCombie
Stephen Shea
Christopher Vakoc

## EXTERNAL ADVISORY COMMITTEE

**Keith Yamamoto (Chair)**
Professor, Cellular Molecular Pharmacology
Vice Chancellor for Science Policy and Strategy
University of California, San Francisco

**Victor Corces**
Professor, Department of Biology
Emory University

**Barbara Meyer**
Professor of Genetics and Development
University of California, Berkeley
Investigator, Howard Hughes Medical Institute

**Jodi Nunnari**
Director, Altos Bay Area Institute of Science
Editor-in-Chief, *Journal of Cell Biology*

**Joanna Wysocka**
Lorry Lokey Professor of Developmental Biology
Stanford University

# SCHOOL OF BIOLOGICAL SCIENCES DIRECTOR'S REPORT

## The School Welcomes New Staff Members

On June 26, we welcomed a new Assistant Director of Academic Programs, Brianna Campmier, to the School of Biological Sciences. Brianna earned a Ph.D. in Chemistry at Stony Brook University, where she worked on the creation of a long-lived fluorescent signal for neural mapping. Brianna has teaching experience at the undergraduate level and, for the past five years, has been the co-director of science, research, and technology at the high school level at Chaminade High School in Mineola, New York. Brianna has already brought new innovative ideas to the School's educational programs and helped launch the new Pedagogy course.

In March, Kimberley Graham, who had been with the School as an Administrative Assistant since 2007, went on medical leave. In her absence, the School hired a temporary replacement, Catherine Perez. Catherine was promoted to Administrative Coordinator of Academic Programs and became a permanent member of our team in August. Catherine is keeping us all organized and has been scheduling the more than 100 thesis committee meetings that take place each year at the School.

We are so fortunate to have Brianna and Catherine as part of our School of Biological Sciences family, and we are already seeing the positive impact of their contributions to the School.

## Entering Class of 2023

On August 21, 2023, the School welcomed the 25th incoming class, consisting of 12 new students: Mia Lin Amato, Emma Courtney, Todor Cvetanovic, Fernanda "Renee" Garcia Flores, Diego Hernandez, Nemanja Kutlesic, Masayuki "Moon" Nagai, Satwik Vasant Pasani, Kristina Shaw, Isabella Valentino, Yunxin Xie, and Chris Zhao.

### ENTERING CLASS OF 2023

**Mia Lin Amato,** Florida Atlantic University: B.S. in Neuroscience, Cellular Track (2023)
**Academic Mentor:** Camila dos Santos

**Emma Courtney,** Minerva University: B.S. in Biochemistry (2023)
**Academic Mentor:** Rebecca Leshan

**Todor Cvetanovic,** University of Belgrade: B.Sc./M.Sc. in Molecular Biology and Physiology (2022/2023)
**Academic Mentor:** Adrian Krainer

**Fernanda "Renee" Garcia Flores,** National Autonomous University of Mexico: B.S. in Genomic Sciences (2022)
**Academic Mentor:** Corina Amor Vegas

**Diego Hernandez,** Cornell University: B.S. in Biological Sciences (2022)
**Academic Mentor:** Jon Preall

**Nemanja Kutlešić,** University of Belgrade: B.Sc. in Molecular Biology and Physiology (2023); University of Camerino: B.Sc. in Biosciences and Biotechnology (2022)
**Academic Mentor:** Gabrielle Pouchelon

**Masayuki "Moon" Nagai,** DePauw University: B.A. in Bioinformatics (2022)
**Academic Mentor:** David Jackson

**Satwik Vasant Pasani,** All India Institute of Medical Sciences: M.B.B.S. in Medicine (2021)
**Academic Mentor:** Hannah Meyer

**Kristina Shaw,** University of Cambridge: B.A./M.S. in Biochemistry (2022)
**Academic Mentor:** Jeremy Borniger

**Isabella Valentino,** Villanova University: B.S. in Biology (2023)
**Academic Mentor:** Bruce Stillman

**Yunxin Xie,** Tsinghua University: B.S. in Biological Sciences (2023)
**Academic Mentor:** Florin Albeanu

**Chris Zhao,** Stony Brook University: B.S. in Biology (2021)
**Academic Mentor:** Alexander Gann

2023 Entering Class (*from left* to *right*): Diego Hernandez, Chris Zhao, Mia Lin Amato, Yunxin Xie, Fernanda "Renee" Garcia Flores, Satwik Vasant Pasani, Kristina Shaw, Masayuki "Moon" Nagai, Emma Courtney, Nemanja Kutlešić, Isabella Valentino, and Todor Cvetanovic.

## Graduation

On May 7, 2023, we celebrated the School's 20th graduating class. Eleven students were awarded Ph.D. degrees: Kathryn O'Neill and Chengxiang (Charlie) Yuan from the Entering Class of 2016; Lyndsey Aguirre, Sara Boyle, and Diogo Maia e Silva from the Entering Class of 2017; Ilgin Ergin, Amritha Varshini Hanasoge Somasundara, Asad Aziz Lakhani, Alexa Pagliaro, and Jenelys Ruiz Ortiz from the Entering Class of 2018; and Julia Wang from the Entering Class of 2019. Honorary degrees were awarded to Dr. Nouria Hernandez and Mr. Stephen Hall, who also gave the commencement address.

Stephen S. Hall is a journalist, author, and science communicator who has made outstanding contributions to the field of science journalism and to the public's understanding of fundamental science and its impact on human health. He holds a Bachelor of Arts degree in English Literature with Honors from Beloit College, where he also completed a minor in History. For more than three decades, Mr. Hall has written extensively on contemporary science, with a special focus on molecular biology, genetics, and the medical implications of new technologies in the life sciences. He has taught science journalism classes at Columbia University and is presently an Adjunct Professor in the Science, Health and Environmental Reporting Program at New York University. In 2012, he was named a Fellow of the John Simon Guggenheim Foundation, and he received the Walter Sullivan Award for Excellence in Science Journalism from the American Geophysical Union.

Nouria Hernandez is a prominent scientist and also a leader in higher education. She obtained a diploma in Biology from the Faculty of Sciences at the University of Geneva and a Ph.D. in Molecular Biology from the University of Heidelberg. Following her postdoctoral training at Yale University, Nouria joined the faculty at Cold Spring Harbor Laboratory, where she was selected as a Howard Hughes Medical Institute Investigator. In 2005, Nouria left Cold Spring Harbor Laboratory to join the University of Lausanne, in Switzerland, as Professor and Director of the Center for Integrative Genomics of the Faculty of Biology and Medicine. In 2015, Nouria was elected President of the University of Lausanne—the first woman to lead the university.

*Left to right*: CSHL Assistant Professor Peter Koo, CSHL President Bruce Stillman, Banbury Center Executive Director Rebecca Leshan, Diogo Maia e Silva, Lyndsey Maray Aguirre, Julia Huiming Wang, Jenelys Ruiz Ortiz, CSHL Chair Marilyn Simons, SBS Honorary Doctor of Science Nouria Hernandez, Asad A. Lakhani, Alexa Hope Pagliaro, Kathryn Shea O'Neill, Amritha Varshini Hanasoge Somasundara, Sara Elizabeth Boyle, Ilgin Ergin, and CSHL Director of Graduate Studies Zachary Lippman.

**2023 DOCTORAL RECIPIENTS**

| Student | Thesis advisor | Academic mentor | Current position |
|---|---|---|---|
| Sara Boyle | Bo Li | Jessica Tollkuhn | Data Analytics Consultant, Northern Trust |
| Ilgin Ergin | Semir Beyaz | Thomas Gingeras | Foundation and Corporate Relations Coordinator, CSHL |
| Amritha Varshini Hanasoge Somasundara | Camila dos Santos | Leemor Joshua-Tor | Senior Scientist, Memorial Sloan Kettering Cancer Center (Advisor: Ross Levine) |
| Yuzhao (Richard) Hu | Ullas Pedmale | Justin Kinney | Postdoctoral Scientist, Corteva Agriscience |
| Asad Aziz Lakhani | Jason Sheltzer | David Tuveson | Associate, Search and Evaluation, Autobahn Labs |
| Diogo Maia e Silva | Christopher Vakoc | Bruce Stillman | Research Fellow in Pathology, Mass General Hospital (Advisor: Keith Joung) |
| Alexa Pagliaro | Stephen Shea | John Inglis | Foundation and Corporate Relations Officer, Whitehead Institute, MA |
| Jenelys Ruiz Ortiz | Camila dos Santos | Molly Gale Hammell | Bioinformatics Analyst, The Rockefeller University |
| Julia Wang | Tatiana Engel | Saket Navlakha | Research Fellow, Albert Einstein College of Medicine (Advisors: Ruben Coen-Cagli and Adam Kohn) |
| Jonathan Werner | Jesse Gillis | Adam Siepel | Postdoctoral Fellow, University of Toronto (Advisor: Jesse Gillis) |
| Cole Wunderlich | Molly Gale Hammell | Adam Siepel | Exploring postdoctoral opportunities |
| Chengxiang (Charlie) Yuan | Je Hyuk Lee | Nicholas Tonks | Postdoctoral Fellow, Institute of Molecular and Cell Biology, Singapore (Advisor: Weimiao Yu) |

## 2023 THESIS DISSERTATION DEFENSES

### ENTERING CLASS OF 2016

**Chengxiang (Charlie) Yuan,** March 1, 2023

*Applications of oligonucleotide ligation in sequencing*

**Thesis Examining Committee**

| | |
|---|---|
| Chair: | **Jesse Gillis** |
| Research Mentor: | **Je Hyuk Lee** |
| Academic Mentor: | **Nicholas Tonks** |
| Committee Member: | **Dan Levy** |
| Committee Member: | **John Moses** |
| External Examiner: | **Ulf Landegren**<br>*Uppsala University* |

### ENTERING CLASS OF 2017

**Diogo Maia e Silva,** March 2, 2023

*Mechanisms of gene transcription by TP63 and the mediator complex in squamous pancreatic cancer*

**Thesis Examining Committee**

| | |
|---|---|
| Chair: | **Molly Gale Hammell** |
| Research Mentor: | **Christopher Vakoc** |
| Academic Mentor: | **Bruce Stillman** |
| Committee Member: | **Jason Sheltzer** |
| External Examiner: | **Leif Ellisen**<br>*Massachusetts General Hospital, Harvard Medical School* |

**Sara Boyle,** April 3, 2023

*The central amygdala encodes nutritional properties and modulates weight gain*

**Thesis Examining Committee**

| | |
|---|---|
| Chair: | **Florin Albeanu** |
| Research Mentor: | **Bo Li** |
| Academic Mentor: | **Jessica Tollkuhn** |
| Committee Member: | **Alexei Koulakov** |
| External Examiner: | **Alfredo Fontanini**<br>*Stony Brook University* |

**Yuzhao (Richard) Hu,** July 17, 2023

*Role of cryptochromes in chromatin remodeling and DNA damage repair*

**Thesis Examining Committee**

| | |
|---|---|
| Chair: | **David Jackson** |
| Research Mentor: | **Ullas Pedmale** |
| Academic Mentor: | **Justin Kinney** |
| Committee Member: | **Christopher Vakoc** |
| External Examiner: | **Fredy Barneche**<br>*l'Ecole Normale Superieure* |

**Cole Wunderlich,** October 10, 2023

*Quantifying transposable element expression in single-cell RNA sequencing data*

**Thesis Examining Committee**

| | |
|---|---|
| Chair: | **Alexander Dobin** |
| Research Mentor: | **Molly Gale Hammell** |
| Academic Mentor: | **Adam Siepel** |
| Committee Member: | **Camila dos Santos** |
| Committee Member: | **Robert Martienssen** |
| Committee Member: | **Hannah Meyer** |
| External Examiner: | **Joshua Dubnau**<br>*Stony Brook University* |

### ENTERING CLASS OF 2018

**Asad Aziz Lakhani,** March 9, 2023

*The role of recurrently observed aneuploidy in tumorigenesis*

**Thesis Examining Committee**

| | |
|---|---|
| Chair: | **Robert Martienssen** |
| Research Mentor: | **Jason Sheltzer** |
| Academic Mentor: | **David Tuveson** |
| Committee Member: | **Christopher Vakoc** |
| External Examiner: | **Neil Ganem**<br>*Boston University* |

**Jenelys Ruiz,** April 19, 2023

*Characterizing models to study cellular composition changes in response to pregnancy hormones in the human mammary gland*

**Thesis Examining Committee**

| | |
|---|---|
| Chair: | **Adam Siepel** |
| Research Mentor: | **Camila dos Santos** |
| Academic Mentor: | **Molly Hammell** |
| Committee Member: | **Semir Beyaz** |
| Committee Member: | **Peter Koo** |
| External Examiner: | **Senthil Muthuswamy**<br>*National Institutes of Health* |

**Ilgin Ergin,** April 24, 2023

*Metabolic regulation of anti-tumor immunity*

**Thesis Examining Committee**

| | |
|---|---|
| Chair: | **Christopher Vakoc** |
| Research Mentor: | **Semir Beyaz** |
| Academic Mentor: | **Thomas Gingeras** |
| Committee Member: | **Camila dos Santos** |
| Committee Member: | **Michael Lukey** |
| External Examiner: | **Brian Sheridan**<br>*Stony Brook University* |

**Alexa Pagliaro,** April 26, 2023

*A disrupted parvalbumin network state impairs maternal behavior in a mouse model of Rett Syndrome*

**Thesis Examining Committee**

| | |
|---|---|
| Chair: | **Florin Albeanu** |
| Research Mentor: | **Stephen Shea** |
| Academic Mentor: | **John Inglis** |
| Committee Member: | **Anthony Zador** |
| External Examiner: | **David Schneider**<br>*New York University* |

**Jonathan Werner,** September 13, 2023

*Transcriptomic approaches for investigating developmental lineage: exploiting the X-chromosome as a marker for lineage specification and quantifying the lineage fidelity of neural organoid systems*

**Thesis Examining Committee**
Chair: **Dan Levy**
Research Mentor: **Jesse Gillis**
Academic Mentor: **Adam Siepel**
Committee Member: **Camila dos Santos**
External Examiner: **Christine M. Disteche**
*University of Washington School of Medicine*

**Danielle Ciren,** September 26, 2023

*Decoding* cis*-regulatory control and evolution of conserved and divergent phenotypes in plants*

**Thesis Examining Committee**
Chair: **Jessica Tollkuhn**
Research Mentor: **Zachary Lippman**
Academic Mentor: **Ullas Pedmale**
Committee Member: **Adam Siepel**
External Examiner: **Nathan Springer**
*Bayer*

### ENTERING CLASS OF 2019

**Julia Wang,** April 17, 2023

*Interpretable brain state manifold for characterizing heterogeneity*

**Thesis Examining Committee**
Chair: **Anthony Zador**
Research Mentor: **Tatiana Engel**
Academic Mentor: **Saket Navlakha**
Committee Member: **Jeremy Borniger**
External Examiner: **Maxim Bazhenov**
*University of California, San Diego*

## DOCTORAL THESIS RESEARCH

| Student | Academic mentor | Research mentor | Thesis research |
|---|---|---|---|
| **ENTERING CLASS OF 2018** | | | |
| **King Hei (Teri) Cheng**<br>*Robert and Theresa Lindsay Fellow* | Adrian Krainer | Robert Martienssen | Investigating the role of RNAi in maintaining genome stability |
| **Marie Dussauze**<br>*Florence Gould Fellow*<br>*Annette Kade Fellow* | Stephen Shea | Florin Albeanu | Sensorimotor neural representations in the olfactory cortex |
| **Connor Fitzpatrick**<br>*Ainslie Family Fellow* | Robert Martienssen | Christopher Vakoc | Investigating the role of OCT1 in the neuroendocrine identity of small-cell lung cancer (SCLC) |
| **Ziyi Mo**<br>*Gladys and Roland Harriman Foundation Fellow* | David McCandlish | Adam Siepel | Scalable and robust deep-learning methods power evolutionary-genetic studies of BioBank-scale population genomic data |
| **ENTERING CLASS OF 2019** | | | |
| **Leah Braviner**<br>*Elizabeth Sloan Livingston Fellow* | Linda Van Aelst | Leemor Joshua-Tor | Structural and biochemical characterisation of CSR-1 and C04F12.1, two closely related WAGO clade Argonautes in *C. elegans* |
| **Patrick Cunniff**<br>*NIH Predoctoral Trainee* | David Jackson | Christopher Vakoc | Aberrantly expressed Krüppel-like factor 5 (KLF5) co-opts RUVBL1/2 to drive PDAC proliferation |
| **Michael Passalacqua**<br>*Hearst Foundation Scholar* | Thomas Gingeras | Jesse Gillis | Conservation of coexpression in plants uncovers shared regulatory programs, enhancing cross-species comparisons |
| **Leonardo Jared Ramirez Sanchez**<br>*Gonzalo Rio Arronte Fellow* | Christopher Hammell | Bo Li | Understanding the neural representations of positive and negative reinforcement in the striosome compartment |
| **Nicole Sivetz**<br>*NIH Predoctoral Trainee* | Camila dos Santos | Mikala Egeblad | Lipopolysaccharide-induced sterile endotoxemia suppresses pancreatic cancer liver metastasis |
| **Ziqi (Amber) Tang**<br>*Charles A. Dana Fellow* | Alea Mills | Peter Koo | Building a feature extraction model for RNA regulatory mechanisms |
| **Shushan Toneyan**<br>*Crick-Clay Fellow* | David Stewart | Peter Koo | Integrative modeling of gene expression using deep learning |
| **ENTERING CLASS OF 2020** | | | |
| **Salomé Carcy**<br>*Starr Centennial Scholar* | John Inglis | Hannah Meyer | Regulation of T-cell development in the murine and human thymus |

(*continued*)

## DOCTORAL THESIS RESEARCH (*continued*)

| Student | Academic mentor | Research mentor | Thesis research |
|---|---|---|---|
| **Jed de Ruiter-Swain**<br>*Leslie C. Quick, Jr. Fellow* | Tobias Janowitz | Michael Lukey | Astrocyte–cancer cell metabolic coupling as a novel therapeutic target in brain-metastatic breast cancer |
| **Iacopo Gentile**<br>*George and Marjorie Anderson Fellow* | Adam Siepel | Zachary Lippman | Dissecting the evolution and redundancy of the CLE multigene family in angiosperms |
| **Emily Isko**<br>*NIH Predoctoral Trainee*<br>*National Science Foundation Fellow* | Stephen Shea | Arkarup Banerjee and Anthony Zador | Evolving neural circuits underlying vocal behaviors in closely related murine species |
| **Jessica Kahng**<br>*NIH Predoctoral Trainee* | Rebecca Leshan | Lucas Cheadle | A molecular characterization of sensory-induced oligodendrocyte precursors cell (OPC) engulfment of thalamocortical synapses in the visual cortex |
| **Sam Kleeman**<br>*David H. Koch Fellow* | Adrian Krainer | Tobias Janowitz | Molecular mechanisms of anti-NMDA receptor autoantibodies |
| **Matty Peacey**<br>*George and Marjorie Anderson Fellow* | Hiro Furukawa | Andrea Schorn | Regulatory interactions between 3′ tRNA-fragments and domesticated LTR-retrotransposons |
| **Xiaoyue (Mike) Zheng**<br>*Goldberg-Lindsay Fellow* | Saket Navlakha | Arkarup Banerjee | The role of the periaqueductal gray in controlling vocal flexibility in singing mice |
| **Xingyu (CiCi) Zheng**<br>*William R. Miller Fellow* | Helen Hou | Saket Navlakha | Statistical analysis and modeling of networks in natural systems |
| **ENTERING CLASS OF 2021** | | | |
| **Hoda Ansari**<br>*Elizabeth Sloan Livingston Fellow* | Semir Beyaz | Stephen Shea and Arkarup Banerjee | Investigating the role of basal amygdala projection neurons to the auditory cortex in maternal behavior |
| **Paul Bunk**<br>*Starr Centennial Scholar* | Christopher Vakoc | Semir Beyaz | The role of PPARδ signaling in governing antitumor immunity |
| **Nikolas Holland**<br>*Ainslie Family Fellow* | Linda Van Aelst | Jeremy Borniger | Characterizing a bidirectional relationship between the locus coeruleus and breast cancer progression via efferent and afferent neurons of the tumor microenvironment |
| **Sessen Daniel Iohannes**<br>*Edward and Martha Gerry Fellow* | Doreen Ware | David Jackson | Uncovering the mechanisms of compensation between paralogous genes in the maize meristem |
| **Emmanuella Nnuji-John**<br>*George and Marjorie Anderson Fellow* | Molly Gale Hammell | Corina Amor | Exploring the role of neuro-regulatory molecules in cellular senescence and aging |
| **Rachel Polfer**<br>*Bristol-Myers Squibb Fellow* | David Jackson | Hiro Furukawa | Physiological representation of the CALHM channel and antigen diversity in NMDA receptor autoantibodies |
| **Yihan (Leonie) Qin**<br>*Farish-Gerry Fellow* | Hiro Furukawa | Michael Lukey | Modeling the interplay of cancer clonal evolution and immunosurveillance in the mouse |
| **Lucía Téllez Pérez**<br>*Fulbright Scholar* | John Inglis | Camila dos Santos | Impact of female hormones on normal and malignant male breast development |
| **ENTERING CLASS OF 2022** | | | |
| **Matías Gleason**<br>*Robert and Theresa Lindsay Fellow* | David McCandlish | David Jackson | Untangling the mechanisms of cell-to-cell mRNA trafficking in plants |
| **Yuriko Nishino**<br>*Goldberg-Lindsay Fellow* | David Tuveson | Leemor Joshua-Tor | Elucidating the molecular mechanism of poly(UG) tail addition to RNAs by the ribonucleotidyltransferase RDE-3 |
| **Kaeli Rizzo**<br>*Edward and Martha Gerry Fellow* | Bruce Stillman | Peter Koo | Uncovering cell type–specific *cis*-regulatory grammar of enhancers |

(*continued*)

**DOCTORAL THESIS RESEARCH** (*continued*)

| Student | Academic mentor | Research mentor | Thesis research |
|---|---|---|---|
| **Stephen Staklinski**<br>*Starr Centennial Scholar* | Bruce Stillman | Adam Siepel | Understanding cancer metastasis through evolvable barcoding |
| **Maha Syed**<br>*George and Marjorie Anderson Fellow* | Benjamin Cowley | Hannah Meyer | Genetics of T-cell development |
| **Yaman Thapa**<br>*Jordan and Thomas A. Saunders III Neuroscience Fellow* | Camila dos Santos | Arkarup Banerjee and Benjamin Cowley | Computational and neural insights into the temporal patterning of songs in singing mice |

## Teaching Award

The School awarded its 16th Winship Herr Award for Excellence in Teaching this year to Dr. Christopher Hammell, an instructor of the Scientific Reasoning and Logic core course, and Dr. Stephen Shea, the lead instructor of the Specialized Disciplines course in Systems Neuroscience. Chris and Steve were chosen by the first-year students for this award. Here is what one student said in their nomination: "I chose Chris because he got me excited about a topic I had minimal interest in. His teaching style was logical rather than factual. He gave us all the facts needed to answer a question and left the critical thinking part for us. He gave us the space to think for ourselves and ask relevant questions." Another said of Steve in their nomination: "Stephen introduced many prior pieces of research to explain key concepts, and he conveyed his passion and enthusiasm by explaining what was cool about that research."

## Faculty Changes

In 2023, no new faculty joined the School. However, we did have three faculty members depart. Tatiana Engel, a faculty member since 2017, moved her laboratory to Princeton University. Mikala Egeblad, a faculty member since 2009, moved her laboratory to The Johns Hopkins University. And Bo Li, a faculty member since 2008, moved his laboratory to Westlake University in China. Tatiana, Mikala, and Bo were all involved in teaching and mentoring. Mikala and Bo were also members of the admissions committee. We wish them luck in their future endeavors.

## Academic Mentoring

The School takes great pride in the mentoring that it offers its students. One example is our two-tiered mentoring approach, whereby each student chooses both an academic and a research mentor. The academic mentor is a critical advisor during the intensive coursework of the first term, during their rotations, and when identifying a suitable research mentor. Furthermore, the academic mentor continues to guide them throughout their doctoral experience, often serving as important advocates for the students. Entering students select, by mutual agreement, a member of the research or nonresearch faculty to serve as their academic mentor. This program continues to receive much support from the faculty who volunteer to be academic mentors, and it has rightfully become a vital ingredient in our success. The Academic Mentors for the Entering Class of 2023 are:

Mia Lin Amato
Emma Courtney
Todor Cvetanovic
Fernanda "Renee" Garcia Flores
Diego Hernandez
Camila dos Santos
Rebecca Lehsan
Adrian Krainer
Corina Amor
Jon Preall

| | |
|---|---|
| Nemanja Kutlešić | Gabrielle Pouchelon |
| Masayuki "Moon" Nagai | David Jackson |
| Satwik Vasant Pasani | Hannah Meyer |
| Kristina Shaw | Jeremy Borniger |
| Isabella Valentino | Bruce Stillman |
| Yunxin Xie | Florin Albeanu |
| Chris Zhao | Alexander Gann |

## Recruiting Efforts

To attract students to the program, administrators, students, and faculty attended large national graduate school fairs and conferences and participated in both virtual recruitment presentations and Q&A sessions arranged by the School. Additionally, emails were sent to personalized contacts

**SCHOOL OF BIOLOGICAL SCIENCES 2023 RECRUITMENT SCHEDULE**

| Event | Location | Date |
|---|---|---|
| University of Maryland McNair Scholars Research Conference Graduate School Fair | Hyattsville, Maryland | March 16–19 |
| National Conference on Undergraduate Research (NCUR) Annual Conference | Eau Claire, Wisconsin | April 12–15 |
| SUNY Undergraduate Research Conference (SURC) Graduate School Fair | Bronx, New York | April 14 |
| American Association of Cancer Research Annual Meeting | Orlando, Florida | April 14–19 |
| University of Washington, Office of Minority Affairs & Diversity Pacific Northwest McNair Undergraduate Research Conference Annual Symposium | Seattle, Washington | April 17–19 |
| The College of William & Mary, Department of Biology Information Session | Virtual Presentation | May 2 |
| SUNY Old Westbury, ICaRE Cancer Symposium Graduate School Fair | Old Westbury, New York | May 4 |
| NIH Graduate and Professional School Fair Annual Event | Bethesda, Maryland | July 19 |
| University at Buffalo McNair Research Conference Graduate School Fair | Virtual Forum | July 20–21 |
| Stony Brook University Simons STEM Scholars Program Information Session | Cold Spring Harbor, New York | July 25 |
| NIH Graduate and Professional School Fair Annual Event | Virtual Presentation | July 26 |
| Adelphi University, STEM Honors College Information Session | Garden City, New York | September 20 |
| Farmingdale State College SUNY, Research Aligned Mentorship (RAM) Program Information Session | East Farmingdale, New York | September 21 |
| Mount Holyoke College Information Session | Virtual Presentation | September 26 |
| LaGuardia Community College Information Session | Long Island City, New York | September 28 |
| NYS Graduate School Information Session for Postbaccalaureates Information Session | Virtual Forum | October 2 |
| Women's Colleges and Universities Diversity Career Expo Information Session | Virtual Forum | October 3 |
| Hunter College MARC and MBRS/RISE Information Session | New York, New York | October 4 |
| CSHL Recruitment Event Open House | Cold Spring Harbor, New York | October 7 |
| Rio Hondo College Information Session | Virtual Presentation | October 20 |
| Cancer Biology Training Consortium (CABTRAC) Annual Retreat | Newport Beach, California | October 22–24 |
| Yale University STARS Program Information Session | Virtual Presentation | October 25 |
| Society for Advancement of Chicanos and Native Americans in Science (SACNAS) Conference National Conference | Portland, Oregon | October 26–28 |
| American Society for Human Genetics Annual Meeting | Washington, D.C. | November 1–5 |
| CSHL Recruitment Event Open House | Virtual Forum | November 3 |
| Metropolitan Association of Colleges and University Biologists (MACUB) at the University of Bridgeport Annual Meeting | Bridgeport, Connecticut | November 4 |
| Society for Neuroscience Annual Meeting Graduate School Fair | Washington, D.C. | November 11–15 |
| Annual Biomedical Research Conference for Minoritized Scientists (ABRCMS) National Conference | Phoenix, Arizona | November 15–18 |
| American Society for Cell Biology Annual Meeting | Boston, Massachusetts | December 2–6 |

and an electronic mailing list of more than 50,000 individuals who receive information from the Cold Spring Harbor Press or have attended Meetings or Courses at the lab. We are grateful to these departments for sharing this contact list.

We received 971 applications for the Entering Class of 2023, the highest number to date. For admissions season, we hosted virtual interview sessions in January followed by an in-person, accepted student campus visit in March.

The School hosted a recruitment open house on October 7 at which approximately 40 undergraduate, post-baccalaureate, and graduate students came to campus to learn about CSHL's Undergraduate Research Program, the PREP postbaccalaureate program, the Ph.D. graduate program, and the postdoctoral program. They had a chance to meet with School staff, faculty, current trainees, and PREP scholars. They also heard faculty research talks and personal stories of our current students' scientific journeys and interacted with our campus affinity groups.

## Students from Other Institutions

Students enrolled in the School of Biological Sciences account for approximately half of the total graduate student population here at CSHL; the other half comprises visiting graduate students from other universities who have decided to conduct some or all of their thesis research in CSHL faculty members' laboratories. A large fraction of these students are from Stony Brook University (SBU), via a program established between CSHL and SBU many years ago. Over the years we have built relationships with other institutions around the world, enabling their students to conduct research here at CSHL—including the Zucker School of Medicine at Hofstra/Northwell. The School provides a contact person for the students and maintains relationships with the administrators from their home institutions. These students are fully integrated into the CSHL community and receive all the necessary assistance as they navigate the complexities of performing doctoral research away from their home institutions. The following students joined us from SBU and Hofstra in 2023:

| STUDENT | CSHL RESEARCH MENTOR | HOME PROGRAM |
|---|---|---|
| Javier Anduaga | Nicholas Tonks | SBU Molecular and Cellular Pharmacology |
| Jack Bauer | Leemor Joshua-Tor | SBU Molecular and Cellular Pharmacology |
| Zarmeena Dawood | Florin Albeanu | Hofstra M.D./Ph.D. |
| Lu Jia | Adrian Krainer | SBU Molecular and Cell Biology |
| Irene Nozal | Helen Hou | SBU Neuroscience |
| Katie Tam | Christopher Vakoc | Hofstra M.D./Ph.D. |
| Kamil Taneja | Christopher Vakoc | SBU Genetics/MSTP |
| Alice Wang | Tobias Janowitz | SBU Genetics/MSTP |
| Jialin Zhang | Peter Westcott | Biomedical Engineering |

## Graduate Student Symposium

Each year the students participate in two Graduate Student Symposia held at the Laboratory's Genome Research Center in Woodbury: one in June, the other in December. Each Symposium consists of senior students giving short talks, while coffee breaks and lunch provide opportunities for more informal interactions. The prize for best talk for the June session was awarded to Vahag Kechejian (Hofstra, Vakoc laboratory), and for the December session was awarded to Patrick Cunniff (SBS, Vakoc laboratory). We are grateful to the student chairs—Lucía Téllez Pérez and Iacopo Gentile (SBS) and Isobel Bolger and Timothy Maher (SBU)—for helping make the Symposium a great success.

## Graduate Student and Postdoctoral Fellow Departures

Each year brings not only new arrivals, but also departures. The following graduate students and postdoctoral fellows left the Lab during 2023:

POSTDOCTORAL FELLOWS

Jose Adrover Montemayor
Cina Aghamohammadi
Aktan Alpsoy
Florencia Campetella Mayoral
Chen Chen
Taemoon Chung
Samantha Cyrill
Jaya Balan Devasahayam
Talitha Forcier
Nikita Francis
Rasmani Hazra
Xueyan He
Aleksander Kaplan
Christopher Langdon
Shujing Li
Ruchi Lohia
Andre Machado Xavier
Evdokia Michalopoulou
Rodrigo Muñoz Castañeda
Sukalp Muzumdar
Karthick Natarajan
Farhad Pashakhanloo
Yanliang Shi
Daniele Silva Pereira Rosado
Dongyan Song
Laura Taylor
Pavel Tolmachev
Anand Vasudevan
Ledong Wan
Juanjuan Xie
Xiaosa Xu

GRADUATE STUDENTS

Seamus Balinth
Isobel Bolger
Sara Boyle
Gabriel Brill
Cristian Cleary
Katherine Denney
Ilgin Ergin
Amritha Varshini Hanasoge Somasundara
Yuzhao Hu
Shruti Iyer
Asad Aziz Lakhani
Tristan Liu
Diogo Maia e Silva
Craig Marshall
Conor McGrory
Alexa Pagliaro
Flaviani Pierdoná
Zhe Qian
Jenelys Ruiz Ortiz
Margaret Shevik
Padmina Shrestha
Jennifer Thalappillil
Pablo Villar
Jonathan Werner
Cole Wunderlich
Xinmeng Xu
Chengxiang Yuan

## Executive Committee

The School's Executive Committee, in its monthly meetings, provides year-round direction for the School and its students through its invaluable policy recommendations. We thank faculty members Arkarup Banerjee, Camila dos Santos, Molly Gale Hammell, Rebecca Leshan, Michael Lukey, W. Richard McCombie, David Spector, and Lloyd Trotman for their service in 2023. We also thank the student representatives Nikolas Holland (SBS) and Steven Lewis (SBU), who contributed to discussions and provided useful suggestions and feedback from their colleagues.

## The School Continues to Benefit from Generous Benefactors

It is the support of generous donors and benefactors that allows us to run a unique and successful graduate program. In 2023, we received a new gift from Simons Foundation International. The summer Undergraduate Research Program (URP) also received a gift from the estate of Fred Goldberg. Fred Goldberg was an URP participant in 1961 and a longtime friend of the program.

## Student and Alumni Achievements

To date, 148 students have received their Ph.D. degree from the School. Forty-three of our graduates have secured tenure track faculty positions (although three have now left these positions for industry). Six have been promoted to Associate Professor (often conferring tenure) and 12 are full Professors. Our graduates have also moved into influential positions in administration, publishing, consulting, science communication, and industry. In 2023, the following graduates either started new jobs or received major promotions:

| Name | Position |
|---|---|
| Brianna Bibel | Visiting Professor of Biochemistry, Saint Mary's College of California |
| Allison Blum | Senior Director of Scientific Affairs, Pfizer Research and Development |
| Sara Boyle | Data Analytics Consultant, Northern Trust |

| | |
|---|---|
| Ilgin Ergin | Foundation and Corporate Relations Coordinator, CSHL |
| Kristina Grigaityte | Associate Director of Data Science, Merck |
| Eyal Gruntman | Assistant Professor, University of Toronto, Scarborough |
| Yuzhao Hu | Postdoctoral Scientist, Corteva Agriscience |
| Christopher Krasniak | Research Data Scientist, Somatus |
| Asad Aziz Lakhani | Associate in Search and Evaluation, Autobahn Labs |
| Katie Liberatore | Executive Director of Operations, Cibus |
| Shaina Lu | Senior Data Scientist in R&D, Miltenyi Biotec |
| Diogo Maia e Silva | Research Fellow in Pathology, Mass General Hospital |
| Marco Mangone | Professor, Arizona State University |
| Katie Meze | Senior Scientist, Structural Biology at Sanofi |
| Elizabeth Nakasone | Junior Researcher in Translational and Clinical Research, University of Hawai'i Cancer Center |
| Hassana Oyibo | Associate Director of Synthetic Biology, Tessera Therapeutics |
| Alexa Pagliaro | Foundation and Corporate Relations Officer, Whitehead Institute |
| Yevgeniy Plavskin | Visiting Clinical Associate Professor, New York University |
| Amy Rappaport | Senior Director of Translational Immunology, Gritstone Bio |
| Jenelys Ruiz Ortiz | Bioinformatics Analyst, The Rockefeller University |
| Catherine Seiler | Director of Operational Excellence in Translation Medicine, AstraZeneca |
| Martyna Sroka | Scientist, Relay Therapeutics |
| Gowan Tervo | Group Leader, HHMI Janelia Research Campus |
| Charles Underwood | Professor, Radboud University in the Netherlands |
| Julia Wang | Research Fellow, Albert Einstein College of Medicine |
| Wei Wei | Professor, University of Chicago |
| Jonathan Werner | Analysis Supplier, University of Toronto |
| Ran Yan | Senior Scientist, Parthenon Therapeutics |
| Chengxiang Yuan | Postdoctoral Fellow, Institute of Molecular and Cell Biology in Singapore |

In 2023, our alumni were successful in receiving the following prestigious awards and fellowships:

- **Jacqueline Giovanniello** won the UCLA Brain Research Institute Scheibel distinguished postdoc award. She was also named a 2023 Early Career Policy Ambassador by the Society for Neuroscience. She was also awarded an NIH K99/R00 Pathway to Independence Award.
- **Elvin Wagenblast** received a 2023 Damon Runyon-Rachleff Innovation Award for early-career scientists.
- **Justus Kebschull** received a Pershing Square Foundation's Maximizing Innovation in Neuroscience Discovery (MIND) Prize.
- **Alberto Corona** was named a 2023 Burroughs Wellcome Fund Postdoctoral Diversity Enrichment Program Recipient.
- **Arkarup Banerjee** received a Klingenstein-Simons Fellowship for Neuroscience.
- **Monica Dus** was selected for the Guggenheim Fellowship and won the Ajinomoto Award for Gustation. She was also appointed as a White House Fellow, United States Department of the Navy.
- **Kristen Delevich** was awarded the Dean's outstanding junior faculty research award by the College of Veterinary Medicine, Washington State University.
- **Niraj Tolia** was named a Fellow of the American Society of Tropical Medicine and Hygiene (FASTMH).

In 2023, our current students were successful in receiving the following prestigious awards and fellowships:

- **Patrick Cunniff** was awarded an NIH Ruth L. Kirschstein National Research Service Award Individual Fellowship.

- **Nicole Sivetz** was awarded an NIH Ruth L. Kirschstein National Research Service Award Individual Fellowship.
- **Lucía Téllez Pérez** was awarded a fellowship from the la Caixa Foundation.

## 2023 STUDENT (CURRENT OR PREVIOUS) PUBLICATIONS


**Aguirre L**, Hendelman A, Hutton SF, McCandlish DM, Lippman ZB. 2023. Idiosyncratic and dose-dependent epistasis drives variation in tomato fruit size. *Science* **382:** 315–320.

Amor C, Fernández-Maestre I, Chowdhury S, Ho YJ, Nadella S, Graham C, Carrasco SE, **Nnuji-John E**, Feucht J, Hinterleitner C, et al. 2023. Prophylactic and long-lasting efficacy of senolytic CAR T cells against age-related metabolic dysfunction. Res Sq doi:10.21203/rs.3.rs-3385749/v1

**Berube B**, Ernst E, Cahn J, Roche B, de Santis Alves C, Lynn J, Scheben A, Siepel A, Ross-Ibarra J, Kermicle J, Martienssen R. 2023. *Teosinte Pollen Drive* guides maize diversification and domestication by RNAi. bioRxiv doi: 10.1011/10:2023.07.12.548689

Carter JA, Matta B, Battaglia J, Somerville C, **Harris BD**, LaPan M, Atwal GS, Barnes BJ. 2023. Identification of pan-cancer/testis genes and validation of therapeutic targeting in triple-negative breast cancer: Lin28a-based and Siglece-based vaccination induces antitumor immunity and inhibits metastasis. *J Immunother Cancer* **11:** e007935.

**Corona A**, Choe J, Muñoz-Castañeda R, Osten P, Shea SD. 2023. A circuit from the locus coeruleus to the anterior cingulate cortex modulates offspring interactions in mice. *Cell Rep* **42:** 112771.

**de Ruiter Swain J**, Michalopoulou E, Noch EK, Lukey MJ, Van Aelst L. 2023. Metabolic partitioning in the brain and its hijacking by glioblastoma. *Genes Dev* **37:** 681–702.

Ferrer M, Mourikis N, Davidson EE, **Kleeman SO**, Zaccaria M, Habel J, Rubino R, Gao Q, Flint TR, Young L, et al. 2023. Ketogenic diet promotes tumor ferroptosis but induces relative corticosterone deficiency that accelerates cachexia. *Cell Metab* **35:** 1147–1162.

Funamizu A, **Marbach F**, Zador AM. 2023. Stable sound decoding despite modulated sound representation in the auditory cortex. *Curr Biol* **33:** 4470–4483.

Gao Y, He XY, Wu XS, Huang YH, **Toneyan S**, Ha T, Ipsaro JJ, Koo PK, Joshua-Tor L, Bailey KM, et al. ETV6 dependency in Ewing sarcoma by antagonism of EWS-FLI1-mediated enhancer activation. *Nat Cell Biol* **25:** 298–308.

Girish V*, **Lakhani AA***, Thompson SL*, Scaduto CM*, Brown LM, Hagenson RA, Sausville EL, Mendelson BE, Kandikuppa PK, Lukow DA, et al. 2023. Oncogene-like addiction to aneuploidy in human cancers. *Science* **381:** eadg4521.

Guillotin B, Rahni R, **Passalacqua M**, Mohammed MA, Xu X, Raju SK, Ramírez CO, Jackson D, Groen SC, Gillis J, Birnbaum KD. 2023. A pan-grass transcriptome reveals patterns of cellular divergence in crops. *Nature* **617:** 785–791.

**Hu Y**, Rosado D, Lindbäck LN, Micko J, Pedmale UV. 2023. Cryptochromes and UBP12/13 deubiquitinases antagonistically regulate DNA damage response in *Arabidopsis.* bioRxiv doi:10.1011/16:2023.01.15.524001

Hur SK, Somerville TDD, Wu XS, **Maia-Silva D**, Demerdash OE, Tuveson DA, Notta F, Vakoc CR. 2023. p73 activates transcriptional signatures of basal lineage identity and ciliogenesis in pancreatic ductal adenocarcinoma. bioRxiv 10.1011/21:2023.04.20.537667.

**Iohannes SD**, Jackson D. 2023. Tackling redundancy: genetic mechanisms underlying paralog compensation in plants. *New Phytol* **240:** 1381–1389.

Itam MO*, **Iohannes SD***, Albertsen M. Andrade M, Bor GA, Atta-Krah K, Bertram R, Danquah E, Horvath DM, Jones T, et al. 2023. Demonstrating the benefit of agricultural biotechnology in developing countries by bridging the public and private sectors. *Nat Plants* **10:** 2–5.

Kawaguchi RK, **Tang Z**, Fischer S. Rajesh C, Tripathy R, Koo PK, Gillis J. 2023. Learning single-cell chromatin accessibility profiles using meta-analytic marker genes. *Brief Bioinform* **24**: bbac541.

**Kahng JA**, Xavier AM, Ferro A, Auguste YSS, Cheadle L. 2023. Integrated high-confidence and high-throughput approaches for quantifying synapse engulfment by oligodendrocyte precursor cells. bioRxiv doi:10.1011/25:2023.08.24.554663

**Kleeman SO**, Thakir TM, Demestichas B, Mourikis N, Loiero D, Ferrer M, Bankier S, Riazat-Kesh YJRA, Lee H, Chantzichristos D, et al. 2023. Cystatin C is glucocorticoid responsive, directs recruitment of Trem$^{2+}$ macrophages, and predicts failure of cancer immunotherapy. *Cell Genom* **3:** 100347.

**Lakhani AA**, Thompson SL, Sheltzer JM. 2023. Aneuploidy in human cancer: new tools and perspectives. *Trends Genet* **39:** 968–980.

Lee NK, **Tang Z**, **Toneyan S**, Koo PK. 2023. EvoAug: improving generalization and interpretability of genomic deep neural networks with evolution-inspired data augmentations. *Genome Biol* **24:** 105.

Lee SC, Adams DW, Ipsaro JJ, Cahn J, Lynn J, Kim HS, **Berube B**, Major V, **Calarco JP**, LeBlanc C, et al. 2023. Chromatin remodeling of histone H3 variants by DDM1 underlies epigenetic inheritance of DNA methylation. *Cell* **186:** 4100–4116.

Lei PJ, Pereira ER, Andersson P, Amoozgar Z, Van Wijnbergen JW, O'Melia MJ, Zhou H, Chatterjee S, Ho WW, …, **Ergin I**, et al. 2023. Cancer cell plasticity and MHC-II-mediated immune tolerance promote breast cancer metastasis to lymph nodes. *J Exp Med* **220:** e20221847.


(*continued*)

## 2022 STUDENT (CURRENT OR PREVIOUS) PUBLICATIONS (*continued*)

Loh L, **Carcy S**, Krovi HS, Domenico J, Spengler A, Lin Y, Torres J, Palmer W, Norman PJ, Stone M, et al. 2023. Unraveling the phenotypic states of human innate-like T cells: comparative insights with conventional T cells and mouse models. bioRxiv doi:10.1011/8:2023.12.07.570707.

**Maia-Silva D**, Schier AC, Skopelitis D, Kechejian V, Alpso, Liverpool J, Taatjes DJ, Vakoc CR. 2023. Marker-based CRISPR screening reveals a MED12–p63 interaction that activates basal identity in pancreatic ductal adenocarcinoma. bioRxiv doi:10.1011/27:2023.10.24.563848

**Meze K**, Axhemi A, Thomas DR, Doymaz A, Joshua-Tor L. 2023. A shape-shifting nuclease unravels structured RNA. *Nat Str Mol Bio* **30:** 339–347.

**Mo Z**, Siepel A. 2023. Domain-adaptive neural networks improve supervised machine learning based on simulated population genetic data. *PLoS Genet* **19:** e1011032.

Rupert DD, **Pagliaro AH**, Choe J, Shea SD. 2023. Selective deletion of *Methyl CpG binding protein 2* from parvalbumin interneurons in the auditory cortex delays the onset of maternal retrieval in mice. *J Neurosci* **43:** 6745–6759.

**Sroka MW**, Skopelitis D, Vermunt MW, Preall JB, El Demerdash O, de Almeida LMN, Chang K, Utama R, Gryder B, Caligiuri G, et al. 2023. Myo-differentiation reporter screen reveals NF-Y as an activator of PAX3-FOXO1 in rhabdomyosarcoma. *Proc Natl Acad Sci* **120:** e2303859120.

**Staklinski SJ**, Scheben A, Siepel A, Kilberg MS. 2023. Utility of AlphaMissense predictions in Asparagine Synthetase deficiency variant classification. bioRxiv doi:10.1011/2:2023.10.30.564808.

**Tang Z**, **Toneyan S**, Koo PK. 2023. Current approaches to genomic deep learning struggle to fully capture human genetic variation. *Nat Genet* **55:** 2021–2022.

**Toneyan S**, Koo, PK. 2023. Interpreting *cis*-regulatory interactions from large-scale deep neural networks for genomics. bioRxiv doi:10.1011/3:2023.07.03.547592

**Werner JM**, Gillis J. 2023. Preservation of co-expression defines the primary tissue fidelity of human neural organoids. bioRxiv doi:10.1011/7:2023.03.31.535112

**Werner JM**, Hover J, Gillis J. 2023. Population variability in X-chromosome inactivation across 9 mammalian species. bioRxiv doi:10.1011/19:2023.10.17.562732.

Xie Y, Huang L, **Corona A**, **Pagliaro AH**, Shea SD. 2023. A dopaminergic reward prediction error signal shapes maternal behavior in mice. *Neuron* **111:** 557–570.

**Yan R**, Moresco P, Gegenhuber B, Fearon DT. 2023. T cell–mediated development of stromal fibroblasts with an immune-enhancing chemokine profile. *Cancer Immunol Res* **7:** OF1–OF11.

**Zhang A**, Zador AM. 2023. Neurons in the primary visual cortex of freely moving rats encode both sensory and non-sensory task variables. *PLoS Biol* **21:** e3002384.

*Authors contributed equally to the work. **Boldface indicates School of Biological Sciences student.**

## SBS GRADUATES IN FACULTY OR INDEPENDENT POSITIONS (IN ORDER OF COMPLETION)

| Name | Faculty Position |
|---|---|
| Ira Hall | Professor, Yale University, Connecticut |
| Niraj Tolia | Senior Investigator, NIAID, National Institutes of Health |
| Patrick Paddison | Professor, Fred Hutchinson Cancer Research Center, Washington |
| Elizabeth Bartom (*nee* Thomas) | Assistant Professor, Northwestern University, Illinois |
| Michelle Heck (*nee* Cilia) | Research Molecular Biologist, USDA-ARS and Adjunct Associate Professor, Cornell University, New York |
| Zachary Lippman | Professor, Cold Spring Harbor Laboratory and Investigator, Howard Hughes Medical Institute, New York |
| Ji-Joon Song | Professor, Korea Advanced Institute of Science and Technology (KAIST), South Korea |
| Elena Ezhkova | Professor, Mount Sinai School of Medicine, New York |
| Masafumi Muratani | Professor, University of Tsukuba, Japan |
| Santanu Chakraborty | Program Director for Sustainable Life Sciences and Associate Professor, Atria University, India |
| Claudia Feierstein | Research Associate, Champalimaud Neuroscience Programme, Portugal |
| Gowan Tervo | Group Leader, HHMI, Janelia Research Campus, Virginia |
| Marco Mangone | Professor, Arizona State University, Arizona |

(*continued*)

## SBS GRADUATES IN FACULTY OR INDEPENDENT POSITIONS (IN ORDER OF COMPLETION) (*continued*)

| Name | Faculty Position |
|---|---|
| Elizabeth Murchison | Professor, Cambridge University, United Kingdom |
| Hiroki Asari | Group Leader, EMBL Monterotondo, Rome |
| François Bolduc | Associate Professor, University of Alberta, Canada |
| Wei Wei | Professor, University of Chicago, Illinois |
| Christopher Harvey | Associate Professor, Harvard University, Massachusetts |
| Tomas Hromadka | Group Leader, Institute of Neuroimmunology, Slovak Academy of Sciences, Bratislava |
| Monica Dus | Associate Professor, University of Michigan, Michigan and White House Fellow, United States Department of the Navy |
| Shu-Ling Chiu | Assistant Professor, Academia Sinica, Taiwan |
| Daniel Chitwood | Assistant Professor, Michigan State University, Michigan |
| Jeremy Wilsuz | Associate Professor, Baylor College of Medicine, Texas |
| Shraddha Pai | Investigator, Ontario Institute for Cancer Research, Toronto, Canada |
| Oliver Fregoso | Assistant Professor, University of California, Los Angeles, California |
| Hiroshi Makino | Assistant Professor, Nanyang Technological University, Singapore |
| Galen Collins | Assistant Professor, Mississippi State University, Mississippi |
| Katherine McJunkin | Stadtman Tenure Track Investigator, National Institutes of Health, Maryland |
| Zhenxun Wang | Adjunct Junior Principal Investigator, Singapore Eye Research Institute and Adjunct Assistant Professor Duke-NUS Medical School, Singapore |
| Eyal Gruntman | Assistant Professor, University of Toronto Scarborough, Canada |
| Petr Znamenskiy | Assistant Professor, The Crick Institute, United Kingdom |
| Yevgeniy Plavskin | Visiting Clinical Associate Professor, New York University, New York |
| Melanie Eckersley-Maslin | Assistant Professor, Peter Mac Cancer Centre at University of Melbourne, Australia |
| Philippe Batut | Assistant Professor, Columbia University (January 2024), New York |
| Elvin Wagenblast | Assistant Professor, Mount Sinai School of Medicine, New York |
| Kristen Delevich | Assistant Professor, Washington State University, Washington |
| Cinthya Zepeda Mendoza | Medical Director, Cytogenetics and Genomic Microarray at ARUP Laboratories and Assistant Professor, Department of Pathology, University of Utah |
| Wee Siong Goh | Assistant Professor Shenzhen Bay Laboratory, China |
| Arkarup Banerjee | Assistant Professor, Cold Spring Harbor Laboratory, New York |
| Charles Underwood | Professor, Radboud University, The Netherlands |
| Justus Kebschull | Assistant Professor, The Johns Hopkins University, Maryland |
| Paul Masset | Assistant Professor, McGill University, Canada (January 2024) |
| Brianna Bibel | Visiting Professor of Biochemistry, Saint Mary's College of California, Moraga |

## SBS GRADUATES IN INDUSTRY POSITIONS (IN ORDER OF COMPLETION)

| Student | Current Position |
|---|---|
| Amy Caudy | Principal, Scientific Operations, Maple Flavored Solutions, New York* |
| Emiliano Rial-Verde | Vice President, Food & Ingredients Strategy, Bunge Limited, New York |
| Rebecca Ewald | International Business Leader, Ventana Medical Systems/Roche, Arizona |
| Catherine Seiler (*nee* Cormier) | Director, Operational Excellence in Translation Medicine, AstraZeneca, Massachusetts |
| Allison Blum | Senior Director, Scientific Affairs, Pfizer Research and Development, New York |
| Yaniv Erlich | CEO, Eleven Therapeutics, United Kingdom* |
| Colin Malone | Co-Founder and Head of Biology at VNV NewCo, New York* |
| Amy Leung | Principal Solution Engineer, DNAnexus, California |
| Amy Rappaport | Senior Director, Translational Immunology, Gritstone bio, Inc., California |
| Claudio Scuoppo | Principal Scientist, Sapience Therapeutics, Inc., Tarrytown, New York |
| Frederick Rollins | Director, Competitive Intelligence, Oncology R&D, AstraZeneca, Washington. D.C. |
| Patrick Finigan | Associate Director, Regulatory Affairs CMC, Gilead Sciences, Foster City, California |
| Kyle Honegger | Data Scientist, Ann & Robert H. Lurie Children's Hospital of Chicago, Illinois |
| Maria Pineda | Co-Founder, CEO, Envisagenics, New York |
| Felix Schlesinger | Bioinformatics Lead, ScaleBio, San Diego, California |

(*continued*)

## SBS GRADUATES IN INDUSTRY POSITIONS (IN ORDER OF COMPLETION) (*continued*)

| Student | Current Position |
|---|---|
| Megan Hogan (*nee* Bodnar) | Team Lead, Neochromosome, Inc., New York |
| Paloma Guzzardo | Head, Target Biology, Variant Bio, Washington |
| Saya Ebbesen | Associate Director, Medical + Scientific Strategy at BluPrint Oncology, London, United Kingdom |
| Hassana Oyibo | Associate Director, Synthetic Biology, Tessera Therapeutics, Somerville, Massachusetts |
| Michael Pautler | Head, Genomics Services, Platform Genetics, Canada* |
| Marek Kudla | Director of Bioinformatics, Ardigen, Poland |
| Joshua Sanders | Founder and C.E.O., Sanworks, L.L.C., New York |
| Katie Liberatore | Molecule Design Director, Cibus, Minnesota |
| Kaja Wasik | Co-Founder and Chief Scientific Officer, Gencove & Variant Bio, New York |
| Stephane Castel | Co-Founder and Chief Technology Officer, Variant Bio, Washington |
| Mitchell Bekritsky | Senior Manager, Bioinformatics, Illumina, Inc., Cambridge, United Kingdom |
| Sang-Geol Koh | Scientist and Entrepreneur, {Mind}, South Korea |
| Ozlem Aksoy (*nee* Mert) | Senior Scientist, Pfizer, California |
| Susann Manchado (*nee* Weissmueller) | Strategic Assistant (Chief of Staff) for President Global Drug Development, Novartis, Switzerland |
| Nilgun Tasdemir | Senior Scientist, Pfizer, New York |
| Silvia Fenoglio | Senior Scientist, Tango Therapeutics, Cambridge, Massachusetts |
| Jack Walleshauser | Head of Analytics, Spotlight Therapeutics, California |
| Lisa Krug | Senior Scientist, Kallyope, New York |
| Robert Aboukhalil | Senior Software Engineer, Chan Zuckerberg Initiative |
| Anja Hohmann | Senior Director, Cell Therapies at Be Biopharma, Massachusetts |
| Matt Koh | Research Engineer, ASAAP, New York |
| Annabel Romero Hernandez | Director of AI for Drug Discovery and Molecular Modeling, SFL Scientific, Massachusetts |
| Maria Nattestad | Senior Software Engineer, Google, California |
| Onyekachi Odoemene | Research Scientist, ML Solutions Lab, Amazon |
| Daniel Kepple | Engineer, Meta, New York |
| Sashank Pisupati | Machine Learning Scientist, Limbic, United Kingdom |
| Kristina Grigaityte | Associate Director, Data Science, Merck |
| Elizabeth Hutton | Senior Computational Biologist, ArtisanBio, Colorado |
| Ngoc (Tumi) Tran | Machine Learning Scientist, Relay Therapeutics, Massachusetts |
| Katie Meze | Senior Scientist, Structural Biology, Sanofi, Massachusetts |
| Benjamin Berube | Research Scientist, Monocot Breeding and Biotechnology, Ohalo Genetics, California |
| Shaina Lu | Senior Data Scientist, R&D, Miltenyi Biotec, Gaithersburg, Maryland |
| Benjamin Harris | Computational Biologist II, Lyell Immunopharma, California |
| Christopher Krasniak | Research Data Scientist, Somatus, New Hampshire |
| Ran (Rena) Yan | Senior Scientist, Parthenon Therapeutics, Massachusetts |
| Martyna Sroka | Scientist, Relay Therapeutics, Massschusetts |
| Sara Boyle | Data Analytics Consultant, Northern Trust, Illinois |
| Asad Aziz Lakhani | Associate, Search and Evaluation, Autobahn Labs, Palo Alto, California |

*Left a faculty position

## SBS GRADUATES IN POSTDOCTORAL OR ACADEMIC RESEARCH POSITIONS (IN ORDER OF COMPLETION)

| **Student** | **Current Position** |
|---|---|
| Charles Kopec | Associate Professional Specialist, Princeton University (Advisor, Dr. Carlos Brody) |
| Oliver Tam | Research Scientist, NYU Institute for Systems Genetics, New York |

(*continued*)

## SBS GRADUATES IN POSTDOCTORAL OR ACADEMIC RESEARCH POSITIONS (IN ORDER OF COMPLETION) (*continued*)

| Student | Current Position |
|---|---|
| Elizabeth Nakasone | Junior Researcher, Translational and Clinical Research, University of Hawai'i Cancer Center |
| Zinaida Perova | Project Lead, Cancer Models, European Bioinformatics Institute (EMBL-EBI), Cambridge, United Kingdom |
| Dario Bressan | Head of IMAXT laboratory, CRUK, Cambridge Institute, United Kingdom |
| Joaquina Delas Vives | Postdoctoral Fellow, Francis Crick Institute, United Kingdom (Advisor: James Briscoe) |
| Abram Santana | Postdoctoral Fellow, Harvard University, Cambridge, Massachusetts (Advisor: Joan Brugge) |
| Fred Marbach | Senior Laboratory Research Scientist, Francis Crick Institute, United Kingdom (Advisor: Andreas Schaefer) |
| Yu-Jui (Ray) Ho | Bioinformatic Specialist II, Memorial Sloan Kettering Cancer Center, New York (Advisor: Scott Lowe) |
| Talitha Forcier | Postdoctoral Fellow, New York University, New York (Advisor: Molly Hammell) |
| Laura Maiorino | Postdoctoral Fellow, Koch Institute, MIT, Cambridge, Massachusetts (Advisor: Darrell Irvine) |
| Giorgia Battistoni | Research Associate, Cancer Research UK (Advisor: Gregory Hannon) |
| Jacqueline Giovanniello | Postdoctoral Fellow, UCLA (Advisor: Kate Wassum) |
| Hamza Giaffar | Postdoctoral Fellow, UCSD (Advisor: Mikio Aoi) |
| Anqi Zhang | Postdoctoral Fellow, Harvard University, Cambridge, Massachusetts (Advisor: Florian Engert) |
| Alexandra Nowlan | Postdoctoral Fellow, UNC Chapel Hill, North Carolina (Advisor: Zoe McElligott) |
| Luqun Shen | Postdoctoral Fellow, Cornell University (Advisor: Alex Kwan) |
| Alberto Corona | Postdoctoral Fellow, Icahn School of Medicine at Mount Sinai, New York (Advisor: Paul Kenny) |
| Kathryn O'Neill | Child Health Equity Postdoc, University of Washington, Seattle |
| David Johnson | Postdoctoral Fellow, Invent at Seattle Children's Postdoctoral Scholars Program (Advisor: Siobhan Pattwell) |
| Bruno Gegenhuber | Postdoctoral Fellow, Harvard University, Cambridge, Massachusetts (Advisor: Michael Greenberg) |
| Lyndsey Aguirre | Biological Sciences Tech, Peanuts and Small Grains Research Unity, USDA-ARS, Oklahoma |
| Amritha Varshini Hanasoge Somasundara | Senior Scientist, Memorial Sloan Kettering Cancer Center, New York (Advisor: Ross Levine) |
| Chengxiang (Charlie) Yuan | Postdoctoral Fellow, Institute of Molecular and Cell Biology, Singapore (Advisor: Weimiao Yu) |
| Diogo Maia e Silva | Research Fellow in Pathology, Massachusetts General Hospital (Advisor: Keith Joung) |
| Yuzhao (Richard) Hu | Postdoctoral Scientist, Corteva Agriscience, Iowa |

## SBS GRADUATES IN NONRESEARCH POSITIONS (IN ORDER OF COMPLETION)

| Student | Current Position |
|---|---|
| Ahmet M. Denli | Associate Editor, *Genome Research*, CSHL Press, New York |
| Beth Chen | Operations Manager, Homer Scientific Holdings Inc., Bellevue, Washington |
| Darren Burgess | Senior Editor, *Nature*, United Kingdom |
| Rebecca Bish-Cornelissen | Head of Discovery and Preclinical Research, The Mark Foundation for Cancer Research, New York |
| Angelique Girard | Director of Finance and Administration, Amplitude Studios, Paris, France |
| Keisha John | Associate Dean for Diversity, Equity and Inclusion, University of Virginia, Virginia |
| David Simpson | Freelance Data Scientist and Strategy Consultant |

(*continued*)

**SBS GRADUATES IN NONRESEARCH POSITIONS (IN ORDER OF COMPLETION)** (*continued*)

| Student | Current Position |
|---|---|
| Ian Peikon | Co-Founder and CSO, Cajal Neuroscience, Seattle, Washington; Venture Partner, Lux Capital, New York |
| Colleen Carlston | Clinical Fellow in Pediatrics, Boston Children's Hospital, Massachusetts |
| Tyler Garvin | VP Operations, Underground Cellar, California |
| Brittany Cazakoff | Associate Lawyer, Cooley, California |
| Lital Chartarifsky | Medical Director, FCB CURE, New York |
| Michael Gutbrod | Senior Associate, Santé Ventures, Massachusetts |
| Emilis Bruzas | Associate Director, Medical Strategy, FCB Health, New York |
| Sanchari Ghosh | Content Specialist for bioRxiv, New York |
| Georgi Yordanov | Equity Research Associate, Cowen and Company, New York |
| Jue Xiang Wang | Consultant, Boston Consulting Group, New York |
| Matt Lee | Oncology Medical Science Liaison, EMD Serono, Inc., Rockland, Massachusetts |
| Sofya Polyanskaya | Consultant, Scitaris, Berlin, Germany |
| Jenelys Ruiz Ortiz | Bioinformatics Analyst, Bioinformatics Resource Center, The Rockefeller University, New York |
| Ilgin Ergin | Foundation and Corporate Relations Coordinator, CSHL, New York |
| Alexa Pagliaro | Foundation and Corporate Relations Officer, Whitehead Institute, Massachusetts |
| Asad Aziz Lakhani | Associate, Search and Evaluation, Autobahn Labs, Palo Alto, California |

## First Annual Faculty Retreat

On February 28, we held the first annual faculty teaching retreat at the Banbury Conference Center. The goal of the retreat was to take a fresh look at the curriculum and explore new ways of teaching within and across research areas in which we excel. The more than 30 faculty members in attendance discussed areas in which shared core principles could be integrated into the curriculum. The retreat led to a lecture series on these shared core principles, which we introduced to the first-year students in fall 2023 (see below).

## New Courses in Writing and Teaching

We have partnered with esteemed science journalist and 2023 honorary degree recipient Steve Hall to teach a writing course for our senior graduate students. From Steve's own words, "As should be universally apparent in the post-pandemic world, science communication is an essential skill for any 21st century scientist. The ability to communicate complex (and often incomplete) knowledge to general audiences has a huge and enduring impact on public health policy, emergency advice, advocacy for scientific initiatives, education about the scientific method, and instilling public understanding of all aspects of the scientific enterprise—what it examines, why it chooses to do so, how it makes discoveries, and how its continual process of fact-finding is organic, dynamic, and self-correcting." The proposed format will be a 10-session program modeled on a course that has been offered to graduate students, postdocs, and faculty at New York University since 2009. It is designed to enhance communication both to audiences outside science and within science.

The School has organized a new course in Undergraduate Teaching and Pedagogy, which was held this fall. Faculty from the SUNY Old Westbury Department of Biological Sciences led sessions on course design, assessment, and inclusive teaching. These faculty teach at a primarily undergraduate institution with years of experience incorporating active learning and research into the classroom. Following the course, student and postdoc attendees have opportunities to acquire undergraduate teaching experience at SUNY Old Westbury and other local colleges.

**Zachary Lippman**
*Professor and Director of Graduate Studies*

# SPRING CURRICULUM

## TOPICS IN BIOLOGY

Each year, invited instructors offer week-long courses at the Banbury Conference Center exploring specialized topics outside the expertise of the Cold Spring Harbor Laboratory faculty. These courses include morning and evening lectures as well as afternoon sessions during which students read assigned papers or work on problem sets and presentations. After a long COVID-induced break, in spring 2023 there were two courses: Physical Biology of the Cell and Microbes in Health and Disease.

### Physical Biology of the Cell

**February 12–18** **Attended by the entering classes of 2020 and 2021**

INSTRUCTOR **Rob Phillips,** Caltech

VISITING LECTURER **Madhav Mani,** Northwestern University

TEACHING FELLOW **Griffin Chure,** Stanford University

The aim of this course was to provide a hands-on experience in the use of quantitative models as a way to view biological problems. The courses began with "order of magnitude biology," showing how simple estimates can be exploited in biology. The students were shown how to construct simple models of a variety of different biological problems, primarily using the tools of statistical mechanics. One of the key themes of the course was to show how physical biology unites and organizes topics in a fundamentally different way, often revealing that topics that are nearby in physical biology seem unrelated when viewed from the vantage point of molecular or cell biology. The instructors guided the students from start to finish on several modeling case studies.

### Microbes in Health and Disease

**March 5–11** **Attended by the entering classes of 2019 and 2022**

INSTRUCTOR **Stanley Maloy,** San Diego State University

VISITING LECTURERS **J. Antonio Ibarra,** Instituto Politécnico Nacional, Mexico City*
**Stefan Pukatzki,** City College of New York
**Brenda Wilson,** University of Illinois, Urbana-Champaign*
**via Zoom*

Understanding how microbes promote health or disease demands a detailed knowledge of the host response as well as the microbe itself. Elucidating these distinct aspects of microbial pathogenesis requires an interdisciplinary approach that integrates microbiology, cell biology, immunology, biochemistry, genetics, and genomics.

This course focused on the role of microbes in human health and disease, discussing questions such as: What is the role of the microbiome in human health, and how does disruption of the microbiome lead to disease? How do environmental cues modulate the interaction of microbes with their host? How do microbes manipulate host cells, and how does the host respond to microbes? How do new microbial pathogens evolve? How can we thwart microbes to promote health and prevent disease? The course integrated interactive lectures, directed readings of research papers, seminars, and student presentations on various aspects of these topics.

# SPECIAL COURSES

## Teaching Experience at the Dolan DNA Learning Center

**Entering Class of 2023**

DIRECTOR **David A. Micklos**

INSTRUCTORS **Amanda McBrien** (Lead)
**Elna Carrasco-Gottlieb**
**Megan Capobianco**
**Sharon Pepenella**
**Cristina Fernandez-Marco**
**Jeffrey Petracca**

As science plays an increasing role in society, there is also an increasing need for biologists to educate nonscientists of all ages about biology. The School of Biological Sciences doctoral program offers its students the opportunity to teach in the Laboratory's Dolan DNA Learning Center (DNALC), where they teach laboratory courses to high school and middle school students. In the process, they learn how to communicate with nonbiologists and to inspire and educate creative young minds.

The teaching module entails pairs of students teaching once a week for 12 weeks. In the initial weeks, experienced DNALC instructors taught the students the didactic process through observation and co-teaching. By the fifth week graduate students began to teach independently, with consistent pedagogical feedback from DNALC staff. At the end of the rotation, they have completed a rigorous classroom immersion—and are very excited about their teaching experience.

## Rotations

**Entering Class of 2022**

The most important element of a doctoral education is learning to perform independent research. After the Fall Term courses, students participated in rotations; these provided students and faculty the opportunity to get to know each other and to explore possibilities for doctoral thesis research. At the end of each rotation, students made short presentations of their studies to the other students and their rotation advisors. These talks gave students an opportunity to share their experiences and to practice giving scientific presentations. This year, 16 faculty members served as rotation mentors, some mentoring more than one student.

ROTATION MENTORS

Arkarup Banerjee
Semir Beyaz
Benjamin Cowley
Camila dos Santos
Helen Hou
David Jackson
Leemor Joshua-Tor
Justin Kinney
Peter Koo
Adrian Krainer
Zachary Lippman
Hannah Meyer
Saket Navlakha
Ullas Pedmale
Adam Siepel
Christopher Vakoc

# FALL CURRICULUM

### Entering Class of 2023

The students started the semester by attending boot camps in Molecular, Cellular, and Quantitative Biology to introduce them to the techniques and terminology that they encounter in subsequent courses. The Molecular and Cellular Biology boot camp featured lectures from Semir Beyaz, Hiro Furukawa, Dick McCombie, Monn Monn Myat, Jon Preall, Lloyd Trotman, and Ericka Wee. A boot camp on how to read a scientific paper was led by Corina Amor Vegas, Richard Sever, and two senior graduate students, Patrick Cunniff and Michael Passalacqua. The boot camp in Quantitative Biology was taught by Justin Kinney and Ivan Iossifov.

## CORE COURSES

## The Leslie C. Quick, Jr. Core Course on Scientific Reasoning and Logic

| | |
|---|---|
| INSTRUCTORS | **Alexander Gann**<br>**Christopher Hammell**<br>**Leemor Joshua-Tor**<br>**Jessica Tollkuhn**<br>**Linda Van Aelst** |
| GUEST LECTURERS | **Adrian Krainer**<br>**Christopher Vakoc** |
| TEACHING ASSISTANTS | **Alexandra Battison**<br>**Deeptiman Chatterjee**<br>**Ankur Garg**<br>**Peipei Wu** |

In this core course, which forms the heart of the curriculum, students (1) acquired a broad base of knowledge about the biological sciences, (2) learned the scientific method, and (3) learned how to think critically. The initial four to five modules were on a different general theme. In each, students read an assigned set of research articles; at the end of the module, they provided written answers to a problem set that guided them through several of the articles.

Twice weekly, students attended lectures related to the module's topic that included concepts and fundamental information as well as experimental methods. The students met among themselves to discuss the assigned papers not covered by the problem set. Each week, students spent an evening discussing the assigned articles with faculty. In the final module of the course, students participated in a mock study section in which funded National Institutes of Health, R01 grants were reviewed and critiqued. This allowed students to evaluate the research questions before discoveries are made, evaluate routes toward discovery before knowing where they will end, and make critical judgments about how to proceed in the face of an uncertain outcome.

In 2023, the module topics for this course were as follows:

| Topic | Instructor(s) |
|---|---|
| Gene Expression | Alex Gann |
| Gene Regulatory Logic and the Construction of Multicellular Organisms: Insights from Flies, Plants, and Worms | Christopher Hammell |
| The Brain: Wiring, Plasticity, and Maladaptation | Jessica Tollkuhn |
| Macromolecular Structure and Function | Leemor Joshua-Tor |
| Study Section | Linda Van Aelst |

## The Darrell Core Course on Scientific Exposition and Ethics

INSTRUCTORS
**Jeremy Borniger**
**David Jackson**
**Rebecca Leshan**
**Hannah Meyer**

GUEST LECTURERS (CSHL)
**Diane Esposito**
**Alyson Kass-Eisler**
**Charla Lambert**
**Rachel Rubino**
**Richard Sever**
**Jan Witkowski**

VISITING LECTURERS
**Stephane Castel,** Variant Bio
**Nyasha Chambwe,** Weill Cornell School of Medicine
**Susan Friedman,** The Innocence Project
**Radha Ganesan,** Alan Alda Center for Communicating Science
**Jaclyn Jansen,** Weill Cornell School of Medicine
**Adam Rutherford,** University College London
**Alex Vaughan,** Meta

This core course offered instruction in the fundamental elements of scientific exposition—writing skills and public speaking—and ethics. The ability to communicate effectively and to appreciate the intricacies of ethical issues are essential skills for biologists; both subjects were taught in a series of example-based lectures and discussion groups. Writing skills included the fundamentals of modern scientific English and the organization and preparation of papers, research abstracts, and grant applications. Oral presentation skills were taught by instructors with different modes of presentation. Together with instructors, students critiqued formal seminar presentations at the Laboratory. Instruction and discussions about ethics included the ethical implications of biological discovery for society as well as the nature and boundaries of ethical behavior of scientists and their rights and responsibilities. A primary objective of the course was that students consider exposition and ethics integral aspects of scientific research.

## Research Topics Series

ORGANIZERS **Alyson Kass-Eisler**
**Catherine Perez**

This series provided students with an in-depth introduction to the fields of research that the scientists investigate. Students and faculty attended a weekly Research Topics seminar at which faculty members and CSHL fellows presented their current research topics, methods of investigation, and mentorship styles each Wednesday evening. The students learned how to approach important problems in biology. These seminars provided students with a basis for selecting laboratories in which to do rotations.

# SPECIALIZED DISCIPLINES COURSES

The students in the Entering Class of 2023 took a total of four Specialized Disciplines courses this fall: Quantitative Biology, Genetics, Cancer, and Systems Neuroscience.

## Quantitative Biology

**Throughout the semester**

INSTRUCTOR **Justin Kinney** (Lead)

CSHL GUEST LECTURERS **Benjamin Cowley**
**Alex Dobin**
**Ivan Iossifov**
**Peter Koo**
**Hannah Meyer**
**Saket Navlakha**

TEACHING ASSISTANTS **John Desmarais**
**Chandana Rajesh**

Quantitative reasoning is a powerful tool for uncovering and characterizing biological principles, ranging from the molecular scale all the way to the ecological. With the advent of high-throughput technologies in genomics and neuroscience, it has become increasingly necessary for biological researchers to be able to analyze and interpret large data sets and frame biological hypotheses quantitatively. To this end, this course aimed to equip the students with a working knowledge of standard statistics and Python programming, as well as provide exposure to more advanced topics in machine learning, genomics, population genetics, neuroscience, and biophysics.

## Genetics and Plant Biology

**September 11–October 2**

INSTRUCTOR **Ullas Pedmale** (Lead)

GUEST LECTURERS **David Jackson**
**Robert Martienssen**
**Sophie Zebell**

TEACHING ASSISTANT **Jason Lynn**

The course offers students a comprehensive understanding of analyzing and interpreting complex genetic data while exploring the realm of plant biology. Through a series of lectures and discussions, students delved into the intricacies of genetic data analysis, equipping them with the necessary skills to navigate complex genetic data sets. The course content encompasses the fundamental principles and techniques required to analyze genetic data effectively.

Moreover, the course emphasizes the integration of plant biology into genetic analysis. Students explored the defining characteristics that make plants unique organisms and gained insights into the genetic mechanisms underlying plant development, physiology, and adaptation.

Additionally, the course incorporated an examination of the genetic basis of human diseases. This component delves into the study of chromosomal anomalies and their implications in human health. By the end of the course, students had gained a solid foundation in analyzing and interpreting complex genetic data, acquired insights into plant biology and its integration with genetics, and developed critical thinking skills necessary for genetic research and applications in various fields.

## Cancer

**October 4–October 30**

INSTRUCTORS **Corina Amor Vegas**
**Christopher Vakoc**
**Peter Westcott**

GUEST LECTURERS **Semir Beyaz**
**Camila dos Santos**
**Tobias Janowitz**
**Michael Lukey**
**David Tuveson**

TEACHING ASSISTANT **Olaf Klingbeil**

Cancer represents an increasing cause of morbidity and mortality throughout the world as health advances continue to extend the life spans of our populations. Our basic understanding of cancer has increased considerably since 1971, when U.S. President Richard Nixon initiated the "War on Cancer." Specific hypotheses developed from our knowledge of cancer biology are being tested in increasingly complex model systems ranging from cell culture to genetically engineered mouse models, and such investigations should prove invaluable in discovering new methodologies for the

detection, management, and treatment of cancer in humans. Importantly, our ability to translate our knowledge of cancer biology into a health benefit for patients is now starting to take form.

At the conclusion of this course, students should have been able to elaborate an understanding of cancer as a pathobiological process that invades our bodies without offering any known benefit to the host; discuss how cancer progresses; and contemplate how to expand on the methods currently used to treat cancer. Students would also be able to design tractable methods to investigate fundamental aspects of cancer biology and would be familiar with translational approaches to defeating cancer. Topics covered in this course included hallmarks of cancer, tumor progression, the cancer genome, microenvironment, tumor immunology, metastasis, and approaches to treating cancer, including targeted therapy.

## Systems Neuroscience

**November 1–December 1**

INSTRUCTORS **Florin Albeanu**
**Arkarup Banerjee**
**Stephen Shea**

TEACHING ASSISTANT **Priyanka Gupta**

This course provided an overview of key aspects of neuroscience, with a focus on learning and plasticity from its cellular basis, through development, to systems and behavior. Both experimental and theoretical viewpoints were explored. There were three main components to the class: lectures, a problem set, and paper presentations.

Although they originated as separate modules from two different courses, in 2023, the SRL module on "The Brain: Wiring, Plasticity, and Maladaptation" and the Specialized Disciplines module on "Systems Neuroscience" were combined into a single module spanning the length of the two original modules to provide better continuity and integration of the concepts being taught.

## Shared Core Principles in Biology

**November 28–December 8**

INSTRUCTORS **Arkarup Banerjee**
**Christopher Hammell**
**Zachary Lippman**
**Saket Navlakha**

Are there common motifs applied across biological systems? Are there networks that are overrepresented in nature? What can we learn from these? These questions formed the basis for a series of interactive discussions on shared core principles in biology. New in 2023, this series took place over four sessions at the end of the fall term: (1) Network Architecture and Function; (2) Decision Making; (3) Robustness—Homeostasis; and (4) How Do You Evolve a Biological Network? These discussions gave students an opportunity to apply what they learned throughout the term to think about the connections between seemingly disparate disciplines (e.g., plant genetics and systems neuroscience).

# POSTDOCTORAL PROGRAM

PROGRAM DIRECTOR **Nicholas Tonks**

PROGRAM ADMINISTRATOR **Alyson Kass-Eisler**

An important measure of our postdoctoral program's success is the ability of postdoctoral fellows to secure positions after they complete their training. Recently our fellows accepted positions at Fresenius Kabi, *Genes & Development*, Government Office for Science UK, Meta, RayzeBio, Stony Brook University, University of California Davis, University of New Haven, Washington University in St. Louis, and Westlake University.

## Postdoctoral Liaison Committee

The Postdoctoral Liaison Committee (PDLC), which is an elected group of postdoctoral fellows who communicate information and ideas between the administration and the postdoctoral community, continues to enhance CSHL's postdoctoral experience. The PDLC is essentially the voice of the community and holds regular meetings and an annual Town Hall with Dr. Bruce Stillman. The current PDLC members are Mackenzie Calloway, Deeptiman Chatterjee, Viet Hang Le, Jason Lynn, and Helene Sertznig.

CSHL endeavors to prepare postdocs to be competitive for the jobs available. It is increasingly becoming the Laboratory's role to introduce the diversity of career opportunities available and to provide the tools postdocs need to prepare for these positions. As a result, a number of events were organized with the assistance of the PDLC and other career development groups. For example, the PDLC organized a chat with Dr. Oliver Bolger, director of the National Institutes of Health (NIH) National Cancer Institute Center for Cancer Training so postdocs could learn about training and career opportunities at the NIH.

The Laboratory also pays special attention to the social needs of the postdoctoral community. The vast majority of CSHL postdocs are not from Long Island or indeed from the United States, and so do not have a built-in social network. To this end, PDLC organizes social activities for the community throughout the year, including an annual postdoc retreat & BBQ, a day at the Banbury pool, brewery visit and fall festival outing to celebrate National Postdoc Appreciation Week, a holiday crawl, ice cream breaks, Pints & Postdocs, and a spring social. A lunar New Year celebration was hosted by the Laboratory. The first annual Postdoctoral Appreciation Banquet Dinner, co-organized by Bruce Stillman and the PDLC, gathered postdocs for a chance to enjoy an evening together to share their scientific and nonscientific interests.

## Bioscience Enterprise Club

The Bioscience Enterprise Club (BEC) disseminates information about nonacademic careers to the CSHL postdoc community. Topic areas include biotechnology, intellectual property, law, regulatory affairs, and venture capitalism.

Recently they organized visits to two incubator sites: one in New York City—Alexandria LaunchLabs—and one on Long Island—Broad Hollow Bioscience Park. They hosted a Fireside Chat with the Scientists of Wall Street; Chris Mortko (Merck), Jim Reddoch (Royalty Pharma), and Graig Suvannavejh (Mizuho Equity Research) shared their experience of working in the finance-facing life-science ecosystem. And they hosted a panel discussion on Paths Beyond Academia,

which featured six speakers from the bioscience sector: Drs. Arijit Bhowmick (Regeneron Pharmaceuticals), Anne Bothmer (Tessera Therapeutics), Qingmin Chen (Atalanta Therapeutics), David Lebwohl (Intellia Therapeutics), Muthiah (Mano) Manoharan (Alnylam Pharmaceuticals), and Omer Ziv (Eleven Therapeutics).

In November, BEC hosted Beyond the Bench, a daylong career symposium with multiple panel discussions covering various bioscience careers, followed by a wine and cheese networking session. Panels included Biotech Entrepreneurship, Program Management, Consulting and Finance, Science Communication, and Industry Research. A Keynote Speech was given by SBS alumna Dr. Becky Bish, Head of Discovery and Preclinical Research at The Mark Foundation for Cancer Research.

## Academic Career Training

Academic career training includes courses, lectures, and workshops on scientific enrichment, career exploration, and transferrable skills such as leadership, mentorship, and communication. This year the series included the Chalk Talk, Mock Chalk Talks, and Navigating the Job Search and Interviewing. In addition, the CSHL faculty gave one-on-one reviews of personal statements and CVs and conducted mock interviews for those postdocs who were planning to go on the job market in 2023.

## WiSE and DIAS

There are two affinity groups on campus, largely run by postdocs and students, with administrative support from Programs Coordinator Stephanie Franco: WiSE (Women in Science & Engineering) and DIAS (Diversity Initiative for the Advancement of STEM). These groups are dedicated to promoting diversity, inclusion, and equity in science. In addition to hosting prominent women and underrepresented minority scientists during the weekly CSHL seminar series, the groups held special events throughout the year.

DIAS hosted a DIAS Trainee Chats series at which participants could discuss their experiences and challenges; a welcome event for first-year graduate students; and a session for undergraduates on how to apply to postbaccalaureate and undergraduate research programs. Along with WiSE and PDLC, they hosted DEI Learning Nights at which they screened and discussed the films *The Immortal Life of Henrietta Lacks*, *From Swastika to Jim Crow*, *Who We Are*, *One Night In Miami*, and *Not the Science Type*.

WiSE hosted their annual in-house education retreat (IHER) to foster discussion and self-education about gender-disparity issues in STEM; a session on Career Launch and Acceleration, presented by COACh; and a DEI Learning Night in honor of International Day of Women and Girls in Science. They held a Special Seminar, Women in STEM: Pushing for Change, presented by Dr. Jennifer Geddes-McAlister, Professor of Molecular and Cellular Biology at University of Guelph, Canada and the founder of Moms in Proteomics, an international initiative established to mentor and support mothers in STEM internationally.

WiSE also conducts a number of outreach activities for local girls who are interested in science. This year they held a coding camp; a brain awareness day with the Girl Scouts; and a *Fun with DNA* event.

The WiSE, DIAS, and PDLC affinity groups also held a leadership retreat. The retreat provided an opportunity for affinity groups to hear from CSHL administrators, share feedback and suggestions, talk to each other about activities that could be implemented collaboratively, and start developing their proposals for institutional funds to run their programming.

## CSHL Cancer Center and CSHL Research Operations

In addition to the career development opportunities organized by the School, there is Lab-wide programming for students and postdocs including the Core Knowledge Series; Writing Resource Center Coffee Chats; the Career Directions series; Biostatistics, Bioinformatics, and Coding and Computational Office Hours; and one-on-one assistance at the Writing Resource Center.

CSHL's annual four-week Science Writing Course took take place in November. This course is designed to help researchers at all career stages improve their written communication skills. The Writing Resource Center also hosted trainings on "Slack: A Quick How-To for Newbies and Pros," "Leveraging LinkedIn for Career Exploration and Scientific Career Transitions," and "Effective (PowerPoint) Presentations."

CSHL's frameSHIFT speaker series examines the social and cultural aspects of working at an academic research institute like CSHL with the goal of fostering inclusion, empathy, and allyship. In 2023, there were talks on "Unsettling Settler-Genomics: Advancing Indigenous Genomics for Indigenous Peoples," by Drs. Krystal Tsosie and Keolu Fox, the 2022–2023 Global Chairs of ENRICH (Equity for Indigenous Research and Innovation—A Co-ordinating Hub), which centers Indigenous rights to develop, control, and govern Indigenous data; "Telling the Story of Modern Science through the History of CSHL," by Dr. Alistair Sponsel (Historian of the Life Sciences at CSHL); "Art & Photography with Shinnecock Indian Nation Artist Jeremy Dennis"; and "Disability in Academia," by Drs. Joel Reynolds (Georgetown University) and Kara Ayers (University of Cincinnati).

## The Science Alliance

Trainees at CSHL are provided with free membership in a special initiative of the New York Academy of Science (NYAS), the Science Alliance. The Science Alliance for graduate students and postdoctoral fellows is a consortium of universities, teaching hospitals, and independent research facilities in the New York City metro area. The Alliance's aim is to provide career and professional development monitoring for postdoctoral fellows and graduate students in science and engineering. Recent programs included a webinar on Transition to Research Independence: Funding and Grantsmanship, a live online course on How to Effectively Communicate Your Science to Any Audience, and a Leadership in STEM Series.

"What Can You Be with a PhD," the longest-running and largest biomedical career symposium in the United States, which is organized and hosted by New York University, with sponsorship from CSHL and other area institutions, took place in October 2023. The program featured a range of panel discussions over two days covering careers inside and outside academia.

# PREP POSTBACCALAUREATE PROGRAM

PROGRAM DIRECTORS **Monn Monn Myat**

PROGRAM ADMINISTRATOR **Alyson Kass-Eisler**

The School was awarded the PREP award from the National Institute of General Medical Sciences (NIGMS) in January 2022 to start a new one-year postbaccalaureate program for underrepresented students to prepare them for matriculation into Ph.D. graduate programs. In June 2023, we welcomed five PREP Scholars. Each Scholar has a research advisor (a CSHL faculty member), a direct mentor (a student or postdoc in their research laboratory), and a near-peer mentor (a CSHL graduate student) to help with their research and career development.

In addition to their research, the Scholars participated in professional development workshops and in select modules of courses in the first-year Fall Curriculum. They attended the Annual Biomedical Research Conference for Minoritized Scientists (ABRCMS) this past November, where they all presented posters, and one PREP Scholar received the poster award in Cancer Biology. They also submitted applications to Ph.D. programs across the country. An anonymous survey conducted at the beginning of the PREP program indicated that all Scholars already felt integrated into the CSHL community.

The following five scholars, selected from 75 applicants, took part in the 2023 program:

**Theresa Clark**
Research Advisor: **David Jackson**
Direct Mentor: **Sessen Daniel Iohannes**
Near-Peer Mentor: **Paul Bunk**
Plant regeneration using morphogenic regulators.

**Daniel DiMartino**
Advisor: **Lucas Cheadle**
Direct Mentor: **Irene Sanchez Martin**
Near-Peer Mentor: **Manojit Swamynathan**
Characterization of microglia in a maternal immunity model.

**Pretty Garcia**
Advisor: **Peter Koo**
Direct Mentor: **Ziqi (Amber) Tang**
Near-Peer Mentor: **Lucía Téllez Pérez**
Uncovering RNA–protein binding preferences using interpretable deep learning.

**Melissa Lozada**
Advisor: **Camila dos Santos**
Direct Mentor: **Dhivyaa Anandan**
Near-Peer Mentor: **Michael Passalacqua**
Characterization of strategies to block mammary tumor development and progression.

**Germaine Smart-Marshall**
Advisor: **Semir Beyaz**
Direct Mentors: **Paul Bunk and Timothy Maher**
Near-Peer Mentor: **Kaeli Rizzo**
Characterizing the interactions of ketogenic diet, microbiome, and colon cancer.

*Left to right:* Daniel DiMartino, Pretty Garcia, Theresa Clark, Germaine Smart-Marshall, and Melissa Lozada.

# UNDERGRADUATE RESEARCH PROGRAM

| | |
|---|---|
| PROGRAM DIRECTORS | **Christopher Hammell**<br>**Monn Monn Myat** |
| PROGRAM ADMINISTRATOR | **Kimberly Creteur** |

Established 64 years ago, the CSHL Undergraduate Research Program (URP) provides undergraduates from around the world with hands-on undergraduate research training in biological sciences. Several activities are implemented to ensure that URP participants transition smoothly into the Laboratory community and research. The URPs work, live, eat, and play among CSHL scientists, and have a very busy academic and social calendar throughout the summer program. The students receive training in scientific research, science communication, career preparation, and bioinformatics and computational biology, all while interacting socially with fellow program participants and members of the CSHL community in formal and informal activities. Some of the 2023 activities included Director's Tea, talks from faculty and program alumni, designing the URP T-shirt, a Broadway show, and the ever-famous URP versus PI volleyball match and BBQ.

The students' scientific development is the most important component of the program. At the beginning of the summer, each URP writes an abstract and presents a talk on their proposed research. The URP participants work alongside scientists and become increasingly independent throughout the summer. Concluding the program in August, each URP student prepares a final report and presents their results in a 15-minute talk at the URP Symposium. As in previous years, the program directors and faculty mentors were highly impressed with the accomplishments of the URP students.

2023 Undergraduate Research Program Participants

The following 20 students, selected from 738 applicants, took part in the 2023 program:

**John Apollo**
Advisor: **Adrian Krainer**
Funding: **Alfred L. Goldberg Fellowship and Burroughs Wellcome Fellowship**
Mechanistic study of pre-mRNA alternative splicing in PDAC.

**Luke Bemish**
Advisor: **Arkarup Banerjee**
Funding: **National Science Foundation Scholar**
Quantitative modeling of vocal behavior in the singing mouse.

**Isabelle Brown-Lyden**
Advisor: **Stephen Shea**
Funding: **National Science Foundation Scholar**
The role of smell in auditory perception in mice based on maternal experience.

**Inle Bush**
Advisor: **Saket Navlakha**
Funding: **National Science Foundation Scholar**
Modeling *Arabidopsis* root morphogenesis in a heterogenous nitrogen environment.

**Emily Davis**
Advisor: **Helen Hou**
Funding: **National Science Foundation Scholar**
Recording facial expressions in mice pups.

**Leah Fitzgerald**
Advisor: **Ullas Pedmale**
Funding: **National Science Foundation Scholar**
Characterization of ISWI-CRAF chromatin remodelers in regulation of CRY2.

**Shane Holmes**
Advisor: **Justin Kinney**
Funding: **National Science Foundation Scholar**
Predicting transcription rates in *E. coli* using artificial neural networks.

**Harper Lowrey**
Advisor: **Rob Martienssen**
Funding: **Libby Fellowship and von Stade Fellowship**
The role of AGO2/3 proteins in RNA-directed stress responses.

**Tianhao Luo**
Advisor: **Peter Koo**
Funding: **Dorcas Cummings Scholar and 30th Anniversary URP Scholar**
Interpreting single-cell chromatin accessibility with scBasset: enhancing performance and unraveling regulatory mechanisms.

**Pablo Mantilla**
Advisor: **Camila dos Santos**
Funding: **Joan Redmond Read Fellowship and William Shakespeare Fellowship**
Investigation of the impact of aging on mammary gland response to pregnancy.

**Abigail O'Meara**
Advisor: **Doreen Ware**
Funding: **National Science Foundation Scholar**
FAIRifing gene expression to support interoperability for analyses and visualization tools.

**Alister Orozco**
Advisor: **Gabrielle Pouchelon**
Funding: **National Science Foundation Scholar**
Transient somatostatin interneuron output in the development of Fragile X syndrome.

**Meredith Ortiz Rivera**
Advisor: **Christopher Vakoc**
Funding: **National Science Foundation Scholar**
ETV6 dependency in Ewing's sarcoma.

**John ReyMartin**
Advisor: **Lloyd Trotman**
Funding: **Katya H. Davey Fellowship and Yakov Gluzman Fellow**
Live cell imaging of a novel, oxidant-dependent cell death pathway.

**Ana Rock**
Advisor: **Corina Amor Vegas**
Funding: **James D. Watson Fellow**
Understanding the interactions between nerves and senescent cells.

**Jean Rodriguez-Rivera**
Advisor: **Semir Beyaz**
Funding: **National Science Foundation Scholar**
Effects of fatty acid metabolism on anti-tumor immunity.

**Daisy Rubio**
Advisor: **Andrea Schorn**
Funding: **National Science Foundation Scholar**
Elucidating primer binding site determinants for expression and silencing of the murine retrotransposon MusD using a massively parallel reporter assay.

**Jadyn Scott**
Advisor: **Benjamin Cowley**
Funding: **National Science Foundation Scholar**
Modeling olfaction to behavior with decision trees.

**Zhiyu Song**
Advisor: **Lingbo Zhang**
Funding: **Garfield Fellowship and Robert H.P. Olney Fellow**
Targeted CRISPR-Cas9 knockout of WT1 and DIPK1B: identify and validate novel drug targets in acute myeloid leukemia (AML).

**Tess Stanley**
Advisor: **Lucas Cheadle**
Funding: **National Science Foundation Scholar**
Cytokine receptor Fn14 up-regulation in epilepsy.

# SUMMER RESEARCH INTERNSHIP FOR MEDICAL STUDENTS (SRIMS)

PROGRAM DIRECTOR **Priya Sridevi,** Ph.D.

PROGRAM ADMINISTRATIVE COORDINATOR **Joanie O'Connor**

Through the CSHL and Northwell Health affiliation, a summer internship program, SRIMS (Summer Research Internship for Medical Students) was created to give first-year Zucker School of Medicine–Hofstra University students basic research experience and the opportunity to spend a summer working in a CSHL laboratory. To date, 21 students have been offered positions in mainly cancer and neuroscience labs at CSHL. Students commit 8–10 weeks (roughly July–September) to work full-time in a CSHL research laboratory during the program. The students work with their host PI to design a research project and present their work both at CSHL and at the annual "Medical Student Research Day" at Hofstra University, the following fall.

The following student took part in the 2023 program:

| STUDENT | CSHL MENTOR |
|---|---|
| Nandan Vithlani | David Tuveson |

# PARTNERS FOR THE FUTURE

PROGRAM DIRECTOR **David Jackson**

PROGRAM ADMINISTRATOR **Bridget Shanley**

The Partners for the Future Program, established in 1990, provides an opportunity for talented Long Island high school students to have hands-on experience in biological and biomedical research at Cold Spring Harbor Laboratory. Applications to this highly competitive program are open to Long Island high school students in their junior year. Each high school science chairperson may nominate up to two students. The top candidates are interviewed by CSHL scientists. Students selected for the program are paired with a scientist mentor and spend a minimum of 10 hours per week, September through March of their senior year, conducting original research. At the conclusion, the students present their projects to an enthusiastic audience of the students, scientific mentors and colleagues, CSHL administrators, parents, and teachers. Although the students learn a great deal about modern biology and state-of-the-art research techniques, the main advantage of the program is that they are exposed to day-to-day life in a laboratory. Interacting with scientists and support staff, the students learn to define and pursue a research goal while solving problems that may occur along the way.

The 2023–2024 Partners for the Future program students were chosen from approximately 60 nominations:

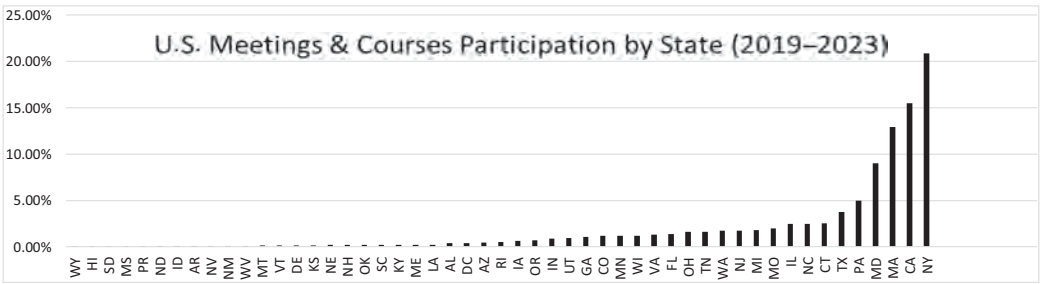
| STUDENT | SCHOOL | MENTOR | LAB |
|---|---|---|---|
| Maialee Bernas | Farmingdale High School | Jon Preall | Jon Preall |
| Deanna Besart | Oyster Bay High School | Nicole Sivetz | Mikala Egeblad |
| Makayla Castillo | Walt Whitman High School | Josh Homer | John Moses |
| Carlos Diaz Sanchez | Wyandanch High School | Lloyd Trotman | Lloyd Trotman |
| Kathleen Engel | Cold Spring Harbor High School | Peter Koo | Peter Koo |
| Alexander Grosh | Cold Spring Harbor High School | Rishvanth Kaliappan Prabakar | Hannah Meyer |
| Griffin Hon | Syosset High School | Semir Beyaz | Semir Beyaz |
| Ania Kelly | Oyster Bay High School | Johannes Yeh | Johannes Yeh |
| Stephanie Neri | Lynbrook Senior High School | David Jackson | David Jackson |
| Kshan Pandey | East Meadow High School | Kyle Swentowsky | David Jackson |
| Aayush Prakash | Half Hollow Hills High School East | Peter Koo | Peter Koo |
| Julia Rodriguez | Farmingdale High School | Helen Hou | Helen Hou |
| Daniel Sanwo | Half Hollow Hills High School East | Gabrielle Pouchelon | Gabrielle Pouchelon |
| Aberam Sriganesh | Syosset High School | Doreen Ware | Doreen Ware |
| Jonathan Zhang | Commack High School | Saket Navlakha | Saket Navlakha |
| Qingyuan Zhang | The Stony Brook School | Christopher Hammell | Christopher Hammell |

# MEETINGS & COURSES PROGRAM

In all, the meeting program of 27 conferences and ancillary events attracted almost 6,900 participants to the CSHL campus, while almost 2,000 additional attendees participated virtually. The course program included 31 courses and attracted an additional 1,400 faculty and trainees. Several meetings were oversubscribed, notably the Genome Engineering: CRISPR Frontiers and Neurobiology of *Drosophila* meetings, which each attracted about 450 participants and a sizeable additional virtual audience. A new meeting on Cell State Conversions assembled leaders in the field, together with junior faculty, postdoctoral fellows, and graduate students, to discuss new, cutting-edge developments in the study of all aspects of cellular reprogramming. The Meetings & Courses program continues to attract significant global participation, with participants from 95 countries and from all 50 U.S. states over the past five years.

**A Five-Year Look at Global Participation in Our Program: 2019–2023**

| *Region* | *Number* | *Percentage* |
|---|---|---|
| North America | 41,852 | 73.9 |
| Europe | 10,755 | 19.0 |
| East Asia | 2,443 | 4.3 |
| Australasia | 571 | 1.0 |
| South East Asia | 451 | 0.8 |
| Central and South America | 350 | 0.6 |
| Africa | 133 | 0.2 |
| Middle East | 90 | 0.2 |
| **Total** | **56,645** | |




Total U.S. 5-Year Count: 40,154 attendees (includes CSHL)

## Demographics from 2023

Course Students

- Totals: 31 courses, 1,454 applicants, 618 accepted.
- Career level: Graduate students 51%/53% of applicants/accepted students, and postdocs 25%/28%.
- Gender: Women represented 54%/56% of the applicants/accepted students. Nonbinary/other gender was indicated by seven applicants and five accepted students.
- Nationality: 67% of the applicants and 78% of the accepted students from U.S. institutions (selection is biased toward the United States because of federal grant requirements).
- Underrepresented minorities (URMs): 13% of all applicants and 17% of all accepted students (19% of U.S. applicants and 21% of accepted U.S. students) come from URM groups (more than doubled in the past decade).

2023 Plate Race

## Nobel News

Katalin Karikó and Drew Weissman, winners of the 2023 Nobel Prize in Physiology or Medicine for their foundational work toward the development of effective COVID mRNA vaccines, both have history with CSHL. Dr. Weissman discussed mRNA vaccines at our COVID/SARS CoV2 Rapid Research Reports virtual conference in July 2020, and Dr. Karikó spoke about the science behind the incredible success of the vaccines at our recent biohistory conference on Recombinant DNA: Fifty Years of Discussion Debate—just 48 hours before the prize was announced on October 2, 2023.

## Cold Spring Harbor Asia

The COVID-19 situation in mainland China in 2022 bled over into 2023, and so plans were made to run early 2023 meetings in our primary Japanese site on Awaji Island, close to Kobe, Osaka, and Kyoto. In the second half of the year, the situation normalized so the majority of our events—including one Strategic Summit, two summer schools, and five conferences—were held at our Suzhou, China headquarters. In total, the program attracted almost 1,500 participants.

## Looking to 2024

Hybrid meetings will continue to be offered in 2024. It is clear that participants with work to present at our conferences benefit greatly from attending in person, and so oral and poster virtual presenters will be kept to a minimum whenever possible.

## Special Thanks

We thank several course instructors who stepped down in 2023 after many years of service, in particular Farran Briggs (vision), Vincent Carey (statistical data analysis), David Chenoweth (single cells), Maitreya Dunham (yeast), Jan Grundemann (ion channels), Angelika Grundling (bacterial genetics), Darryl Pappin (proteomics), Cary Lai (molecular neuroscience), Julie Law (plants), Mark Reimers (neural data analysis), and Phil Tsai (imaging in the nervous system).

## Program Staff

The Meetings & Courses program staff comprises a diverse team of talented professionals who handle the complexities of database design, programming, web and multimedia design, educational

grants management, marketing and recruitment, conference and course administration, audiovisual and digital design services, and other activities. Their hard work, enthusiasm, and willingness to develop and adapt to new and untried virtual operations was truly remarkable, all the more so because so much was done while working remotely. In 2023, we said our goodbyes to a number of staff, including IT coordinator Adam Crespo and conference coordinators Sam Mastronardi and Amy Traina.

## Program Funding

We appreciate the major financial support for our courses from the following: Helmsley Charitable Trust, Howard Hughes Medical Institute, National Institutes of Health, National Science Foundation, and Regeneron. The course program is normally supported by major equipment and reagent companies.

Contributions from the following companies provide core support for the Cold Spring Harbor meetings program:

## Corporate Benefactors

Estée Lauder Companies
Regeneron

## Corporate Sponsors

Agilent Technologies
Biogen
Bristol-Myers Squibb
Calico Labs
Merck & Co., Inc.
New England BioLabs
Novartis Institutes for Biomedical Research

## Corporate Partners

Alexandria Real Estate

**David Stewart**
*Executive Director, Meetings & Courses Program*
*President, Cold Spring Harbor Asia*

**Terri Grodzicker**
*Academic Guidance, Dean of Academic Affairs*

# 87TH COLD SPRING HARBOR LABORATORY SYMPOSIUM ON QUANTITATIVE BIOLOGY

## Stem Cells

**May 31–June 4** 267 Participants (262 in person, 5 virtual [invited speakers only])

ARRANGED BY **Terri Grodzicker, David Stewart, and Bruce Stillman,** Cold Spring Harbor Laboratory

The 87th Cold Spring Harbor Symposium focused on stem cells and provided a current synthesis of the enormous progress in our understanding of stem cell biology since the last Cold Spring Harbor Symposium on this topic in 2008. Topics addressed at the Symposium included germ cells; embryonic and trophoblast stem cells; early embryo including retrotransposons; differentiation and reprogramming; epigenetics; organ stem cells; hematopoietic/immune system; neural stem cells; adult stem cells/regeneration; cancer stem cells; induced pluripotent stem cells; organoids and multicellular organoids; disease models; and clinical therapies.

M-E. Torres-Padilla, N. Gogia

The Symposium attracted 267 participants and provided an extraordinary five-day synthesis of current understanding in the field. Opening night talks setting the scene for later sessions included Maria-Elena Torres-Padilla addressing epigenetic mechanisms of cellular plasticity and reprogramming to totipotency; Takanori Takebe on understanding interconnectedness in liver development and disease; Jürgen Knoblich, who presented the modeling of neural network pathology in cerebral organoids; and Sean Morrison, who addressed how leptin receptor$^{+}$ cells promote bone marrow innervation and regeneration by synthesizing nerve growth factor.

J. Wells, J. Inglis, and B. Stillman at the Dorcas Cummings Lecture

A. Sánchez Alvarado, E. Fuchs

D. Torre, S. Cheloufi

Z. Lippman, P. Benfey

I. Weissman

In a new "fireside format" for the Dorcas Cummings lecture, CSHL's John Inglis moderated a discussion with James Wells, Cincinnati Children's Hospital, on the subject of "Growing human organs in the lab—using organoids to understand human development, model diseases, and discover new therapies" for the Laboratory's friends and neighbors. Interviews with leading scientists captured during the Symposium provide a snapshot of the state of current research and are available on the CSHL Leading Strand channel (https://www.youtube.com/user/LeadingStrand).

Support for this meeting was provided in part by Genentech, Creative Biolabs, STEMCELL Technologies, Stem Genomics, and Thermo Fisher Scientific.

## PROGRAM

**Introduction**
Bruce Stillman, *Cold Spring Harbor Laboratory, New York*

**Germ Cells and Early Development**
Joanna Wysocka, *Stanford University, California*
Maria-Elena Torres-Padilla, *Helmholtz Centre Münich, Germany*

**Organ Stem and Progenitor Cells and Organoid Models**
Kenneth Zaret, *University of Pennsylvania, Perelman School of Medicine, Philadelphia*
Meritxell Huch, *Max Planck Institute of Molecular Cell Biology and Genetics, Dresden, Germany*

J. Wysocka, J. van Wolfswinkel

K. Zaret, J. Munera

S. Rainsford, N. Dias

M. Zernicka Goetz, K. Zaret, A. Schaefer, A. Schorn, S. Morrison

**Neural Stem and Progenitor Cells and Organoid Models**
Jürgen Knoblich, *Institute of Molecular Biotechnology of the Austrian Academy of Sciences, Vienna*
Magdalena Götz, *University of Münich, Germany*

**Regeneration and Repair**
Elaine Fuchs, *The Rockefeller University, New York, New York*
Alejandro Sánchez Alvarado, *Stowers Institute for Medical Research / HHMI, Kansas City, Missouri*

**Dorcas Cummings Lecture:** Growing human organs in the lab—using organoids to understand human development, model diseases, and discover new therapies
James Wells, *Cincinnati Children's Hospital*

**Hematopoiesis**
Stephen Smale, *UCLA School of Medicine, Los Angeles, California*
Nancy Speck, *University of Pennsylvania, Perelman School of Medicine, Philadelphia*

**Stem Cells and Disease**
Christine Mummery, *Leiden University Medical Center, the Netherlands*
Guo-li Ming, *University of Pennsylvania, Philadelphia*

**Stem Cells and Regulatory Mechanisms**
Ellen Rothenberg, *California Institute of Technology, Pasadena*

A. Tajsharghi, L. Yiangou, Mitheera V, D. Sridharan, N. Gogia

# MEETINGS

## Probabilistic Modeling in Genomics

**March 8–11** 326 Participants (228 in-person, 98 virtual)

ARRANGED BY **Carolin Kosiol,** University of St. Andrews, Scotland
**Molly Przeworski,** Columbia University, New York, New York
**Adam Siepel,** Cold Spring Harbor Laboratory

M. Przeworski, C. Kosiol, A. Siepel

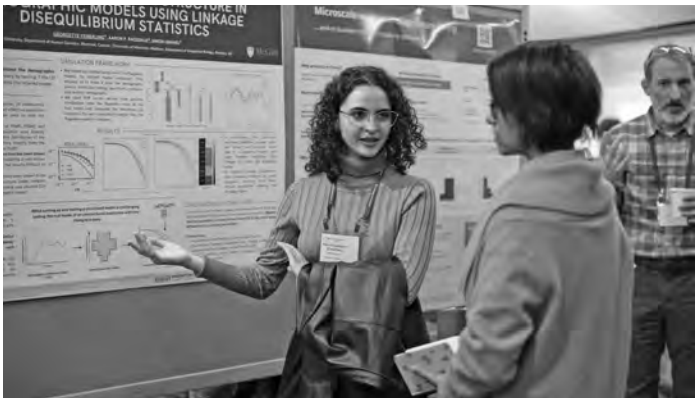
This was the eighth (mostly) annual Probabilistic Modeling in Genomics (PROBGEN) conferences and the fourth one hosted at Cold Spring Harbor. The main goal of this meeting, which grew out of two ad hoc meetings on a similar topic in 2013 at Janelia Farm and in 2014 at Merton College, Oxford, is to provide a forum for presentation and exchange of ideas among researchers who are working in the general area of genomics but are particularly focused on the development of new probabilistic and machine-learning models, algorithms, and methods for inference. These researchers come from a variety of backgrounds, including computer science, statistics, applied mathematics, and physics. Initially, the meeting strongly emphasized population genetics, but over time it has branched out to include topics such as functional genomics, systems biology, and quantitative genetics.

Session topics included Population and Statistical Genetics; Machine Learning in Genomics; Demographic Inference; Quantitative Genetics and Association Mapping; Systems Biology; and Phylogenetics and Phylodynamics. There were six oral presentations per session for a total of 36 presentations. Two invited session chairs were presented in each session, and four additional talks were selected from submitted abstracts. Talks were 20 minutes long plus five minutes for questions

S. Tittes, S. Temple, A. Buerkle

P. Faux, A-S. Malaspinas

H. Louw




M.G. Femerling, S. Musharoff

and answers. The quality of the presentations was very high overall, with considerable mathematical sophistication combined with a clear focus on biological relevance. We were honored to recruit two distinguished senior scientists to attend the meeting and present keynote lectures: Kathryn Roeder from Carnegie Mellon University and Richard Durbin from Cambridge University.

Notably, the diversity of the meeting has steadily improved over time. This time, half the speakers and 41% of attendees were women. The meeting continues to be well represented by trainees, with 36% of attendees being grad students and another 20% being postdocs. In addition, about one-half (160) of attendees came from outside the United States, with representation from 22 countries across Europe, the Americas, the Middle East, Asia, and Africa.

Overall, the meeting appeared to be highly successful, with several attendees commenting favorably about the quality of the content. The next meeting will be held in Vienna on April 10-12, 2024. It is being arranged by Magnus Nordorg, Nick Barton, Carolin Kosiol, and Joachim Hermisson.

## PROGRAM

**Population and Statistical Genetics**
A. Hinch, *Oxford University, United Kingdom*; S. Sunyaev, *Harvard Medical School, Boston, Massachusetts*

**Machine Learning in Genomics**
E. Azizi, *Columbia University, New York, New York*; D. Kelley, *Calico Life Sciences, South San Francisco, California*

**Demographic Inference**
A. Hobolth, *Aarhus University, Denmark*; A-S. Malaspinas, *Swiss Institute of Bioinformatics, Lausanne, Switzerland*

**Keynote Speaker**
K. Roeder, *Carnegie Mellon University, Pittsburgh, Pennsylvania*

**Quantitative Genetics and Association Mapping**
G. Sella, *Columbia University, New York, New York*; K. Swarts, *Gregor Mendel Institute, University of Vienna, Austria*

**Session 5: Systems Biology**
T. Lieberman, *Massachusetts Institute of Technology, Cambridge*; Simon Myers, *Oxford University, United Kingdom*

**Keynote Speaker**
R. Durbin, *University of Cambridge, United Kingdom*

**Phylogenetics and Phylodynamics**
T. Heath, *Iowa State University, Ames*; C.B. Ogbunu, *Yale University, New Haven, Connecticut*

# Network Biology

**March 14–18** 202 Participants (134 in-person, 68 virtual)

ARRANGED BY **Anne-Ruxandra Carvunis,** University of Pittsburgh, Pennsylvania
**Fritz Roth,** University of Toronto, Ontario, Canada
**Roded Sharan,** Tel Aviv University, Israel
**Michael Springer,** Harvard Medical School, Boston, Massachusetts

F. Roth, A.-R. Carvunis, M. Springer, R. Sharan

In cells and organisms, genomic information is translated into phenotypes by complex and highly dynamic molecular networks formed by proteins, nucleic acids, and small molecules. A systems-level understanding of biological systems, as well as the design of rational biotechnological or pharmaceutical interventions in human, crops, and microbes, hinges on our knowledge of these networks.

As molecular networks still remain largely incomplete, an important goal of biological network science is to experimentally map or computationally infer the wiring of cells. A second major goal is the mechanistic characterization of smaller network modules, translating large-scale network connectivity into molecular mechanisms as a basis for development of quantitative predictive models. Finally, network science also aims to develop and apply statistical tools to extract insights from known biological networks to identify disease-causing genes and modules, identify targets for intervention, and decipher the fundamental principles that underlie biological systems and their evolution. These goals have been aided by the rapid advances in high-throughput techniques, synthetic biology, and organism editing, which have led to an explosion not just in the quantity but also in the types of data available. Two key goals of the Network Biology meeting remain to bridge these three aspects of network biology and to cover a diversity of biological systems—from humans

J. Reimand, Y. Ivarsson

H. Yuan, G. Coppin

S. Alkhairy, N. Manosalva

M. Polychronidou, M. Walhout

G. Trave, J. Rogers

and model organisms to plants and microbes—while continuing to highlight new experimental and computational opportunities and approaches. The most recent meeting successfully achieved these goals and continues to serve as the main international meeting for the network biology community.

We continued the open panel discussions, which were highly successful for community reflection in the previous meetings. This year's panel, led by Dr. Marian Walhout, discussed the Future of Network Biology. This highly interactive session involved panelists and plenty of audience participation, including a mix of scientific questions and career development advice. Two Networking for Network Biologists half-hour sessions were organized on the first and second day of the meeting, each engaging groups of four people who had never met before to get to know each other and find their "network edges" (session I) and a possible collaboration topic (session II). This format was intended to facilitate and catalyze networking and help young scientists get in direct contact with senior scientists and PIs. Finally, Slack was used for continuing discussions and including the virtual audience. These elements were praised by attendees and resulted in continued discussions throughout the meeting.

The scientific program opened on the morning of March 15. There were 16 invited presentations and 23 short talks selected from submitted abstracts, all of them outstanding and many given by postdocs and Ph.D. students, with good gender balance among the presenters (42.5% were women presenters). The talks covered a wide range of concepts, ranging from the impact of disordered protein ensembles on protein interaction networks to computational data representations, and addressed diverse questions from the circadian clock to viral infections. This highlights how network biology brings together people from different fields of biology. The presentations were followed by dynamic and lively discussions.

Two Keynote Addresses inspired junior and senior scientists alike by showcasing the power of collaborative research and dreaming big. Dana Pe'er (Memorial Sloan Kettering Cancer Center) opened the meeting with an overview of her computational research on transcriptomics. Nevan Krogan (University of San Francisco) summarized how years of network biology and virus–host interactome research enabled lifesaving breakthroughs during the COVID-19 pandemic.

This meeting was funded in part by the National Human Genome Research Institute, a branch of the National Institutes of Health.

## PROGRAM

**Keynote Speaker**
D. Pe'er, *Memorial Sloan Kettering Cancer Center, New York, New York*

**Protein Networks I**
S. Mukhtar, *University of Alabama, Birmingham*

**Network Evolution**
A.-R. Carvunis, *University of Pittsburgh, Pennsylvania*

**Keynote Speaker**
N. Krogan, *University of California, San Francisco*

**Regulatory and Genetic Networks I**
M. Springer, *Harvard Medical School, Boston, Massachusetts*

**Disease Networks**
I. Overton, *Queen's University Belfast, United Kingdom*

**Regulatory and Genetic Networks II**
A. Skirycz, *Boyce Thompson Institute, Ithaca, New York*

**Computational Methods for Network Biology**
M. Chikina, *University of Pittsburgh, Pennsylvania*

**Protein Networks II**
M. Vidal, *Dana-Farber Cancer Institute, Harvard Medical School, Boston, Massachusetts*

**Panel Session: The Future of Network Biology**
T. Ideker, *University of California, San Diego*; C. Myers, *University of Minnesota-Twin Cities*; M. Vidal, *Dana-Farber Cancer Institute, Boston, Massachusetts/Harvard Medical School, Boston, Massachusetts*; M. Walhout, *University of Massachusetts Medical School, Worcester, Massachusetts*

# Nucleic Acid Therapies

**March 22–25** 212 Participants (131 in-person, 81 virtual)

ARRANGED BY **Shalini Andersson,** AstraZeneca, Sweden
**Masad Damha,** McGill University, Montréal, Québec, Canada
**Anastasia Khvorova,** University of Massachusetts Medical School, Worcester, Massachusetts
**Matthew Stanton,** Generation Bio, Cambridge, Massachusetts

M. Damha, S. Andersson, A. Khvorova, M. Stanton

This was the sixth Cold Spring Harbor conference on Nucleic Acid Therapeutics. The meeting was focused on development of nucleic acids as drugs, covering different nucleic acid modalities (siRNA, mRNA, RNase H, splice modulation, and CRISPR-Cas9), and involving different aspects of nucleic acid therapy development (i.e., chemistry, delivery, and preclinical and clinical studies). Numerous nucleic acid therapies have received regulatory approval, including the dramatically rapid development of COVID-19 vaccines. The meeting contained a keynote by Dr. David Lebwohl of Intellia Therapeutics describing initial successes with CRISPR-based in vivo genome editing for treatment of patients with severe illness. Other sessions showed a diversity of nucleic acid modalities being developed preclinically and clinically. Furthermore, there were sessions on important aspects such as safety, chemistry and delivery, and gene editing.

The participants came from 96 companies and from universities and research institutions from the United States and abroad. The five scientific sessions featured 26 platform talks, 66 posters, and a panel discussion and included 212 registered attendees. Animated and insightful exchanges continued throughout the sessions, during breaks, and during social events. Most participants indicated the meeting was excellent and expressed interest in attending the next edition.

The next Nucleic Acid Therapies meeting will take place in March 2025.

X. Wang, M. Mullari

N. Batistatou, J. Grindley

A. Wolfson, M. Daniels

Attendees with caricatures

## PROGRAM

**Keynote Speaker**
D. Lebwohl, *Intellia Therapeutics, Cambridge, Massachusetts*

**Editing**
M. Stanton, *Generation Bio, Cambridge, Massachusetts*; W. Yan, *Arbor Technologies, Cambridge, Massachusetts*

**Delivery**
M. Stanton, *Generation Bio, Cambridge, Massachusetts*; A. Wittrup, *Lund University, Sweden*

**Preclinical**
S. Andersson, *Astra Zeneca, Gothenburg, Sweden*; M. Maier, *Alnylam Pharmaceuticals, Cambridge, Massachusetts*

**Clinical**
A. Khvorova, *University of Massachusetts Medical School, Worcester*

**Chemistry**
M. Damha, *McGill University, Montréal, Québec, Canada;* K. Yamada, *University of Massachusetts Medical School, Worcester*

# Brain Barriers

**March 28–April 1** 265 Participants (175 in-person, 90 virtual)

ARRANGED BY **Maria Lehtinen,** Boston Children's Hospital, Massachusetts
**Julie Siegenthaler,** University of Colorado, Boulder
**Benoit Vanhollebeke,** Université Libre de Bruxelles, Belgium
**Joy Zuchero,** Denali Therapeutics, South San Francisco, California

M. Lehtinen, J. Siegenthaler, B. Vanhollebeke, J. Zuchero

The Brain Barriers Cold Spring Harbor Laboratory Meeting highlighted the latest scientific advances linked to the biology of the central nervous system barriers, including the blood–brain barrier (BBB), the blood–retinal barrier, and the blood–cerebrospinal fluid (CSF) barrier. By virtue of its hybrid format, the conference attendance was particularly high, with 175 attending in person and 90 attending virtually. Close to 65% of the submitted abstracts were from graduate students or postdoctoral fellows, who were given an important role in giving oral presentations (49.1%). Moreover, women scientists were well represented at the meeting. There were 55.1% female attendees, and many of them were session chairs and oral presenters (47.3%).

Besides the significant time allocated to emerging technologies and model organisms, the conference highlighted more than in the past the important role played by central nervous system (CNS) barriers other than the BBB in regulating molecular and cellular traffic in and out of the CNS. Another focus of the conference was the molecular definition of the human brain barriers and their comparison to the model systems structures. Due time was also allocated to the clinical implications of the brain barriers, balancing between BBB dysfunction in disease and emerging therapeutic strategies for CNS drug delivery and BBB repair.

The Keynote lecture, titled Mechanisms of Angiogenesis and Neurovascular Barrier Development, Breakdown, and Repair in the Central Nervous System, was given by Dritan Agalliu (Columbia University). The meeting oral sessions were (1) Emerging Concepts and Hot

T. Kiniwa, M. Zawadzki

S. Nozohouri

D. Jensen, T. Toft

P. Kalugin, M. Zawadzki

Topics in CNS Barriers, (2) Development, Plasticity, and Specialization of CNS Barriers, (3) CNS Barrier Damage and Repair in Injury, Infection, and Diseases, (4) Barrier-Omics in Health and Disease, (5) CNS Barrier Access for Therapeutics, (6) Imaging Approaches to Study CNS Barriers, (7) Bioengineering a Better In Vitro Barrier, and (8) Barriers at Choroid Plexus, Meninges, and Circumventricular Organs. The abstracts were grouped across each of the above topics. The selection of material for oral and poster presentations was made by the organizers on the basis of scientific merit.

Support for this meeting was provided in part by the National Institute of Neurological Disorders and Stroke and the National Institute of Aging, both branches of the National Institutes of Health; BioArctic; Genentech; NeuVasQ Biotechnologies; Science Advances; and Stemcell Technologies.

## PROGRAM

**Keynote Speaker**
D. Agalliu, *Columbia University Irving Medical Center, New York, New York*

**Emerging Concepts and Hot Topics in CNS Barriers**
M. Bernabeu, *EMBL, Barcelona, Spain*; R. Fame, *Stanford University, Palo Alto, California*

B. Kim, I. Kelly

T. Kiniwa, S. Funes, N. MacAulay

**Development, Plasticity, and Specialization of CNS Barriers**
J. Kipnis, *Washington University School of Medicine, St. Louis, Missouri*; S. Liebner, *Goethe University, Frankfurt am Main, Germany*

**CNS Barrier Damage and Repair in Injury, Infection, and Disease**
B. Engelhardt, *University of Bern, Switzerland*; C. Menard, *Universite Laval / CERVO Brain Research Center, Quebec City, Canada*

**Barrier-Omics in Health and Disease**
C. Betzholtz, *Uppsala University, Sweden*; R. Daneman, *University of California, San Diego*

**CNS Barrier Access for Therapeutics**
V. Gradinaru, *Caltech, Pasadena, California*; M. Pizzo, *Denali Therapeutics, South San Francisco, California*

**Panel Discussion: Barriers to Success for Women and Minorities in Academia and Industry—Impact and Solutions**
B. Engelhardt, *University of Bern, Switzerland*; C. Menard, *Université Laval, Québec City, Canada*; M. Pizzo, *Denali Therapeutics, South San Francisco, California*

**Imaging Approaches to Study CNS Barriers**
A. Ben-Zvi, *Hebrew University of Jerusalem, Israel*; H. Junge, *University of Minnesota, Minneapolis*

**Bioengineering a Better In Vitro Barrier**
N. MacAulay, *University of Copenhagen, Denmark*; L. Pelligrini, *MRC Laboratory of Molecular Biology, Cambridge, United Kingdom*

**Workshop: How to Frame and Write Constructive, Fair Peer Review**
J. Zuchero, *Denali Therapeutics, South San Francisco, California*

**Barriers at Choroid Plexus, Meninges, and Circumventricular Organs**
D. McGavern, *NINDS, National Institutes of Health, Bethesda, Maryland*; J. Strahle, *Washington University School of Medicine in St. Louis, Missouri*

F. Garcia, N. Khoury

Poster discussions

# Systems Immunology

**April 18–22** 348 Participants (237 in-person, 111 virtual)

ARRANGED BY **Menna Clatworthy,** University of Cambridge, United Kingdom
**Kathryn Miller-Jensen,** Yale University, New Haven, Connecticut
**Harinder Singh,** University of Pittsburgh, Pennsylvania
**John Tsang,** Yale University, New Haven, Connecticut

J. Tsang, K. Miller-Jensen, M. Clatworthy, H. Singh

The third Cold Spring Harbor Laboratory meeting on Systems Immunology was held on campus from April 18 to April 22, 2023. There were 348 participants, of whom 45% were female, 29% were graduate students, and 18% were postdocs. Notably, 45 corporate scientists and four journal staff were also among the attendees. The meeting included 55 talks and 126 posters. The biennial meeting continues to bring together scientists working at the interface of experimental immunology and computational and systems biology. The meeting is helping to nucleate and foster a vibrant community of systems immunologists, with many in the next generation (assistant professors, postdocs, and graduate students) viewing it as the main forum in the field. The scientific program featured sessions on single-cell multiomic analyses, modeling of signaling and gene regulatory networks, high-resolution cellular and molecular profiling of immune cell states including cell–cell communication, engineering of immune cells with therapeutically beneficial capabilities, exploration of B- and T-cell antigen repertoires and their specificities, and systems-level dissection of human immunity at different scales of organization. Given the continuing progress, two sessions were devoted to Human Systems Immunology. The format of the meeting included oral presentations from invited speakers as well as those selected from abstracts. The majority of the selected talks from abstracts were given by

T. Aurich, C. Kaur

S. Gao, S. Freeman

A. Hoffmann, Y. Zhang

R. Thomas, J. Thakar

graduate students or postdocs. Based on the vigorous discussions spawned at the oral and poster presentations and the informal feedback received, the meeting appeared to be a resounding success. We expect the next meeting to be held in April 2025 to sustain and nurture this vital and rapidly evolving field.

Support for this meeting was provided in part by the National Institute of Allergy and Infectious Diseases, a branch of the National Institutes of Health.

## PROGRAM

**Single-Cell Multi-Ome Analysis**
C. Leslie, *Memorial Sloan Kettering Cancer Center, New York, New York*; H. Singh, *University of Pittsburgh, Pennsylvania*

**Human Systems Immunology I**
John Tsang, *Yale University, New Haven, Connecticut/National Institute of Allergy and Infectious Diseases, NIH, Bethesda, Maryland;* Shai Shen-Orr, *Technion-Israel Institute of Technology, Haifa, Israel*

**Modeling of Immune Signaling and Gene Regulatory Networks**
G. Altan-Bonnet, *National Cancer Institute, NIH, Bethesda, Maryland*; A. Hoffmann, *University of California, Los Angeles*

**Cellular Dynamics, Interactions, and Communication**
R. Germain, *National Institute of Allergy and Infectious Diseases, NIH, Bethesda, Maryland*; K. Miller-Jensen, *Yale University, New Haven, Connecticut*

M. Hale, G. Kundu

D. Amsen, C. Gerlach

Y. Feng, G. Altan-Bonnet

**Engineered Cells and Systems**
M. Davis, *HHMI, Stanford University School of Medicine, California*; R. Pompano, *University of Virginia, Charlottesville*

**Tissue Systems Immunology**
M. Clatworthy, *University of Cambridge, United Kingdom*; D. Fowell, *Cornell University, Ithaca, New York*

**Immunoreceptors: Specificities and Signaling**
J. Jiang, *University of Pennsylvania, Philadelphia*

**Human Systems Immunology II**
D. Farber, *Columbia University, New York, New York*; E. McKinney, *University of Cambridge, United Kingdom*

Blackford lawn

# Ubiquitins, Autophagy, and Disease

**April 25–29** 239 Participants (178 in-person, 61 virtual)

ARRANGED BY **Anne Bertolotti,** MRC Laboratory of Molecular Biology, United Kingdom
**Michael Rape,** University of California at Berkeley
**Nicolas Thomä,** Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland

M. Rape, A. Bertolotti, N. Thomä

This is the 13th Ubiquitin, Autophagy, and Disease Meeting (previously titled Ubiquitin Family Meeting) following the successful inauguration of this series in 2003. The ubiquitin field has evolved from groundbreaking enzymatic and structural work to the physiology of ubiquitin-dependent signaling in development and disease, particularly its function in autophagy, to biomedical applications that alter the activity of ubiquitin-dependent signaling in disease. To reflect the latter, two sessions were now dedicated to the therapeutic modalities of induced protein degradation and ubiquitin-based drug discovery. As in previous years, this meeting discussed fundamental questions of ubiquitin-dependent signaling, including how the specificity of ubiquitin or ubiquitin-like conjugation is established; how the many ubiquitylation enzymes control crucial signaling events in protein quality control, the cell cycle, or development; how aberrant ubiquitylation contributes to disease; and how the ubiquitin system can be leveraged to fight disease states. As such, the Cold Spring Harbor meeting brings together researchers from diverse areas of biology, fostering deeply informed, open, and creative discussions.

An important highlight of the 2023 ubiquitin meeting was two exciting keynote lectures. The opening keynote lecture was delivered by Ingrid Wertz, who spearheaded the discovery and translation of ubiquitin-based therapies, detailing the small-molecule modulators of the ubiquitin system that have recently entered the clinic. The second keynote lecture was delivered by David

J. Tatzelt, K. Winklhofer

D. O'Hara, L. Rieger

P. Brzovic, D. Ivanov

Komander, a department head at Wehi, Australia, who has made groundbreaking discoveries detailing the biology of deubiquitinases and their role in disease.

The meeting had 239 participants, and more than half were international guests from 22 countries. These witnessed many scientific premieres with unpublished studies that underscored the rapid pace of discovery and the complexity of the ubiquitin field. To name but a few, we saw the first molecular mechanism of arsenic trioxide in the treatment of promyeloid leukemia, one of the rare events when cancers can be cured; the first structure of the fully assembled enzyme complex behind cancer-prone human papillomavirus; new ubiquitin ligases that can be harnessed in drug discovery; novel quality control mechanisms and protein modifications that trigger degradation; exciting new players and mechanisms in autophagy; and novel mechanisms demonstrating how molecular glues can inhibit deubiquitylates.

We are now seeing ubiquitin researchers in companies and academic centers working hand-in-hand to develop ubiquitin-directed drugs. These active and engaged discussions about basic ubiquitin biology and its medical uses were continued in the poster session, over dinner, and at social events. Many talks were presented by graduate students and postdoctoral researchers introducing the next generation of leaders in this field to the community. The Cold Spring Harbor meeting is without a doubt the key scientific event in the field in 2023. The field in its collaborative and interdisciplinary nature is very much alive and fast-moving—propelled to new heights by exciting drug discovery angles.

S. Elledge, H.-C. Yen

I. Bezsonova, E. Korchak

I. Koren, W. Harper

C. Wolberger, N. Thomä

Y. Zhang, W. Harper

## PROGRAM

**Keynote Speaker**
I. Wertz, *Lyterian Therapeutics, South San Francisco, California*

**Therapy I**
G. Winter, *CeMM Research Center for Molecular Medicine of the Austrian Academy of Science, Vienna*; N. Zheng, *University of Washington, Seattle*

**Autophagy**
H. An, *Memorial Sloan Kettering Cancer Center, New York, New York*; R. Zoncu, *University of California, Berkeley*

**Structure and Mechanism I**
C. Lima, *Sloan Kettering Institute/HHMI, New York, New York*; C. Wolberger, *Johns Hopkins University School of Medicine, Baltimore, Maryland*

**Therapy II**
E. Fischer, *Dana-Farber Cancer Institute, Harvard Medical School, Boston, Massachusetts*; C. Woo, *Harvard University, Cambridge, Massachusetts*

**Signaling I**
A. Martin, *University of California, Berkeley/HHMI*; J. Zuber, *Research Institute of Molecular Pathology (IMP), Vienna, Austria*

**Signaling II**
A. Werner, *National Institutes of Health, Bethesda, Maryland*; H-C. Yen, *Academia Sinica, Taipei, Taiwan*

**Keynote Speaker**
David Komander, *WEHI, Parkville, Victoria, Australia*

**Structure and Mechanism II**
B. Schulman, *Max Planck Institute of Biochemistry, Martinsried, Germany*; R. Voorhees, *California Institute of Technology, Pasadena*

**Signaling III**
I. Dikic, *Goethe University Frankfurt, Germany*; E. Holzbauer, *University of Pennsylvania, Philadelphia*

# Telomeres and Telomerase

**May 2–6** 309 Participants (223 in-person, 86 virtual)

ARRANGED BY **Peter Baumann,** IMB, Johannes Gutenberg University, Germany
**Jan Karlseder,** The Salk Institute, La Jolla, California
**Deborah Wuttke,** University of Colorado, Boulder

J. Karlseder, P. Baumann

This was the thirteenth conference in the CSHL Telomeres and Telomerase series, which has been held every two years since 1999. The conference consisted of eight sessions of talks and two poster sessions. As for every one of these meetings, the format was to invite two chairs per session, choosing a mix of established scientists in the field and younger scientists who had already made their mark as independent investigators. Session chairs were given the choice between giving a 15-minute presentation themselves or having a member of his/her laboratory give a talk. The remainder of the talks (also 15 minutes with five minutes of discussion) were chosen from submitted abstracts, allowing as many presentations as possible. Oral presentations were primarily given by graduate students and postdoctoral fellows. There were 223 in-person participants, a large fraction of whom presented the 98 posters and 63 talks.

The talks and posters covered diverse aspects of telomere and telomerase biology, including regulation of telomerase expression and activity, telomerase biogenesis and structure, telomere replication, mechanisms of alternative lengthening of telomeres (ALT), mechanisms of telomere protection, telomere protein functions at telomeres and throughout the genome, telomere shortening and mechanisms of senescence and aging, and the role of telomeres in human health and disease.

The quality and novelty of scientific content were very high throughout the conference in both the talks and the posters. Most of the presented data were unpublished and extensively discussed in an open fashion. Formal and informal discussions were lively and informative. The conference was judged to be highly successful based on verbal and email communications received by the organizers. There is strong enthusiasm for another meeting on the same topic in 2025.

T. Cech, R. Wellinger

O. Carlund, P. Osterman

E. Morales, A. Mojiri, T. Schmidt

G. Aubert, S. Agarwal

T. Hoang, D. Nguyen, Z. Sekne, T. Goncalves, E. Zanella

Support for this meeting was provided in part by the National Institute on Aging, a branch of the National Institutes of Health.

## PROGRAM

**Chromosome and Genome Stability**
T. de Lange, *The Rockefeller University, New York, New York*; M. Hayashi, *Kyoto University, Japan/IFOM ETS, Milan, Italy*

**Replication Stress**
H. Pickett, *Children's Medical Research Institute, Westmead, New South Wales, Australia*; G. Raffa, *Sapienza Università di Roma, Italy*

**Telomerase Biogenesis and Regulation**
T. Cech, *HHMI/University of Colorado, Boulder*; T.H.D. Nguyen, *MRC Laboratory of Molecular Biology, Cambridge, United Kingdom*

**Telomeropathies, Premature Aging, and Cancer Predisposition**
M. Armanios, *Johns Hopkins School of Medicine, Baltimore, Maryland*; A. Bertuch, *Baylor College of Medicine, Houston, Texas*

**Telomere Protection and DNA Damage Signaling**
S. Artandi, *Stanford University, California*; C. Greider, *University of California, Santa Cruz*

**Telomere Function during the Cell Cycle and Cell Death Regulation**
R. Greenberg, *University of Pennsylvania, Philadelphia*; D. Hockemeyer, *University of California, Berkeley*

**Regulation of Immortality by Telomerase**
J. Kanoh, *University of Tokyo, Japan*; J. Lingner, *EPFL, Lausanne, Switzerland*

**Regulation of Immortality by ALT**
R. Reddel, *Children's Medical Research Institute, Westmead, New South Wales, Australia*; R. Wellinger, *Université de Sherbrooke, Québec, Canada*

S. Ahmed, C. Greider

Y. Lin, M.T. Teixeira

# The Biology of Genomes

**May 9–13** 522 Participants (425 in-person, 97 virtual)

ARRANGED BY **Christina Curtis,** Stanford University, California
**Hopi Hoekstra,** Harvard University, Cambridge, Massachusetts
**Tuuli Lappalainen,** New York Genome Center, KTH Royal Institute of Technology & SciLifeLab
**John Marioni,** European Bioinformatics Institute, Hinxton, United Kingdom

The 2023 Biology of Genomes meeting brought the community back to Cold Spring Harbor in full force, with the hybrid format allowing attendees from across the world to participate. As always, the meeting was a highlight of the genomics year, featuring world-class talks from senior academics as well as showcasing the work of postdoctoral and predoctoral trainees. The community was also happy to have the opportunity to congratulate Svante Pääbo in person for his 2022 Nobel Prize; he is a long-term Biology of Genomes attendee and a former keynote speaker and was this year given a poster presentation slot.

T. Lappalainen, C. Curtis

The meeting started on Tuesday, May 9, with an outstanding session on Population Genomics, chaired by Ian Mathieson, who described how ancient DNA provides insights into natural selection in humans, and Shamil Sunyaev, who discussed the gap in our knowledge of regulatory mechanisms of genome-wide association studies (GWAS) loci and potential answers. Altogether, the session represented diverse approaches to understanding evolution, architecture, function, and phenotypic impact of genetic variation in different species.

The second session, on Developmental and Single-Cell Genomics, focused on molecular and cellular processes of developmental trajectories of cells and their complex functions as tissues, with Dan Landau and Marta Mele providing powerful demonstrations of these processes. An after-lunch panel discussion hosted by NHGRI provided insights into collaborative research and consortia that have a major role in the genomics field. The afternoon session on Wednesday addressed computational genomics, with outstanding talks on new computational methods for processing large genomic data sets and extracting biological insights (e.g., into regulatory circuits)

S. Reilly, S. Gosai

J. Kribelbauer, L. Pinello

S. Agrawal, T. Gingeras

X. Guitart, A. Goncalves

and tissue structure in single-cell and spatial data, with another major theme being large-scale whole-genome analysis. This was followed by the Wine and Cheese Party in sunny and warm spring weather and the first poster session.

Thursday, May 11, kicked off with a session on Complex Traits and Microbiome, with Ran Blekhman as the sole chair on behalf of Melina Claussnitzer, who was not able to attend. He discussed human microbiome diversity. Other major topics included the functional characterization of GWAS loci, with single cells continuing to be a major topic. The day continued with a session on Functional Genomics and Epigenetics, with Doug Fowler describing high-throughput assays of protein-coding variant function, and Lars Velten giving a presentation on single-cell technologies—their topics representing the major themes of the session.

The final full day of the meeting began with a shorter session on Systems Genetics, with Ben Lehner giving an exciting talk about genetic variant effects on protein structure. Another shorter session on Evolutionary and Non-Human Genomics was chaired by Doric Bachtrog. This was a tour de force into genome evolution with multiple excellent talks. The final poster session was followed by the keynote speakers, Evan Eichler and Erich Jarvis, who each described different aspects of long-read genomes and their interpretation, drawing upon the landmark pangenome papers published just before the meeting.

Support for this meeting was provided in part by the National Human Genome Research Institute (NHGRI), a branch of the National Institutes of Health; PacBio; Oxford Nanopore Technologies; Complete Genomics; and the James P. Taylor Foundation for Open Science Scholarship Fund.

M. Cuoco, S. Pääbo

M. Corominas, P. Llorens-Giralt

S. Kundu, J. Schreiber

M. Murphy, L. Sweet, S. Song

## PROGRAM

**Population Genomics**
I. Mathieson, *University of Pennsylvania, Philadelphia*; S. Sunyaev, *Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts*

**Developmental and Single-Cell Genomics**
D. Landau, *Weill Cornell Medicine, New York, New York*; M. Mele, *Barcelona Supercomputing Center, Spain*

**NHGRI Lunchtime Panel Discussion Session: Trainees and Early Career Researchers: Tips for Team Science**
A. Felsenfeld, *NHGRI, National Institutes of Health, Bethesda, Maryland*; D. Fowler, *University of Washington, Seattle*; C. Hutter, *NHGRI, National Institutes of Health, Bethesda, Maryland*; D. Kaufman, *NHGRI, National Institutes of Health, Bethesda, Maryland*; T. Lappalainen, *New York Genome Center, New York*

**Computational Genomics**
X. Lin, *Harvard T.H. Chan School of Public Health, Boston, Massachusetts*; J. Zou, *Stanford University, California*

**Complex Traits and Microbiome**
R. Blekhman, *University of Chicago, Illinois*; M. Claussnitzer, *Broad Institute, MGH/Harvard Medical School, Boston, Massachusetts*

**Functional Genomics and Epigenetics**
D. Fowler, *University of Washington, Seattle*; L. Velten, *Centre for Genomic Regulation (CRG), Barcelona, Spain*

**ELSI Panel and Discussion: Scientists' Roles and Responsibilities Combatting the Misuse of Genomic Research**
K. Dasgupta, *UCLA Institute for Society and Genetics*; D. Kaufman, *NHGRI, National Institutes of Health, Bethesda, Maryland*; D. Martschenko, *Stanford Center for Biomedical Ethics, California*; A. Smart, *Massachusetts Institute of Technology, Cambridge*; R. Wedow, *Purdue University*, *West Lafayette, Indiana*

**Systems Genetics**
B. Lehner, *Wellcome Sanger Institute, Cambridge, United Kingdom*

**Evolutionary and Nonhuman Genetics**
D. Bachtrog, *University of California, Berkeley*

**Guest Speakers**
E. Eichler, *University of Washington, Seattle/HHMI*; E. Jarvis, *Rockefeller University, New York/HHMI*

**Cancer and Medical Genomics**
E. Khurana, *Weill Cornell Medicine, New York, New York*; P. van Loo, *The University of Texas MD Anderson Cancer Center, Houston*

# Mechanisms of Metabolic Signaling

**May 16–20** 313 Participants (234 in-person, 79 virtual)

ARRANGED BY **Bart Deplancke,** EPFL, Lausanne, Switzerland
**Jared Rutter,** HHMI/University of Utah School of Medicine, Salt Lake City
**Kathryn (Katy) Wellen,** University of Pennsylvania, Philadelphia

K. Wellen, B. Deplancke, J. Rutter

This was the fifth Cold Spring Harbor Laboratory meeting on Metabolic Signaling & Disease, held from May 16 to May 20, 2023, with 313 participants, 234 of whom were in-person attendees and 79 of whom attended the meeting virtually through Zoom. This meeting followed a highly successful Mechanism of Metabolic Signaling meeting held fully virtually one and a half years earlier.

The main goal of this meeting was to bring together researchers from diverse fields to explore how principles of cellular metabolism manifest in different cell types; how metabolic regulation underlies the functions of specialized tissues; and how these differences impact both normal physiology and diseases such as mitochondrial disease, diabetes, and cancer.

The 24 invited speakers were leaders in the various aspects of metabolic research from all over the world, and the majority had not been part of the previous versions of this meeting in 2017, 2019, or 2021 programs. The meeting included two inspiring keynote addresses by Susan Kaech and Mitch Lazar. Dr. Kaech is a world leader in the burgeoning field of immunometabolism and discussed the metabolic underpinnings of T-cell activation. Dr. Lazar presented a combination of historical perspective and new findings on the functions of nuclear hormone receptors in metabolic regulation. There were seven sessions, all of which featured oral presentations selected from the submitted abstracts, all highlighting unpublished research and focused on key areas in the field of metabolism. Emphasis included, but was not limited to, genomic and epigenomic mechanisms, signaling pathways, lipid flux and storage, and mitochondrial function, as well

A. Walker, J. Baskin

C. Taylor, P. Zahradka

C. Thompson, M. Lazar, S. Mandrup

A. Dillin, B. Manning, S. Kaech

as emerging technologies to study metabolism—with an accent on comparing and contrasting normal and pathologic metabolic states. Short talks were chosen from abstracts to increase the exposure of younger investigators and to highlight hot topics that complemented and extended the exciting program. In all, there were 53 talks by speakers from Canada, Europe, Australia, and Asia, in addition to the United States. Twenty of the talks were given by women.

In addition to the seven oral sessions, there were two poster sessions featuring a total of 130 posters. There was engagement from most of the attendees for the poster sessions and the interactions were robust throughout. All seven sessions were characterized by open and wide-ranging discussions, both immediately after each talk and during the breaks, meals, and poster sessions. A dedicated Slack channel also facilitated further discussion opportunities. The meeting provided a unique forum for the exploration of the commonalities and differences in metabolic principles and details across different laboratories, systems, and diseases. All attendees gained in-depth exposure to the remarkable cell, organ, and disease specificity of metabolic flux and its regulation. Indeed, a great success of the meeting was its interactive nature, whereby stimulating questions and discussion led to new concepts and future collaborations.

This meeting was funded in part by the National Institute of Diabetes and Digestive and Kidney Diseases, a branch of the National Institutes of Health.

R. Perera, G. Rademaker

C. Ribeiro, Y-Y. Pesch

K. Biedsoe, J. Mellor

## PROGRAM

**Opening Keynote and Metabolism in Different Cell Types**
K. Wellen, *University of Pennsylvania, Philadelphia*

**Keynote Speaker**
S.M. Kaech, *Salk institute for Biological Study, San Diego, California*

**Emerging Technologies and Metabolism**
J. Rutter, *University of Utah School of Medicine, Salt Lake City*

**Organellar Metabolism**
J. Ellis, *East Carolina University, Greenville, North Carolina*

**Interorgan Metabolic Cross Talk**
A. Rose, *Monash University, Clayton, Victoria, Australia*

**Metabolism, Growth, and Proliferation**
J.Y. Guo, *Rutgers Cancer Institute of New Jersey, New Brunswick*

**Metabolite Signaling**
I. Harris, *University of Rochester Medical Center, New York*

**Closing Keynote and Metabolic Cross Talk with Epigenome/Genome**
B. Deplancke, *École Polytechnique Fédérale de Lausanne, Switzerland*

**Keynote Speaker**
M.A. Lazar, *University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania*

# Retroviruses

**May 22–27** 430 Participants (335 in person, 95 virtual)

ARRANGED BY **Andrés Finzi,** Université de Montréal, Québec, Canada
**Viviana Simon,** Icahn School of Medicine Mount Sinai, New York, New York

D. Ivanov, A. Finzi

The 2023 CSH Retroviruses meeting was arranged by Drs. Viviana Simon and Andrés Finzi. The meeting is considered the best for basic biology of retroviruses including HIV, bringing together scientists from around the world, fostering friendships and collaborations. It has frequently been the venue where major scientific breakthroughs have been announced prior to peer-review publication. The 2023 CSH Retrovirus meeting offered a hybrid mode of participation, with 335 individuals attending in person and 95 virtually. Forty-one percent of the attendees were female and 48% of the attendees were trainees. This year, 10 sessions were grouped on the viral life cycle and key viral or host proteins important in retroviral infection. There was an emphasis on pathogenesis with the presentation of a new animal model for HIV vaccine development and latency. Several exciting presentations covered new high-resolution structures of HIV-1 Env, integrase, Nef, retroviral capsid lattice, and restriction factors.

The keynote speakers were Drs. Jaquelin Dudley and Frank Kirchhoff. Dr. Dudley provided a comprehensive historical review of research on the mouse mammary tumor virus (MMTV) and how this research modeled current retrovirology. She also spoke about the additional hurdles that female scientists face to do research. Dr. Kirchhoff gave an overview of his work on SIV and HIV accessory proteins and how they impact pathogenesis. He also presented innovative approaches to identify new restriction factors as well as retroviral inhibitors. Both keynote presentations were very well received by all attendees. Many trainees expressed to the organizers how special and influential these two presentations were, particularly for young female scientists.

S. Kharytonchyk, I. Rouzina

Z. Williams, O. Cingoz, J. Coffin, M. Thali

G. Melikian, G. Raghunath

A. Engelman, Guest, H. Levin, H. Veler

The oral sessions were notable for their mix of presentations by both junior and senior researchers, an important feature of this meeting. A very limited number of oral presentations were given virtually, which worked well thanks to the excellent technical AV setup and know-how of the CSHL team. The poster presentations were spread out over three sessions. These poster sessions were very well attended and characterized by lively discussions. A novelty of the 2023 CSH Retrovirus meeting was the introduction of the "Retrocards." The idea behind this ludic activity is to facilitate exchanges between trainees and principal investigators (PIs): Trainees have to reach out to PIs, ask any question related to training, research, career development, etc., and in exchange for their question they get a Retrocard. The trainee with the most Retrocards wins a prize at the end of the meeting. This activity was very well received by both PIs and trainees and fostered communication between them. Several PIs and trainees told the organizers that this activity facilitated meeting new people, starting discussions with senior investigators, etc. They strongly suggested to keep this "icebreaker" activity in future Retrovirus meetings.

As in past years, awards were presented to a distinguished PDF (Andy Kaplan Prize), a distinguished senior graduate student (Uta von Schwedler Prize), and the best poster presentation of the meeting (Daniel Wolf Prize).

This meeting was funded in part by the National Institute of Allergy and Infectious Diseases, a branch of the National Institutes of Health, and by ViiV Healthcare.

Y. Kayode, M. Abajorga

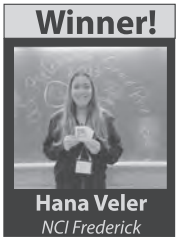
H. Gottlinger, J. Luban










# Retrocards 2023

Trainees at the 2023 Retroviruses meeting competed to win the most Retrocards.

*Concept:* **Viviana Simon** *and* **Andres Finzi** *Design:* **Meredith Cassuto**

T. Nyalile, C. Santos da Costa

## PROGRAM

**Entry**
E. Freed, *National Cancer Institute, Frederick, Maryland*; M. Lu, *University of Texas at Tyler Health Science Center*

**Post-Entry Events**
G. Melikian, *Emory University, Atlanta, Georgia*; A. Zamborlini, *INSERM, CEA-University Paris Saclay, France*

**Restriction Factors, Innate Immunity, and Nuclear Entry**
M. Ohainle, *University of California, Berkeley*; S. Yoh, *Scripps Research, La Jolla, California*

**Keynote Speaker**
J. Dudley, *University of Texas at Austin*

**Integration**
A. Engelman, *Dana-Farber Cancer Institute, Boston, Massachusetts*; M. Lusic, *Heidelberg University Hospital, Germany*

**Pathogenesis**
F. Maldarelli, *National Cancer Institute, Frederick, Maryland*; L. Manganaro, *INGM-Istituto Nazionale di Genetica Molecolare, Milan, Italy*

**Latency**
E. Browne, *University of North Carolina, Chapel Hill*; C. Goffinet, *Charite-Universitatsmedizin Berlin, Germany*

**RNA, Packaging and Assembly**
C. Carter, *Renaissance School of Medicine, Stony Brook, New York*; S. Stavrou, *SUNY University at Buffalo, New York*

***Seventeenth Annual* Andy Kaplan Prize**
Presented by A. Ono, *University of Michigan Medical School*
Awarded to J. Johnson, *University of Utah School of Medicine*

**Keynote Speaker**
F. Kirchhoff, *Ulm University Medical Center, Ulm, Germany*

**Capsid and Maturation**
C. Aiken, *Vanderbilt University Medical Center, Nashville, Tennessee*; C. Jolly, *University College London, United Kingdom*

***Twelfth Annual* Uta von Schwedler Prize for Retrovirology**
Presented by W. Mothes, *Yale University, New Haven, Connecticut*
Awarded to J. Xu, *University of Utah, Salt Lake City*

**Accessory Genes**
J. Dikeakos, *Western University, London, Ontario, Canada*; L. Etienne, *ENS-Lyon, INSERM, CNRS, France*

***Thirteenth Annual* Daniel Wolf Prize**
Presented by Stephen Goff, *Columbia University, New York, New York*
Awarded to best poster presentation

**Transcription, Splicing, and ERV**
L. Parent, *Penn State College of Medicine, Hershey, Pennsylvania*; R. Taube, *Ben Gurion University, Beer Sheva, Israel*

# Genome Engineering: CRISPR Frontiers

**August 16–20** 720 Participants (440 in person, 280 virtual)

ARRANGED BY **Brittany Adamson,** Princeton University, New Jersey
**Maria Jasin,** Memorial Sloan Kettering Cancer Center, New York, New York
**David R. Liu,** Broad Institute of Harvard and MIT, Boston, Massachusetts
**Sam Sternberg,** Columbia University, New York, New York

Emeriti organizers **Jennifer Doudna** and **Jonathan Weissman**

D.R. Liu, S. Sternberg, J. Doudna, J. Weissman, M. Jasin, B. Adamson

The 2023 Genome Engineering: CRISPR Frontiers meeting was the ninth consecutive conference in the series held at Cold Spring Harbor. The attendance was again robust, reaching the limit of in-person attendees. The abstract book cover highlighted the 50th anniversary of the generation of recombinant DNA, a precursor to genome editing. The emeriti organizers participated in the meeting as speakers, including one as a keynote and the other providing a "regular" talk and the introduction.

The meeting this year was a half day longer than in previous years, with a full day on Saturday, to decompress the meeting. It was agreed that this format worked well to make a quite intense meeting more relaxed. Attendees were also offered time to participate in volleyball or other sports, and to go on tours of CSHL, which was very popular (175 participants in campus tours).

The goal of this meeting continued to be met in terms of bringing together researchers working in diverse fields to stimulate discussions and ideas to further exploit CRISPR and related technologies for biological discovery and medical applications. Session titles were CRISPR Biology, Disease Biology, Therapeutics, CRISPR Technologies I and II, DNA Repair, and Emerging Technologies. Twenty-five speakers were invited to cover these diverse topics, including some that had not previously attended this meeting (e.g., Kiran Musunuru, Wasim Qasim, Tanya Paull, Paula Rio, and Jack Wang) and young faculty who had not previously spoken (e.g., Stan Brouns, Teresa

D. Sashital, N. Craig

F. Mohammad, K. Agger

M. Saito, Z. Chen

P. Matheus Carnevali, D. Goltsman

Davoli, and Peter Ly). Another approximately 33 speakers were chosen from submitted abstracts. One speaker was virtual because of visa issues.

Highlights included the keynote presentations by Jonathan Weissman (epigenetic modifiers) and Fyodor Urnov (facilitating CRISPR cures) and a presentation by Jack Wang on tree editing, the first talk on that topic at the meeting. Jonathan spoke in the last session, which likely helped retain a good portion of the attendees until the last talk. The other talks highlighted a diverse array of topics, ranging from basic CRISPR biology to in vivo gene editing for correcting human disease. A large number of talks focused on technology development, with a particular focus on prime editing. The latter reflects the strong interest in developing this platform for diverse applications, although it may have been overrepresented based on the comments received by the organizers.

The meeting was highly interactive as evidenced by many questions after the talks and well-attended poster sessions. The number of posters was very large (168) and the sessions relatively short (five hours in total), perhaps making it difficult for attendees to view many of the posters.

Although there was an attempt to reach parity, the percentage of female speakers still needs improvement, and although speakers represented diverse ethnic backgrounds, few were from diverse racial groups.

The meeting continued to have a large number of corporate attendees; included within the meeting were 2 workshops hosted by companies (Cell Microsystems and IDT).

It was planned to hold the meeting again in the summer of 2024, keeping with the half-day-longer format.

J. Boucher, E. Sontheimer

T. Loveless, L. Mayer

Support for this meeting was provided in part by Agilent Technologies; BEX Co., Ltd.; Cell Microsystems; Horizon Discovery; Integrated DNA Technologies; Ionis; and Synthego.

## PROGRAM

**CRISPR Biology**
J. Bondy-Denomy, *University of California, San Francisco*; D. Sashital, *Iowa State University, Ames*

**Keynote Speaker**
F. Urnov, *University of California, Berkeley*

**Disease Biology**
M. Maeder, *Chroma Medicine, Cambridge, Massachusetts*; F. Sanchez-Rivera, *MIT/Koch Institute, Cambridge, Massachusetts*

**CRISPR Technologies I**
T. Loveless, *Rice University, Houston, Texas*; E. Sontheimer, *UMass Chan Medical School, Worcester*

**DNA Repair**
A. Ciccia, *Columbia University Irving Medical Center, New York, New York*; T. Davoli, *Institute for Systems Genetics, New York, New York*

**CRISPR Technologies II**
L. Gilbert, *University of California, San Francisco*; N. Sanjana, *New York Genome Center, New York*

A. Pascual-Garrigos, F. Neumann

**Therapeutics**
K. Musunuru, *University of Pennsylvania Perelman School of Medicine, Philadelphia*; P. Rio, *CIEMAT/CIBERER/IIS-FJD, UAM, Madrid, Spain*

**Keynote Speaker**
J. Weissman, *Whitehead Institute, Massachusetts Institute of Technology/HHMI, Cambridge*

**Emerging Technologies**
A. Ke, *Cornell University, Ithaca, New York*; A. Komor, *University of California, San Diego*

CRISPR volleyball competition

# Eukaryotic mRNA Processing

**August 22–26** 433 Participants (374 in-person, 59 virtual)

ARRANGED BY **Javier F. Caceres,** MRC Human Genetics Unit, University of Edinburgh, United Kingdom
**Tracy Johnson,** University of California, Los Angeles
**Yongsheng Shi,** University of California, Irvine

Y. Shi, T. Johnson, J. Caceres

The 14th Eukaryotic mRNA processing meeting was held this summer to present and discuss recent developments in mRNA metabolism in eukaryotes. Following a fully virtual meeting in 2021 because of the pandemic, this year the meeting was held in person but also included virtual participation. Eighty-six percent of all participants attended in person. The meeting consisted of sessions on Mechanisms of pre-mRNA Splicing, Alternative Splicing, RNA Processing and Disease, 3′ End Processing, Gene Regulation by RNA Modification, RNA–RNA and RNA–Protein Interactions, Co-Transcriptional RNA Processing, and RNA Turnover and Quality Control. Throughout the meeting, participants were introduced to cutting-edge technologies to study the expression, regulation, and biochemical properties of RNA.

As per tradition for this meeting, oral presentations were solely selected from submitted abstracts focusing on unpublished work primarily from graduate students, postdoctoral researchers, and junior faculty. This meeting continues to serve a critical role in the field as a place for early-career researchers to present their research to an international audience and to discuss new and exciting unpublished findings. This year, there were 76 selected talks, two Keynote talks, and 186 posters, with 433 participants in total. Organizers and participants felt that the use of Slack helped promote spontaneous 1:1 and group interactions that would otherwise have been nearly impossible for remote participants.

In the session on Mechanisms of RNA Splicing, highlights included studies to reveal the structure of the intron branch point intermediates and real-time visualization of spliceosome assembly in

S. Brenner, L.-L. Chen

E. Bechara

A. Krainer, M. Hagiwara

A. Mayeda, A. Srebrow

living cells. Moreover, a number of exciting translational studies were presented that discussed a range of disorders associated with mutations in splicing factors. These were detailed in this session as well as the session on RNA Processing in Disease, providing tantalizing hints at opportunities for RNA therapeutics. In sessions on Alternative Splicing, there were presentations on mechanisms and networks of alternative splicing and alternative exon production as well as their biological consequences. Highlights include the new concept of hybrid exons, generated from alternative transcription start-site usage, back-splicing mechanisms that yield circular RNAs, and evidence for coordination of splicing and polyadenylation. The application of massively parallel assays combined with machine-learning approaches provided new insights into the regulatory mechanisms of splicing.

The 3′ End Processing session provided structural, molecular, and cellular insights into the mechanism and regulation of 3′ end formation of mRNAs and small noncoding RNAs. New, sophisticated techniques to globally identify RNA modifications and structures in cells were presented in the RNA Modification sessions. Sessions on co-transcriptional RNA Processing focused on the timing and coordination between transcription and RNA processing events. There was an important exploration of how strict the requirement is for splicing to be co-transcriptional and a discussion of how this knowledge could contribute to splicing-correcting therapeutic strategies for human hereditary disease. In the session on RNA Turnover and Quality Control, new insights into mechanisms of both nuclear and cytoplasmic mRNA quality control pathways were presented, as well as new findings on the central role of deadenylation in mRNA decay. A session on RNA–RNA and RNA–Protein Interactions covered a number of areas of eukaryotic mRNA processing in which RNA–protein complexes and RNA structures play central roles.

P. Gajduskove, K. Ruzicka

A. Fiszbein, K. Neugebauer

J. Manley, D. Rio

I. Gainetdinov, J. Wilusz

The keynote presentation by Dr. Kornblihtt highlighted previous and recent developments in co-transcriptional RNA processing and the use of epigenetic modulators in improving the treatment of spinal muscular atrophy. The second keynote speaker, Dr. Ling-Ling Chen, presented the latest research in her laboratory that uses circular RNAs as a therapeutic tool for treating autoimmune diseases such as psoriasis.

Research into RNA processing continues to yield important and surprising insights into cellular processes and disease and reveals extraordinary opportunities for new therapeutic strategies!

Support for this meeting was provided in part by the National Cancer Institute, a branch of the National Institutes of Health.

**PROGRAM**

**Keynote Speaker**
A. Kornblihtt, *University of Buenos Aires, Argentina*

**RNA Processing in Disease**
O. Anczukow, *Jackson Laboratory, Farmington, Connecticut*; O. Abdel-Wahab, *Memorial Sloan Kettering Cancer Center, New York, New York*

**Mechanism of RNA Splicing**
W. Galej, *EMBL, Grenoble, France*; K. Lynch, *University of Pennsylvania, Philadelphia*

**RNA Turnover and Quality Control**
S. Cherry, *University of Pennsylvania School of Medicine, Philadelphia*; H. Le Hir, *Institut de Biologie de l'ENS, Paris, France*

**Alternative Splicing**
A. Fiszbein, *Boston University, Massachusetts*; F. Heyd, *FU Berlin, Germany*

**3′ End Processing**
D. Bentley, *University of Colorado School of Medicine, Aurora, Colorado*

**Keynote Speaker**
L-L. Chen, *Chinese Academy of Sciences, Shanghai Institute of Biochemistry and Cell Biology*

**RNA–Protein Interactions/RNP Complexes**
D. Dominguez, *University of North Carolina, Chapel Hill*; S. Rouskin, *Harvard Medical School, Cambridge, Massachusetts*

**Co-Transcriptional RNA Processing**
D. Larson, *NCI, National Institutes of Health, Bethesda, Maryland*; K. Sträßer, *University of Giessen, Germany*

**RNA Modification and Technologies**
R. Gregory, *Boston Children's Hospital, Massachusetts*; N. Martinez, *Stanford University, California*

# Mechanisms of Eukaryotic Transcription

**August 29–September 2** 498 Participants (377 in-person, 121 virtual)

ARRANGED BY **Wendy Bickmore,** MRC Human Genetics Unit, University of Edinburgh, United Kingdom
**Michael Levine,** Princeton University, New Jersey
**Eva Nogales,** University of California, Berkeley

E. Nogales, M. Levine, W. Bickmore

The 18th biennial Mechanisms of Eukaryotic Transcription was convened at CSHL from August 29 to September 2, 2023. There were 67 oral presentations in addition to three poster sessions. The talks were organized into eight plenary sessions: enhancers, promoters and Pol II elongation, structure, development, transcription factors, imaging, chromatin, and 3D genome. Most of the talks were selected from submitted abstracts (about 43 of 67), and both invited and selected speakers were allocated 15 + 5 minutes for their talks and Q&A. This combination of invited and selected speakers ensured participation by a broad spectrum of scientists, from graduate students and postdocs to Assistant Professors and senior PIs. The talks were generally clear and concise and held to schedule. It is an open issue whether this "one size fits all" format is better than the more traditional distinction between invited and selected speakers (e.g., 20 + 5 minutes and 12 + 3 minutes, respectively). Nonetheless, as is typical for this meeting, Grace Auditorium was packed and energetic, and the discussions were lively and fun.

The meeting covered a broad range of disciplines and technologies—all united by a focus on elucidating mechanisms of transcription. Highlights included visualizing transcription using state-of-the-art imaging methods; the use of clever high-throughput assays to study transcription factors; combinations of transcription factors, chromatin structure, and function; and even a glimpse into the cellular organization of the mammalian brain. Structural studies were incisive and instructive and included compelling examples of how chromatin modifications are coupled to Pol II transcription through nucleosomes, or how gene silencing complexes like the Polycomb

L. Pereira, A. Dall'Agnese

M. Guertin

P. Whitney, L. Denes

A. Shilatifard

J.C. Andran, L. Simpson

PRC2 complex are down-regulated when they encounter regions of active transcription as marked by certain histone modifications. Higher-order associations of such transcription-associated complexes are generally accepted in the field, and the nature of these weak, multivalent associations (clusters, hubs, condensates, etc.) was actively discussed throughout the meeting.

Past meetings emphasized promoters and associated transcriptional machineries required for gene expression. These topics were nicely covered at this meeting as well, but there was also an emphasis on enhancers, celebrating the 40th anniversary of the discovery of the first cellular enhancer (*IgH* heavy chain gene). Some enhancers were shown to possess fixed arrangements of transcription factor binding motifs (grammar or syntax), to coordinate the binding of clusters of activators, to disengage from their promoters during transcription elongation, and to work over large genomic scales. However, despite these amazing properties, we still do not know exactly what they convey to their target promoters to trigger transcription.

The first Mechanisms of Eukaryotic Transcription meeting was held in 1989, and the latest incarnation sustained the energy and excitement that has long characterized this field of study. Of special note is the excellence of many of the talks and posters from the youngest members of the community. The future appears to shine bright for the field of transcription. In 2025 Eva Nogales will organize the meeting along with Wendy Bickmore and the newest member of the organizing committee, Joanna Wysocka. There is little doubt that it, too, will be a blast.

R. Kingston, S. Churchman, K. Adelman

M. Xie, R. Yi

J. Wysocka, M. Timmers, L. Tora

## PROGRAM

**Enhancers**
M. Levine, *Princeton University, New Jersey*

**Promoters and Pol II Elongation**
E. Nogales, *University of California, Berkeley*

**Structure**
W. Bickmore, *MRC Human Genetics Unit, Edinburgh, United Kingdom*

**Development**
K. Adelman, *Harvard Medical School, Boston, Massachusetts*

**Transcription Factors**
C. Kadoch, *Dana-Farber Cancer Institute, Harvard Medical School, Boston, Massachusetts*

**Imaging**
J. Wysocka, *Stanford University, California*

**Chromatin**
I. Cissé, *Max Planck Institute of Immunobiology and Epigenetics, Freiburg, Germany*

**3D Genome**
M. Harrison, *University of Wisconsin, Madison*

# Eukaryotic DNA Replication and Genome Maintenance

**September 5–9** 446 Participants (345 in-person, 101 virtual)

ARRANGED BY **Karlene Cimprich,** Stanford University, California
**John Diffley,** The Francis Crick Institute, London, United Kingdom
**Bob Duronio,** University of North Carolina, Chapel Hill

J. Diffley, K. Cimprich, B. Duronio

The 19th biennial meeting on Eukaryotic DNA Replication and Genome Maintenance was held at Cold Spring Harbor, September 5–9, 2023. This meeting provides an important forum for discussion and exchange of ideas in the field of chromosome replication and genome stability. The strategy adopted several years ago of broadening the conference scope to encompass the intersection of replication processes with genome stability mechanisms continues to work well. It has expanded the range of the research covered and brought together biochemists, cell biologists, and geneticists in new and meaningful ways, and this year the intersection of the fields was more evident than ever. We saw strong attendance with 446 participants in total and 345 participating in person. The percentage of female participants was 44.5% in person (57% virtual) and participation of students and postdocs was also strong with 183 registered in person (47 virtual). The meeting featured 80 platform talks, and because the number of posters was greater than usual, a third poster session was added. These sessions were extremely well attended and there was ample discussion that extended past the scheduled session times.

The 10 platform sessions reflected the great breadth of the field, covering a wide range of topics from detailed molecular mechanisms to global regulation, genomic tools, and human disease. Approaches ranged from high-resolution cryo-electron microscopy (cryo-EM) and single-molecule dynamics through to genome-wide analysis of replication origin use, histone deposition, and replication timing. Similarly, talks involved a wide range of model organisms complementing talks involving human cells. Talks drew many stimulating questions from the audience and provoked ongoing discussion that continued throughout the meeting. The meeting had sessions covering Replication Initiation Factors and Origin Activation, Replication Origin Choice, and Control and Fork Progression and Repair. Highlights of these sessions included a series of talks on

D. Cortez, B. Shen

J. Haber, C. Wiese

V. Ellison, S. Bell

V. Rojat, D. Ciardo

how replication is initiated and describing a new initiation factor, DONSON, which was studied using genetics, cryo-EM, and ensemble biochemistry. High-resolution cryo-EM structures were presented showing how different clamp loaders acted on leading and lagging strands and how the replicative helicase handled damaged DNA (e.g., G4 quadruplex structures). The dynamics of replication initiation were also explored using single-molecule approaches, and the use of artificial intelligence–based structure and docking predictions (AlphaFold) to build and test hypotheses was a common theme in these and other sessions. There were several sessions covering various aspects of processes critical for ensuring stable genome maintenance: Fork Stalling, Stabilization and Checkpoint Activation, Replication-Coupled Genome Instability and Disease, and both Causes and Consequences of Replication Stress. Among the highlights from these sessions was an increasingly clear view of mechanisms involved in responding to and recovering from fork stalling. The use of break mapping techniques revealed new insights into endogenous sources of DNA damage in cells and damage that results from deregulating the DNA damage response or silencing. Secondary DNA structures ranging from G4s to i-motifs and triplexes were also a focus of a number of talks that explored the factors involved in their resolution as well as their impact on fork progression. The use of reconstituted replication systems to study these was also elegantly demonstrated and reflects the continuing intersection of these fields and increased mechanistic understanding of that interface.

An additional session covered Developmental Control of Replication and Replication Timing, and the meeting ended with one on Chromatin and Genome/Nuclear Organization. These sessions highlighted noncanonical DNA replication programs of specialized cell types, including how hormone signaling contributes to early replication timing and how lagging chromosome fragments are recovered during polyploid mitoses. High-resolution temporal imaging of rapid embryonic cycles revealed how replication and transcription are coordinated to avoid genome destabilizing conflicts. Examining how the replication fork moves through chromatin revealed

J. Walter, K. Labib

N. Rhind, A. Bedalov

how the replisome component Mrc1 contributes to the movement of parental histones into chromatin behind the fork, and how other chaperoning activities of the replisome ensure the relatively equal distribution of parental and newly synthesized histones to both the leading and lagging strands, an activity important for correctly propagating epigenetic information.

This meeting continues to be the preeminent meeting in the field. The quality of the presentations and discussions indicates this will continue to be the case going forward.

Support for this meeting was provided in part by the National Cancer Institute, a branch of the National Institutes of Health.

## PROGRAM

**Replication Initiation Factors**
J. Berger, *Johns Hopkins University School of Medicine, Baltimore, Maryland*; H. Masai, *Tokyo Metropolitan Institute of Medical Science, Japan*

**Replication Origin Choice and Control**
O. Aparicio, *University of Southern California, Los Angeles*; J. Cook, *University of North Carolina at Chapel Hill*

**Replication Fork Stalling, Stabilization, and Checkpoint Activation**
P. Knipsheer, *Hubrecht Institute, Utrecht, Netherlands*; H. Ulrich, *Institute of Molecular Biology, Mainz, Germany*

**Developmental Control of Replication and Genome Organization**
D. Fox, *Duke University, Durham, North Carolina*; C. Gutierrez, *Centro de Biología Molecular Madrid, Spain*

**Replication-Coupled Genome Instability, Cancer, and Disease**
A. Sfeir, *Memorial Sloan Kettering Cancer Center, New York, New York*; A. Smogorzewska, *The Rockefeller University, New York, New York*

**Causes and Consequences of Replication Stress**
S. Cantor, *University of Massachusetts Chan Medical School, Worcester*; T. Halazonetis, *University of Geneva, Switzerland*

**Fork Progression and Repair**
N. Dekker, *Delft University of Technology, Netherlands*; D. Remus, *Memorial Sloan Kettering Cancer Center, New York, New York*

**Chromatin and Replication Timing**
S. Polo, *University Paris Cité/CNRS, France*; Z. Zhang, *Columbia University Irving Medical Center, New York, New York*

# Microbial Pathogenesis and Host Response

**September 11–15** 385 Participants (335 in-person, 50 virtual)

ARRANGED BY **Denise Monack,** Stanford University, California
**Anita Sil,** University of California, San Francisco
**Victor Torres,** New York University School of Medicine, New York

A. Sil, V. Torres, D. Monack

Despite advances in modern healthcare, infectious diseases continue to be major causes of human morbidity and mortality. The co-evolution of microbial pathogens with humans has resulted in complex interactions that impact the struggle between the infectious invader and the susceptible host. Increased understanding of these interactions with the goal of developing new therapeutics and preventive strategies requires collaborative and interdisciplinary scientific approaches. The virtual Cold Spring Harbor meeting on Microbial Pathogenesis and Host Response held on September 11–15, 2023, brought together a diverse group of international scientists who approach the study of bacterial, parasitic, and fungal pathogens from a broad range of perspectives. Investigators from the disciplines of molecular microbiology, eukaryotic cell biology, immunology, and genomics and representing academia, scientific publishing, industry, and the public health sector shared recent findings concerning microbial and host aspects of infectious diseases.

The meeting had a strong focus on the interaction between microbial pathogens and the host, which helps us better understand how our bodies respond to and defend against invading microbes. The discussions were organized into oral sessions, each centered around specific topics and involving a wide range of studies on various organisms. These sessions covered several areas, including toxins, how virulence is regulated, the mechanisms of innate immunity, interactions between different types of organisms and within hosts (trans-kingdom interactions), and a session dedicated to Visualizing and Analyzing Host–Pathogen Interactions. This last session emphasized the use of cutting-edge imaging and analysis techniques to enhance our understanding of how bacteria, fungi, viruses, and mammalian hosts interact. The speakers for these sessions were

A. Brakhage, B. Geier

H. Darwin, J. Seeliger, L.M. Nisbett

M. Keestra-Gounder, R. Vance

S. Almagro-Moreno, C. Moreira

a mix of well-established leaders in the field and emerging investigators. What is particularly noteworthy is that half the speakers, including postdoctoral fellows and graduate students, were selected to give oral presentations based on their submitted abstracts, highlighting the importance of fresh perspectives and innovative research. The keynote address was delivered by Dr. Arturo Zychlinsky, a globally recognized expert in the field of molecular mechanisms of innate immunity. His presentation, titled Neutrophil Extracellular Traps—The Immune Function of Chromatin, provided a captivating insight into his groundbreaking discovery of neutrophil extracellular traps (NETs). He discussed the mechanisms behind this unique form of cell death, involving elastase and myeloperoxidase. Dr. Zychlinsky also explored the various ways in which NETs kill microbes, one of which involves chromatin and histone-dependent antimicrobial activities. His findings have profound implications for our understanding of the role of chromatin in the immune defenses of mammalian hosts.

The informal atmosphere, combined with the broad perspectives of the meeting participants, resulted in a free flow of novel and refreshing ideas on pathogenesis and clinical treatment, with the atmosphere of a small meeting. Extensive question and answer sessions followed each oral presentation, including Zoom questions. The poster sessions were engaging with vibrant discussions. We strongly encouraged submission of abstracts by junior researchers in the field, and many young investigators were in attendance. These interactions have already produced fruitful scientific collaborations. This year also included opportunities for trainees to interact more formally with speakers and organizers at lunches.

J. Liu, C. Hawkes

J. McKenna, C. Bollinger

K. Orth, J. Berman

S. Shames, T. Van Opijnen

This meeting was partially supported by funds from the National Institute of Allergy and Infectious Diseases, a branch of the National Institutes of Health.

## PROGRAM

**Microbial Toxins and Effectors**
S. Mukherjee, *University of California, San Francisco*

**Visualizing and Analyzing Host–Pathogen Interactions**
G. Jensen, *Brigham Young University, Provo, Utah*

**Molecular Genetics of Host–Pathogen Interactions**
N. Elde, *University of Utah School of Medicine, HHMI, Salt Lake City*

**Immunity during Infection**
R. Vance, *University of California, Berkeley*

**Regulation of Host–Microbial Interactions**
M. Hamon, *Institut Pasteur, Paris, France*

**Keynote Speaker**
A. Zychlinsky, *Max Planck Institute for Infection Biology, Berlin, Germany*

**Cell Biology of Hosts and Microbes**
S. Almagro-Moreno, *University of Central Florida, Orlando*

**Microbial Communities in Health and Disease**
L. Hooper, *UT Southwestern Medical Center, HHMI, Dallas*

# Biology of Cancer: Microenvironment and Metastasis

**September 19–23** 385 Participants (329 in-person, 56 virtual)

ARRANGED BY **Scott Lowe,** Memorial Sloan Kettering Cancer Center, New York, New York
**M. Celeste Simon,** University of Pennsylvania Medical School, Philadelphia
**Valerie Weaver,** University of California, San Francisco

V. Weaver, S. Lowe, M.C. Simon

Solid tumors are multicellular "ecosystems" comprised of malignant cells adjacent to "stromal" components consisting of endothelial cells, immune cells, cancer-associated fibroblasts, neuronal axons, adipocytes, and lymphatics. Other features of the tumor microenvironment include extensive extracellular matrices, soluble factors (cytokines, chemokines, and hormones), hypoxia, and nutrient scarcity. The 2023 Biology of Cancer: Microenvironment and Metastasis meeting was the first in-person event in four years, given that the 2021 meeting was held virtually because of constraints imposed by the COVID-19 pandemic. The 2023 meeting attracted a large number of participants who were mostly in-person and focused on dynamic interactions between transformed cells and other microenvironmental constituents that promote tumor evolution, progression, metastasis, and interplay with the host. This meeting aimed to capture emerging topics currently under vigorous investigation: microenvironmental regulation of tumor behavior, metabolism, tumor epigenetics, tumor–stroma interactions, and systemic influences on metastasis, as well as reciprocal effects on the host by growing tumors (e.g., cachexia). All presentations by senior scientists, junior faculty, postdoctoral fellows, and graduate students were of exceptionally high caliber, resulting in extended discussion with the audience. Although 56 individuals participated online, there was no convenient mechanism for fielding their questions. It is hoped the online participants will communicate with speakers outside of the meeting format. Pancreatic ductal adenocarcinoma was a cancer subtype emphasized at the 2023 conference. The first session, New Insights into Pancreas Cancer, featured a keynote presentation by David Tuveson. The second keynote in the Tumor Evolution session was delivered by

F. Thege and guest

M. Goncalves, M. Paddock

S. Nelson, A. Jefferson

J. Bargonetti, S. Lowe

Kornelia Polyak. Other highlights included exciting new findings presented by Mikala Egeblad on the relationship between host organismal stress and tumor growth. Elaine Fuchs discussed how squamous cell carcinoma exhibits altered communication between Ras oncogene–transformed stem cells and the immune microenvironment. Mara Sherman further elaborated on the exceptional complexity of cancer-associated fibroblasts. The concept of tumor innervation was discussed by Humsa Venkates, who described a key role for cortical neurons in fostering the growth and survival of small-cell lung carcinoma brain metastases. Katherine Alexander (a postdoctoral fellow in Shelley Berger's laboratory) described exciting work on how oncogenes like HIF-2α can up-regulate target genes in renal cancer by targeting DNA to nuclear "speckles," subnuclear membrane-less structures that enhance transcription and RNA processing. Other discussion surrounded senescence-induced epigenetic programs in solid tumors, highlighting their connections to treatment failure, tissue tropism, and metastasis. Both evening poster sessions were well attended and provided an additional informal platform to increase interactions and networking for trainees. We are confident that overall the participants enjoyed the diversity, timeliness, and depth of the topics covered and were proud to be part of this vibrant community.

Support for this meeting was provided by the Cold Spring Harbor Laboratory/Northwell Health Alliance; the National Cancer Institute, a branch of the National Institutes of Health; and by the Life Science Alliance.

J. Houthuijzen, K. Hilgendorf

S. Morrison, D. Yang

J. Ablain, D. Tuveson

A. Habowski, M. Callaway

## PROGRAM

**New Insights into Pancreas Cancer**
S. Lowe, *Memorial Sloan Kettering Cancer Center, New York, New York*; M. Pasca di Magliano, *University of Michigan, Ann Arbor*

**Keynote Speaker**
D. Tuveson, *Cold Spring Harbor Laboratory*

**Microenvironmental Regulation of Tumor Behavior**
M. Egeblad, *Johns Hopkins University, Baltimore, Maryland*; E. Fuchs, *The Rockefeller University, New York, New York*

**Immunometabolism/Cachexia**
E. Pearce, *Johns Hopkins University School of Medicine, Baltimore, Maryland*; J. Rathmill, *Vanderbilt University Medical Center, Nashville, Tennessee*

**Tumor Evolution**
J. Guerriero, *Brigham and Women's Hospital, Boston, Massachusetts*; C. Simon, *University of Pennsylvania Perelman School of Medicine, Philadelphia*

**Keynote Speaker**
K. Polyak, *Dana-Farber Cancer Institute, Boston, Massachusetts*

**Tumor–Stroma Interactions I**
E. Puré, *University of Pennsylvania, Philadelphia*; M. Sherman, *Memorial Sloan Kettering Cancer Center, New York, New York*

**Tumor Epigenetics and Metabolism**
S. Berger, *University of Pennsylvania, Philadelphia*; A. Kimmelman, *NYU Grossman School of Medicine, New York, New York*

**Tumor–Stroma Interactions II**
D. Gutmann, *Washington University School of Medicine in St. Louis, Missouri*; V. Weaver, *University of California, San Francisco*

**Systemic Influences on Metastasis**
A. Boire, *Memorial Sloan Kettering Cancer Center, New York, New York*; J. DeGregori, *University of Colorado Anschutz Medical Campus, Aurora*

# Recombinant DNA: Fifty Years of Discovery and Debates

**September 27–30** 189 Participants (161 in-person, 28 virtual)

ARRANGED BY **Richard Mulligan,** Sana Biotechnology, Seattle, Washington
**Ludmila Pollock,** Cold Spring Harbor Laboratory
**Shirley Tilghman,** Princeton University, New Jersey

From September 27 to October 1, 2023, the Cold Spring Harbor Laboratory Center for Humanities & History of Modern Biology, together with the Meetings & Courses Program, convened a major international conference, Recombinant DNA: Fifty Years of Discovery and Debates. The meeting was arranged by Richard Mulligan (Sana Biotechnology), Shirley Tilghman (Princeton), and Mila Pollock (Cold Spring Harbor Laboratory).

W. Gilbert, S. Tilghman

First synthesized in a laboratory half a century ago, recombinant DNA (rDNA) is produced when genetic material from one organism is transferred to another one, such as a bacterium or yeast. This technique spurred wide debate over genetic modification even as it made possible the biotechnology revolution of the late twentieth century.

This topic is so significant, and so timely, that our speakers were receiving the world's most prestigious research awards almost literally while the meeting took place. Just days before we gathered, Carl June and Michel Sadelain were announced as the winners of the upcoming 2024 Breakthrough Prize for developing CAR T-cell immunotherapy. And the day after the meeting ended, we awoke to the wonderful news that Katalin Karikó would receive the Nobel Prize in Physiology or Medicine in recognition of her work creating methods that allow the delivery of mRNA-based vaccines and therapies.

The event was also unique in combining the format of a typical Wednesday-to-Saturday CSHL scientific meeting with a first-of-its-kind workshop for historians on Saturday afternoon and Sunday. Thus, during the first four days, we heard talks by notable participants in the first decades of the revolution in rDNA and biotechnology, including researchers, clinicians, regulators, and

M. Rosbash, W. Bender

F. Grosveld, B. Stillman

J. Beggs, P. Sharp

F. Zhang, D. Jiang

investors. The speakers included five Nobel Laureates (with a sixth who was about to be named, and perhaps more to follow!), young investigators at the cutting edge of the field, and scientist–administrators who have directed institutions and projects such as the U.S. National Institutes of Health (NIH), the Human Genome Project, and the U.S. Food and Drug Administration (FDA).

Over the weekend, we heard from an equally international group of historians and archivists who have devoted their careers to documenting and analyzing the recombinant revolution. By bringing scientists and historians together to hear each other's work and converse during banquets, campus walks, and evening drinks, this meeting exemplified and extended CSHL's remarkable tradition as a meeting place for critical discussions of the life sciences.

Support for this meeting was provided in part by Cold Spring Harbor Laboratory; Illumina; New England Biolabs; Promega; and QIAGEN.

**PROGRAM**

**The Discovery of Recombinant DNA**
T. Maniatis

**Panel: Reflection on the Recombinant DNA Controversy…**
D. Kevles

**From Genes to the Genome—Time Flies and Fruit Flies**
D. Glover

**Modifying the Genome and its Expression**
M. Capecchi

**Panel: The History of the Human Genome Project**
E. Lander

**From Genes to Human Disease**
F. Collins

M. Sadelain, R. Mulligan

S. Cohen, J. Mertz

G. Montgomery, M. Capecchi

G. Morgan, T. Maniatis

**From the Human Genome Project to Complex Traits**
D. Altshuler

**Delivering Genes into Humans**
R. Mulligan

**Modalities for Genetic Interventions**
F. Zhang

**From Cloning to Gene Therapy—The Globin Gene Paradigm**
S. Tilghman

**The RNA Revolution**
P. Sharp

**Panel: Recombinant DNA Creates the Biotech Industry**
S. Papadopoulos

**The Biotech Industry Creates Therapies**
S. Gottleib

**Exhibition: Documenting the History of Recombinant DNA: From CSHL Archives**
Curator: C. Tonks, *CSHL*; Consultant: M. Pollock, *CSHL*

## POSTMEETING WORKSHOP ON THE HISTORY OF RECOMBINANT DNA

**Framing the History of rDNA Roundtable**
A.N.H. Creager, *Princeton University, New Jersey*

**rDNA in Europe and Asia**
A. Sponsel, *Cold Spring Harbor Laboratory*

**Safety and Governance in a Recombinant World**
D.J. Kevles, *Yale University, New Haven, Connecticut*

**Biotech Abroad Roundtable**
L. Campos, *Rice University, Houston, Texas*

C. June, K. Karikó

E. Lander, P. Sharp

C. Mello

# Neurobiology of *Drosophila*

**October 3–7** 548 Participants (460 in-person, 88 virtual)

ARRANGED BY **C. Andrew Frank,** University of Iowa, Iowa City
**Avital Rodal,** Brandeis University, Waltham, Massachusetts

A. Rodal, C.A. Frank

The 2023 Neurobiology of *Drosophila* meeting provided a forum for the discussion of new discoveries, techniques, and advances in *Drosophila* neurobiology. Fourteen sessions ran in series over four days with alternating platform and poster presentations. These included two professional development sessions: a panel discussion on diversity, equity, and inclusion focusing on "growth mindset," and a speaker from industry sharing how skills developed in *Drosophila* research are translatable to the pharma and biotech world.

In addition to the traditional, in-person conference, there were virtual components, too. The virtual options associated with the meeting allowed access to formal presentations and also promoted informal interactions. The meeting's online Slack site had multiple channels that facilitated collaborations, exchange of reagents (e.g., antibodies, clones, mutants, and other stocks), methods (genetic, physiological, and optical), and ideas among trainees and investigators. In-person speakers received numerous questions from virtual participants. A premeeting meetup of virtual participants was hosted by virtually attending faculty via Zoom. Finally, the meeting organizers held a Zoom session to meet with virtual attendees from Africa who use *Drosophila* as a model organism. This combination of formal and informal interactions was important for scientists new to *Drosophila* and/or neurobiology because they can benefit from learning of new advances in the field and opportunities to build their professional networks.

The nine platform session topics were chosen to reflect the areas in which cutting-edge advances are being made—neurodevelopment, neuronal cell biology, excitability and synaptic plasticity, disease and injury models, circadian-controlled processes, sensory modalities, how circuits govern behavior, "neuromic" approaches, and technological innovations. Attendees commented positively


X. Chen, L. Amadio

M. Fenckova, S. Tomchik

T. Clandinin, L. Griffith

C. O'Brien, S. Rumpf

on the diversity of presentations and balance of research investigating the nervous system at different levels. Session chairs worked alongside the meeting organizers to select presenters for these platform sessions from submitted abstracts, whereas the remaining abstracts were presented as posters. Most of the speakers were graduate students and postdoctoral fellows. Slightly more than half the presenters were female, and approximately one-third of abstracts submitted by members of historically underrepresented groups were selected for talks, which matches with the overall talk selection average. Overall, ∼10% of the talks were presented by attendees who self-identified as Hispanic, Black, American Indian/Native American, and/or Native Hawaiian/Pacific Islander.

The Seymour Benzer Lecture was presented by Dr. Marc Freeman (Vollum Institute, Oregon Health & Science University). He shared how his laboratory has used forward genetic screens to uncover conserved factors needed for the interplay between glia and neurons. The Elkins Award Memorial lecture is presented at each meeting by a graduate student whose dissertation exemplifies the finest work in our field. This year's Elkins Lecture was presented by John Vaughen, who is completing his training with Dr. Thomas Clandinin at Stanford University. He presented his exceptional work on how sphingolipids control neural circuitry. The Neurobiology of *Drosophila* meeting was extremely well attended, with presentations spanning the breadth of modern neurobiology. The interactions and career development opportunities fostered by this meeting are sure to enhance this vibrant field.

Support for this meeting was provided by the National Institute of Neurological Disorders and Stroke, a branch of the National Institutes of Health.

J. Dubnau, A. Rothenfluh

K. Fisher, T. Clandinin

T. Jay, V. von Saucken

G. Sterne, J.H. Kim

M. Freeman, H. Gupta

## PROGRAM

**Neuromics**
W-C.A. Lee, *Harvard Medical School, Boston, Massachusetts*

**Seymour Benzer Lecture**
M. Freeman, *Oregon Health & Science University, Portland*

**Neurodevelopment**
T. Mosca, *Thomas Jefferson University, Philadelphia, Pennsylvania*

**Panel Session: Diversity, Equity, Inclusion**
K. Wright, *Assistant Dean for Diversity and Inclusion, Stony Brook University, New York*

### Elkins Memorial Lecture Award

**Circadian and Sleep**
M. de la Paz Fernandez, *Columbia University, New York, New York*

**Disease and Injury**
M. Pearce, *St. Joseph's University, Philadelphia, Pennsylvania*

**Circuits and Behavior I**
Y. Fisher, *University of California, Berkeley*

**Cell Biology of the Neuron**
R. Insolera, *Wayne State University School of Medicine, Detroit, Michigan*

**Excitability, Transmission, and Plasticity**
T. Wang, *Georgetown University, Washington, D.C.*

**Career Session: Roles for *Drosophila* Neuroscientists in Industry**
S. Sanyal, *Calico Labs, San Francisco, California*

**Sensory Modalities**
R. Padinjat, *National Centre for Biological Sciences-TIFR, Bangalore, India*

**Technological Innovations/Circuits and Behavior II**
A. Barber, *Rutgers University, Piscataway, New Jersey*

# Cell State Conversions

**October 10–14** 193 Participants (139 in-person, 54 virtual)

ARRANGED BY **Magdalena Götz,** Munich University & Helmholtz Center Münich, Germany
**Thomas Graf,** Center for Genomic Regulation, Barcelona, Spain
**Deepak Srivastava,** Gladstone Institutes and University of California, San Francisco

T. Graf, M. Götz, D. Srivastava

The inaugural Cold Spring Harbor Laboratory meeting on Cell State Conversions was held from October 10 to 14, 2023, at the Grace Auditorium on the campus of Cold Spring Harbor Laboratory. The meeting united the community of the exciting and ever-growing field of direct reprogramming, be it naturally such as in disease conditions, or forced as when a cell type is converted into another very different cell type. The goal of these approaches is to study transcription factor function and fate restriction barriers at the chromatin level, aid cell repair by replacing specific cell types (such as neurons in the brain or cardiomyocytes in the heart), and to apply reprogramming as a tool to fight cancer. The meeting brought together about 140 researchers in person and more than 50 others online, representing U.S., Canadian, European, and Asian participants. The goals of the meeting were to highlight the major recent insights into fate conversion, discuss some pitfalls in the field due to viral vector artifacts, and discuss the emerging translational outcomes presented by revolutionary advances in cancer and heart attack treatments in a special session at the end. The opening keynote address was given by Ken Zaret from the University of Pennsylvania, who presented a novel concept on the function of reprogramming transcription factors—namely, to act less in a lineage-specific manner, but rather across lineages by increasing chromatin fluidity. Forty-one talks were given by leading scientists and physicians—12 by women, focusing on the themes of forced fate conversion, natural fate, and state conversions in development and disease as

M. Karow, B. Berninger, M. Wernig

S. Stricker, B. Tursun, J. Brickman

D. Yucel, M. Puglisi

S. Ruiz Macias, B. Di Stefano

well as the use of direct reprogramming technology for therapy and cell replacement in vivo. The second keynote presentation was given by Juan Carlos Izpisua Belmonte. He presented his revolutionary discovery of transient expression of pluripotency reprogramming factors for rejuvenating organs and organisms and the progress in bringing this approach to the clinic. The discussions after the talks and during the poster sessions were very lively and resulted in a dynamic engagement between junior and senior researchers. This resulted in several collaborative projects proposed between researchers who interacted for the first time. The participants expressed much enthusiasm for repeating the meeting, and, given the success of the inaugural meeting, it will be a great series of CSHL meetings—with the next one taking place in September 2025.

## PROGRAM

**Keynote Speaker**
K. Zaret, *Perelman School of Medicine, University of Pennsylvania, Philadelphia*

**Reprogramming Mechanisms I**
M. Wernig, *Stanford University School of Medicine, California*

**Reprogramming Mechanisms II**
T. Graf, *Center for Genomic Regulation, CRG, Barcelona, Spain*

**Developmental and Oncogenic Cell State Transitions**
R. Majeti, *Stanford University, California*

R. Rock, R. Sridharan

J. Brickman, F. Zoppi

T. Reh, E. Halitzki

S. Ruiz Macias, Y. Karpova

**Reprogramming and Modeling**
M-E. Torres-Padilla, *Helmholtz Centre Münich, Germany*

**In Vivo Reprogramming**
M. Karow, *Friedrich-Alexander-Universität Erlangen, Germany*

**Keynote Speaker**
J.C. Izpisua Belmonte, *Altos Labs, Tempe, Arizona*

**Regeneration and Cancer**
D. Srivastava, *Gladstone Institutes, San Francisco, California*

**Therapeutic Applications of Reprogramming**
H. Deng, *Peking University, Beijing, China*

# Single-Cell Analyses

**November 8– 11** 413 Participants (295 in-person, 118 virtual)

ARRANGED BY **Scott Fraser,** University of Southern California, Los Angeles
**Junhyong Kim,** University of Pennsylvania, Philadelphia
**Tatjana Sauka-Spengler,** University of Oxford, United Kingdom
**Aaron Streets,** University of California, Berkeley

A. Streets, J. Kim, T. Sauka-Spengler, S. Fraser

The goal of this seventh Workshop on Single-Cell Techniques was to bring together scientists who analyze and engineer single cells using a wide variety of experimental paradigms to discuss the progress that is being made. This was the first in-person meeting since the pandemic, when in 2021 we had an all-virtual meeting, and the first in-person meeting organized by Aaron Streets and Tatjana Sauka-Spengler. The in-person meeting generated much more energy and excitement amongst the attendants, and we had a record number of participants. Following previous years' format, we had two keynote speakers, 12 invited speakers, 19 contributed speakers, 24 two-minute flash talks, and more than 150 posters. The first keynote speaker was Nancy Allbritton from the University of Washington, who gave the opening lecture of the meeting. Nancy was one of the founding organizers of the CSHL single-cell meeting, serving until 2019. She talked about the synergy between bioengineering and single-cell biology and the ongoing opportunity for technologies to play a key role in biomedical sciences. The second keynote speaker, Lior Pachter from Caltech, anchored the last day of the meeting with his keynote. Lior talked about the broad landscape of computational biology, some of the problems in data analyses, and the need for mechanism-based models. As in previous years, much of the meeting concentrated on covering new technologies and methods loosely organized under the topics of New Methods in Single-Cell Analysis (I and II), New Applications of Single-Cell Analysis, Single-Cell Spatial Analyses, and Interpreting Single-Cell Analyses. The talks covered

T. Attenborough, A.S. Booeshaghi

A. Fomitcheva Khartchenko, N. Nouri

Streets Lab

continued advances in single-cell methods, especially the continuing development of new measurement modalities and multimodal measurements. The development of spatial technologies has been an increasingly important trend, and we had exciting talks including the development of 3D imaging, multimodal spatial analysis, and spatial proximity analysis. One of the more interesting applications of single-cell technologies came from multiple groups using massively multiplexed assays to discover rare events (cells, mutations, etc.). This year we finished the meeting with a panel discussion featuring Lior Pachter, Caroline Uhler, Smita Krishnaswamy, Sina Booeshaghi, and Josh Orvis. The panel discussion centered on the topic of challenges in single-cell computational analysis. We solicited questions from the audience using shared web document and presented moderated questions to the panelists. As mentioned above, with the first in-person meeting since 2019, there was an unusual level of energy among the participants, with many stating this was the best meeting they have been to in recent years. The meeting continued to solidify single-cell biology as one of the most exciting areas of biomedical sciences with unabated growth in technologies, impact, and interest. There was an incredible outpouring of positive feedback, and many of the invited speakers, including Dr. Pachter, expressed an interest in returning to this meeting regularly

N. Slavov, S. Madler

S. Li, L. Pacther

with their laboratory members. Lastly, Scott Fraser stepped down from the organizing committee after more than 10 years of service with a great round of appreciation from the audience, and we will be seeking another co-organizer in preparation for the 2025 meeting.

Support for this meeting was provided in part by CellSorter; the Chan Zuckerberg Initiative; PacBio; and Scale Biosciences. The Early Career Researcher Spotlight Award was presented by the Life Science Alliance.

N. Patikas, L. Alessandri

## PROGRAM

**Keynote Speaker**
N. Allbritton, *University of Washington, Seattle*

**New Methods in Single-Cell Analyses I**
R. Cotrim Chaves, *University of California, Berkeley*

**New Methods in Single-Cell Analyses II**
S. Yang, *University of California, Berkeley*

## FLASH TALKS

**New Applications of Single-Cell Analyses**
A.S. Booeshaghi, *California Institute of Technology, Pasadena*

**Single-Cell Spatial Analyses**
I. Clark, *University of California, Berkeley*

**Interpreting Single-Cell Analyses**
M. Rozenwald, *University of California, Berkeley*

**Keynote Speaker**
L. Pachter, *California Institute of Technology, Pasadena*

**Panel Discussion: Critical Considerations in Analyzing and Interpreting Single-Cell Data**
A.S. Booeshaghi, *California Institute of Technology, Pasadena*; Junhyong Kim, *University of Pennsylvania, Philadelphia*; J. Orvis, *University of Maryland School of Medicine, Baltimore*; L. Pachter, *California Institute of Technology, Pasadena*; Tatjana Sauka-Spengler, *Stowers Institute for Medical Research & University of Oxford, United Kingdom*

# Zebrafish Neurobiology

**November 15–18** 280 Participants (250 in-person, 30 virtual)

ARRANGED BY **Rainer Friedrich,** Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland
**David Shoppik,** NYU Grossman School of Medicine, New York
**Vatsala Thirumalai,** National Center for Biological Science, Bangalore, Kamataka, India

R. Friedrich, D. Schoppik

This was the second Cold Spring Harbor Laboratory meeting on Zebrafish Neural Circuits and Behavior, held from November 15 to 18, 2023, at Grace Auditorium on the campus of Cold Spring Harbor Laboratory. This meeting brought together an international group of neuroscientists using the zebrafish as a model system to understand fundamental problems in neurobiology.

Prior to the meetings in Cold Spring Harbor, the community had come together in smaller venues annually over the past decade to share the latest findings, disseminate innovative technologies, and begin new collaborations. This year, in keeping with the tradition of the best CSHL meetings, speakers were selected from submitted abstracts—allowing the meeting to highlight the most promising ideas, organized into sessions that reflected the state of the field. As a result, there were nine sessions featuring short talks on a broad range of topics including Development and Regeneration, Disease and Drug Development, Sensory Systems, Motor Control and Behavior, and Learning and Cognition. Sessions were chaired by faculty members charged with facilitating discussion and introducing speakers. Across the meeting, gender and rank were well balanced: five out of nine session chairs and 24 of 39 speakers were female, and 24 out of the 39 speakers were either graduate students or postdoctoral fellows. The poster sessions were very lively: Highlights ranged from the latest in unpublished design improvements of microscopes for in vivo imaging to advances in the molecular underpinnings of complex behaviors in health and disease. There was ample time for discussion both at social events and later in the evenings, where trainees and faculty mingled. A number of new collaborations were proposed among laboratories that had not worked

F. Del Bene, M. Halpern

K. Arena, K. Liu

K. Slangewal, F. Kampf

R. Hindges, J. Semmelhack

together, and reagents were shared, moving the field forward. There was enthusiasm to continue this series of meetings.

Partial support for this meeting was provided by Aquatic Enterprises; ViewPoint Behavior Technology; and Zantiks.

**PROGRAM**

**Keynote Speaker**
J. Fetcho, *Cornell University, Ithaca, New York*

**Cognitive/Learning and Memory**
B. Judkewitz, *Charité Universitätsmedizin Berlin, Germany*

**Sensory Systems**
K-H. Huang, *Academia Sinica, Taipei, Taiwan*

**Prey Capture**
F. Kubo, *National Institute of Genetics, Mishima, Japan*

**Social Behavior**
R. Portugues, *Technical University of Münich, Germany*

**Diseases and Drug Development**
S. Kucenas, *University of Virginia, Charlottesville*

**Neuromodulators and State**
C. Satou, *HHMI Janelia Research Campus, Ashburn, Virginia*

**Development and Regeneration**
K. Monk, *Oregon Health & Science University, Portland*

**Keynote Speaker**
V. Ruta, *The Rockefeller University, New York, New York*

**Motor Control**
N. Jurisch-Yaksi, *Norwegian University of Science and Technology, Trondheim*

**Atlases and Tools**
T. Kawashima, *Weizmann Institute of Science, Rehovot, Israel*

F. Voigt, M. Hancock

J. Dallmen, G. Downes

D. Prober, H. Zwaka

# Plant Genomes and Biotechnology: From Genes to Networks

**November 29–December 2** 156 Participants (132 in-person, 24 virtual)

ARRANGED BY **Julia Bailey-Serres,** University of California, Riverside
**Elizabeth Sattley,** Stanford University, California
**Cyril Zipfel,** Zürich University, Switzerland

C. Zipfel, J. Bailey-Serres, E. Sattley

The 13th Plant Genomes meeting at Cold Spring Harbor Laboratory highlighted how genome-enabled methods are accelerating fundamental discoveries into plant development, environmental interactions, and natural product synthesis. The unpublished research presented demonstrated the value of pan-genomes and advances in defining spatial–temporal cellular processes including gene regulatory circuitry, protein condensates, and cellular sensors. Talk highlights included successful manipulation of plant immunity and climate resilience. Impressive was the range of plant species studied (moss to citrus), as well as the breadth in background and career stage of the presenters. A well-received addition to the program was a series inviting short talks delivered by poster presenters.

Each thematic session spanned from new insights into fundamental mechanisms to the path to impact. The conference opened with a Keynote Lecture presented by Sarah O'Connor from the Max Planck Institute for Chemical Ecology. This set the stage for a rich session on natural product chemistry that included insights into the tinkering through natural selection with genes that underlie families of bioactive compounds synthesized for protection from or attraction of other organisms. Not unexpectedly, we learned that single-cell technologies accelerate recognition of the specific cells and enzymes involved in secondary metabolite synthesis. Single-cell and -nucleus methods also powered new discoveries in plant development, nutrient uptake, adaptation to drought, and the gene regulatory control underlying cell-type compartmentation of C3 and C4 photosynthetic metabolic processes.

U. Pedmale, D. Wagner

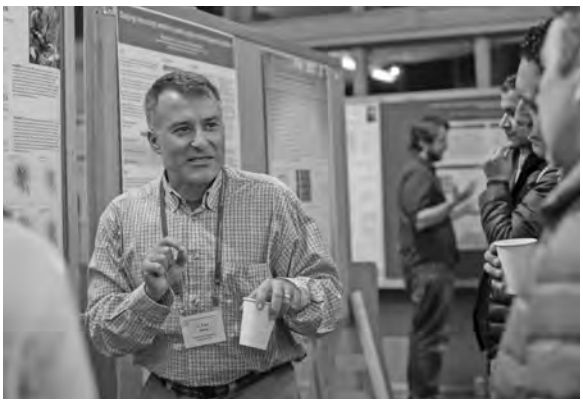
C. Zipfel, S. Kamoun

L. Banda

L. Luginbuehl, J. Swift

An overarching topic was the integration of plant growth with one or more environmental cues. There were multiple demonstrations that genetic control of adaptative plasticity to drought is stage-/cell type–specific and that different allelic combinations can be beneficial. Progress was made in understanding gene regulatory circuitry of cells and cell layers within root systems in the context of nutrient acquisition, protection from water deficit, and response to soil compaction. Gene editing has accelerated discovery to translation, with new base-editing tools promising to expand the realm of possibility. In addition to abiotic cues, a diversity of biotic interactions was considered: plant–plant, plant–insect, plant–virus, and plant–microbe. The latter included pathogenic and symbiotic associations.

Keynote presenter Giles Oldroyd of Cambridge University focused on beneficial interactions between root systems and soil microbes—particularly the nutrient exchange partnership with rhizobial bacteria that fix soil nitrogen and the arbuscular mycorrhizal fungi that exchange soil phosphate for plant carbon. Whereas establishment of rhizobial nodulation in cereals remains a challenge, the phosphate starvation requirement for mycorrhizal fungal colonization can be overridden in rice and barley. A remaining challenge is to bolster fungal colonization of roots in the field for maximal benefit. A superb final example of bridging from pan-genome and mechanistic knowledge to translation was provided by Sophien Kamoun of the Sainsbury Laboratory in Norwich, who described a protein-based "plant vaccination" and recapturing "retired" resistance genes that can effectively squelch the attack of pathogens and certain nematodes. The research presented at the CSHL Plant Genomes 2023 provides fundamental knowledge and opportunities



U.U. Sezen

C. Gao, C. Zipfel

to benefit human health, reduce plant species and crop loss associated with climate change, and enhance sustainable agriculture across the globe.

Support for this meeting was provided in part by Corteva Agriscience; Inari; and Pairwise.

## PROGRAM

**Keynote Speaker**
S. O'Connor, *Max Planck Institute for Chemical Ecology, Jena, Germany*

**Development and Productivity**
J. Bailey-Serres, *University of California, Riverside*; E. Sattley, *Stanford University/HHMI, California*; C. Zipfel, *University of Zürich, Switzerland, and The Sainsbury Laboratory, Cambridge, United Kingdom*

**Natural Products and Synthetic Biology**
S. O'Connor, *Max Planck Institute for Chemical Ecology, Jena, Germany*

**Abiotic Responses**
J. Bailey-Serres, *University of California, Riverside*

## FLASH TALKS

**Genetic Variation**
D. Weijers, *Wageningen University, Netherlands*

**Keynote Speaker**
G.E.D. Oldroyd, *University of Cambridge, United Kingdom*

**Nutrition**
C. Zipfel, *University of Zürich, Switzerland, and The Sainsbury Laboratory, Cambridge, United Kingdom*

**Biotic Stress and Plant–Plant Relations**
S. Cutler, *University of California, Riverside*

# Genome Informatics

**December 6–9** 337 Participants (252 in-person, 85 virtual)

ARRANGED BY **Joanna Kelley,** Washington State University, Pullman
**Pall Melsted,** University of Iceland, Reykjavik
**Nicola Muldar,** University of Cape Town, South Africa
**Oliver Stegle,** German Cancer Research Center, Heidelberg

K. Manpearl, P. Hicks

The 20th Cold Spring Harbor Laboratory/Wellcome Trust conference on Genome Informatics was held at Cold Spring Harbor, New York, as well as being offered virtually. The conference series continues to demonstrate robust attendance, both from the United States and from abroad. A total of 179 abstracts were submitted from 240 in-person participants and 84 virtual. There were 10 invited talks by session chairs and two keynote addresses. The remaining 37 talks were selected for presentation by session chairs from submitted abstracts, and 129 posters were presented during the conference.

The abstracts covered a wide area of bioinformatics and genomic analysis, with special emphasis on innovations in the analysis of long reads, pangenomics, and analysis of single-cell data sets. The sessions in the conference were centered around the following topics: Microbial and Metagenomics, Pangenome, Single-Cell and Spatial Omics, Functional Genomics, Variant Discovery, and Genome Assembly and Sequencing Algorithms.

The first keynote address was delivered by Dr. Sean Eddy from Harvard University. His talk focused on the computational problems arising from analyzing secondary structures of RNA molecules and the insights they can give into the role of long noncoding RNAs. The second keynote was given by Dr. Karen Miga from UC Santa Cruz. Her talk gave an overview of the computational and technological background of constructing complete T2T diploid assemblies and how complete genomes from multigenerational families enable studying the genetic and epigenetic inheritance of large repetitive regions.

W. Timp, S. Salzberg

P. Melsted, V. Makinen

X. Zheng, Y. Sakamoto

C. Semple, B. Erdogdu

In addition to the scientific work presented, the conference hosted a panel discussion for NIH Early Stage Investigators, moderated by Anton Nekrutenko from Penn State with Sean Eddy from Harvard University and Shurjo Sen, NHGRI, as panelists.

The meeting was actively discussed on Twitter/X (using hashtag #gi2023), and a Slack channel dedicated to the conference facilitated asking questions during talks and the poster session. Additionally, two talks were delivered virtually and virtual participants asked several questions during the Q&A after each talk.

This meeting was funded in part by the National Human Genome Research Institute, a branch of the National Institutes of Health, and the James P. Taylor Foundation for Open Science Scholarship Fund.

## PROGRAM

**Microbial and Metagenomics**
E. Segal, *Weizmann Institute of Science, Rehovot, Israel*

**Pangenome**
V. Makinen, *University of Helsinki, Finland*

**Keynote Speaker**
S. Eddy, *HHMI/Harvard University, Cambridge, Massachusetts*

**Single-Cell and Spatial Omics**
A. Carpenter, *Broad Institute of Harvard and MIT, Cambridge, Massachusetts*; A. Tanay, *Weizmann Institute, Rehovot, Israel*

**Functional Genomics**
B. Lehner, *Centre for Genomic Regulation, Barcelona, Spain*; E. Zeggini, *Helmholtz Zentrum München, Germany*

H. Clawson, K. Chougule

M. Schatz, A. Nekrutenko

A. Carpenter, O. Stegle


D. Ware, R. Davuluri

S. Koren, S. McCarthy

**Panel Discussion: NIH Early-Stage Investigators**
Sean Eddy, *Harvard University/HHMI, Cambridge, Massachusetts*; A. Nekrutenko, *Pennsylvania State University, University Park*; S. Sen, *NHGRI, National Institutes of Health*

**Variant Discovery**
I. Moltke, *University of Copenhagen, Denmark*; J. Simpson, *University of Toronto, Ontario, Canada*

**Keynote Speaker**
K. Miga, *University of California, Santa Cruz*

**Genome Assembly and Sequencing Algorithms**
R. Chikhi, *Institut Pasteur, Paris, France*; J. Kelso, *Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany*

# POSTGRADUATE COURSES

## Cryo-Electron Microscopy

**March 23–April 5**

| | |
|---|---|
| INSTRUCTORS | **J. Cash,** University of California, Davis<br>**J. Kollman,** University of Washington, Seattle<br>**G. Lander,** The Scripps Research Institute, La Jolla, California |
| CO-INSTRUCTOR | **M. Vos,** Institut Pasteur, Paris, France |
| ASSISTANTS | **L. Anderson,** University of California, Davis<br>**K. Hvorecny,** University of Washington, Seattle<br>**C. Zerio,** Scripps Research, San Diego, California |

Cryo-electron microscopy (cryo-EM) is a rapidly developing technique in structural biology wherein the biological sample of interest is flash-frozen under cryogenic conditions. The utility of cryo-electron microscopy stems from the fact that it allows specimens to be observed under "near-to-native" conditions without the need for staining or fixation. This contrasts with X-ray crystallography, which requires crystallizing the specimen, which can be a long and challenging process and often involves the introduction of biomolecules into nonphysiological environments that can occasionally lead to functionally irrelevant conformational changes. Cryo-EM is now

routinely applied to study the structures of viruses, ribosomes, ion channels, transcription and splicing machinery, and many other protein and nucleoprotein complexes. The spiraling number of publications that incorporate cryo-EM methodologies is evidence of this technique's importance to the structural community: Since 2017, single-particle cryo-EM has been used to solve the structures of more than 1,800 molecules, nearly half of which are resolved to better than 4 Å resolution. The resolution of single-particle cryo-EM maps is improving steadily, with recent improvements in processing methodologies yielding structures at better than 2 Å resolution. This powerful technique additionally enables researchers to study the conformational landscape of a biological specimen from a single flash-frozen sample, to deduce the mechanism by which it works.

The course covered the theory, practice, and application of single-particle cryo-EM. Participants in the course learned to perform all steps involved in solving high-resolution cryo-EM structures, including sample prep, microscope alignment, data collection, image processing, and model building. Students had supervised access to CSHL's Titan Krios and K3 direct electron detectors, as well as the Talos L120C. This hands-on course included lectures by leading experts who will discuss practical and conceptual approaches to structure determination using these techniques, as well as covering a wide range of state-of-the-art applications of cryo-EM in the biological sciences.

Applications were open to individuals at universities and colleges, medical research institutions, and industry, both from within the United States and from overseas, and to individuals at any stage in their postgraduate (PI, postdoc, grad student, etc.) career. Applicants demonstrated that they were committed to applying cryo-electron microscopy directly to their own research, and that they work in an environment where such application is feasible and realistic.

The techniques that were taught were

- Electron microscope optics and alignment
- Negative stain and cryo-EM sample preparation
- Sample screening and optimization
- Automated single-particle data acquisition
- Single-particle analysis and 3D reconstruction
- EM density visualization, interpretation, and validation
- Atomic model building and validation.

This course was supported in part by grants from the Helmsley Charitable Trust and the Howard Hughes Medical Institute. Additional scholarship support was provided by Regeneron.

## PARTICIPANTS

Akkermans, O., Ph.D., MIT, Cambridge, Massachusetts
Bobrovskyy, M., Ph.D., The University of Chicago, Illinois
Borza, R., M.S., Netherlands Cancer Institute, Amsterdam, Netherlands
Bruguera, E., Ph.D., Stanford University, California
Crowe, C., M.Sc., University of Dundee, United Kingdom
Fadel, F., Ph.D., Institute of Molecular and Clinical Ophthalmology, Basel, Switzerland
Hollmann, N., Ph.D., HHMI, Baltimore, Maryland
Howard, D., B.S., University of Cincinnati, Ohio
Katti, S., Ph.D., Texas A&M University, College Station
Perry, K., B.S., Arizona State University, Tempe
Samara, N., Ph.D., National Institutes of Health, Bethesda, Maryland
Simanshu, D., Ph.D., Frederick National Laboratory for Cancer Research, Maryland

## SEMINARS

Afonine, P., Lawrence Berkeley National Laboratory, Berkeley, California: Phenix for cryo-EM
Eng, E., New York Structural Biology Center, New York: How to run an efficient cryo-EM facility

Frank, J., Columbia University, New York, New York: History of single-particle cryo-EM.
Grant, T., Morgridge Institute/UW-Madison, Wisconsin: Single-particle processing in cisTEM.
Kollman, J., University of Washington, Seattle: Intro to single-particle image processing; Maximizing resolution for cryo-EM structure determination/advanced image analysis techniques.
Lander, G., The Scripps Research Institute, La Jolla, California: Course overview: structure determination of biological macromolecules using electron microscopy; Introduction to EM sample preparation; Microscope optics—the basics; Reconstruction principles; Maximizing resolution for cryo-EM structure determination/advanced image analysis techniques.
Naydenova, K., MRC Laboratory of Molecular Biology, Cambridge, United Kingdom: Electron detectors; Data processing in RELION; Latest cryo-EM tech
Richardson, J., Duke University, Durham, North Carolina: Atomic model validation
Rosenthal, P., The Francis Crick Institute, London, United Kingdom: Validating cryo-EM structures
Russo, C., MRC Laboratory of Molecular Biology, Cambridge, United Kingdom: Cryo-EM: upcoming innovations and future opportunities
Turner, J., EMBL's European Bioinformatics Institute, Hinxton, United Kingdom: EMDB & EMPIAR: data in, data out, and everything in-between

# The Genome Access Course

**April 2–4**

INSTRUCTORS **D. Fagegaltier**, Merck Research, New York, New York
**E. Hodges**, Vanderbilt University School of Medicine, Nashville, Tennessee
**B. King**, University of Maine, Orono
**S. Munger**, The Jackson Laboratory, Farmington, Connecticut

The Genome Access Course (TGAC) was an intensive two-day introduction to bioinformatics. Participants are expected to arrive by 6 p.m. on the first day (Sunday, April 2) with the course running two full days until 5 p.m. on the third day (Tuesday, April 4).

TGAC was broken into modules that were each designed to give a broad overview of a given topic, with ample time for examples chosen by the instructors. Each module featured a brief lecture describing the theory, methods, and tools, followed by a set of worked examples that students complete. Students were encouraged to engage instructors during the course with specific tasks or problems that pertain to their own research.

The core of the course is the analysis of sequence information framed in the context of completed genome sequences. Featured resources and examples primarily come from mammalian species, but concepts can be applied to any species. The course also features methods to assist the analysis and prioritization of gene lists from large-scale microarray gene expression and proteomics experiments. The topics covered in each two-day iteration of TGAC were taken from the following list.

NCBI and Model Organism Database Resources

- NCBI sequence, gene, and protein resources
- Gene Ontology
- Model organism databases: Mouse Genome Informatics, ZFIN

UCSC and Ensembl Genome Browsers

- Overview and comparison of resources
- Adding custom tracks
- Bulk genome analysis tools: BioMart, UCSC Table Browser

Sequence Variation

- Types of sequence and structural variation
- Large-scale biobanks: UK BioBank, All of Us
- SNP resources: dbSNP, gnomAD
- GWAS resources: GWAS Catalog
- Phenome-wide association resources: PheWeb, Open Targets
- Prioritizing variants by predicting variant effects: Ensembl VEP

Functional Genomic Elements and ENCODE

- Functional genomic resources: ENCODE
- Viewing ENCODE data using UCSC Genome Browser TrackHubs

High-Throughput Sequence Data Analysis

- Reference-based analysis workflows
- Common file formats: FASTQ, SAM, BAM
- Quality control and diagnostic analyses
- Repositories of high-throughput sequence data: GEO and SRA

RNA Sequencing Data Analysis

- Experimental design
- Analysis workflows

Hands-On Analysis of Sequence Variation Data Using Galaxy

- Mapping reads to a reference sequence
- Generating read counts to infer gene expression levels
- Generating lists of differentially expressed genes

Hands-On Introduction to the R Statistical Computing Environment

- Overview of R packages, Bioconductor and R Studio
- Basic R syntax for data analysis and plotting

Hands-On Analysis of Bulk RNA Sequencing Data using the R/DESeq Package

- Analyze bulk RNA sequencing read count data using the R/DESeq2 package in R Studio
- Generate diagnostic plots, MA plot, volcano plot, and heatmaps

Gene Set Enrichment and Pathway Analysis

- Gene set enrichment analysis tools: DAVID
- Pathway resources: KEGG
- Protein interaction resources: STRING

Major support was provided by the Helmsley Charitable Trust, Howard Hughes Medical Institute, and Regeneron.

## PARTICIPANTS

AbuMaziad, A., University of Arizona, Tucson
Amos, S., MIT, Cambridge, Massachusetts
Arduini, B., Regenerative Research Foundation, Rensselaer, New York
Arnone, J., Le Moyne College, Syracuse, New York
Barua, A., Tufts University, Boston, Massachusetts
Baumel-Alterzon, S., Icahn School of Medicine at Mount Sinai, New York, New York
Bian, J., Stanford University, California
Bonney Arku, C., University of Massachusetts, Amherst
Cabana, C., MIT, Somerville, Massachusetts
Corbali, O., Brigham and Women's Hospital, Boston, Massachusetts
Coronel, J., Columbia University, New York, New York
Dafalias, T., Boston University, U.S. Virgin Islands
Dantas, E., Weill Cornell Medical College, New York, New York
Davis, B., Ph.D., Johns Hopkins University, Baltimore, Maryland
Graham, M., Memorial Sloan Kettering Cancer Center, New York, New York
Grudda, T., Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland
Hodges, J., Sangamo Therapeutics, Walnut Creek, California
Korada, S., Stony Brook University, New York
Krey, W., Metagenomi, Emeryville, California
Lenert, M., University of Texas at Dallas, Richardson
Lin, Y., Cold Spring Harbor Laboratory
Marquez-Gomes, P., Georgia Institute of Technology, Atlanta
Mathenge, P., Dana-Farber Cancer Institute, Boston, Massachusetts
O'Toole, B., M.S., Paratus Sciences, New York, New York
Panoyan, M.A., M.Sc., University of Toronto, Mississauga, Ontario, Canada
Park, P., Temple University Lewis Katz School of Medicine, Philadelphia, Pennsylvania
Patel, A., Georgia Institute of Technology, Atlanta
Ramos Freitas, L., University of Texas at Dallas, Richardson
Reid, A., Ph.D., The City College of New York, New York
Sahin, Y., Cold Spring Harbor Laboratory
Simi, A., Istituto Italiano di Tecnologia, Genova, Italy
Song, S., GESC at Washington University in St. Louis, Missouri
Trevisan, A., St. Jude Children's Research Hospital, Memphis, Tennessee
Vaidya, H., Coriell Institute for Medical Research, Camden, New Jersey
van Heule, M., Ghent University, Merelbeke, Belgium
Wu, X., Columbia University, New York, New York

## SEMINAR

Tollkuhn, J., Cold Spring Harbor Laboratory: Hormonal control of sex differences in neural gene expression

# Quantitative Imaging: From Acquisition to Analysis

**April 11–25**


INSTRUCTORS

**H. Elliott,** PathAI, Inc., Boston, Massachusetts
**F. Jug,** Fondazione Human Technopole, Milan, Italy
**T. Lambert,** Harvard Medical School, Boston, Massachusetts
**J. Waters,** Harvard Medical School, Boston, Massachusetts

CO-INSTRUCTOR

**S. Manley,** École Polytechnique Fédérale de Lausanne (EPFL), Switzerland

ASSISTANTS

**D. Dalle Nogare,** Fondazione Human Technopole, Milan, Italy
**F. Gasparoli,** Harvard Medical School, Boston, Massachusetts
**L. Horin,** Harvard Medical School, Boston, Massachusetts
**S. Nørrelykke,** Harvard Medical School, Boston, Massachusetts
**R. Walsh,** Harvard Medical School, Boston, Massachusetts


Combining careful image acquisition with rigorous computational analysis allows extraction of quantitative data from light microscopy images that is far more informative and reproducible than what can be seen by eye. This course will focus on advanced quantitative fluorescence microscopy techniques used for imaging a range of biological specimens, from tissues to cells to single molecules. The course is designed for quantitative cell and molecular biologists, biophysicists, and bioengineers.

We provided a thorough treatment of the complete process of quantitative imaging, from the photons emitted from the sample to the extraction of biologically meaningful measurements from digital images. Material is covered in lectures, discussion groups, and hands-on quantitative exercises using commercial microscopes and open-source image analysis tools.

The concepts covered include:

- Widefield fluorescence microscopy
- Laser scanning and spinning disk confocal microscopy
- CCD, EM-CCD, and sCMOS cameras
- Total internal reflection fluorescence microscopy (TIRF)
- Light sheet microscopy
- Super-resolution microscopy (structured illumination, STED, and localization microscopy)
- Imaging and analyzing ratiometric "biosensors" (including FRET)
- Fluorescent proteins and live sample imaging
- Image processing (filtering, denoising, corrections, deconvolution)
- Image segmentation
- Quantitative shape and intensity measurements
- Object detection and tracking
- Machine learning
- Designing and troubleshooting quantitative imaging experiments
- And more!

The course also included a series of seminars from guest speakers who applied the methods discussed.

This course was supported in part by a grant from the National Cancer Institute. Scholarship support was provided by the Helmsley Charitable Trust, the Howard Hughes Medical Institute, Regeneron, and BioImaging North America (BINA).

## PARTICIPANTS

Ashesh, A., M.Tech., Fondazione Human Technopole, Milan, Italy
Avivi Kela, S., M.Sc., Tel Aviv University, Rishon Letsiyon, Israel
Baster, Z., Ph.D., National Institutes of Health, Bethesda, Maryland
Blake, T., Ph.D., University of Cambridge, United Kingdom
Christophers, B., B.A., Tri-Institutional MD-PhD Program, New York, New York
Flores Aldama, L., Ph.D., University of Wisconsin, Madison
Fragasso, A., Ph.D., Stanford University, California
John, M., M.S., University of Geneva, Switzerland
Kenry, F.N.U., Ph.D., Harvard University, Allston, Massachusetts
Kraus, J., Ph.D., Princeton University, New Jersey
Kumar, G., Ph.D., HHMI-UCSD, San Diego, California
Moravec, C., Ph.D., University of Wisconsin, Madison
Pathni, A., B.Tech., University of Maryland, College Park
Quansah, E., Ph.D., Salk Institute for Biological Studies, La Jolla, California
Shepherd, A., B.S., University of Utah, Salt Lake City
Troughton, L., Ph.D., Loyola University, Chicago, Illinois

## SEMINARS

Elliott, H., PathAI, Inc., Boston, Massachusetts: Basics of image processing and digital microscopy: resolution, SNR, and diffraction-limited objects; Image corrections and basic segmentation; Image correlation: colocalization, registration, and stitching; Image corrections and basic segmentation; Basic concepts in machine learning; 3D image analysis and deconvolution.
Jug, F., Fondazione Human Technopole, Milan, Italy: Intro to neural networks and deep learning; Machine learning in bioimage analysis; Image time series analysis: tracking.
Lambert, T., Harvard Medical School, Boston, Massachusetts: Digital imaging: Cameras, Parts I and II; Confocal microscopy: theory and hardware; Super-resolution microscopy: patterned illumination.
Manley, S., École Polytechnique Fédérale de Lausanne (EPFL), Switzerland: Super-resolution microscopy: localization.
Shaner, N., The Scintillon Institute, San Diego, California: Fluorescent proteins
Waters, J., Harvard Medical School, Boston, Massachusetts: Course overview; Introduction to quantitative microscopy;

Objective lenses; Koehler illumination and image formation; Phase, darkfield and DIC microscopy; Fluorescence microscopy; Quantifying fluorescence: image acquisition and controls; Live cell imaging; TIRF microscopy; Live confocal microscopy and intensity measurements over time; Multiphoton microscopy; Limitations on quantitative imaging of thick samples; Light sheet microscopy, Parts I and II.

Witkowski, J., Cold Spring Harbor Laboratory: Ethics, rigor, and reproducibility.

# Cell and Developmental Biology of *Xenopus*: Gene Discovery and Disease

**April 12–25**

INSTRUCTORS

**C. Chang,** University of Alabama at Birmingham
**L. Davidson,** University of Pittsburgh, Pennsylvania

ASSISTANTS

**G. Bastos Ventura**, University of Copenhagen, Denmark
**K. McCluskey**, University of California, San Francisco
**R. Stephenson**, University of Michigan, Ypsilanti
**G. Ventura**, University of Copenhagen, Denmark
**T. Yamamoto**, University of Zürich, Switzerland
**J. Yang**, University of Pittsburgh, Pennsylvania

In vivo animal models are an important tool for the understanding of human development and disease. Studies using the frog *Xenopus* have made remarkable contributions to our understanding of fundamental processes such as cell cycle regulation, transcription, translation, and many other topics. *Xenopus* is remarkable for studying development and disease, including birth defects, cancer, and stem cell biology. Because *Xenopus* are easy to raise, producing many thousands of eggs per day, these frogs have emerged as a premiere model for understanding human biology from the fundamental building blocks to the whole organism.

The recent development of CRISPR-Cas9 technology has made it easy to target genes of interest using *Xenopus*. This course has been designed with that in mind. Our goal was for each student to design a set of experiments focusing on their genes or biological interest. Prior to starting the course, students chose gene(s) of interest, and the instructors generated single-guide RNAs (sgRNAs) targeting these genes. These could be the students' own genes or chosen from a bank provided by the instructors. The gene-targeting experiments were combined with other manipulations, such as tissue explants and transplants and live imaging, to analyze the function of the genes.

*Xenopus* is increasingly being used as an imaging test bed to investigate the roles of cytoskeleton and intracellular trafficking in cell biological and morphogenetic contexts. The course maintained stock messenger RNAs (mRNAs) for targeting fluorescent proteins to specific structures to study cell shape and cytoskeletal dynamics, but students were encouraged to bring or suggest additional tools, including fluorescent biosensors, tension sensors, etc. The power of *Xenopus* can be leveraged when live-cell fluorescence imaging is combined with microsurgery, grafting, and dissociated cell culture.

During the course, the students analyzed any phenotypes generated from CRISPR-Cas9–based gene depletion while learning the diverse array of techniques available in *Xenopus*. In previous courses, we have guided students in the ablation of a wide variety of genes and helped them design suitable assays for their biological interests. Most recently, students have targeted autism genes, thyroid genes, and immune modulators, several of which have already led to publications. Approaches covered included microinjection and molecular manipulations such as CRISPR-Cas9 knockouts, antisense morpholino-based depletions, transgenics, and mRNA overexpression. In addition, students combined these techniques with explant and transplant methods to simplify or test tissue-level interactions. Additional methods included mRNA in situ hybridization and protein immunohistochemistry as well as basic bioinformatic techniques for gene comparison and functional analysis. Biochemical approaches such as proteomics and mass spectrometry and biomechanical concepts were also discussed. Finally, to visualize subcellular and intercellular activities, a variety of sample preparation and imaging methods including time-lapse, fluorescent imaging, optical coherence tomography, and confocal microscopy were introduced. These were facilitated by state-of-the-art equipment from Nikon, Leica, Thorlabs, and Bruker.

This course was supported in part by the National Institute of Child Health and Human Development. Scholarship support was provided by the Helmsley Charitable Trust, the Howard Hughes Medical Institute, and Regeneron.

## PARTICIPANTS

Adhikary, B., M.Sc., University of Michigan, Ann Arbor
Berns, H., Faculty of Medicine, University Hospital Freiburg, Germany
Boitel, C., M.A., Institut de Biologie du Développement, Marseille, France
Hansen, M., M.S., University Freiburg Medical Center, Freiburg Im Breisgau, Germany
Jayaprakashappa, D., M.S., Institute of Molecular Biology, Mainz, Germany
Kostyanovskaya, E., B.S., University of California, San Francisco
Liu, X., Ph.D., University of California, Berkeley, Berkeley, CA
Mosqueda, N., Ph.D., Stony Brook University, New York
Seal, S., M.Sc., Institut Curie, Orsay, France
Shirouzu, H., B.S., Waseda University, Shinjuku-Ku, Japan
Thurner, L.A., M.S., Novo Nordisk Center for Stem Cell Medicine reNEW, Copenhagen, Denmark
Walker, B., M.S., UTHealth McGovern Medical School, Houston, Texas
Xia, B., Ph.D., Broad Institute of MIT and Harvard, Cambridge, Massachusetts
Youmans, L., M.D., UTHealth Houston, Texas
Zhou, C., M.S., Institut Curie, Orsay, France

SEMINARS

Davidson, L., University of Pittsburgh, Pennsylvania: Leveraging *Xenopus* mechanics and morphogenesis.
Deniz, E., Yale School of Medicine, New Haven, Connecticut: OCT imaging application in *Xenopus*.
Funabiki, H., Rockefeller University, New York, New York: Using *Xenopus* egg extracts for chromatin studies.
Houston, D., University of Iowa, Ames: Maternal control of development.
Khokha, M., Yale University, New Haven, Connecticut: Behold the power of *Xenopus*: the organizer, oxygen, and mitochondria.
Miller, R., UTHealth Houston, McGovern Medical School, Texas: Advancing *Xenopus* as a model of kidney development and birth defects; Kidney development.
Mitchell, B., Northwestern University, Chicago, Illinois: Cytoskeleton and ciliogenesis.
Monsoro-Burq, A.-H., Université Paris Saclay, Institut Curie, France: The neural crest: exploring a multifaceted vertebrate innovation in *Xenopus* embryos from tissue to single cells.
Nascone-Yoder, N., North Carolina State University, Raleigh, North Carolina: Left–right asymmetric gut morphogenesis.
Peshkin, L., Marine Biological Laboratory, Woods Hole, Massachusetts: Tabula rasa—single-cell transcriptomics and cell type atlas of *Xenopus*.
Willsey, H., University of California San Francisco: Using frogs to dissect psychiatric disorders.
Witkowski, J., Cold Spring Harbor Laboratory: Ethics, rigor, and reproducibility.
Woolner, S., University of Manchester, United Kingdom: Using *Xenopus* to investigate how mechanical force regulates cell division.
Zahn, N., Independent Creative, East Calais, Vermont: Workshop: scientific Illustration.

# Expression, Purification, and Analysis of Proteins and Protein Complexes

**April 12–25**


INSTRUCTORS

**M. Marr,** Brandeis University, Waltham, Massachusetts
**S. Nechaev,** University of North Dakota School of Medicine, Grand Forks
**A. Quigley,** Diamond Light Source, Didcot, United Kingdom

ASSISTANTS

**B. Bodley-Gomes,** Brandeis University, Waltham, Massachusetts
**C. Boyle,** University of North Dakota, Grand Forks
**W. Dahl,** Brandeis University, Waltham, Massachusetts
**P. Harrison,** Diamond Light Source, Didcot, United Kingdom
**M. Olejnik,** Diamond Light Source, Didcot, United Kingdom


This course was for scientists, including graduate students, postdoctoral scholars, staff scientists, and principal investigators, who wanted a rigorous introduction to expression and purification of proteins as well as analysis of protein structure and function.

Through hands-on experience in the laboratory as well as extensive lecture and discussion, each student became familiar with key approaches in expression, purification, and analysis of soluble and membrane proteins and protein complexes from both natural sources and overexpression systems.

The emphasis of the course was on the following.

- Approaches in protein expression: choosing the best bacterial or eukaryotic expression system tailored for the particular protein and experimental problem; determining how to optimize expression; and understanding protein tagging: the advantages and pitfalls of various affinity and solubility tags.

- Approaches in protein purification: choosing the best strategy for a given protein, including solubilization, bulk fractionation, and liquid chromatography, encompassing conventional methods (ion exchange, size exclusion, reverse phase, etc.) and affinity methods (e.g., MAC, DNA affinity, immunoaffinity), as well as FPLC/HPLC.
- Approaches in protein analysis: Introduction to common approaches for characterization of proteins including binding assays, activity assays, mass spectroscopy to identify protein interaction partners, and posttranslational modifications.

In addition to purification, students gained exposure to fundamental analytical approaches such as mass spectroscopy and protein structure determination (e.g., X-ray crystallography, cryo-EM).

Major support for this course was provided by the National Cancer Institute. Scholarship support was provided by the Helmsley Charitable Trust, the Howard Hughes Medical Institute, and Regeneron.

## PARTICIPANTS

Argiles Martinez, G., M.D., Memorial Sloan Kettering Cancer Center, New York, New York
Avad, K., B.S., University of Tennessee Health Science Center, Memphis, Tennessee
Baboolall, K., Ph.D., University of Chicago, Illinois
Choudhary, S., B.E., Virginia Tech, Blacksburg
Diggs, M., M.S., Georgia Institute of Technology, Atlanta
Hsu, C.-I., M.S., Virginia Tech, Washington, D.C.
Kim, H.S., Cold Spring Harbor Laboratory
Langmack, L., B.A., University of Iowa, Iowa City
Le, A., B.A., University of Colorado Anschutz Medical Campus, Aurora
Lizardi-Ortiz, J., B.S., University of Puerto Rico, San Juan
Pelzer, B., M.D., Weill Cornell Medical College, New York, New York
Rolli, S., B.S., Marquette University, Milwaukee, Wisconsin
Stevens, R., B.S., University of South Alabama, Mobile
Strandback, E., Ph.D., Karolinska Institutet, Solna, Sweden
Tran, H., B.S., Rutgers University, Piscataway, New Jersey
Wang, Y., Ph.D., Cold Spring Harbor Laboratory

## SEMINARS

Adamala, K., University of Minnesota, Minneapolis: Protein expression without cells: in vitro translation systems.
Harrison, P., Diamond Light Source, Didcot, United Kingdom: Electron microscopy of small membrane proteins.
Jarvis, D., University of Wyoming, Laramie: Overview of the baculovirus–insect cell system for recombinant protein production.
Marr, M., Brandeis University, Waltham, Massachusetts: Protein purification.
Nechaev, S., University of North Dakota School of Medicine, Grand Forks: Cellular responses to stress reveal new roles for Pol II transcription machinery components.
Okafor, D., Pennsylvania State University, University Park: Ligand interactions that determine transcriptional outcomes.
Pappin, D., Cold Spring Harbor Laboratory: Intro to MS of proteins: quantitative approaches in protein MS.
Quigley, A., Diamond Light Source, Didcot, United Kingdom: The A–Z of membrane protein structural biology.
Wen, F., University of Michigan, Ann Arbor: Multiplexed protein detection at the single-cell level using cytometry.

# Advanced Bacterial Genetics

**June 6–24**


INSTRUCTORS **M. Erhardt,** Humboldt-Universität zu Berlin, Germany
**A. Grundling,** Imperial College London, United Kingdom
**S. Pukatzki,** The City College of New York, New York

TEACHING ASSISTANTS **E. Donohue,** City College of New York, New York
**M. Giralt Zuniga,** Humboldt-Universität zu Berlin, Germany
**P. Popp,** HU Berlin, Germany
**C. Weiss,** Imperial College London, United Kingdom


The Advanced Bacterial Genetics course presented logic and methods used in the genetic dissection of complex biological processes in diverse bacteria.

Laboratory methods included:

- Classical and cutting-edge mutagenesis using transposons and allelic exchange
- Recombineering with single- and double-stranded DNA
- CRISPR-Cas genome editing
- Genome sequencing and assembly
- Mapping mutations using genetic and next-generation sequencing techniques
- Generation of genomic DNA libraries and use for genetic screens
- Modern approaches to the generation and analysis of targeted gene disruptions and reporter gene fusions

- Fluorescence microscopy
- Key components of the course will be the use of sophisticated genetic methods in the analysis of model bacteria (including *Escherichia coli*, *Salmonella*, *Staphylococcus aureus*, and *Vibrio cholerae*), and the use of the wealth of new genomic sequence information to motivate these methods.

This course was supported in part by grants from the National Science Foundation. Scholarship support was provided by the Helmsley Charitable Trust, the Howard Hughes Medical Institute, and Regeneron.

## PARTICIPANTS

Accilien, K., B.S., University of Kansas, Lawrence
Appell, K., B.S, University of Arkansas for Medical Sciences, Little Rock
Broux, K., M.S., KULeuven, Belgium
Chlebek, J., Ph.D., Lawrence Livermore National Laboratory, Livermore, California
Coe, L., B.S., University of Florida, Gainesville
Diakogianni, A., B.Sc., Institute of Science and Technology, Austria, Klosterneuburg
Hernandez, G., B.S., Wake Forest University School of Medicine, Winston-Salem, North Carolina
Hustus, K., B.S., Virginia Commonwealth University, Richmond
Koehler, R., M.Sc., Max Planck Institute for Terrestrial Microbiology, Marburg, Germany
Künnecke, J., B.Sc., University of Basel, Switzerland
Marciniak-Haynosch, J.-W., B.S., University of Illinois at Urbana-Champaign, Urbana
Mi, W., Ph.D., Yale University School of Medicine, New Haven, Connecticut
Phelps, G., Pharm.D., St. Jude Children's Research Hospital, Memphis, Tennessee
Prasad, A., A.B., University of California, Berkeley
Roux, I., Ph.D., University of Cambridge, United Kingdom
Wende, M., M.Sc., Helmholtz Centre for Infection Research, Braunschweig, Germany

## SEMINARS

Berkmen, M., New England Biolabs, Ipswich, Massachusetts: Bridging the gap with bacterial art.
Burton, B., University of Wisconsin, Madison: The ins and outs of natural transformation.
Fiebig, A., Michigan State University, East Lansing: Many applications of BARseq: lessons learned using barcoded transposons.
Jenal, U. University of Basel, Switzerland: Surface colonization of a human pathogen revisited.
Wakeman, C., Texas Tech University, Lubbock: Malonate utilization alters metal homeostasis and community structure in *Pseudomonas aeruginosa*.
Wang, J., University of Wisconsin, Madison: Stress signaling by alarmones in Gram-positive bacteria.
Witkowski, J., Cold Spring Harbor Laboratory: Ethics in science research.

# Ion Channels in Synaptic and Neural Circuit Physiology

**June 6–25**


INSTRUCTORS

**T. Giraldez,** Universidad de La Laguna, San Cristobal de La Laguna, Spain
**C. Grienberger,** Brandeis University, Waltham, Massachusetts
**J. Grundemann,** DZNE, Bonn, Germany
**A. Scimemi,** SUNY Albany, New York
**N. Wanavezrbcq,** Aix Marseille University INT, France

COURSE TEACHING ASSISTANTS

**A. Erickson,** RWTH Aachen University, Germany
**A. Galloni,** Rutgers University, Piscataway, New Jersey
**K. Kalary,** Scimemi Laboratory, Albany, New York
**O. Rauh,** Technische Universität Darmstadt, Germany
**M. Vivian,** Brandeis University, Waltham, Massachusetts


Ion channels are the fundamental building blocks of activity in the nervous system. The primary goal of this course was to demonstrate, through lectures and laboratory work, the different biophysical properties of ion channels that enable neurons to perform unique physiological functions in various neural systems.

Areas of particular interest included:

- Voltage- and ligand-gated ion channels at central and peripheral synapses
- Synaptic integration and plasticity
- Neural circuit function in vitro and in vivo
- Optogenetic strategies for circuit manipulation.

This intensive laboratory and lecture course introduced the participants to state-of-the-art electrophysiological approaches to studying ion channels in their native environments, specifically focusing on understanding different biophysical phenomena from first principles. A typical day consisted of morning lectures followed by hands-on laboratory practical sessions in the afternoons and evenings, with guest lecturers available to give one-on-one practical advice. The course participants also learned how to troubleshoot their experiments and discuss and evaluate their data.

The course provided students with hands-on experience in:

- Using patch-clamp electrophysiology to examine single-channel activity in cultured cells
- Ion channel biophysics in acutely dissociated neurons and synaptic integration
- Plasticity and circuit dynamics in slice and in vivo preparations
- The use of different recording configurations and modes (e.g., cell-attached and whole-cell configurations, voltage- and current-clamp modes, somatic and dendritic recordings from neuronal and nonneuronal cells) and the advantages and limitations of each method were discussed in the context of cutting-edge scientific questions. The course also provided practical experience applying cellular and circuit manipulation techniques in vitro (e.g., pharmacology and optogenetics).

This course was supported in part by grants provided by the Helmsley Charitable Trust, the Howard Hughes Medical Institute, and Regeneron.

## PARTICIPANTS

Aquino Miranda, G., Ph.D., University of Texas Health Science Center at Houston
Barlow, S., Ph.D., University of Maryland School of Medicine, Baltimore, Maryland
Barraza, M., B.A., Northwestern University, Chicago, Illinois
George, A., Rutgers University, Piscataway, New Jersey
Gonzalez, K., B.S., Columbia University, New York, New York
Gonzalez Velandia, K., Ph.D., University of Rochester, New York
Hernandez, R., B.A., Florida Atlantic University, Jupiter
Kwan, W., B.Sc., Monash University, Clayton, Victoria, Australia
Park, E., B.A., Stanford University, California
Roll, F., M.Sc., RWTH Aachen, Germany
Segura Covarrubias, M., Ph.D., Case Western Reserve University, Cleveland, Ohio
Viswanathan, J., B.S., Stanford University, California

## SEMINARS

Baranovic, J., University of Edinburgh, United Kingdom: Single channels.
Beeton, C., Baylor College of Medicine, Waco, Texas: Potassium channels.
Cohen, J., Allen Institute, Seattle, Washington: Principles of extracellular electrophysiology in vivo.
Colecraft, H., Columbia University, New York, New York: Voltage-gated calcium channels.
Couto, J.P.dO., University of California, Los Angeles: High-capacity electrophysiology: how we got here and where can we go.
Duguid, I., University of Edinburgh, United Kingdom: Patch-clamp recording in vivo.
Dulla, C., Tufts University School of Medicine, Boston, Massachusetts: Ion channels in glia.
Fenno, L., UT Austin/Dell Medical School, Texas: Optogenetics.
Gasparini, S., LSU Health Sciences Center New Orleans, Louisiana; Giraldez, T., Universidad de La Laguna, Spain; Grienberger, C., Brandeis University, Boston, Massachusetts; Grundemann, J., DZNE, Göttingen, Germany; and Scimemi, A., SUNY Albany, New York: Principles of biophysics and recordings modes.
Harris, T., HHMI Janelia Research Campus, Ashburn, Virginia: High-channel-count electrophysiology.
Huguenard, J., Stanford University, California: "Modern" approaches to study ion channel function.
Lampert, A., RWTH Aachen University, Germany: Voltage-gated sodium channels.
Mallarino, R., Princeton University, New Jersey: The developmental basis of biological diversity.

Overstreet-Wadiche, L., University of Alabama at Birmingham: Synaptic inhibition.
Posfai, E., Princeton University, New Jersey: Genome engineering and live imaging of the early mouse embryo.
Rancz, E., Aix-Marseille University, France: Synaptic integration and dendritic recording.
Ruta, V., Rockefeller University, New York, New York: Closing plenary lecture.
Daniel Schramek, D., Lunenfeld-Tanenbaum Research Institute/University of Toronto, Ontario, Canada: *In vivo veritas*—using CRISPR in mice to reveal novel cancer genes and vulnerabilities.
Sjostrom, J., McGill University, Montréal, Québec, Canada: Mapping visual cortex microcircuits with two-photon optogenetics.
Wadiche, L., University of Alabama at Birmingham: GABAergic inhibition.
Witkowski, J., Cold Spring Harbor Laboratory: Ethics lecture.
Wollmuth, L., Stony Brook University, New York: Introduction to ion channels.
Matthew Xu-Friedman, M., SUNY Buffalo, New York: Long-term consequences of abnormal activity on auditory nerve synapses.

# Mouse Development, Stem Cells, and Cancer

**June 6–26**


INSTRUCTORS
**S. Vokes,** The University of Texas at Austin
**Y. Yamanaka,** Goodman Cancer Institute, McGill University, Montréal, Québec, Canada

CO-INSTRUCTORS
**D. Devenport,** Princeton University, Princeton, New Jersey
**L. Jerome-Majewska,** McGill University, Montréal, Québec, Canada

COURSE TEACHING ASSISTANTS
**M. Cortez,** Princeton University, Princeton, New Jersey
**M. Cowan,** McGill University (MICAM), Montréal, Québec, Canada
**Y. Dong,** McGill University, Montréal, Québec, Canada
**K. Harwalkar,** McGill University, Montréal, Québec, Canada
**D. Konjusha,** McGill University, Montréal, Québec, Canada
**H. Lee,** University of Texas at Austin
**C. Millet-Boureima,** McGill University Health Centre, Montréal, Québec, Canada
**S. Paramore,** Princeton University, New Jersey
**P. Sil,** Princeton University, New Jersey


This intensive lecture and laboratory course was designed for scientists, typically at the early principal investigator, postdoctoral, or graduate stages, interested in using mouse models to study mammalian development, stem cells, and cancer. The lecture portion of the course taught by leaders in the field provided the conceptual basis for contemporary research in embryogenesis;

organogenesis in development and disease; embryonic, adult, and induced pluripotent stem cells; and cancer biology.

The laboratory and workshop portions of the course provided a hands-on introduction to engineering of mouse models, stem cell technologies, and tissue analyses.

Experimental techniques included:

- CRISPR-Cas9 editing and allele design
- Zygote microinjection and embryo transfer
- Isolation, culture, and manipulation of pre- and post-implantation embryos
- Zygote and embryo electroporation
- In vivo electroporation for somatic cancer modeling
- Embryo roller bottle culture
- In vitro fertilization
- Embryonic stem cell–directed differentiation
- Fluorescent RNA in situ hybridization
- Immunofluorescence
- Skeletal preparation
- Live imaging and confocal microscopy

This course was supported in part by a grant from the National Cancer Institute. Scholarship support was provided by the Helmsley Charitable Trust, the Howard Hughes Medical Institute, and Regeneron.

## PARTICIPANTS

Acevedo, D., M.S., St. Jude Children's Research Hospital, Memphis, Tennessee
Adkins, G., B.A., St. Jude Children's Research Hospital, Memphis, Tennessee
Aziz, F., B.S., Michigan State University, East Lansing
English, K., B.S., University of Alabama at Birmingham
Gandhi, S., Ph.D., University of California, Berkeley
Gonzalez, C., B.S., City of Hope Medical Center, Duarte, California
Kataruka, S., Ph.D., Yale University, New Haven, Connecticut
Kneuss, E., Ph.D., European Molecular Biology Laboratory, Heidelberg, Germany
Lee, H., B.A., Cincinnati Children's Hospital Medical Center, Ohio
Pathak, M., Ph.D., Carnegie Institute of Science, Baltimore, Maryland
Picech, F., Ph.D., Columbia University, New York, New York
Pompan, T., B.S., National Cancer Institute, Frederick, Maryland
Staneva, D., Ph.D., University of Cambridge, United Kingdom
Taylor, G., The Francis Crick Institute, London, United Kingdom
Yeh, S.-Y., M.S., University of Toronto, Ontario, Canada
Zhou, J., Ph.D., Columbia University, New York, New York, and China

## SEMINARS

Anderson, M., NCI, Frederick, Maryland: Whole-mount mouse embryo hybridization chain reaction mRNA in situ workshop.
Ripla Arora, R., Michigan State University, East Lansing: Uterine 3D structure facilitates early embryo–uterine interactions.
Devenport, D., Princeton University, New Jersey: Spatial and directional patterning of the epidermis.
Egeblad, M., Cold Spring Harbor Laboratory: Tumors in living color: intravital imaging in cancer research.
Francois, P., Universite de Montréal, Québec, Canada: Mathematical modelling of embryonic development.
Han, X., Cold Spring Harbor Laboratory: Window implantation surgery for intravital imaging.
Laird, D., University of California, San Francisco: Development and heterogeneity in the germline.
Lickert, H., Helmholtz Zentrum Munchen, Germany: Deciphering endoderm and pancreas development for beta cell regeneration.
Maitre, J.-L., Institut Curie, Paris, France: Mechanics of blastocyst morphogenesis.

Mallarino, R., Princeton University, New Jersey: The developmental basis of biological diversity.

Munger, S., Jackson Laboratory, Bar Harbor, Maine: The stories that SNPs tell: genetic diversity as a tool for discovery.

Murray, S., Jackson Laboratory, Bar Harbor, Maine: Large-scale discovery and characterization of mammalian essential genes.

Papaioannou, V., Columbia University Medical Center, New York, New York: In the beginning: a brief history of mouse embryology.

Eszter Posfai, E., Princeton University, New Jersey: Genome engineering and live imaging of the early mouse embryo.

Schramek, D., Lunenfeld-Tanenbaum Research Institute/ University of Toronto, Ontario, Canada: *In vivo veritas*—using CRISPR in mice to reveal novel cancer genes and vulnerabilities.

Srinivas, S., University of Oxford, United Kingdom: Early heart development.

Stottmann, R., Nationwide Children's Hospital Medical Center, Columbus, Ohio: Gene discovery and craniofacial structural birth defects.

Sumigray, K., Yale School of Medicine, New Haven, Connecticut: Principles of intestinal morphogenesis and patterning.

Sun, X., University of California, San Diego: Using mouse to study the lung.

Sutherland, A., University of Virginia, Charlottesville: Axial elongation and neural tube morphogenesis.

Tarchini, B., Jackson Laboratory, Bar Harbor, Maine: How cell polarity mechanisms shape sensory perception in the inner ear.

Trainor, P., Stowers Institute for Medical Research, Kansas City, Missouri: Neural crest cells and their fundamental roles in development, evolution, and disease.

Vokes, S., UT Austin, Texas: Transcriptional interpretation of Hedgehog signaling.

Wellik, D, University of Wisconsin, Madison: Hox genes in development and beyond.

Yamanaka, Y., Goodman Cancer Institute, McGill University, Montréal, Québec, Canada: Overview of mouse development: the heterogeneity of oviduct luminal epithelial cells and ovarian cancer initiation.

# Metabolomics

**June 10–25**

INSTRUCTORS

**J. Cross**, Memorial Sloan Kettering Cancer Center, New York, New York
**A. Rosebrock**, Chemistry of Biology LLC, Indianapolis, Indiana

CO-INSTRUCTORS

**I. Abramovich**, Technion Institute of Technology, Haifa, Israel
**A. Caudy**, Maple Flavored Solutions, LLC, Zionsville, Indiana
**D. Sumpton**, Beatson Institute for Cancer Research, Glasgow, United Kingdom

COURSE TEACHING ASSISTANTS

**T. Banjo**, Memorial Sloan Kettering Cancer Center, New York, New York
**A. Huerta Uribe**, Beatson Institute for Cancer Research, Glasgow, United Kingdom
**A. Kyaw**, Memorial Sloan Kettering Cancer Center, New York, New York
**C. Pohl**, Memorial Sloan Kettering Cancer Center, New York, New York
**R. Ramos**, Memorial Sloan Kettering Cancer Center, New York, New York
**M. Saoi**, Memorial Sloan Kettering Cancer Center, New York, New York
**E. Shokry**, The Beatson institute for Cancer Research, Glasgow, United Kingdom
**C. Strozek**, Memorial Sloan Kettering Cancer Center, New York, New York
**S. Violante**, Memorial Sloan Kettering Cancer Center, New York, New York
**J. Yeh**, Cold Spring Harbor Laboratory

Advances in genomics, transcriptomics, and proteomics have enabled both broad and deep analysis of genomes and their encoded proteins. Metabolomics focuses on measuring the biochemical contents and activities of cells, tissues, and organisms. Biochemical phenotypes represent a unique view into the dynamic state of biological systems and are relevant to a range of fields, from model organisms to patients, from bioprocess to bedside.

CSHL's Metabolomics course combined theoretical and practical training including hands-on experience with a range of cutting-edge approaches to interrogate biochemical state. Mass spectrometry (MS) is currently the most powerful and flexible approach in the metabolomics toolbox; students become proficient in the generation and analysis of both gas-chromatography (GC) and liquid-chromatography (LC) mass spectrometry data. New biochemistry awaits discovery, even in well characterized systems. Participants learned how to quantitate known metabolites in complex biological samples and how to discover and characterize unknown compounds using LC and GC mass spectrometry.

Metabolomics makes use of many complementary tools in addition to mass spectrometry. Students gained hands-on experience in:

- Multiple LC and GC mass spectrometers to directly measure metabolites and metabolic flux
- Developing practical chromatographic separations
- Measuring metabolic state in live cells using the Agilent Seahorse XF platform.

The Metabolomics course integrated practical lab sessions, hands-on data analysis, and lecture-based learning. Students interacted with instructors and TAs as well as a diverse panel of field-leading guest speakers who presented both formal talks and a nuts-and-bolts view of metabolomics in their laboratories.

The objectives for students were to:

- Become proficient in quantitative and qualitative analysis of GC- and LC-MS data using vendor and open-source tools
- Understand the use cases and limitations of currently available metabolomics instrumentation and be able to identify the right approach for a given question
- Learn key factors in experimental design and sample preparation that enable collection of interpretable and actionable metabolomics data
- Gain the core knowledge and vocabulary required to fruitfully interact with other researchers in the metabolomics field.

The goal of this course was to make students proficient in metabolomic analysis. Applicants did not need to have prior experience in metabolomics or mass spectrometry. Successful course participants come from a range of backgrounds: Alumni range from senior graduate students through full professors and established industrial scientists.

This course was supported in part by grants from the Helmsley Charitable Trust, the Howard Hughes Medical Institute, and Regeneron.

## PARTICIPANTS

Awad, D., Ph.D., University of Michigan, Ann Arbor
Azubuike, C., Ph.D., Oak Ridge National Laboratory, Tennessee
Dekina, S., Ph.D., The European Molecular Biological Laboratory, Heidelberg, Germany
Dias de Medeiros, H.C., Ph.D., Michigan State University, East Lansing,
Edel, M., Ph.D., Hamburg University of Technology, Germany
Heckhausen, Z., M.Sci., Imperial College London, United Kingdom
Henze, E., Ph.D., Cornell University, Ithaca, New York
Jiang, H., Stanford University, California
Kober, M., Ph.D., University of California, Berkeley
Leite Dias, S., M.Sc., IPK Gatersleben, Seeland Ot Gatersleben, Germany
McNeil, C., Ph.D., Finless Foods, Emeryville, California
Shalev, O., Ph.D., The Hebrew University of Jerusalem, Israel
Sharma, P., Ph.D., St. Jude Children's Research Hospital, Memphis, Tennessee
Sun, Z., M.S., Pfizer, Cambridge, Massachusetts
Tan, J., University of Cambridge, United Kingdom
Wedan, R., B.S., Van Andel Institute, Grand Rapids, Michigan
Wei, P., Ph.D., Vertex Pharmaceuticals, Boston, Massachusetts
Wolosker, H., Ph.D., Technion-Israel Institute of Technology, Haifa

SEMINARS

Aron, A., University of Denver, Colorado: Computational tools for nontargeted mass spectrometry. Analysis

Bennett, B., Calico, South San Francisco, California: Metabolism of the integrated stress response.

Caudy, A., Maple Flavored Solutions, LLC, Stony Brook, New York: Rewriting textbooks—using metabolomics to discover new enzymes.

Clasquin, M., Pfizer, Cambridge, Massachusetts: Metabolic fluxes in chronic disease and lessons from pharmacological intervention.

Dexter, A., National Physical Laboratory, Teddington, United Kingdom: Exploring tumor heterogeneity using multimodal MS imaging strategies.

Evans, A., Metabolon, Morrisville, North Carolina: Precision metabolomics—a single technology for understanding human health.

Fan, T., University of Kentucky, Lexington: NMR profiling of metabolites and their 13C labeling patterns in stable isotope-resolved metabolomics studies.

Gottlieb, E., University of Texas MD Anderson Cancer Center, Houston: Metabolic vulnerabilities of cancer.

Gross, S., Weill Cornell Medical College, New York, New York: What a long strange trip it's been: from deep dive pharmacology of NO to being a NYC metabolomics guy.

Keshari, K., Memorial Sloan Kettering Cancer Center, New York, New York: Metabolic imaging using hyperpolarized magnetic resonance.

Sheldon, R., Van Andel Institute, Grand Rapids, Michigan: In vivo stable isotope labeling.

Vander Heiden, M., Massachusetts Institute of Technology, Cambridge: How metabolomics can inform cancer biology.

Vander Voorde, J., Beatson Institute for Cancer Research, Glasgow, United Kingdom: Metabolic mapping of colorectal cancer for tissue stratification and target identification.

# Vision: A Platform for Linking Circuits, Perception, and Behavior

**June 16–July 1**


INSTRUCTORS **F. Briggs**, University of Rochester, New York
**J. Carroll**, Medical College of Wisconsin, Milwaukee
**K. Nielsen**, Johns Hopkins University, Baltimore, Maryland

COURSE TEACHING ASSISTANT **A. Murphy**, University of Rochester, New York


The purpose of this course was to bring together students and faculty for in-depth and high-level discussions of modern approaches for probing how specific cell types and circuits give rise to defined categories of visual perception and behavior. It was also designed to address novel strategies aimed at overcoming diseases that compromise visual function.

The visual system is the most widely studied sensory modality. In recent years, emerging technological advances have encouraged exploration of visual function across a wider array of model systems using diverse experimental approaches. For example, the tractability of genetic manipulation and imaging in mice has led to an increase in the use of the mouse as a model system for exploring how specific cells and circuits underlie visual and multisensory processing and cognition. Additionally, advances in genetic and viral methods have enabled similar cell- and circuit-centric explorations of visual function in a variety of model systems including insectivores,

carnivores, and primates. Finally, the field of visual neuroscience is at the forefront of technological and therapeutic advances in clinical/translational work to restore visual function in humans.

The time was ripe to build on the classic paradigms and discoveries of visual system structure, function, and disease in order to achieve a deep, mechanistic understanding of how neuronal populations encode sensory information, how different circuits can induce defined categories of percepts and behaviors, and how modulations of cells and circuits may restore visual function in the diseased brain.

This course was supported in part by grants from the National Eye Institute, part of the National Institutes of Health, and the Fighting Blindness Foundation. Additional scholarship support was provided by the Helmsley Charitable Trust, the Howard Hughes Medical Institute, and Regeneron.

## PARTICIPANTS

Adusei, M., B.S., University of Rochester, New York
Baumann, M., M.S., University of Tübingen, Germany
Bucklaew, A., B.S., University of Rochester, West Henrietta, New York
Comeaux, P., B.S., University of Utah, Salt Lake City
D'Angelo, J., B.S., University of California, Berkeley
Fars, J., Ph.D., WIN, FMRIB, University of Oxford, United Kingdom
Gonzalez-Amoretti, J., B.S., University of Rochester, New York
Haak, R., B.Sc., Netherlands Institute for Neuroscience, Amsterdam
Hellmer, C., B.S., University of Louisville, Kentucky
Hong, F., B.S., New York University, New York
Jones, E., B.A., University of California, San Francisco
Khoussine, J., B.S., University of Wisconsin, Madison
Kramer, M., M.A., Carnegie Mellon University, Pittsburgh, Pennsylvania
Kreis, J., B.A., Medical College of Wisconsin, Wauwatosa
Li, C., B.S., University of Virginia, Charlottesville
Meyer, E., B.S., University of Pennsylvania, Philadelphia
Nagy, J., B.S., University of Wisconsin, Madison
Nguyen, T., B.S., Duke University, Durham, North Carolina
Poudel, S., B.S., State University of New York, College of Optometry, New York
Quiroli, E., M.Sc., Institute of Science and Technology Austria (ISTA), Klosterneuburg
Sanchez, A., B.A., University of California, Davis
Schleufer, S., B.S., University of Washington, Seattle
Whitley, J., B.S., University of Louisville, Kentucky
Yip, H.M.K., B.Soc.S., Monash University, Clayton, Victoria, Australia

## SEMINARS

David Brainard, D., University of Pennsylvania, Philadelphia: Optics and visual psychophysics.
Bridge, H., University of Oxford, United Kingdom: Prospects for rehabilitation of the human visual system following stroke.
Briggs, F., University of Rochester, New York: Course overview.
Callaway, E., Salk Institute, La Jolla, California: Functional organization for color appearance mechanisms in primary visual cortex.
Carroll, J., Medical College of Wisconsin, Milwaukee: Visual system overview.
Chen, C., Boston Children's Hospital, Harvard Medical School, Massachusetts: Retinogeniculate development and plasticity.
Cooper, E., University of California, Berkeley; Read, J., Newcastle University, United Kingdom: 3D vision.
Dalkara, D., Institut de la Vision, Paris, France: Restoring vision using gene and cell therapy.
Dunn, F., University of California, San Francisco: Retina.
Field, G., University of California, Los Angeles: Information processing in the retina.
Fitzpatrick, D., Max Planck Florida Institute for Neuroscience, Jupiter: Circuits for feature computations.
Horton, J., University of California, San Francisco: Binocular function, amblyopia, and strabismus.
Kastner, S., Princeton University, New Jersey: The cognitive thalamus.
Krauzlis, R., National Eye Institute, Bethesda, Maryland; Moore, T., Stanford University, California: Attention and goal-directed behavior.
Leopold, D., National Institute of Mental Health, Bethesda, Maryland: Single-unit fMRI mapping: new perspectives on brain organization.
Mason, C., Columbia University, New York, New York: Retinal ganglion cell development in disease.
Movshon, J.A., New York University, New York; Pasupathy, A., University of Washington, Seattle: Extrastriate cortical computations.

Neitz, J., University of Washington, Seattle: Color vision.
Nielsen, K., Johns Hopkins University, Baltimore, Maryland: Development of higher-level vision.
Poletti, M., University of Rochester, New York: Active vision.
Roe, A., Zhejiang University, China: Modern methods to visualize complex visual activity.
Roorda, A., University of California, Berkeley: Hacking human vision.
Sommer, M., Duke University, Durham, North Carolina: Eye movements.
Usrey, W.M., University of California, Davis: Feedforward and feedback circuit dynamics for vision.
Wilke, M., University Medicine Goettingen/DPZ, Germany: Thalamic contributions to spatial decisions and action selection.
Witkowski, J., Cold Spring Harbor Laboratory: Ethics lecture.

# Statistical Analysis of Genome Scale Data

**June 30–July 13**


INSTRUCTORS

**V. Carey**, Harvard Medical School, Boston, Massachusetts
**S. Davis**, University of Colorado Anschutz School of Medicine, Denver

COURSE TEACHING ASSISTANTS

**F.R. Garcia-Flores**, Lieber Institute for Brain Development, Baltimore, Maryland
**D. Gonzalez-Padilla**, Lieber Institute for Brain Development, Baltimore, Maryland
**S. Liu**, Columbia University, New York, New York


Over the past decade, high-throughput assays have become pervasive in biological research because of both rapid technological advances and decreases in overall cost. To properly analyze the large data sets generated by such assays and thus make meaningful biological inferences, both experimental and computational biologists must understand the fundamental statistical principles underlying analysis methods. This course was designed to build competence in statistical methods for analyzing high-throughput data in genomics and molecular biology.

Topics included:

- The R environment for statistical computing and graphics
- Introduction to Bioconductor
- Review of basic statistical theory and hypothesis testing
- Experimental design, quality control, and normalization
- High-throughput sequencing technologies

- Expression profiling using RNA-seq and microarrays
- *In vivo* protein binding using ChIP-seq
- High-resolution chromatin footprinting using DNase-seq
- DNA methylation profiling analysis
- Integrative analysis of data from parallel assays
- Representations of DNA binding specificity and motif discovery algorithms
- Predictive modeling of gene regulatory networks using machine learning
- Analysis of posttranscriptional regulation, RNA binding proteins, and microRNAs.

This course was supported in part by grants from the National Human Genome Research Institute, part of the National Institutes of Health. Additional scholarship support was provided by the Helmsley Charitable Trust, the Howard Hughes Medical Institute, and Regeneron.

## PARTICIPANTS

Chriss, A., B.S., Icahn School of Medicine at Mount Sinai, New York, New York
Chung, T.T., M.S., Chungnam National University, Daejeon, Korea
Como, C., B.A., University of Colorado Anschutz Medical School, Aurora
Elias, R., M.D., Johns Hopkins University, Baltimore, Maryland
Fausto, C.-J., B.S.+M.S., University of Southern California, Los Angeles
Gonilski Pacin, D., M.S., IBioBA, Caba, Argentina
Guptan, A., M.S., Children's National Medical Center, Washington, D.C.
Hull, B., B.S., Princeton University, New Jersey
Lee, C-Y.J., B.S., Yale University, New Haven, Connecticut
Liu, J., Rutgers University, Piscataway, New Jersey
Mawson, T., M.D., Columbia University Irving Medical Center, New York, New York
Mirizio, G., M.S., Cincinnati Children's Hospital, Ohio
Morlot, L., M.S., University of Copenhagen, Denmark
Persad, G., B.S., University of Toronto, Ontario, Canada
Phelan, K., B.S., Cincinnati Children's Hospital Medical Center, Ohio
Ripin, N., Ph.D., University of Colorado, Boulder
Sherman, J., Ph.D., Nationwide Children's Hospital, Columbus, Ohio
Siantoputri, M.E., B.Eng., Rockefeller University, New York, New York
Smith, L., B.S., University of Florida, Gainesville
Zein, J., Ph.D., Cleveland Clinic, Ohio

## SEMINARS

Bussemaker, H., Columbia University, New York New York: Motif discovery: dissecting *cis*-regulatory logic with deep sequencing and biophysically interpretable machine learning; Mapping inferred transcription factor activity as a quantitative genetic trait.
Carey, V., Harvard Medical School, Boston, Massachusetts: Probability and simulation; EDA and visualization; Inference, p-values; Overview of Bioconductor; Linear regression and GLMs.
Collado Torres, L., Lieber Institute for Brain Development, Baltimore, Maryland: (Material based on https://lcolladotor.github.io/rnaseq_LCG-UNAM_2023/index.html#course-schedule and https://lmweber.org/OSTA-book/) Visualization of expression of data with iSEE + mean/centering heatmaps; Downloading RNA-seq data using recount3; Statistical models with ExploreModeMatrix; Recount3 data + hands-on exercise overview; Downloading public Vosium spatial data with spatialLIBD + visualizing it with (shiny app + R ufunctions); SPEAQeasy/smokingMouse study introductions; Introduction to the world of spatially resolved transcriptomics with Visium Data (SingleCellExperiment, SpatialExperiment, spatialLIBD); SpatialLIBD: re-use of bulk RNA-seq methods for spatial data (pseudobulk + PCA) + hands-on exercise overview; Hands-on with making plots with pseudobulked spatial data.
Davis, S., University of Colorado Anschutz School of Medicine, Denver: Introduction to R; *t*-statistic and basic power calculations; Public data and data access; Rmarkdown; Summarized experiments; ATAC-seq; Machine learning; ML or ATAC.
DeMeo, D., Brigham and Women's Hospital/Harvard Medical School, Boston, Massachusetts: Network medicine through a sex and gender lens.
Geistlinger, L., Harvard Medical School, Boston, Massachusetts: Functional enrichment analysis of high-throughput omics data.

Gonzalez-Padilla, D., Lieber Institute for Brain Development, Baltimore, Maryland: SPEAQeasy/smokingMouse study introductions; Overview of smokingMouse plots (colData + some genes) + hands-on exercise overview; Choosing variables for your DE analysis with variancePartition

Rube, H., University of California, Merced: Dissecting *cis*-regulatory logic with deep sequencing and biophysically interpretable machine learning.

Zhang, M., University of California, Irvine: Statistical methods for genome-wide association studies; Causal inference of genome-wide gene regulatory networks.

# Advanced Techniques in Molecular Neuroscience

**June 30–July 15**

INSTRUCTORS

**C. Lai**, Indiana University, Bloomington
**J. LoTurco**, University of Connecticut, Storrs
**A. Schaefer**, Icahn School of Medicine at Mount Sinai and MPI, New York, New York

COURSE TEACHING ASSISTANTS

**A. Battison**, Yale University, New Haven, Connecticut
**M. Challman**, Icahn School of Medicine at Mount Sinai, New York, New York
**J. Crowley**, Icahn School of Medicine at Mount Sinai, New York, New York
**E. Hays**, Icahn School of Medicine at Mount Sinai, New York, New York
**P. Hwang**, Icahn School of Medicine at Mount Sinai, New York, New York
**N. Khatri**, Indiana University, Bloomington
**N. Lin**, Indiana University, Bloomington
**P. Ruksee**, Indiana University, Bloomington

The course curriculum was divided into three sections: an extensive and up-to-date set of laboratory exercises, daily lectures covering the theoretical and practical aspects of the various methods used in the laboratory, and a series of evening research seminars.

The informal and interactive evening lectures were given by leading molecular neuroscientists and served to illustrate the ways in which the various experimental approaches have been used to advance specific areas of neurobiology.

In this year's course, the laboratory portion included the topics:

- Assessment of open chromatin in specified neural populations by ATAC-seq
- Assays of mRNA expression in specified neural subtypes by TRAP
- Introduction to overall strategies, use, and design of BAC transgenic vectors
- Methods for preparing purified cell cultures of defined neuronal and glial cells
- Preparation methods, isolation, and analysis of proteins for proteomics

- Preparation and stereotaxic delivery of lentiviral vectors
- Identification of nuclear factor binding sites in brain by Cut & Run
- Tissue clearing
- CRISPR genome editing and construction of CRISPR vectors for targeting genes in neurons and glia
- Gene delivery to neural cells in vivo by viral infection and plasmid electroporation techniques.

This course was supported in part by a grant from the National Institute of Mental Health, part of the National Institutes of Health. Additional scholarship support was provided by the Helmsley Charitable Trust, the Howard Hughes Medical Institute, and Regeneron.

## PARTICIPANTS

Argañaraz, C.A., M.S., IFIBYNE/CONICET-UBA, Buenos Aires, Argentina
Clarin, J., B.S., Drexel University College of Medicine, Philadelphia, Pennsylvania
Colón Ortiz, C., Ph.D., Duke University, Durham, North Carolina
Dernst, A., M.Sc., University Clinic Bonn, Germany
Jenks, J., B.S., Oregon Health and Science University, Portland
Lappalainen, A., B.S., Umeå University, Sweden
Lopez de Boer, R., B.S., Case Western Reserve University, Cleveland, Ohio
Ly, T., Ph.D., University of California, San Francisco
Medlock-Lanier, T., B.S., University of Georgia, Athens
Mills III, W., Ph.D., University of Virginia, Charlottesville
Mukherjee, U., B.S., University of Iowa, Iowa City
Murphy, F., M.Sc., Center for Neurogenomics and Cognitive Research, Amsterdam, Netherlands
Nair, K., Ph.D., Stanford University, Palo Alto, California
Nelson, N., B.S., University of California, Irvine
Padovano, V., M.S., Sapienza Università di Roma, Italy
Ruiz, D., Ph.D., Emory University, Atlanta, Georgia

## SEMINARS

Ayata, P., CUNY Advanced Science Research Center, New York, New York: TRAP.
Battison, A., Yale University, New Haven, Connecticut: Proximity labeling proteomics.
Blanchard, J., Icahn School of Medicine at Mount Sinai, New York, New York: Dissecting the pathological complexities of Alzheimer's disease with in vitro models of the human brain.
Chang, Y.J., University of California, San Francisco; Turrero Garcia, M., University of California, San Francisco: Single-cell RNA-seq and RNAscope.
Eberwine, J., University of Pennsylvania, Philadelphia: Dynamics of single-cell single-organelle 'omics and emergent behaviors.
Haas, K., University of British Columbia, Vancouver, Canada: Investigating relationship of genetic variation in disease using a nongenetic model system.
Hogg, P., University of British Columba, Vancouver, Canada: Single-cell electroporation.
Kenny, P., Icahn School of Medicine at Mount Sinai, New York, New York: A septal circuit that links environmental threats to defensive strategies.
Lai, C., Indiana University, Bloomington: BAC recombineering.
Liddelow, S., NYU Grossman School of Medicine, New York: Purification and culture of CNS glia.
LoTurco, J., University of Connecticut, Storrs: CRISPR and gRNA design and synthesis.
Marshall, K., Baylor College of Medicine, Waco, Texas: Under pressure: the role of PIEZO ion channels in interoception.
Nowakowski, T., University of California, San Francisco (virtual presentation): Uncovering genetic, cellular, and intercellular strategies of human brain development.
Rufen-Blanchette, U., NYU Langone Medical Center, New York: Purification of rat astrocytes by immunopanning.
Sakaki, K., Scientifica Ltd., Uckfield, United Kingdom: Advanced imaging modalities in biology (multiphoton).
Schafer, D., UMass Chan Medical School, Worcester: Studying microglia function and dysfunction within neural circuits.
Schmidt, E., Rockefeller University, New York, New York: Molecular phenotyping of discrete cell types using TRAP.
Silver, D., Duke University Medical Center, Durham, North Carolina: Techniques in brain development.
Tollkuhn, J., Cold Spring Harbor Laboratory: Application of molecular genomic approaches to understand brain sex differences.

Van Aelst, L., Cold Spring Harbor Laboratory: Chandelier cells.

Vogt, D., Michigan State University, Ann Arbor: Monogenic syndrome advantages to uncover complicated molecular and cellular phenotypes of complex disorders; RAS/MAPK mutations converge on common GABAergic neuron programs.

Williams, W., Michigan State University, East Lansing: Probing the genetics of synaptic connectivity by viral vectors.

Witkowski, J., Cold Spring Harbor Laboratory: Ethics in science research.

Zador, A., Cold Spring Harbor Laboratory: Molecular connectomics and in situ sequencing.

# Single-Cell Analysis

**June 30–July 15**

| | |
|---|---|
| INSTRUCTORS | **D. Chenoweth**, University of Pennsylvania, Merion Station<br>**L. Martelotto**, University of Adelaide, Seacliff, South Australia, Australia<br>**M. McConnell**, Lieber Institute for Brain Development, Charlottesville, Virginia<br>**J. Vlassakis**, Rice University, Houston, Texas |
| CO-INSTRUCTOR | **H. Zhang**, Carnegie Mellon University, Pittsburgh, Pennsylvania |
| COURSE TEACHING ASSISTANTS | **M. Chatterjee**, Harvard Medical School, Boston, Massachusetts<br>**C. Deng**, University of Pennsylvania, Philadelphia<br>**N. Lopez**, University of California, San Diego, La Jolla<br>**D. Schaff**, University of Pennsylvania, Philadelphia<br>**A. Valejo**, University of Adelaide, Southampton, United Kingdom<br>**R. Zhao**, Carnegie Mellon University, Pittsburgh, Pennsylvania |
| COMPUTATIONAL TEACHING ASSISTANTS | **Trent Gomberg**, University of California, San Diego, La Jolla<br>**Adam Klie**, University of California, San Diego, La Jolla<br>**Nicole A Lopez**, University of California, San Diego, La Jolla<br>**Deepak Pant**, University of California, San Diego, La Jolla<br>**Brian Yee**, University of California, San Diego, La Jolla (attended virtually) |

The goal of this two-week course was to teach students cutting-edge wet-lab approaches for single-cell analysis, and to provide familiarity with basic bioinformatic approaches to single-cell data. The modules of the course were taught by scientists with expertise in distinct areas of single-cell analysis.

The topics discussed were:

- Droplet- and microwell-based single-cell isolation
- SplitSeq device-free single-cell RNA library preparation
- Single-molecule FISH
- Single-cell western blot
- Photoactivatable single-cell probes
- Single-cell mass spectrometry
- Introductory single-cell sequencing analysis

The course was supported in part by grants from the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, part of the National Institutes of Health. Additional scholarship support was provided by the Helmsley Charitable Trust, the Howard Hughes Medical Institute, and Regeneron.

## PARTICIPANTS

Albao, D., B.S., M.S., The Scripps Research Institute, Jupiter, Florida
Alvarez, R., Ph.D., Icahn School of Medicine at Mount Sinai Hospital, New York, New York
Anosike, N., B.S., University of Calgary, Calgary, Alberta, Canada
Antone, J., B.A., B.S., Translational Genomics Research Institute, Phoenix, Arizona
Bainbridge, R., Ph.D., National Institute of Environmental Health Science, Research Triangle Park, North Carolina
Bassal, M., Ph.D., Beth Israel Deaconess Medical Center, Boston, Massachusetts
Brachman, R., Ph.D., Cornell Tech (Cornell University), New York, New York
Ciobanu, M., Ph.D., Paris-Saclay University, CNRS, Gif-sur-Yvette, France
Dash, S., Ph.D., Stowers Institute for Medical Research, Kansas City, Missouri
Eakins, B., B.S., University of California, Berkeley
Fernandez, S., Ph.D., New York Genome Center, New York, Spain
Fish, M., B.S., Boston University, Massachusetts
Mandujano, E., Ph.D., University of Calgary, Alberta, Canada
Quesada, L., B.S., University of Virginia, Charlottesville
Sayed Issa, I., B.S., Yale University, New Haven, Connecticut
Srivastava, A., New York Genome Center, Jersey City, New Jersey
Swarovski, M., B.S., University of Utah, Salt Lake City
Wootan, C., B.S., University of Minnesota, St. Paul
Zambrano, M., B.S., University of Illinois Urbana-Champaign, Urbana
Zolkin, K., B.S., Stanford University, California

## SEMINARS

Ahmed, N., University of California, San Diego: Next-generation spatial transcriptomics with subcellular phenotyping.
Allbritton, N., University of Washington, Seattle: Separating single cells and the contents of single cells.
Chenoweth, D., University of Pennsylvania, Philadelphia: Single-cell photoactivatable probes for protein localization.
Eberwine, J., University of Pennsylvania, Philadelphia: Organization of emergent properties by single organelles in single cells
Fan, J., Johns Hopkins University, Baltimore, Maryland: Single-cell spatial transcriptomics data analysis.
Gebhart, S., Cell Microsystems, Inc., Durham, North Carolina: Fluidic-free isolation of single cells using CellRaft Technology.
Lam, T., University of California, Berkeley; Overton, M., University of California, Berkeley: Single-cell western blot.
Lareau, L., University of California, Berkeley: Alternative splicing in single cells.
Mah, C., University of California, San Diego: Spatial transcriptomics analyses.
Meltzer, R., Fluent BioSciences, Watertown, Massachusetts: PIPseq: Single-cell biology made simple.
Roco, C., Parse Biosciences, Seattle, Washington.
Romanova, E., University of Illinois at Urbana-Champaign, Urbana: Principles of mass spectrometry data analysis.
Rubakhin, S., University of Illinois at Urbana-Champaign, Urbana: Multiple lectures during four-day period.
Sims, P., Columbia University, New York, New York: Scalable co-sequencing of single-cell genomes and transcriptomes.
Slavov, N., Northeastern University, Boston, Massachusetts and Broad Institute, Cambridge, Massachusetts: Driving

biology with single-cell proteomics: new data acquisition and interpretation methodologies.

Tan, L., Stanford University, California: Probing the single-cell 3D genome architectural basis of neurodevelopment and aging in vivo.

Valejo, A., University of Adelaide, United Kingdom: Multiome Bioinformatics Module.

Vlassakis, J., Rice University, Houston, Texas: Single-cell analysis of protein complexes in pediatric cancers.

# *Drosophila* Neurobiology: Genes, Circuits, and Behavior

**June 30–July 20**


INSTRUCTORS

**R. Carrillo**, University of Chicago, Illinois
**T. Mosca**, Thomas Jefferson University, Philadelphia, Pennsylvania
**T. Reis**, University of Colorado, Anschutz Medical Campus, Aurora

COURSE TEACHING ASSISTANTS

**K. Davis**, Thomas Jefferson University, Philadelphia, Pennsylvania
**J. Humenik**, Thomas Jefferson University, Philadelphia, Pennsylvania
**R. Salazar**, University of Chicago, Illinois


This laboratory/lecture course was intended for researchers at all levels from beginning graduate students through established primary investigators who want to use *Drosophila* as an experimental system for nervous system investigation. This three-week course was designed to introduce students to a wide variety of topics and techniques, including the latest approaches for studying nervous system development, activity, and connectivity, as well as complex behaviors and disease models.

Daily research seminars presented comprehensive overviews of specific subfields of nervous system development or function or focused on state-of-the-art techniques and approaches in *Drosophila* neuroscience. Expert guest lecturers discussed their findings and approaches and brought along their own assays and techniques for students to learn in the laboratory part of the course.

The hands-on portion of the course was centered on inquiry-based projects, utilizing the different morphological and physiological measurements and behavioral paradigms learned at the course. This included molecular-genetic analyses, immunocytochemistry, recording of

activity using electrophysiology and genetically encoded calcium indicators, optogenetic and thermogenetic control of neural activity, and numerous quantitative behavioral measures.

Collectively, the course provided a comprehensive and practical introduction to modern experimental methods for studying the neural basis of behavior in *Drosophila*.

This course was supported in part by a grant from the National Science Foundation. Additional scholarship support was provided by the Helmsley Charitable Trust, the Howard Hughes Medical Institute, and Regeneron.

## PARTICIPANTS

Blaum, N., M.Sc., University of Copenhagen, Denmark
Delgado-Seo, H., B.S., Baylor College of Medicine, Houston, Texas
Droste, E., B.S., NC State University, Raleigh, North Carolina
Falt, T., M.Sc., Max Planck Institute for Biological Intelligence, Martinsried, Germany
Huang, T.-T., Ph.D., Tohoku University, Sendai, Japan
Iyer, A., Ph.D., Barnard College, Columbia University, New York, NY
Jana Joha, Ph.D., University of Oxford, United Kingdom
Kan, S., B.S., Duke University, Durham, North Carolina
Leier, H., B.A., Case Western Reserve University School of Medicine, Cleveland, Ohio
Li, K., B.S., Stowers Institute for Medical Research, Kansas City, Missouri
Mudunuri, A., B.S.-M.S., University of Konstanz, Germany
Nunn, O., B.S., Vanderbilt University, Nashville, Tennessee
Shah, P., Ph.D., Georgetown University, Washington, D.C.
Wilson, A., D.Phil., University College London, United Kingdom

## SEMINARS

Sarah Ackerman, S., Washington University School of Medicine in St. Louis, Missouri: Glial control of neural circuit structure and function.
Ashley, J., University of Chicago, Illinois: Building and maintaining a larval NMJ; NMJ and larval development; "Applied" NMJ metrics and analysis.
Bhattacharya, M., University of Arizona, Tucson: Nerve injury modeling in *Drosophila*.
Borjon, L., Indiana University, Bloomington: Nociception/larval rolling.
Carrillo, R., University of Chicago, Illinois: Embryo dissections/larval brain pulls; NMJ and larval development; Boot camp: anatomy, genetics, fly rigor; "Applied" NMJ metrics and analysis; NMJ physiology; Physiology data analysis.
Frank, C.A., University of Iowa, Iowa City: *Drosophila* NMJ physiology; Physiology data analysis.
Hassan, B., ICM-Paris Brain Institute, France: Specificity and its discontents: time and variation in brain and behavior.
Heckscher, E., University of Chicago, Illinois: Neural stem cells.
Kayser, M., University of Pennsylvania School of Medicine, Philadelphia: Building brains in our sleep: the regulation and function of sleep during development.
Kim, S.S., University of California, Santa Barbara: Two-photon microscopy for *Drosophila* neurobiology.
Lee, W.-C., Harvard Medical School, Boston, Massachusetts: *Drosophila* connectomics.
Lembke, K., University of Iowa, Iowa City: Navigating teaching and *Drosophila* research at primarily undergraduate institutions.
Loesche, F., Janelia HHMI, Ashburn, Virginia: Vision in flies.
Louis, M., University of California, Santa Barbara, CA Olfactory navigation in the *Drosophila* larva.
Craig Montell, University of California, Santa Barbara: Receptors, channels, and insect behavior.
Mosca, T., Thomas Jefferson University, Philadelphia, Pennsylvania: Microscopy; NMJ and larval development; Boot camp: anatomy, genetics, fly rigor; "Applied" NMJ metrics and analysis.
Muraro, N., Biomedicine Research Institute of Buenos Aires, Argentina: Ephys in the *Drosophila* CNS; CNS physiology II.
Palavicino-Maggio, C., Harvard Medical School, Boston, Massachusetts: Aggression and sex-specific behaviors.
Powell, A., University of Arizona, Tucson: Axon injury models in *Drosophila*.
Ramdya, P., EPFL, Lausanne, Switzerland: Approaches for reverse engineering *Drosophila* action selection and limb control.
Reis, T., University of Colorado, Anschutz Medical Campus, Aurora: Neuronal control of energy balance; Boot camp: anatomy, genetics, fly rigor.
Reiser, M., HHMI; Janelia Research Campus, Virginia: From structure to function in the fly visual system.
Rister, J., University of Massachusetts, Boston: Vitamin A deprivation as a novel approach to identify neuroprotective molecules.

Rodal, A., Brandeis University, Boston, Massachusetts: Routing and remodeling membranes at the synapse.
Tadres, D., Stanford University, California: Chemotaxis in the *Drosophila* larva.
Thompson-Peer, K., University of California, Irvine: Neuronal cell biology.
Tracey, D., Indiana University, Bloomington: *Drosophila* nociception.
Vasconcelos, M.L., Champalimaud Foundation, Lisbon, Portugal: Innate behavior.
Cale Whitworth, C., Indiana University, Bloomington: Resources of the fly.

# Frontiers and Techniques in Plant Science

**June 30–July 20**

INSTRUCTORS

**S. Brady**, University of California, Davis
**N. Geldner**, University of Lausanne, Switzerland
**J. Law**, Salk Institute for Biological Studies, La Jolla, California

COURSE TEACHING ASSISTANTS

**T. Dunivant**, University of California, Riverside
**G. Ferreras Garrucho**, University of Cambridge, United Kingdom
**J. Menconi**, Sant'Anna School of Advanced Studies, Pisa Pi, Italy

The Frontiers and Techniques in Plant Science course provided an intensive overview of topics in genomics, genetics, physiology, biochemistry, development, and evolution and hands-on experiences in molecular, imaging, computational, and high-throughput approaches to understanding plant biology. It emphasized recent results from model organisms including *Arabidopsis* and maize as well as a variety of other plants and provided an introduction to current methods used in basic and applied plant biology, both theoretically and practically.

The seminar series included plant morphology and anatomy, development, evolution, light and mechanical biology, hormones, small RNAs and epigenetic inheritance, biotic and abiotic interactions, plant biochemistry, crop domestication, and applications addressing current

agronomic problems. Speakers provided expert overviews of their fields, followed by in-depth discussions of their own work.

The laboratory sessions provided exposure to cutting-edge experimental and computational techniques currently used in plant research. These included approaches for studying plant development, regulatory networks, transient gene expression, cell type–specific gene expression analysis, computational large-scale data analysis, applications of fluorescent proteins including live imaging, genome editing, and chromatin immunoprecipitation.

The students gained hands-on experience in:

- Computational tools and programming
- Plant imaging and image analysis
- Synthetic biology
- Transcriptomics
- Single-cell sequencing
- Gene network analysis and data visualization tools
- Mathematical modeling of development and hormone action
- Purification of cell type–specific nuclei (INTACT)
- Genome editing
- Scientific publishing.

The course was supported in part by a grant from the National Science Foundation. Additional scholarship support was provided by the Helmsley Charitable Trust, the Howard Hughes Medical Institute, and Regeneron.

## PARTICIPANTS

Aardening, Z., M.Sc., Technion–Israel Institute of Technology, Haifa
Baruah, S., M.Sc., University of Hyderabad, India
Cavallini, E., M.S., University of Geneva, Switzerland
Elhazzime, K., Ph.D., OEB Research Program, Helsinki, Helsinki, Finland
LaStella, S., B.A., University of Texas at Austin
Le, D., B.S., University of California, Riverside
Min, J., B.S., University of Washington, Seattle
Morimoto, K., B.S., University of California, Davis
Nicolet, J., M.S., University of Lausanne, Switzerland
Okinedo, U., B.S., University of Massachusetts, Boston
Singh, D., M.Sc., Institute of Plant Genetics (IGR, PAN), Poznan, Poland
Steinberger, A., B.A., University of Minnesota, Saint Paul
Strout, R., B.A., University of California, Riverside
Wootan, C., B.S., University of Minnesota, St. Paul
Zambrano, M., B.S., University of Illinois at Urbana-Champaign, Urbana
Zolkin, K., B.S., Stanford University, California

## SEMINARS

Bailey-Serres, J., University of California, Riverside: INTACT lab and plant abiotic stress.
Bartlett, M., University of Massachusetts, Amherst: The development and evolution of grass flowers.
Bayer, E., CNRS/University of Bordeaux, France: Electron tomography to study plasmodesmata.
Brady, S., University of California, Davis: Exodermis differentiation and function.
Brophy, J., Stanford University, California: Engineering features outside the reach of evolution: an overview of plant synthetic biology.
Buell, C.R., University of Georgia, Athens: Genome-enabled breeding.
Cody, J., University of Minnesota, Minneapolis: Gene editing lab module.
Cutler, S., University of California, Riverside: Plant hormone sensors: molecular glues, allosteric switches, and scaffolds for protein engineering.
Deal, R., Emory University, Atlanta, Georgia: Applications of the INTACT method for cell type–specific profiling.
Dolan, L., Gregor Mendel Institute, Vienna, Austria: Plant evolution.

Dong, X., Duke University/HHMI, Durham, North Carolina: Plant Immune mechanisms.
Fontanez, K., Fluent BioSciences, Watertown, Massachusetts: PIPseq: Single-cell biology made simple.
Geldner, N., University of Lausanne, Switzerland: The building of diffusion barriers in animals and plants.
Husbands, A., University of Pennsylvania, Philadelphia: Proteins and their ligands.
Jackson, D., Cold Spring Harbor Laboratory: Shoot apical meristem development.
Jones, A., University of Cambridge, United Kingdom: Bringing cellular hormone dynamics in focus using fluorescent biosensors.
Kessler, S., Purdue University, West Lafayette, Indiana: Gametophyte interactions during plant reproduction.
Law, J., Salk Institute for Biological Studies, La Jolla, California: Chromatin, epigenetics, and DNA methylation.
Mendoza Cozatl, D., University of Missouri, Columbia: Plant nutrition: the daily struggle of sufficiency and toxicity of essential elements.
Paszkowski, U., University of Cambridge, United Kingdom: The art and design of harmony: molecular genetics of arbuscular mycorrhizal symbiosis in cereals.
Pedmale, U., Cold Spring Harbor Laboratory: Plant photosensory system and signaling.
Provart, N., University of Toronto, Ontario, Canada: Raising the BAR for hypothesis generation in plant biology using open big data.
Sinha, N., University of California, Davis: Plant structure and function.
Surridge, C., Springer Nature, London, United Kingdom: Publishing without tears.
Ursache, R., Centre for Research in Agricultural Genomics (CRAG), Barcelona, Spain: Imaging plant cell walls: the beauty is in the details.
Voytas, D., University of Minnesota, Minneapolis: Precise genome engineering with sequence-specific nucleases.
Williams, J., Cold Spring Harbor Laboratory: Computational skills for plant sciences.

# Neural Data Science

**July 11–24**


INSTRUCTORS **M. Reimers**, Michigan State University, East Lansing
**P. Wallisch**, New York University, New York

CO-INSTRUCTORS **H. Ben-Esti**, Technion, Haifa, Israel
**J. Sun**, University College London, United Kingdom

SCHOLAR-IN-RESIDENCE **M.X. Cohen**, Donders Centre for Medical Neuroscience, Radboud University, the Netherlands

COURSE TEACHING ASSISTANTS **A. Lebedeva**, University College London, United Kingdom
**E. Marachlian**, ENS Paris, France


Today's technologies enable neuroscientists to gather data in previously unimagined quantities. This necessitates—and allows for—the development of new analysis methods to address dynamic systems function of brain networks.

This course was designed to help neuroscience practitioners develop the conceptual and practical capabilities to meet the challenges posed by the analysis of these hard-won and large data sets. We emphasized statistical issues such as the preprocessing of data, sampling biases, estimation methods, and hypothesis testing, as well as data wrangling (in MATLAB and Python). We worked with data from a variety of recording technologies including multi-electrode array recordings, local field potentials, and EEG, as well as two-photon and wide-field optical imaging.

The course gave a solid conceptual and technical grounding in widely applicable methods such as:

- Data processing for each recording technique
- Spectral methods
- Neural population analysis
- Behavioral analysis
- How to integrate neural data with behavioral data.

The workshop proceeded in a seminar style, guided by leading neural data analysts, with demonstrations and practical lab data analysis exercises supervised by instructors.

The course was supported in part by grants provided by the Helmsley Charitable Trust, the Howard Hughes Medical Institute, and Regeneron.

## PARTICIPANTS

Butola, T., Ph.D., NYU Grossman School of Medicine, New York
Curreli, S., Ph.D., Fondazione Istituto Italiano di Tecnologia, Genova, Italy
Deane, K., B.S., University California, Riverside
Fekos, C., B.Sc., University of Sussex, Brighton, United Kingdom
Fernandes Henriques, C., M.Sc., City University of New York, New York
Gómez, L., Ph.D., Johns Hopkins University, Baltimore, Maryland
Gonzales-Hess, N., M.S., University of Oregon, Eugene
Jia, C., B.S., Salk Institute, La Jolla, California
Li, M., B.S., RIKEN, the Institute of Physical and Chemical Research, Wako, Japan
Martinez de Paz, J.M., M.Sc., MPI Neurobiology, Planegg, Germany
Natesan, D., Ph.D., University of California, Santa Barbara
Nietz, A., Ph.D., University of Minnesota–Twin Cities, Minneapolis
Ning, W., B.A., University of California, Irvine
Philipsberg, P., B.E., Icahn School of Medicine at Mount Sinai, New York, New York
Rakymzhan, A., M.S., University of Pittsburgh, Pennsylvania
Rodriguez Diaz, J., B.S., University of Michigan, Ann Arbor
Rogers, S., B.S., University of Pennsylvania, Philadelphia
Sandoval Ortega, R.A., B.S. + M.S., University of Bern, Switzerland
Voigt, F., Ph.D., Harvard University, Cambridge, Massachusetts
Walder, K., Ph.D., Duke University, Durham, North Carolina
Wu, M., B.S., University of California, Los Angeles
Xiong, Y., B.S., B.A., Vanderbilt University, Nashville, Tennessee
Yin, C., B.S., University of California, Los Angeles

## SEMINARS

Benisty, H., Technion, Haifa, Israel: Dynamics of correlations: why, how, and what is it good for.
Engel, T., Princeton University, New Jersey: A unifying perspective on neural manifolds and circuits for cognition.
Kennedy, A., Northwestern University, Chicago, Illinois: Dissecting the neural population dynamics underlying motivated behavior.
Pachitariu, M., HHMI Janelia, Ashburn, Virginia: Data science with large-scale recordings.
Sun, Y.J., University College London, United Kingdom: Probing visual cortical plasticity with two-photon imaging.
Williams, A., New York University/Flatiron Institute, New York: Unsupervised analysis methods of large-scale neural data.
Yatsenko, D., DataJoint, Houston, Texas: Building shared data workflows.
Yu, B., Carnegie Mellon University, Pittsburgh, Pennsylvania: Neuronal population interactions between brain areas.

# Autism Spectrum Disorders

**July 26–August 1**


INSTRUCTORS

**D. Fallin**, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland
**J. McPartland**, Yale University, New Haven, Connecticut
**S. Pasca**, Stanford University, California
**J. Veenstra-Vander Weele**, Columbia University, New York, New York

COURSE TEACHING ASSISTANT

**L. Grosvenor**, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland


Autism spectrum disorders (ASDs) are developmental disorders with complex phenotypes defined by a triad of symptoms that include disrupted social abilities, verbal and nonverbal communication skills, and restricted interests with repetitive behaviors. Co-occurring neurological and medical conditions often characterize this disorder. The underlying etiology remains a mystery, but ASD is one of the most highly heritable of neuropsychiatric disorders.

This course examined dimensions of ASDs on various levels, including sessions on characteristics of the clinical syndrome; the neuropathology, imaging, and cognitive neuroscience studies that implicate circuits and systems involved in ASD; the current state of findings from human genetics; concepts regarding the developmental neurobiological basis; the use of experimental models; and current etiological theories and hypotheses of ASDs.

In addition to learning about the most recent research in these areas, we explored and debated controversial topics and challenges of basic assumptions in the field. An exceptional faculty with diverse interests brought the most up-to-date results and theories to the students, making this workshop a valuable resource for young researchers starting out in this fast-moving and expansive field. Not only will it help them build the foundation for their future research, it also introduced them to many potential collaborators working to understand ASD from different disciplines.

The course had hands-on exercises to complement the featured intense lecture sessions. Most importantly, students had free time for reading, informal discussions, and recreation on the beautiful campus of the Banbury Center, which includes a beach, a pool, and tennis courts.

The course was supported with funds provided by the Helmsley Charitable Trust, the Howard Hughes Medical Institute, and Regeneron.

## PARTICIPANTS

Balafkan, N., Ph.D., Yale University, New Haven, Connecticut
Cattel, S., B.A., Rutgers University, New Brunswick, Edison, New Jersey
Chhabra, S., Masters, Johannes Gutenberg University, Mainz, Germany
Damilou, A., M.S., University of Zürich/ETH, Switzerland
Eyoh, E., B.E., University of Minnesota, Minneapolis
Ferguson, A., B.S., University of North Carolina at Chapel Hill, Carrboro
Fram, N., Ph.D., Vanderbilt University Medical Center, Nashville, Tennessee
Franklin, A., B.Sc., University of Exeter, United Kingdom
Gaenswein, T., B.S., M.S., ETH Zürich, Basel, Switzerland
Gorkin, D., Ph.D., Emory University, Atlanta, Georgia
Gungor Aydin, A., Ph.D., Rutgers University, Piscataway, New Jersey
Isenstein, E., Ph.D., University of Rochester, New York
Jankovic, M., B.S., UT Southwestern, Dallas, Texas
Kim, K.H., B.A., Seoul National University, South Korea
Klibaite, U., Ph.D., Harvard University, Cambridge, Massachusetts
Koirala, S., B.A., University of Minnesota, Minneapolis
Kozlova, E., B.S., University of California, Riverside
Manessis, A., B.A., Teachers College, Columbia University, New York, New York
McGarry, L., Ph.D., Children's Hospital of Philadelphia/ University of Pennsylvania
Nardi, L., M.D., University Medical Center, Mainz, Germany
Phillips, H., Ph.D., Harvard University, Cambridge, Massachusetts
Schroder, A., B.A., Northwestern University, Chicago, Illinois
Seczon, D., B.S., University of Washington, Seattle
Tornini, V., Ph.D., Yale University, New Haven, Connecticut
Turkalj, L., M.D., Rutgers University, Piscataway, New Jersey

## SEMINARS

Allen, N., Salk Institute for Biological Science, La Jolla, California: Glial cells in neurodevelopmental disorders.
Crawley, J., MIND Institute, University of California, Davis: Mouse behavioral assays relevant to the symptoms of autism spectrum disorder; Animal models, part 2.
Elison, J., University of Minnesota, Minneapolis: Neuroimaging methods and findings, part 1.
Fallin, D., Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland: Epidemiology—descriptive; Epidemiology—analytic.
Geschwind, D., David Geffen School of Medicine, University of California, Los Angeles: Introduction to human genetic principles and ASD genetics; Genetics and functional genomics of ASD.
Kanne, S., Weill Cornell Medical College, New York, New York: The history and clinical presentation of autism.
Kasari, C., University of California, Los Angeles: Outcome measures and early intervention.
Littman, D., New York University, New York: Neuroimmunology of ASD.
Lasalle, J., University of California, Davis: Epigenetics of ASD.
Martin, K., Simons Foundation Autism Research Initiative, New York, New York: Funding and organizing innovative research.
McPartland, J., Yale University, New Haven, Connecticut: EEG; Electrophysiology methods and findings.
Pasca, S., Stanford University, California: Human cell models.
Sahin, M., Boston Children's Hospital, Massachusetts: Translational research in rare genetic disorders.
Schumann, C., University of California, Davis, MIND Institute (attended virtually): Trajectory of brain aging across the lifespan in ASD.
Shah, N., Stanford University, California: Neural circuits for social behaviors. New York, New York: Family perspectives on autism.
Spence, S., Boston Children's Hospital/Harvard University, Massachusetts (attended virtually): Medical comorbidities in autism spectrum disorder.
Veenstra-Vander Weele, J., Columbia University, New York, New York: Pharmacological interventions.
Walsh, C., Boston Children's Hospital, Harvard Medical School, Howard Hughes Medical Institute, Massachusetts (attended virtually): Somatic mosaicism in human neuropsychiatric disease.

# Synthetic Biology

**July 26–August 8**

INSTRUCTORS

**K. Haynes**, Emory University, Atlanta, Georgia
**V. Noireaux**, University of Minnesota, Minneapolis
**E. Young**, Worcester Polytechnic Institute, Massachusetts

COURSE TEACHING ASSISTANTS

**I. Gispert Contamina**, Imperial College London, United Kingdom
**C. Kamm**, Helmholtz Institute Germany, Würzburg
**A. Khakimzhan**, University of Minnesota, Minneapolis
**E. Nakamura**, University of California, Los Angeles

Cells are the world's most sophisticated chemists, and their ability to adapt to changing environments offers enormous potential for solving modern engineering challenges. Nonetheless, biological systems are noisy, massively interconnected, and nonlinear, and they have not evolved to be easily engineered. The grand challenge of synthetic biology is to reconcile the desire for a predictable, formalized biological design process with the inherent "squishiness" of biology.

This course focused on how the complexity of biological systems can be combined with traditional engineering approaches to result in new design principles for synthetic biology. The centerpiece of the course was an immersive laboratory experience in which students worked in teams to learn the practical and theoretical underpinnings of synthetic biology research. Broadly, the course explored how cellular regulation (transcriptional, translational, posttranslational, and epigenetic) can be used to engineer cells that accomplish well-defined goals.

Laboratory modules covered the following areas:

- Cell-free transcription and translation systems to characterize genetic circuits and RNA regulators
- Modeling gene expression using ordinary differential equations
- DNA assembly and design of expression cassettes
- CRISPR technologies for genome editing and gene regulation.

Students learned essential synthetic biology techniques in a four-day "boot camp" at the beginning of the course. After the boot camp, they rotated through research projects in select areas. Students also interacted closely with a panel of internationally recognized speakers who collectively provided a broad overview of synthetic biology applications, including renewable chemical production and therapeutics, state-of-the-art techniques, case studies in human practices, and socially responsible innovation

This course was supported in part by grants from the National Science Foundation. Additional scholarship support was provided by the Helmsley Charitable Trust, the Howard Hughes Medical Institute, and Regeneron.

## PARTICIPANTS

Dankwa, D., B.Sc., Purdue University, West Lafayette, Indiana
Dykstra, C., B.S., Concordia University, Montréal, Québec, Canada
Fang, D., B.S., University of Colorado at Boulder
Harmer, Z., B.S., University of Wisconsin, Madison
He, H., M.S., University of Pittsburgh, Pennsylvania
Howley, E., Ph.D., University of Arizona, Tucson
Hwa, S.-W., B.S., University of Michigan, Ann Arbor
Lee, K., Ph.D., University of Chicago, Illinois
Maklouf, G., M.S., University of Campinas, Brazil
Manna, S., M.S., Ph.D., Helmholtz Institute of RNA-Based Infection Research, Würzburg, Germany
Marlhens, J., M.S., Technische Universität Darmstadt, Germany
Ness, J., M.S., Boston University, Massachusetts
Robins, K., B.S., Zoetis, Fort Collins, Colorado
Sajtovich, V., B.Sc., Max Planck Institute for Terrestrial Microbiology, Marburg, Germany
Seagrave, S., B.S., University of California, Santa Barbara
Standeven, J., B.A., Georgia Institute of Technology, Atlanta

## SEMINARS

Beisel, C., Helmholtz Institute for RNA-Based Infection Research, Germany: From CRISPR biology to versatile technologies.
Diggans, J., Twist Bioscience, South San Francisco, California: Biosecurity in the age of AI.
Green, L., Purdue University, West Lafayette, Indiana: Engineering the host microbiome for regulating chronic disease.
Keasling, J., University of California, Berkeley/Lawrence Berkeley National Laboratory: Production of supply-limited natural product therapeutics using engineered yeast.
Mayalu, M., Stanford University, California: Control theoretic applications for biomedical therapeutics.
Nakamura, E., University of California, Los Angeles: Hybrid modeling and experiments of gene circuits.
Nugent, R., Rebecca Nugent, California: Access to DNA drives innovation.
Stewart, J., University of California, Los Angeles: Harnessing the molecular code of RNA for structure, function, and applications.
Styczynski, M., Georgia Institute of Technology, Atlanta: Synthetic biology field applications in low-resource environments.
Vickers, C., Queensland University of Technology, Brisbane, Australia: Synthetic biology and metabolic engineering in yeast.

# Chromatin, Epigenetics, and Gene Expression

**July 26–August 15**


INSTRUCTORS

**J. Dowen**, University of North Carolina at Chapel Hill
**M. Guertin**, University of Connecticut, Farmington
**A. Johnson**, University of Colorado School of Medicine, Aurora
**M. Mendillo**, Northwestern University School of Medicine, Chicago, Illinois

COURSE TEACHING ASSISTANTS

**A. Klein**, Northwestern University, Chicago, Illinois
**F. Perez**, UC Anschutz Medical Campus, Aurora, Colorado
**N. Rittenhouse**, University of North Carolina at Chapel Hill
**T. Scott**, University of Virginia School of Medicine, Charlottesville


The Chromatin, Epigenetics, and Gene Expression course was designed for students, postdocs, and principal investigators who have recently ventured into the exciting area of gene regulation. Emphasis was placed on exposing students to a broad array of methodologies to study gene regulation, chromatin structure, and dynamics, including both state-of-the-art and well-developed methods.

Students performed widely used techniques such as:

- Chromatin immunoprecipitation (ChIP)
- ChIP coupled with sequencing (ChIP-seq)
- Reporter assays of enhancer activity
- RNA expression analysis (RT-qPCR, RNA-seq)

- Electrophoretic Mobility Shift Assay (EMSA)
- Chromatin Biochemistry

Students performed ChIP-seq and ATAC-seq and applied a basic pipeline to analyze their genomic results.

Students learned about state-of-the-art genetic perturbation strategies. They performed two of these methods to reduce or eliminate the expression of a gene of interest: RNA interference (RNAi), rapidly inducible degrons (dTAG), and CRISPR-Cas9-targeted disruption. Further, students compared how each method affects gene expression and function.

Students learned how to assemble recombinant chromatin with modified histones and test specificity of epigenetic "reader" proteins and enzymes that modify chromatin. Quantitative methods were used to analyze activity and selectivity for specific substrates.

This course also provided the basic concepts behind different methods to analyze the chromatin architecture of the genome. Moreover, we discussed the computational methods required to analyze data concerning three-dimensional chromatin architecture.

Experience with basic recombinant DNA and molecular biology techniques was a prerequisite for admission to this course. Lectures by the instructors covered the current state of the gene expression and epigenetics fields, theoretical aspects of the methodology, and broader issues regarding strategies for investigating the regulation of gene expression in eukaryotes. Emphasis was placed on advantages and limitations of specific techniques and on data interpretation. Each evening, an invited speaker, an expert in the field, presented their work and interacted with students. The students were encouraged and expected to actively participate in these discussions and to take advantage of the many opportunities to network and receive input on their projects and future plans.

This course was supported in part by grants from the National Cancer Institute, part of the National Institutes of Health. Additional scholarship support was provided by the Helmsley Charitable Trust, the Howard Hughes Medical Institute, and Regeneron.

## PARTICIPANTS

Agoro, R., Ph.D., Indiana University, Indianapolis
Bhatt, A., M.S., Dartmouth College, Hanover, New Hampshire
Brossier, N., M.D., Ph.D., Washington University in St. Louis, Missouri
Bujnowska, M., B.S., University of Virginia, Charlottesville
Cain, T., B.S., St. Jude Graduate School of Biomedical Sciences, Memphis, Tennessee
Carew, F., B.Sc., Humboldt Universität zu Berlin, Germany
Colon-Caraballo, M., Ph.D., UT Southwestern Medical Center, Dallas, Texas
Côté, N., B.S., University of Sherbrooke, Sherbrooke, Quebec, Canada
Gorostieta Salas, E., Ph.D., Salk Institute for Biological Studies, La Jolla, California
Mello, R., B.S., The Scripps Research Institute, La Jolla, California
Rodriguez Teran, E., B.S., Rice University, Houston, Texas
Shah, S., Ph.D., Weill Cornell Medicine, New York, New York
Shim, J., M.D., Emory University, Atlanta, Georgia
Sreelal, T., Int.M.S., University of Illinois Urbana–Champaign, Urbana
Stevens, B., B.S./M.S., St. Jude Children's Research Hospital, Memphis, Tennessee
Young, D., B.A., Albert Einstein College of Medicine, Bronx, New York

## SEMINARS

Adelman, K., Harvard Medical School, Boston, Massachusetts: Gene control.
Berger, S., University of Pennsylvania, Philadelphia.
Franco, H., University of North Carolina at Chapel Hill: Enhancer identification and probing using single-cell genomics.
Hammell, C., Cold Spring Harbor Laboratory: A circadian-like gene network programs the timing and dosage of heterochronic miRNA transcription during *C. elegans* development.
He, Y., Northwestern University, Chicago, Illinois: Structural visualization of chromatin regulatory complexes using cryo-EM.

Joyce, E., University of Pennsylvania, Philadelphia: 3D genome organization and regulation.
Larschan, E., Brown University, Providence, Rhode Island: X marks the spot: targeting dosage compensation to the X chromosome.
Lei, E., National Institutes of Health, Bethesda, Maryland: 3D genome organization.
Levine, M., Princeton University, New Jersey.
Martienssen, R., Cold Spring Harbor Laboratory/HHMI: Chromatin remodeling of histone H3 variants by DDM1 underlies epigenetic inheritance.
Pai, A., University of Massachusetts Chan Medical School, Worcester, Massachusetts: mRNA splicing: endless isoforms most beautiful.
Roeder, R., Rockefeller University, New York, New York: Transcriptional regulatory mechanisms in animal cells.
Shilatifard, A., Northwestern University Feinberg School of Medicine, Chicago, Illinois: Epigenetic moonlighting: histone modifiers' catalytic-independent functions in the regulation of transcription, providing new insight for cancer therapeutics.
Strahl, B., UNC School of Medicine, Chapel Hill, North Carolina: Chromatin mechanisms and nonchromatin functions of histone code writers.
Taatjes, D., University of Colorado, Aurora: Understanding transcription regulation through transcriptomics and biochemical reconstitution.
Tollkuhn, J., Cold Spring Harbor Laboratory: Epigenetic regulation of brain sexual differentiation.
Witkowski, J., Cold Spring Harbor Laboratory: Ethics lecture.
Zeitlinger, J., Stowers Institute for Medical Research, Kansas City, Missouri: Deciphering *cis*-regulatory code for development using interpretable deep learning.

# Imaging Structure and Function in the Nervous System

**July 26–August 15**

| | |
|---|---|
| INSTRUCTORS | **E. Hillman**, Columbia University, New York, New York<br>**R. Portugues**, Technical University, Münich, Germany<br>**P. Tsai**, University of California, San Diego |
| CO-INSTRUCTORS | **A. Charles**, Johns Hopkins University, Baltimore, Maryland<br>**H. Dana**, Cleveland Clinic, Ohio |
| ASSOCIATE INSTRUCTORS | **J. Donovan**, Max Planck Institute of Neurobiology, Planegg, Germany<br>**C. Whitmire**, Max Delbrück Center for Molecular Medicine, Berlin, Germany |
| COURSE TEACHING ASSISTANTS | **W. Best**, Cold Spring Harbor Laboratory<br>**A. Das**, Lerner Research Institute, Cleveland Clinic, Ohio<br>**A. Gruzdeva**, Cornell University, Ithaca, New York<br>**H. Lavian**, Technical University of Münich, Germany<br>**D. Maharjan**, Cold Spring Harbor Laboratory<br>**E. Ozen**, Columbia University, New York, New York<br>**O. Prat**, Technical University of Münich, Germany |

Advances in lasers, light microscopy, advanced data analysis techniques and the development of powerful optical indicators and actuators and model organisms present expanding opportunities for investigating the nervous system, from synaptic spines to networks in the brain. This intensive laboratory and lecture course provided participants with theoretical and practical knowledge to leverage and combine emerging imaging technologies for neuroscience research. The primary focus of the course was on in vivo applications of light microscopy, particularly functional imaging with genetically encoded indicators.

Methods taught included:

- Multiphoton microscopy
- Light-sheet microscopy
- Use of spatial light modulators and digital holography
- Combination of imaging with optogenetics
- Head-mounted microscopes and fiber-optic methods
- Analysis of imaging data sets

Lectures and hands-on laboratory modules overseen by leading experts progressed through basic concepts to cutting-edge imaging methods. Students learned the fundamentals of optics, lasers, spectroscopy, and microscopy; laser scanning systems; camera-based systems; methods for quantifying and optimizing signal to noise and resolution; in vivo preparations in mice and zebrafish larvae; and image processing and analysis approaches. Hands-on building exercises were a fundamental component of the course, enabling students to develop an intuitive understanding of optical principles and assemble their own two-photon microscope. The course also hosted a range of state-of-the-art commercial imaging systems from a range of vendors who actively participated in the course and provided ample opportunities to explore, compare, and gain experience with these advanced systems.

We encouraged applications from diverse interdisciplinary researchers (Ph.D. students, postdoctoral fellows, and early-career faculty or equivalent) seeking to expand their knowledge, skillsets, and experience, including those with expertise in experimental and computational neuroscience and neural engineering. Attendees working on a variety of model organisms—including mice, fish, fly, worm, and organoid systems—were welcome.

This course was supported in part by a grant from the National Institute of Mental Health, part of the National Institutes of Health. Additional scholarship support was provided by the Helmsley Charitable Trust, the Howard Hughes Medical Institute, and Regeneron.

## PARTICIPANTS

Aguilar Pérez, C.G., B.Sc., Max-Planck-Institute for Biological Intelligence, Martinsried Münich, Germany
Alageswaran, S., M.S., McGill University, Montréal, Québec, Canada
Chen, P.S., B.S., University of California, Los Angeles
Deb, D., B.S., Janelia Research Campus, Ashburn, Virginia
Dyson, M., B.Sc., VIB-KU Leuven Center for Brain and Disease Research, Belgium
Kirchgessner, M., Ph.D., NYU School of Medicine, New York, New York
Kwon, J.S., B.S., Seoul National University, South Korea
Mughal, A., Ph.D., University of Vermont, Burlington
Phillips, M., Ph.D., NYU School of Medicine, New York, New York
Raiders, S., Ph.D., Institute of Molecular Pathology, Vienna, Austria
Ramirez Sanchez, L.J., B.S., Cold Spring Harbor Laboratory
Razzauti, J., B.Sc., Rockefeller University, New York, New York
Salvati, K., Ph.D., University of California, San Francisco
Selfe, J., Hons., University of Cape Town, South Africa
Uddin, S.A., B.S., Northwestern University, Evanston, Illinois
Wang, C., B.A., Baylor College of Medicine, Houston, Texas

## SEMINARS

Cai, D., Icahn School of Medicine at Mount Sinai, New York, New York: Developing open source miniscopes and analysis pipelines for in vivo calcium imaging.
Charles, A., Johns Hopkins University, Baltimore, Maryland: Computer set-up and data intro; Consider the data: the computational side of neural imaging: Data after dinner: the signal; Data after dinner: the noise; Data after dinner: analysis pipelines.
Dana, H., Cleveland Clinic Main Campus, Ohio: Recording from freely moving mice; Opto/Grin lens basics.
Dieudonne, S., INSERM, Paris, France: Optical recording and actuation of neuronal voltage in 3D with acousto-optic based two-photon microscopy.
Donovan, J., Max Planck Institute of Neurobiology, Planegg, Germany: Practical optical system design; Holographic photostimulation system design/control.

Hillman, E., Columbia University, New York, New York: Microscopy; Light sheet basics.

Ji, N., University of California, Berkeley: Imaging the brain at high spatiotemporal resolution.

Ki, C., Carnegie Mellon University, Pittsburgh, Pennsylvania: Using suite2p for functional segmentation of neural imaging data.

Lichtman, J., Harvard University, Cambridge, Massachusetts: Confocal lecture.

Murthy, M., Princeton University, New Jersey: The ABCs of *Drosophila* neuroscience: activity, behavior, connectome.

Pesaran, B., California Institute of Technology, Pasadena: Multiregional calcium imaging in the nonhuman primate.

Portugues, R., Technical University, Münich, Germany: 2p basics; Table of resolution.

Soudagar, Y., Bruker Ltd., Milton, Ontario, Canada: Development of optical preclinical CNS imaging systems as product.

Stringer, C., HHMI Janelia Research Campus, Ashburn, Virginia (attended virtually): Functional and anatomical image processing with suite2p and cellpose.

Svoboda, K., Allen Institute, Seattle, Washington: Calcium imaging to spy on neurons in their native habitat: past, present, and future.

Tian, L., University of California, Davis: Imaging dynamics of neurotransmitter release with genetically encoded indicators.

Tsai, P., University of California, San Diego: Basic optics lecture; Intro to Fourier optics; Fourier materials; Confocal lecture; Confocal build.

Vaziri, A., Rockefeller University, New York, New York: Two-photon and light-sheet microscopy.

Waters, J., Harvard Medical School, Boston, Massachusetts: Noise.

# Yeast Genetics and Genomics

**July 26–August 15**


INSTRUCTORS

**G. Brown**, University of Toronto, Ontario, Canada
**M. Dunham**, University of Washington, Seattle
**S. Lacefield**, Geisel School of Medicine at Dartmouth, Hanover, New Hampshire
**G. Lang**, Lehigh University, Bethlehem, Pennsylvania

COURSE TEACHING ASSISTANTS

**J. Armstrong**, University of Washington, Seattle
**M. Ferguson**, University of Toronto, Ontario, Canada
**P. Gadgil**, Dartmouth University, Hanover, New Hampshire
**H.-Y. Jhuang**, Lehigh University, Bethlehem, Pennsylvania


The Yeast Genetics and Genomics course was a modern and intensive laboratory course that taught students the full repertoire of genetic and genomic approaches needed to dissect complex problems using the yeast *Saccharomyces cerevisiae*. Both classical and modern approaches were emphasized, including the isolation and characterization of mutants, tetrad analysis, and complementation. Synthetic biology was explored through CRISPR-Cas9-directed engineering of heterologous biosynthetic pathways in yeast.

Students learned genome-based methods of analysis facilitated by the *Saccharomyces* Genome Database, yeast genome sequences, the gene deletion collection, and other genomic resources available to the community. Molecular genetic techniques, including yeast transformation, gene replacement by polymerase chain reaction (PCR), construction and analysis of gene fusions, and generation of mutations, were also emphasized.

Students combined classical approaches with whole-genome sequencing to gain experience in identifying and interpreting genetic interactions, including suppression and synthetic lethality. Students performed genome-scale screens using the synthetic genetic array (SGA) methodology, were immersed in yeast genomics and computation methods for analyzing genome-scale data, and performed and interpreted experiments using PCR-amplicon and whole-genome sequencing.

Students gained firsthand experience in modern cytological approaches such as epitope tagging and imaging yeast cells using fluorescence microscopy with GFP–protein fusions and fluorescent indicators for different subcellular structures and organelles. Lectures on fundamental aspects of yeast genetics and genomics were presented along with seminars given by prominent experts in the field on topics of current interest.

This course was supported in part by grants from the National Science Foundation. Additional scholarship support was provided by the Helmsley Charitable Trust, the Howard Hughes Medical Institute, and Regeneron.

## PARTICIPANTS

Addorisio, S., B.S., Tufts University, Medford, Massachusetts
Adler, M., B.S., University of Tennessee, Knoxville
Chou, L., B.S., University of Pittsburgh, Pennsylvania
DeMouth, M., B.A., Albert Einstein College of Medicine, Bronx, New York
dos Santos, C., B.S., Duke University, Durham, North Carolina
Eke, E., B.Sc., University of New Mexico, Albuquerque
Flagg, L. B.S., University of Chicago, Illinois
Haffner, J.E.L., M.Sc., ETH Zürich, Basel, Switzerland
Heerschop, S., Ph.D., Institute of Molecular Biology, Mainz, Germany
Jalihal, A., Ph.D., University of North Carolina, Chapel Hill
Lunke, M., Ph.D., University of Victoria, British Columbia, Canada
Mathur, S., B.S., M.S., Stanford University, California
Ono, J., M.Sc., Ph.D., University of Exeter, Penryn, United Kingdom
Premkumar, T., M.Tech., UT Health/MD Anderson Cancer Center, Houston, Texas
Tallarico, G., M.S., IFOM-ETS, Milano, Italy
Vasquez Ibarra, M., P.G.Dip., Instituto Gulbenkian de Ciência (IGC), Oeiras, Portugal

## SEMINARS

Burke, M., Oregon State University, Corvallis: Yeast experimental evolution and the study of complex traits.
Dedek, I., Meier's Creek Brewing Company, Cazenovia, New York: The science of brewing.
Haber, J., Brandeis University, Boston, Massachusetts (attended virtually): Repairing broken chromosomes.
Ideker, T., University of California, San Diego: Towards a complete structure/function map of yeast (and cancer).
Landry, C., Université Laval, Quebec City, Quebec, Canada: Combining mutation and protein interaction studies.
Miller, M., University of Utah, Salt Lake City: Mechanisms ensuring accurate chromosome segregation.
Murray, A., Harvard University, Cambridge, Massachusetts: How yeast hedges its bets to manage uncertainty.
Nash, R., *Saccharomyces* Genome Database, Stanford University, California: Navigating data at SGD using YeastMine.
Silva, G., Duke University, Durham, North Carolina: Stress response and proteostasis.
Singer, J., Singer Instrument Company Ltd., Somerset, United Kingdom: Tetrad dissection.
Thompson, D., LifeMine Therapeutics, Cambridge, Massachusetts: Yeasts are major drivers of the bio-economy.
Vickers, C., Queensland University of Technology, Brisbane, Australia: Synthetic biology and metabolic engineering in yeast.
Wittkopp, P., University of Michigan, Ann Arbor: Molecular and evolutionary processes generating variation in gene expression.
Zanders, S., Stowers Institute for Medical Research, Kansas City, Missouri: Tn-Seq; wtf evolution.

# Workshop on Modernizing Neuropathology: Integrating -Omics, Bioinformatics, Advanced Imaging, and AI Approaches into Future Neuropathology Pipelines

**August 2–5**

ARRANGED BY

**B. Benedetti**, CZI Neurodegeneration Program
**K. Brose**, CZI Neurodegeneration Program
**K. Caetano-Anolles**, Mayo Clinic Florida, Jacksonville
**M. Gorham**, CZI Neurodegeneration Program
**M. Hauge Pedersen**, New York Genome Center, New York
**M.E. Murray**, Mayo Clinic Florida, Jacksonville
**H. Phatnani**, Columbia University/New York Genome Center, New York

The CSHL Modernizing Neuropathology workshop held at the Banbury Center in August 2023 proved to be a monumental gathering of thought leaders and emerging talents across the fields of neuropathology, -omics, radiology, and AI/deep learning. The primary objective was to foster collaboration, share insights, and devise solutions to pressing issues in brain banking infrastructure, digital pathology innovations, tissue optimization for -omics, ongoing harmonization efforts in neuroimaging, and the future research landscape with the implementation of AI/deep learning.

Two major outcomes from the workshop are particularly noteworthy. First, the workshop yielded recommendations for key gaps in knowledge, providing a roadmap for exciting progress in human-based tissue studies. This includes strategies to enhance the capacity of brain banks to meet the rising demand for tissues and facilitating groundbreaking research in the field. Second, the workshop identified the need for targeted training opportunities and support for postbaccalaureate cohorts from underrepresented groups. The goal is to enable informatics training through didactic summer coursework and placement in laboratories conducting neurodegeneration research that leverages human tissue. This initiative is poised to contribute significantly to nurturing the next generation of experts in the field. To maintain the momentum gained at the CSHL workshop,

the knowledge generated was shared with key stakeholders, including nonprofit and government funding agencies, via webinar. This dissemination aimed to inform these agencies about the valuable insights gained and to ensure that the collective recommendations of the group are heard and can positively influence decision-making.

We are in the process of drafting a comprehensive perspectives paper and aim to have it completed this year. This paper will encapsulate the rich discussions and outcomes of the workshop, providing key stakeholders and the research community with valuable insights. We delve into the foundational importance of neuropathology in driving pivotal discoveries and diagnostics, emphasizing the often overlooked and underfunded framework supporting brain banking. The paper advocates for integrated leadership and collaborative data flow as key elements in fostering innovative and impactful human-relevant research. The critical unmet needs in neurodegeneration research that were identified during the workshop underscore the importance of collaborative recommendations involving researchers, clinicians, institutions, and policymakers. The perspectives article aims to highlight solutions for enhancing brain bank representation across socioeconomic and ethnoracially diverse communities, emphasizing the integral role of brain banks in advancing modern techniques utilizing human brain tissue.

The workshop was generously supported by the Silicon Valley Community Foundation–CZI Neurodegeneration Challenge Network.

## DISCUSSANTS/MODERATORS

**Welcoming Remarks**
M.E. Murray, *Mayo Clinic Florida, Jacksonville*; H. Phatnani, *Columbia University/New York Genome Center, New York*

**Keynote Lecture**
E. Lein, *Allen Institute, Seattle, Washington*; C.D. Keene, *University of Washington, Seattle*

**Session 1: Neuropathology I: Brain Banking Presentations**
Session lead: C. Smith: *University of Edinburgh, United Kingdom*; A. Kapasi, *Rush University Medical Center, Chicago, Illinois*; M. Flanagan, *UT Health Science Center at San Antonio, Texas*

**Session 1: Neuropathology I: Interactive Session Brain Banks and Clinical Protocols**
Co-moderators: C. Smith, *University of Edinburgh, United Kingdom*; E. Lein, *Allen Institute, Seattle, Washington*; C.D. Keene, *University of Washington, Seattle*; E. Lee, *University of Pennsylvania, Philadelphia*

**Session 2: -Omics and Bioinformatics Presentations**
Session lead: V. Menon, *Columbia University, New York, New York*; B. Zhang, *Icahn School of Medicine at Mount Sinai, New York, New York*; J. Steen, *Harvard Medical School, Boston, Massachusetts*; Ö. Gökçe, *University of Bonn, Germany*

**Session 2: -Omics and Bioinformatics Interactive Session**
Co-moderators: V. Menon, *Columbia University, New York, New York*; P. De Jager, *Columbia University, New York, New York*; C. Karch, *Washington University at St. Louis, Missouri*; S. Vicković, *New York Genome Center/Columbia University, New York*

**Session 3: Neuropathology II: Presentations Molecular and Morphological Phenotyping**
Session lead: B. Dugger, *University of California, Davis*; A. Aguzzi, *University of Zürich, Switzerland*; S. Grant, *University of Edinburgh, United Kingdom*; I. Cobos, *Stanford University, California*

**Session 3: Neuropathology II: Interactive Session**
Co-moderators: B. Dugger, *University California, Davis*; J. Crary, *Icahn School of Medicine at Mount Sinai, New York, New York*; D. Irwin, *University of Pennsylvania, Philadelphia;* T. Lehner, *New York Genome Center, New York*

**Session 4: AI and Imaging Presentations**
Session lead: M. Keiser, *University California, San Francisco*; Z. Huang, *Stanford University, California*; D. Gutman, *Emory University, Atlanta, Georgia*; J. Vogel, *Lund University, Sweden*; M. Hawrylycz, *Allen Institute, Seattle, Washington*

**Session 4: AI and Imaging Interactive Session**
Co-moderators: M. Keiser, *University California, San Francisco*; D. Jones, *Mayo Clinic, Rochester, Minnesota*; E. Hillman, *Columbia University, New York, New York*; D. Pritchett, *Howard University, Washington, D.C.*

**Session 5: Meeting Synthesis, Future Directions, and Outcomes**
Synthesis reports from each of the 4 session leads

# Neuroscience of Addiction

**August 9–15**

INSTRUCTORS

**D. Belin**, University of Cambridge, United Kingdom
**C. Evans**, University of California, Los Angeles
**H. Kober**, Yale University, New Haven, Connecticut
**M. von Zastrow**, University of California, San Francisco

COURSE TEACHING ASSISTANTS

**N. Harp**, Yale University, New Haven, Connecticut
**J. Innes**, University of Cambridge, United Kingdom

Drug addiction is the costliest neuropsychiatric disorder faced by our nation. Acute and repeated exposure to drugs produces neuroadaptation and long-term memory of the experience, but the cellular and molecular processes involved are only partially understood.

The primary objective of the workshop was to provide an intense dialogue on the fundamentals, state-of-the-art advances, and major gaps in the cell and molecular biology of drug addiction. Targeted to new or experienced investigators, the workshop combined formal presentations and informal discussions to convey the merits and excitement of cellular and molecular approaches to drug addiction research. With the advent of genomics and proteomics, an extraordinary opportunity now exists to develop comprehensive models of neuroadaptive processes fundamental to addiction, withdrawal, craving, relapse to drug use, and general brain function.

A range of disciplines and topics were represented, including:

- Noninvasive brain imaging to identify drug targets and adaptive processes
- Neuroadaptative processes at the molecular and cellular level
- Neural networks and their modulation
- Relevance of genotype to susceptibility and drug response
- Tolerance and adaptation at the cellular level
- Approaches to exploiting the daunting volume of data generated by neuroinformatics.

This workshop provided an integrated view of current and novel research on neuroadaptive responses to addiction, fostered discussion on collaboration and integration, and provided critical information needed to construct a model of addiction as a disease and for novel molecular targets for biological treatments. Beyond the plane of scientific endeavor, information is vital for formulating public policy and for enlightening the public on the neurobiological consequences of drug use and addiction.

This workshop was designed to generate interest in this level of analysis, open conduits for collaborations, and present novel routes to investigating the neurobiology of addictive drugs.

This course was supported in part by grants provided by the Helmsley Charitable Trust, the Howard Hughes Medical Institute, and Regeneron.

## PARTICIPANTS

Acuna, A., B.S., Arizona State University, Tempe
Bluitt, M., B.S., University of North Carolina at Chapel Hill, Chapel Hill, NC
Belle Buzzi, B.S., Virginia Commonwealth University, Richmond
Dagunts, A., B.S., Oregon Health and Science University, Portland
Delawalla, C., B.A., Emory University, Atlanta, Georgia
Donovan, A., Ph.D., Charles R. Drew University of Medicine and Science, Los Angeles, California
Ehinger, Y., Ph.D., University of California, San Francisco
Escobedo Lozoya, Y., Ph.D., Harvard Medical School, Boston, Massachusetts
Linville, R., Ph.D., Massachusetts Institute of Technology, Cambridge
Maretich, P., B.A., Massachusetts Institute of Technology, Cambridge
Mesa, J., B.S., University of Florida, Gainesville
Modrak, C., B.S., University of Florida, Gainesville
Mohammad Aghaei, A., M.D., Yale University, New Haven, Connecticut
Murlanova, K., Ph.D., SUNY, Buffalo, New York
Ramirez, L., Ph.D., University of Illinois, Chicago
Savage, J., Ph.D., Vrije Universiteit Amsterdam, Netherlands
Schrock, J., Ph.D., Northwestern University, Chicago, Illinois
Testen, A., Ph.D., Medical University Of South Carolina, Charleston
Tonetto, S., M.S., Mental Health Services, Copenhagen, Denmark
White, A., B.S., West Virginia University, Morgantown
Wilkinson, C., B.S., B.A., University of Florida, Gainesville
Williford, K., B.A., Duke University, Durham, North Carolina
Winchester, S., B.A., Binghamton University, New York
Yousuf, H., Ph.D., Yale University, New Haven, Connecticut
Zhang, H., Ph.D., Cold Spring Harbor Laboratory

## SEMINARS

Belin, D., University of Cambridge, United Kingdom: Ethics.
Banghart, M., University of California, San Diego: Peptidergic neuromodulation.
Cheer, J., University of Maryland School of Medicine, Baltimore: Endogenous cannabinoids and the pursuit of reward.
Evans, C., University of California, Los Angeles: Ethics.
Kalivas, P., Medical University of South Carolina, Charleston: Back to the future to understand substance use disorders.
Kenny, P., Icahn School of Medicine at Mount Sinai, New York, New York: Paracrine regulation of nicotine craving.
Kober, H., Yale University, New Haven, Connecticut: Human neuroscience of addiction.
Koob, G., National Institute on Alcohol Abuse and Alcoholism, Bethesda, Maryland (attended virtually): The neurobiology of alcohol addiction.
Lobo, M.K., University of Maryland School of Medicine, Baltimore: Molecular and mitochondrial mechanisms in addiction.

Lovinger, D. National Institute on Alcohol Abuse and Alcoholism, Bethesda, Maryland: Basal ganglia circuits in addiction.

Morales, F.M., National Institute on Drug Abuse, Bethesda, Maryland: Deciphering the complexity of the ventral tegmental area.

Moron Concepcion, J., Washington University in St. Louis, Missouri: Pain-induced plasticity and opioid-seeking behavior.

Palmer, R., Emory University, Atlanta, Georgia: Genetic etiology of substance use and disorders: progress, gaps, and future directions.

Picciotto, M., Yale University, Hartford, Connecticut: Molecular basis of nicotine addiction.

Wassum, K., University of California, Los Angeles: Actions, habits, stress, and the brain.

# Proteomics

**August 10–22**

INSTRUCTORS
**R. Chalkley**, University of California, San Francisco
**D. Pappin**, Cold Spring Harbor Laboratory
**C. Weisbrod**, National High Magnetic Field Laboratory, Tallahassee, Florida
**C. Wong**, Mount Sinai Hospital, Toronto, Ontario, Canada

CO-INSTRUCTOR
**P. Cifani**, Cold Spring Harbor Laboratory

This intensive laboratory and lecture course focused on cutting-edge proteomic approaches and technologies. Students gained practical experience in sample preparation with in-solution digestion, and then were trained using high-sensitivity nano LC-ESI-MS and tandem mass spectrometry. An introduction to manual interpretation of data was followed by training in different search engines and bioinformatic approaches for data evaluation. Students used label-free and covalent isotopic-labeling quantitative approaches to profile changes. This included a section using Skyline to quantify peptides at the MS1 level, development of PRM targeted assays for quantification at the MS2 level, and an introduction to DIA analysis for more comprehensive fragmentation analysis. Students were shown how to recognize unexpected posttranslational modifications. Diverse techniques for PTM peptide enrichment, including affinity chromatography and immune enrichment, were carried out and the characterizations of the resulting complex mixtures, including site assignments, were performed. For training in protein complex and interaction analysis, students processed and analyzed proximity-labeled samples and learned how to evaluate the specificity of results.

A section on intact protein or large peptide analysis demonstrated the benefits of this approach; cross-linking analysis for structural studies was also discussed.

A series of outside lecturers discussed various proteomics topics including:

- de novo sequence analysis
- Intact protein analysis
- Cross-linking analysis for structural studies
- Data-independent analysis for comprehensive results across many samples
- Protein complex analysis.

Finally, an industrial lecture component by drug discovery scientists allowed participants to understand how chemoproteomics techniques can be routinely used in industry to profile compounds and potential protein targets.

For all sections of the course, a strong emphasis was placed on data analysis. There were opportunities to discuss and provide feedback on individual research projects, and students had the opportunity to learn to process their own data (acquired outside the course) if desired.

The aim of the course was to provide each student with the fundamental knowledge and hands-on experience necessary for performing and analyzing proteomic experiments. The overall goal was to train students to identify new opportunities and applications for proteomic approaches in their biological research.

This course was supported in part by grants from the National Institute of Child Health & Human Development. Additional scholarship support was provided by the Helmsley Charitable Trust, the Howard Hughes Medical Institute, and Regeneron.

## PARTICIPANTS

Bautista, O., B.A., Case Western Reserve University, Cleveland, Ohio
Belliveau, N., Ph.D., University of Washington/HHMI, Seattle
Bok, I., Ph.D., Washington University in St. Louis, Missouri
Brose, N., B.A., Duke University, Durham, North Carolina
Carrillo, G., Ph.D., Boston Children's Hospital/Harvard Medical School, Massachusetts
Dhiantravan, S., B.S., Northwestern University, Chicago, Illinois
Dia, V., Ph.D., University of Tennessee, Knoxville
Faunce, C., B.S., University of Pennsylvania, Philadelphia
Ling, Z., B.S., Carnegie Mellon University, Pittsburgh, Pennsylvania
Mueller, H., Memorial Sloan Kettering Cancer Center, New York, New York
Olanrewaju, G., B.Sc., Ohio University, Athens
Pessa, J., M.Sc., Abo Akademi, Turku, Finland
Peterson, K., B.S., University of Minnesota–Twin Cities, Minneapolis
Rivera-Serrano, M., M.S., University of Puerto Rico, San Juan
Velez-Irizarry, D., Ph.D., ARS-USDA, East Lansing, Michigan
Wahlang, B., Ph.D., University of Louisville, Kentucky

## SEMINARS

Andres, A., AstraZeneca, Waltham, Massachusetts: Industrial proteomics.
Chalkley, R., University of California, San Francisco: Software for MS/MS-based protein identification; PTM enrichment—other than antibodies; How to search for PTMs: how to interpret, evaluate MS/MS data of PTM-bearing peptides.
Clauser, K., Broad Institute of MIT and Harvard, Cambridge, Massachusetts: Manual de novo peptide MS/MS interpretation for evaluating database search results.
Geddes-McAlister, J., University of Guelph, Ontario, Canada: Proteomics of fungal disease in One Health.
Hristova, V., AstraZeneca, Gaithersburg, Maryland: Clinical proteomics applications in industry.
Kitaygorodsky, J., Lunenfeld-Tanenbaum Research Institute, Toronto, Ontario, Canada: Proximity-dependent biotinylation experiments and analysis.
Kumar, P., AstraZeneca, Waltham, Massachusetts: Bioinformatics for proteomics.

Nelson, A., Cell Signaling Technology, Danvers, Massachusetts: Posttranslational modification analysis techniques.
Pappin, D., Cold Spring Harbor Laboratory: Stable isotope quantification.
Pino, L., Talus Bioscience, Seattle, Washington: Skyline for quantitative LC-MS proteomics.
Samson, R., Sinai Health, Toronto, Ontario, Canada: Proximity-dependent biotinylation experiments and analysis.
Shannon, A., Ohio State University, Columbus: Label-free quantification: MS1, SRM/PRM, DIA.
Singh, M., AstraZeneca, Waltham, Massachusetts: Chemical biology and proteomics at AstraZeneca.
Trnka, M., University of California, San Francisco: Cross-linking mass spectrometry for exploring protein architecture and interactions.
Witkowski, J., Cold Spring Harbor Laboratory: Ethics, rigor, and reproducibility.
Weisbrod, C., National High Magnetic Field Laboratory, Tallahassee, Florida: MS 101 "Fundamentals of mass spectrometry for proteomics"; High-throughput intact protein characterization.
Wong, C., Mount Sinai Hospital, Toronto, Ontario, Canada: Intro to PDB methods.

# Antibody Engineering and Display Technologies

**October 15–31**

INSTRUCTORS
**G. Silverman**, NYU Grossman School of Medicine, New York, New York
**G. Veggiani**, Louisiana State University, Baton Rouge

COURSE CO-INSTRUCTORS
**S. Briand**, University of Zürich, Switzerland
**C. Rader**, The Scripps Research Institute, Jupiter, Florida
**J. Yeh**, Cold Spring Harbor Laboratory

COURSE TEACHING ASSISTANTS
**S. Furler**, University of Zürich, Switzerland
**J. Tang**, University of Toronto, Ontario, Canada

Advances in the generation and selection of antibodies from combinatorial libraries allow for the rapid production of antibodies from immune and nonimmune sources. This intensive laboratory/lecture course focused on the construction of combinatorial antibody libraries expressed on the surface of filamentous phage and the subsequent selection of desired antibodies from the library. Students learned the theoretical and practical aspects of constructing combinatorial libraries from immune and nonimmune sources as well as the construction of synthetic antibody libraries. Antigen-specific recombinant monoclonal antibodies were selected from the library. Production, purification, and characterization of antibody fragments expressed in *Escherichia coli* were also covered.

The lecture series, presented by course faculty and several invited speakers, emphasized theory and practice of antibody display technologies, expression of antibodies in *E. coli* and mammalian cells, antibody structure and function, bacterial display of antibodies and other ligand-binding domains, the immunobiology of the antibody response, and the use of monoclonal antibodies for therapy—including the design of chimeric antigen receptor T cells. We also discussed principles and protocols for generation and analysis of immune repertoires determined by next-generation sequencing.

This course was supported in part by grants provided by the Helmsley Charitable Trust, the Howard Hughes Medical Institute, and Regeneron.

## PARTICIPANTS

Aghassizadeh, A., B.Sc., Duke University, Durham, North Carolina
Aguilar, E., M.S., Johns Hopkins University School of Medicine, Baltimore, Maryland
Atanda, H., B.Tech., International Institute of Tropical Agriculture, Ibadan, Nigeria
Diaz-Perlas, C., Ph.D., Institut Quimic de Sarria (IQS-URL), Barcelona, Spain
Facchini, F.A., Ph.D., DiaSorin Italia S.p.A, Bresso, Italy
Flemming Svedmark, S., M.S., KTH Royal Institute of Technology, Stockholm, Sweden
He, C., B.S., University of California, San Francisco
Ibitoye, O., M.S., Kampala International University, Ishaka, Uganda
Kelly, K., Ph.D., University of Virginia, Charlottesville
Lie-Andersen, O., Ph.D., Immunitrack ApS, Copenhagen, Denmark
Murrell, I., Ph.D., Antiverse Ltd., Cardiff, United Kingdom
Ronning, S., M.S., KTH Royal Institute of Technology, Stockholm, Sweden
Saliba, D., Ph.D., University of Malta, Msida
Tharp, C., B.S., University of California, San Francisco
Umthong, S., Ph.D., Chulalongkorn University, Bangkok, Thailand
Zhu, S., Ph.D., Institute for Protein Innovation, Boston, Massachusetts

## SEMINARS

Briand, S., University of Zürich, Switzerland; Furler, S., University of Zürich, Switzerland; Plueckthun, A., University of Zürich, Switzerland (attended virtually): Ribosome display.
DeKosky, B., MIT, Cambridge, Massachusetts: High-throughput screening of immune receptors.
Howarth, M., University of Cambridge, United Kingdom: Evolution of bacterial superglues: from infinite affinity to disease applications.
Rader, C., Aethon Therapeutics, Long Island City, New York: Phagemid family pComb3 and its utilization for antibody drug and target discovery in cancer.
Rakonjac, J., Massey University, Palmerston North, New Zealand: Biology and structure of filamentous bacteriophage.
Sidhu, S., University of Waterloo, Ontario, Canada: Synthetic antibodies.
Siegel, D., University of Pennsylvania, Philadelphia: Phage display tools for chimeric antigen receptor–T cell development.
Silverman, G., NYU Grossman School of Medicine, New York, New York: Molecular characterization of B-cell epitopes in a bacterial pathogen by phage display.
Smith, G., University of Missouri, Virgin Islands: Principles of affinity selection.
Stahl, S., KTH AlbaNova University Center, Stockholm, Sweden: Affibodies and sequestrins for targeted therapy.
Wilson, M., University of California, San Francisco: Antibody discovery in neuroinflammatory diseases with phage display.
Yarmarkovich, M., NYU School of Medicine, New York, New York: CAR T cell libraries for targeting novel antigens.
Yeh, J., Cold Spring Harbor Laboratory: Antibody and protein engineering.

# Macromolecular Crystallography

**October 16–31**


INSTRUCTORS
**P. Adams**, Lawrence Berkeley National Laboratory, California
**D. Liebschner**, Lawrence Berkeley National Laboratory, California
**J. Newman**, University of New South Wales, Kensington, Australia
**J. Pflugrath**, Rigaku Americas (Retired), The Woodlands, Texas

COURSE CO-INSTRUCTOR
**T. Peat**, University of New South Wales, Sydney, Australia

COURSE TEACHING ASSISTANTS
**C. Fan**, California Institute of Technology, Pasadena
**B. Poon**, Lawrence Berkeley National Lab, California


X-ray crystallography has been the cornerstone of structural biology for half a century and remains the technique of choice for atomic-resolution understanding of macromolecules and for structure-guided drug discovery. This intense course combined laboratory and computational instruction to train course participants in the major techniques used to determine three-dimensional structures. It was designed for scientists with a working knowledge of protein structure and function, but who are new to macromolecular crystallography or who wish to increase their in-depth knowledge of macromolecular crystallography. The advent of methods for the accurate prediction of protein structures is accelerating crystallography experiments, and the course trained researchers in how to best use these models.

Topics covered included:

- Basic diffraction theory
- Structure presentation
- Coordinate deposition
- Structure validation

- Model building and refinement
- Electron density maps improvement
- Using predictive models in structure solution and completion
- Protein structure prediction
- Structure solution by experimental phasing methods (SAD, MAD, MIR, and others) and molecular replacement
- Data collection and processing
- Synchrotron X-ray sources and optics
- Crystallization (proteins, nucleic acids, complexes, and membrane proteins)

Participants had extensive hands-on training in well-equipped labs in how to crystallize proteins and determine crystal structures by several methods, while learning through lectures on theory and methods. Both basic and advanced subjects were covered during lectures, which were given by leaders in the field. Informal discussions behind the techniques were frequent and students were expected to pose questions to be answered in interactive sessions.

Applicants needed to be familiar with the creation and editing of simple text files on Linux workstations.

This course was supported in part by grants provided by the Helmsley Charitable Trust, the Howard Hughes Medical Institute, and Regeneron.

## PARTICIPANTS

Cui, J., B.S., Pennsylvania State University, State College
Dubey, G., M.Sc., Texas A&M University, College Station
Dubey, S., M.S., Purdue University, West Lafayette, Indiana
Gulkis, M., B.S., University of Florida, Gainesville
Hardeman, K., Ph.D., University of North Carolina at Chapel Hill
Larus, I., B.S., Stanford University, California
Maddumage, J., B.S., La Trobe University, Bundoora, Victoria, Australia
Maurakis, S., Ph.D., National Institutes of Health–NIDDK, Bethesda, Maryland
McCord, J., B.S., Texas Tech University Health Sciences Center, Lubbock
Nagarajan, B., Ph.D., Virginia Commonwealth University, Richmond
Naik, M., Ph.D., Brown University, Providence, Rhode Island
Neilsen, G., B.S., Emory University, Atlanta, Georgia
Nishino, Y., Cold Spring Harbor Laboratory
Perez i Rafols, A., Ph.D., University of Dundee, Great Britain
Samkutty, A., M.S., Schrödinger, Inc., Natick, Massachusetts
White, I., B.S., Janssen Research & Development, Spring House, Pennsylvania

## SEMINARS

Adams, P., Lawrence Berkeley National Laboratory, California: Crystallography basics.
AlQuraishi, M., Columbia University, New York, New York (attended virtually): The state of protein structure prediction and friends.
Bergfors, T., Uppsala University, Sweden: Basics of crystallization; optimization for crystallization.
Borek, D., UT Southwestern Medical Center, Dallas Texas: Diffraction data processing.
Caffrey, M., Trinity College Dublin, Ireland: Membrane protein crystallisation using lipidic system.
Clarke, O., Columbia University, New York, New York: Map interpretation and model building.
Garman, E., University of Oxford, United Kingdom: Radiation damage in MX: background and prospects.
Holton, J., University of California, San Francisco/Lawrence Berkeley National Laboratory/SLAC National Accelerator Laboratory, Menlo Park, California: Beamline basics and radiation damage.
Joosten, R., Netherlands Cancer Institute, Amsterdam (attended virtually): Model validation and fixing.
Liebschner, D., Lawrence Berkeley National Laboratory, California: Introduction to maps.
Liu, Q. Brookhaven National Laboratory, New York: Membrane protein crystallization.
McCoy, A., University of Cambridge, United Kingdom: Molecular replacement.
McPherson, A., University of California, Irvine: Symmetry, periodicity, unit cells, space groups, and lattices; Scattering by atoms and crystals, Fourier transform, relationships

between real and reciprocal space; The phase problem and the Patterson function.

Meilleur, F., Oak Ridge National Laboratory/North Carolina State University, Tennessee (attended virtually): Neutron macromolecular crystallography.

Peat, T., University of New South Wales, Kensington, Australia: Designing proteins; Ligands in structures.

Pflugrath, J., Rigaku Americas (Retired), The Woodlands, Texas: Crystallization practice: how to prepare crystals for data collection.

Read, R., University of Cambridge, United Kingdom: Lectures and tutorials on phaser.

Richardson, J., Duke University, Durham, North Carolina: Model validation: MolProbity.

Rodriguez, J., University of California, Los Angeles: Electron crystallography.

Smith, C., Stanford University, California: Femtosecond crystallography at synchrotrons and XFELs.

Terwilliger, T., New Mexico Consortium, Los Alamos (attended virtually): Automated model-building and structure solution: AlphaFold changes everything (and nothing).

Williams, C., Duke University, Durham, North Carolina: Guide to MolProbity validation, validation horror show.

Witkowski, J., Cold Spring Harbor Laboratory: Ethics lecture.

# Programming for Biology

**October 16–31**


INSTRUCTORS

**S. Prochnik**, Serenity Comp Bio, San Francisco, California
**S. Robb**, Stowers Institute for Medical Research, Kansas City, Missouri

COURSE TEACHING ASSISTANTS

**J. Bredeson**, DOE JGI Lawrence Berkeley National Laboratory, California
**K. Gotting**, University of Wisconsin, Madison
**R. Kellermeyer**, Stowers Institute for Medical Research, Kansas City, Missouri
**E. Ross**, Stowers Institute for Medical Research, Kansas City, Missouri


More often than not, today's biologist is studying data sets that are too complex or large to be analyzed without a computer and, even so, existing software tools provide only boilerplate analyses. Questions specific to the data set require novel analysis pipelines to be devised and written in computer code. Designed for laboratory biologists with little or no programming experience, this course gave students the bioinformatics and scripting skills they needed to derive biological insights from this abundance of data. The only prerequisite for the course was a strong commitment to learning basic UNIX and a scripting language. Lectures and problem sets from previous years are available online, and students were welcome to study this background material before starting the course.

We used Python, an easy-to-learn scripting language with an established code base and community of users. The course began with one week of introductory coding, continued with practical topics in bioinformatics, with plenty of coding examples, and ended with a group coding project. Formal instruction was provided on every topic by the instructors, teaching assistants, and invited experts. Students solved problem sets covering common scenarios in the acquisition, validation, processing, and analysis of biological data. They learned how to design, construct, and run powerful and extensible analysis pipelines in a straightforward manner. Final group projects

were chosen from ideas proposed by students and were guided by faculty. Students were provided with a library of Python reference print and e-books that they could bring home with them.

Note that the primary focus of this course was to provide students with practical programming experience, rather than to present a detailed description of the algorithms used in computational biology. For the latter, we recommend the Foundation of Computational Genomics course.

This course was supported in part by grants from the National Human Genome Research Institute. Additional scholarship support was provided by the Helmsley Charitable Trust, the Howard Hughes Medical Institute, and Regeneron.

## PARTICIPANTS

Angus-Hill, M., Ph.D., Allen Institute, Seattle, Washington
Barcia Durán, J.G., Ph.D., NYU Grossman School of Medicine, New York
Blanco Salazar, M., B.S., Stowers Institute for Medical Research, Kansas City, Missouri
Cahir, C., B.S., Rockefeller University, New York, New York
Dalipovic, R., B.S., Binghamton University, New York
Dobhal, A., M.D., Vaccine Research Center, National Institutes of Health, Bethesda, Maryland
Fort, G., B.A., University of Utah, Salt Lake City
Gana, T., Ph.D. James Cook University, Cairns, Queensland, Australia
Gebert, T., B.S., Baylor College of Medicine, Houston, Texas
Halvorsen, T., Ph.D., Lawrence Livermore National Laboratory, California
Jin, D., Ph.D., Yale University, New Haven, Connecticut
MacFawn, I., Ph.D., University of Pittsburgh, Pennsylvania
Mankouski, A., M.D., University of Utah, Salt Lake City
Muhoya, B., B.Sc., Princeton University, New Jersey
Paraschivescu, C., Ph.D., Feinstein Institutes for Medical Research, New York, New York
Paulding, D., Ph.D., Cincinnati Children's Hospital Medical Center, Ohio
Rossi, R., B.S., University of "Milano Statale," Milan, Italy
Saha, S., Ph.D., University of Cambridge, United Kingdom
Sengul, E., M.Sc., University of Oxford, United Kingdom
Silverman, S., B.S., Yale University, New Haven, Connecticut
Toth, H., B.E., Ohio State University, Columbus
VandenBosch, L., Ph.D., Seattle Children's Research Institute, Washington
Vazquez Echegaray, C., Ph.D., Lund University, Sweden

## SEMINARS

Bredeson, J., DOE JGI Lawrence Berkeley National Laboratory, California: Bioinformatics file formats.
Cain, S., Ontario Institute for Cancer Research, Toronto, Canada (attended virtually): A modern genome browser in an age of MANY genomes.
Haas, B., Broad Institute, Cambridge, Massachusetts: Python programming for transcriptomics.
Pearson, W., University of Virginia, Charlottesville: Similarity searching/what BLAST does/why BLAST works.
Perera, A., Stowers Institute, Kansas City, Missouri: Practical wet-lab advice to generate good quality data.
Prochnik, S., Serenity Comp Bio, San Francisco, California; and Robb, S., Stowers Institute for Medical Research, Kansas City, Missouri:: Unix 1 intro; Unix 2, vi, git; python1 overview, running python, syntax, data types, and variables; python 2 operators, truth, logic, numbers; python 3 sequences, strings; python 4 loops, lists, tuples; python 6: I/O and files; python 7: regular expressions; python 8: data structures; python 9: exception; python 10: functions, scope, modules; python 11: classes, biopython.
Schorn, A., Cold Spring Harbor Laboratory: Analysis of noncoding RNA and repetitive sequences in the genome.
Thomas, P., University of Southern California, Los Angeles: Gene function annotation and gene set analysis.
Triant, D., University of Virginia, Charlottesville: Genome assembly.
Unruh, J., Stowers Institute for Medical Research, Kansas City, Missouri: Image processing in Python.

# High-Throughput Neuroanatomy

**October 20–31**

INSTRUCTORS

**X. Chen**, Allen Institute for Brain Science, Seattle, Washington
**L. DeNardo**, University of California, Los Angeles
**J. Kebschull**, Johns Hopkins University, Baltimore, Maryland
**L. Schwarz**, St. Jude Children's Research Hospital, Memphis, Tennessee

COURSE TEACHING ASSISTANTS

**M. Gongwer**, David Geffen School of Medicine at UCLA, Los Angeles, California
**H. Kim**, School of Medicine, Johns Hopkins University, Baltimore, Maryland
**D. Ravens**, Cold Spring Harbor Laboratory
**M. Rue**, Allen Institute for Brain Science, Seattle, Washington
**Y-C. Sun**, New York University Grossman School of Medicine, New York
**C. Zhan**, Cold Spring Harbor Laboratory

Modern high-throughput neuroanatomical tools—including but not limited to barcode-based tracing methods—allow unprecedented insight into the structure of the nervous system. Their use, however, is currently restricted to a small set of expert laboratories. This limited uptake is not due to an inherent cost or difficulty of these techniques, but rather due to a lack of training opportunities. In this new CSHL course, we taught the core skills required for successfully applying a range of modern neuroanatomical techniques. We covered viral tracing tools, including anterograde, retrograde, and trans-synaptic tracers, brain clearing and whole-brain analysis of axonal data, MAPseq and BARseq, and RNA in situ sequencing technologies. In practicals, students learned how to prepare viruses, how to acquire and analyze whole-brain volumetric axon tracing data, and how to perform a BARseq2 experiment. After completing this course, we expect students to have the skills to implement a range of cutting-edge neuroanatomical techniques in their own laboratories.

This course was supported in part by grants provided by the Helmsley Charitable Trust, the Howard Hughes Medical Institute, and Regeneron.

## PARTICIPANTS

Megan Anderson, M., B.S., University of California, San Diego/Salk Institute, La Jolla, California
Arzua, T., Ph.D., Columbia University, New York, New York
Bellafard, A., Ph.D., University of California, Los Angeles
Cola, R., B.Sc., University of Zürich, Switzerland
Gagnon, M., B.S., Baylor College of Medicine, Houston, Texas
Guerrero, I., B.S., Florida State University, Tallahassee
Heckman, E., Ph.D., University of Michigan, Ypsilanti
Juengling, M., M.S., Max Planck Institute for Brain Research, Frankfurt Am Main, Germany
Karlupia, N., Ph.D., Harvard University, Cambridge, Massachusetts
Kaur, N., Ph.D., Yale University, New Haven, Connecticut
Li, M., B.Sc., University of Toronto, Ontario, Canada
Pineda, C., B.S., University of California, Davis
Quiniou, M., M.Sc., University of Zürich, Switzerland
Yang, L., B.S., Washington University in St. Louis, Missouri
Yu, B., B.S., University of California, San Diego
Zhu, J., B.S., Baylor College of Medicine, Houston, Texas

## SEMINARS

Bagnall, M., Washington University, St. Louis, Missouri: New insights into vestibular organization by connectomics.
Boyden, E., Massachusetts Institute of Technology, HHMI, Cambridge: Expansion microscopy.
Greely, H., Stanford University, California (attended virtually): Neuroethical issues.
Laurent, G., Max-Planck-Institute for Brain Research, Frankfurt, Germany: Single-cell RNA-seq approaches in brain studies of nonclassical model systems.
Luo, L., Stanford University, California: An introduction to genetic dissection of neural circuits.
Wang, X., Broad Institute of MIT and Harvard, Cambridge, Massachusetts: Translating spatial cell atlas to tissue function.
Wu, Z., Weill Cornell Medicine, New York, New York: Volumetric imaging with tissue clearing to investigate brain complexity.
Zador, A., Cold Spring Harbor Laboratory: Barcoded connectomics and NeuroAi.
Zhan, C., Cold Spring Harbor Laboratory: MAPseq/BARseq.
Zhang, L., Yale University, New Haven, Connecticut: A multidimensional coding architecture of the vagal interoceptive system.
Znamenskiy, P., Francis Crick Institute, London, United Kingdom: High-throughput molecular connectomics with barcoded rabies viruses.

# Advanced Sequencing Technologies and Bioinformatics Analysis

**November 6–19**

INSTRUCTORS

**F. Gomez**, Washington University School of Medicine in St. Louis, Missouri
**M. Griffith**, Washington University School of Medicine, Richmond Heights, Missouri
**O. Griffith**, Washington University School of Medicine in St. Louis, Missouri
**E. Mardis**, Nationwide Children's Hospital, Powell, Ohio
**W. McCombie**, Cold Spring Harbor Laboratory

COURSE CO-INSTRUCTORS

**C. Miller**, Washington University in St. Louis, Missouri
**J. Preall**, Cold Spring Harbor Laboratory

COURSE TEACHING ASSISTANTS

**M. Cannon**, Nationwide Children's Hospital, Columbus, Ohio
**A. Danos**, Washington University in St. Louis, Missouri
**S. Goodwin**, Cold Spring Harbor Laboratory
**M. Hoang**, Washington University in St. Louis, Missouri
**M. Khanfar**, Washington University in St. Louis, Missouri
**J. Kunisaki**, University of Utah, Salt Lake City
**C. Regan**, Cold Spring Harbor Laboratory
**K. Singhal**, Washington University in St. Louis, Missouri

Over the last decade, massively parallel DNA sequencing has markedly impacted the practice of modern biology and is being utilized in the practice of medicine. The constant improvement of these platforms means that costs and data generation timelines have been reduced by orders of magnitude, enabling investigators to conceptualize and perform sequencing-based projects that heretofore were time-, cost-, and sample number–prohibitive. Furthermore, the application of these

technologies to answer questions previously not experimentally approachable is broadening their impact and application. However, data analysis remains a complex and often vexing challenge, especially as data volumes increase.

This intensive two-week course explored the use and applications of massively parallel sequencing technologies, with a focus on data analysis and bioinformatics. Students were instructed in the detailed operation of several short- and long-read platforms, including library construction procedures, general data processing, and in-depth data analysis. A diverse range of the types of biological questions enabled by massively parallel sequencing technologies was explored such as variant calling, transcriptome analysis, single-cell analysis, metagenomics, epigenomics, and others that were tailored to the student's research areas of interest.

Cloud-based computing was also explored. Guest lecturers highlighted unique applications of these disruptive technologies.

We encouraged applicants from a diversity of scientific backgrounds including molecular evolution, development, neuroscience, medicine, cancer, plant biology, and microbiology.

This course was supported in part by grants from the National Human Genome Research Institute. Additional scholarship support was provided by the Helmsley Charitable Trust, the Howard Hughes Medical Institute, and Regeneron.

## PARTICIPANTS

Arefiev, I., M.S., University of Sherbrooke, Québec, Canada
Bahraoui, S., M.S., Ph.D., UT Southwestern, Dallas, Texas
Blue, N., M.D., University of Utah Health, Salt Lake City
Chen, C-H., M.S., Washington State University, Spokane
Clark, M., B.S., Duke University, Durham, North Carolina
DeBose-Scarlett, E., B.S., Duke University School of Medicine, Durham, North Carolina
Ebert, L., Ph.D., University Hospital Cologne, Germany
Falter, R., B.S., University of Utah, Salt Lake City
Fennell, K., Ph.D., The Peter MacCallum Cancer Centre, Parkville, Victoria, Australia
Grill, S., Ph.D., Whitehead Institute for Biomedical Research, Cambridge, Massachusetts
Hong, V., M.S., University of Texas at Dallas, Richardson, Texas
Jackson, L., B.S., Mayo Clinic, Rochester, Minnesota
LaFlamme, C., B.S., St. Jude Children's Research Hospital, Memphis, Tennessee
Lee, B., M.S.E., Columbia University, New York, New York
Molotkova, A., B.S., Georgetown University Medical Center, Washington, D.C.
Pass, C., B.S., University of Florida, Gainesville
Tran, L., B.S., University of Colorado-Anschutz Medical Campus, Aurora
Wedemeyer, M., M.D./Ph.D., Nationwide Children's Hospital and The Ohio State, Columbus
Wilde, S., Ph.D., St. Jude Children's Research Hospital, Memphis, Tennessee
Zhu, W., M.S., Duke University, Durham, North Carolina
Zigackova, D., Ph.D., Yale University, New Haven, Connecticut

## SEMINARS

Bedrosian, T., Nationwide Children's Hospital, Columbus, Ohio (attended virtually): Elucidating epilepsy genetics through advanced sequencing techniques.
Goar, W., Nationwide Children's Hospital, Columbus, Ohio: Variant representation and interpretation.
Gomez, F., Washington University School of Medicine, St. Louis, Missouri: Data visualization.
Griffith, O., Washington University School of Medicine, St. Louis, Missouri: Alignment and visualization concepts; Expression analysis.
Griffith, M., Washington University School of Medicine, St. Louis, Missouri: Introduction to bulk RNA sequencing.
Handley, S., Washington University School of Medicine, St. Louis, Missouri: Metagenomics and viromics.
Koo, P., Cold Spring Harbor Laboratory: Deep learning for regulatory genomics.
Li, Y., Washington University in St. Louis, Missouri: Single-cell ATAC-seq analysis.
Lippman, Z., Cold Spring Harbor Laboratory/HHMI: A pan-genome of pan-genomes in nightshade plants drives fundamental and applied biology.
Mardis, E., Nationwide Children's Hospital, Powell, Ohio: Overview of next-generation short-read sequencing technologies; Emerging next-generation sequencing platforms.
McCombie, W. Cold Spring Harbor Laboratory: Long-read sequencing technology and research applications.
McPherson, J., UC Davis School of Medicine, California: The genomic era is alive and well!

Miller, C., Washington University in St. Louis, Missouri: An overview of bioinformatics; Long-read sequencing and alignment; Introduction to germline variant calling; Somatic variant calling; Introduction to epigenomics—ATAC, ChIPseq, and bisulfite approaches; Introduction to R; Data visualization.

Mozersky, J., Washington University in St. Louis, Missouri (attended virtually): Ethics of genomics.

Phillippy, A., National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland (attended virtually): The human genome is finally finished. What have we been missing?

Preall, J., Cold Spring Harbor Laboratory: Single-cell analysis: soup to nuts.

Pugh, T., University Health Network, Toronto, Ontario, Canada (attended virtually): Introduction to applications of single-cell sequencing.

Quinlan, A., University of Utah, Salt Lake City (attended virtually): Intro to probability for genomics.

Sunkel, B., Nationwide Children's Hospital, Columbus, Ohio: Sequencing approaches for profiling the epigenome.

Vickovic, S., NYGC and Columbia University, New York, New York (attended virtually): Spatial transcriptomics.

# Tutorials in Genomics and Bioinformatics: RNA-seq Analysis

**November 12–14**

INSTRUCTORS

**D. Fagegaltier**, Merck, New York, New York
**E. Hodges**, Vanderbilt University School of Medicine, Nashville, Tennessee
**B. King**, University of Maine, Orono
**S. Munger**, Jackson Laboratory, Bar Harbor, Maine

Tutorials in Genomics and Bioinformatics: RNA-seq is an intensive two-day introductory course to genomics and bioinformatics. Participants were expected to arrive by 6 p.m. on the first day (Sunday, November 12), with the course running two full days until 5 p.m. on the third day (Tuesday, November 14).

Tutorials in Genomics and Bioinformatics (TGB) was modeled on Cold Spring Harbor Laboratory's Genome Access Course, a two-day course normally offered in person at CSHL and other locations. TGB was broken into modules that are each designed to give a broad overview of a given topic, with ample time for examples chosen by the instructors. Each module features a brief lecture describing the theory, methods, and tools, followed by a set of worked examples that students completed. Ample opportunities are provided for students to engage instructors during the course with specific tasks or problems that pertain to their own research.

The core of the course was the analysis of bulk RNA sequencing data. Featured resources and examples primarily came from mammalian species, but concepts can be applied to any species with a reference genome assembly. TGB provided hands-on experience by re-analyzing a published bulk RNA-sequencing data set from mammalian tissues.

Topics included designing RNA-seq studies (best practices in the design of bulk RNA-seq studies and caveats in analysis workflows); analysis of high-throughput sequence data using Galaxy (importing FASTQ files, importing reference genomes and annotation, read quality control and diagnostics, read trimming, and read mapping and read count estimation); an introduction to R (basic syntax, data structures, reading input and writing input, and plotting basics); analysis of RNA-seq read counts using R/DESeq2 (diagnostic analyses, normalization, model fitting, testing for differentially expressed genes, and data visualization—heatmaps, volcano plots); genome

browser resources (genome annotation, functional genome data, and bulk genome analysis); and gene set enrichment and pathway analysis using Gene Ontology and pathway annotations.

## PARTICIPANTS

Akoh-Arrey, T., Princeton University, East Windsor, New Jersey
Clayton, S., Washington University in St. Louis, Missouri
Hahn, P., Herbert Wertheim UF Scripps, Jupiter, Florida
Helms, K., Columbia University Medical Center, New York, New York
Hofford, R., Ph.D., Wake Forest School of Medicine, Winston-Salem, North Carolina
Hossain, M.K., Columbia University, New York, New York
Kalin, T., Cincinnati Children's Hospital, Ohio
Kwan, G., Endicott College, Beverly, Massachusetts
Justin Lee, J., University of California, Berkeley
Lee, P., University of Chicago, Illinois
Li, L., M.S., Merck, Cambridge, Massachusetts
Lynch, W., Boston University, Massachusetts
Marin-Rodero, M., Harvard Medical School, Boston, Massachusetts
Martinez-Lozada, Z., The Children's Hospital of Philadelphia, Pennsylvania
Matthews, A., Stony Brook University, New York
Mitchell, M., University of South Carolina, Tullahoma, Tennessee
Murphy, T., MRC Prion Unit at University College London, United Kingdom
Nelson, J., University of Chicago, Illinois
O'Hagan, D., The Wertheim UF Scripps Institute, Jupiter, Florida
Ondatje, B., Barrow Neurological Institute/Arizona State University, Phoenix
P Thirumalaikumar, V., Cornell University, Ithaca, New York
Paraiso, H., IU School of Medicine, Fort Wayne, Indiana
Paul, S., Ph.D., Johns Hopkins School of Medicine, Baltimore, Maryland
Pelham, J., Washington University in St. Louis, Missouri
Portillo, J., Duke University, Durham, North Carolina
Poznanski, P., Plant Breeding and Acclimatization Institute, Błonie, Poland
Singh, M., Ph.D., University of Iowa, Iowa City
Smith, M., Loyola University Chicago, Maywood, Illinois
Smith, M., Hospital for Special Surgery, New York, New York
Tersey, S., University of Chicago, Illinois
Yalcinbas, E., B.S., Lieber Institute for Brain Development, Baltimore, Maryland
Yamamoto, K., City University of New York, New York
Zhou, H., Massachusetts General Hospital, Boston

## SEMINARS

Fagegaltier, D., Merck, New York, New York: RNA-seq analysis—step 1; RNA-seq analysis—step 2; RNA-seq analysis—step 3.
Hodges, E., Vanderbilt University School of Medicine, Nashville, Tennessee: Introduction to R and R Studio; Analyze read counts using R/DESeq2 in R Studio; UCSC Genome Browser; Functional genomic elements and ENCODE.
Howell, G., Jackson Laboratory, Bar Harbor, Maine: Leveraging mouse genetic diversity to understand neurodegenerative diseases.
King, B., University of Maine, Orono: Course introduction; Accessing RNA-seq data from the NCBI and EBI; RNA-seq analysis—step 1; RNA-seq analysis—step 2; RNA-seq analysis—step 3; Introduction to R and R Studio; Analyze read counts using R/DESeq2 in R Studio; Model organism databases; Q&A session and course summary.
Munger, S., Jackson Laboratory, Bar Harbor, Maine: RNA-seq experimental design and analysis introduction; Ensembl genome browser; Pathway analysis; M structure solution

# Scientific Writing Retreat

**November 14–19**

INSTRUCTORS **C. Lambert**, Cold Spring Harbor Laboratory
**S. Matheson**, PLoS, Tucson, Arizona

COURSE
CO-INSTRUCTORS **M. Bao**, Harvard Medical School, Boston, Massachusetts
**J. Jansen**, Weill Cornell Medicine, New York, New York

The CSHL Scientific Writing Retreat is designed for postdoctoral fellows and junior faculty in all areas of biology who are actively working on professional pieces of writing such as manuscripts, grant proposals, job applications, or research/teaching/personal statements. Applications from staff scientists and more senior independent investigators were also welcome. English proficiency was assumed; this retreat was not appropriate for scientists who are in the initial stages of learning to read, write, or speak English.

The goal of this retreat was to have participants progress significantly on writing projects while improving their professional communication skills. The retreat included a mix of formal sessions and less structured writing time. The formal sessions covered:

- Publication writing for scientific journals from the perspectives of Cell Press, PLoS, and Cold Spring Harbor Press
- Grant writing from the perspective of the National Institutes of Health
- Writing clearly and conversationally about your research in ways that engage diverse audiences, a skill particularly useful when developing lay summaries for NIH and NSF proposals
- Style tips and considerations for clear professional writing in all forms.

The less structured sessions included small writing groups and dedicated individual writing time. For the small group sessions, participants were preassigned to groups of three to four people for the purpose of soliciting peer feedback on their writing. For the individual writing sessions, coaches were on hand to work with participants one-on-one.

This course was supported in part by grants provided by the Howard Hughes Medical Institute and Regeneron.

## PARTICIPANTS

Benjamin, K., Ph.D., Lieber Institute for Brain Development, Baltimore, Maryland
Chuard, A., Ph.D., Boston College, Chestnut Hill, Massachusetts
Grossman, Y., Ph.D., Duke University, Durham, North Carolina
Manuel, J., Ph.D., Universite de Sherbrooke, Québec, Canada
Mondello, P., M.D., Mayo Clinic, Rochester, Minnesota
Pan, H., Ph.D., Columbia University Irving Medical Center, New York, New York
Phan, A., Ph.D., Icahn School of Medicine at Mount Sinai, New York, New York
Pizzurro, G., Ph.D., Yale University, New Haven, Connecticut
Porter, D., Ph.D., University of Michigan, Ann Arbor
Prabakar, R., Ph.D., Cold Spring Harbor Laboratory
Rockfield, S., Ph.D., St. Jude Children's Research Hospital, Memphis, Tennessee
Sanford, S., Ph.D., University of Pittsburgh, Pennsylvania
Scharf, A., Ph.D., Missouri University of Science and Technology, Rolla
Singh, G., Ph.D., Brown University, Providence, Rhode Island
Wang, Y-C., Ph.D., Washington University in St. Louis, St. Louis, MO
Xiong, J., Ph.D., Johns Hopkins University, St. Petersburg, Florida
Xu, X., Ph.D., Cold Spring Harbor Laboratory

## SEMINARS

Bao, M., Harvard Medical School, Boston, Massachusetts: Grant writing and grantsmanship.
Connell, L., Cold Spring Harbor Laboratory.
Jansen, J., Weill Cornell Medicine, New York, New York: Writing for nonexpert audiences.
Lambert, C., Cold Spring Harbor Laboratory: Top 10 tips for strong scientific writing.
Matheson, S., PLoS, Tucson, Arizona: Top 10 tips for strong scientific writing; Publications and manuscripts.

## WRITING COACHES

Pavlovich, M., Cell Press, Cambridge, Massachusetts.
Rubin, J., Columbia University, Tenafly, New Jersey.
Sirois, C., Cell Press, Cambridge, Massachusetts.

# Computational Genomics

**November 29–December 6**


INSTRUCTORS

**D. Hawkins**, University of Washington School of Medicine, Seattle
**D. Miller**, University of Washington, Seattle
**L. Mills**, University of Minnesota, Minneapolis

COURSE TEACHING ASSISTANTS

**N. Damaraju**, University of Washington, Seattle
**S. Gibson**, University of Washington, Seattle
**R. Moss**, University of Minnesota, Minneapolis
**A. Oliveira de Lima**, University of Washington, Seattle


This course presented a comprehensive overview of the theory and practice of computational methods for the characterization of functional elements in DNA and RNA sequence data.

The course helped students achieve a deep, algorithmic understanding of the technologies and methods used to reveal genome function; the goal was to push beyond basic data analysis and into experimental design and the development of new analysis strategies.

The course enabled students to extract the maximum amount of correct information from data by developing a broad understanding of genomic analysis approaches and their shortcomings.

Topics included:

- Protein and DNA sequence similarity, comparisons, multiple alignments, and database searches
- Alignment and analysis of high-throughput sequencing data, with applications from RNA-seq and ChIP-seq experiments
- Analysis environments including Galaxy, RStudio, and the UNIX command line, with a strong focus on reproducible research
- Statistical considerations in the design and analysis of genomic experiments
- Regulatory element and motif identification from conserved signals in aligned and unaligned sequences

- Integration of genetic and sequence information in biological databases
- Genome browsers and features

The course combined lectures with hands-on exercises; students were encouraged to pose challenging problems using their own data. It was designed for biologists seeking advanced training in sequence and genome analysis, computational biology core resource directors and staff, and individuals in other disciplines (e.g., computer science) who wished to survey current research problems in biological sequence analysis. Advanced programming skills were not required.

This course was supported in part by a grant from the National Human Genome Research Institute. Additional scholarship support was provided by the Howard Hughes Medical Institute and Regeneron.

## PARTICIPANTS

Brooke, G., B.S., Kansas State University, Manhattan
Burt, T., M.D., Duke University School of Medicine, Durham, North Carolina
Chan, T., B.S., Texas A&M University, College Station
Deosthale, P., M.S., University of California, San Francisco
Ghosh, R., B.S., M.S., New York Genome Center, New York
Gouveia, M., Ph.D., National Institutes of Health, Bethesda, Maryland
Hager, E., Ph.D., Boston University, Massachusetts
Janto, N., B.S., Emory University, Atlanta, Georgia
Jiao, G., B.S., Purdue University, West Lafayette, Indiana
Kerslake, R., Ph.D., University of California, Santa Cruz
Kobiowu, A., M.Sc., Virginia Tech, Blacksburg
Lotthammer, J., B.S., Washington University in St. Louis, Missouri
Pizzagalli, M., B.S., Brown University, Providence, RI
Ramdass, A., B.Sc., The University of the West Indies, St. Augustine, Trinidad and Tobago
Redpath, K., B.Sc., University of Otago, Dunedin, New Zealand
Rico, J., B.S., Stanford University, Palo Alto, California
Seitz, E., Ph.D., Cold Spring Harbor Laboratory
Swan, S., B.S., University of Chicago, Illinois
Testa-Silva, G., Ph.D., Institute of Molecular and Clinical Ophthalmology, Basel, Switzerland
van 't Erve, I., M.Sc., Stanford University School of Medicine, California
Vouzas, A., B.S., Cold Spring Harbor Laboratory
Wang, H., B.S., New York University, New York
Yusuf, N., B.S., Rutgers University, Highland Park, New Jersey

## SEMINARS

Damaraju, N., University of Washington, Seattle: Haplotype analysis.
Garrison, E., University of Tennessee Health Science Center, Memphis (attended virtually): Pangenomic methods and applications.
Gibson, S., University of Washington, Seattle: Shiny?
Gifford, C., Stanford University, California (attended virtually): Designing context-specific functional genomics assays; Networks/multimodal data combinatorics, CRISPR screening.
Hawkins, D., University of Washington School of Medicine, Seattle: Introduction and overview; Regulatory genomics; Chromatin states 1: analysis of histone modifications; Chromatin states 2: overlapping data sets; Probing higher-dimension chromatin structure.
Hicks, S., Johns Hopkins, Baltimore, Maryland (attended virtually): Introduction to R; Reproducible research; Tidyverse and data viz; Introduction to Bioconductor; Overview of single-cell analysis in R/Bioconductor; Scaling up single-cell and spatial genomic data science.
Miller, D., University of Washington, Seattle: Long-read sequencing; Introduction and overview; IGV for visualizing sequencing data.
Mills, L., University of Minnesota, Minneapolis: Introduction and overview; Intro to genomic sequencing technology—the short and long of it all; RNA sequencing; Remote computing basics: SLURM, writing scripts and submitting jobs; Transcriptomics; Accessing public sequence data sets.
Krishnaswamy, S., Yale University, Madison, Connecticut (attended virtually): Analyzing single-cell data—PHATE/multiscale PHATE/MELD; Single cell analysis II—MAGIC/Pseudo time and trajectory net.
Pearson, W., University of Virginia, Charlottesville: Protein evolution/similarity searching; Practical sequence similarity searching; PSSMs and HMMs—customized scoring matrices.
Quinlan, A., University of Utah, Salt Lake City (attended virtually): An Introduction to bedtools; Discovery and analysis of genetic variation; More genetic variation.

# SEMINARS

## INVITED SPEAKERS PROGRAM ("CSHL SEMINAR SERIES")

Each year, Cold Spring Harbor Laboratory invites speakers from outside the institution to present their findings. These weekly seminars keep the CSHL staff current on the latest scientific developments and broaden their perspectives. Graduate students and postdoctoral fellows meet with the seminar speakers for lunch immediately after the seminar, providing for the exchange of ideas in an informal setting.

| Speaker | Title | Host |
|---|---|---|
| **January** | | |
| Gul Dolen, Ph.D., Associate Professor of Neuroscience, Brain Science Institute, Wendy Klag Center, Johns Hopkins University | Reopening critical periods with psychedelics: basic mechanisms and therapeutic opportunities | Stephen Shea |
| Samantha A. Morris, Ph.D., Associate Professor of Genetics and Developmental Biology, Washington University School of Medicine in St. Louis | Multi-omic lineage tracing: insights into reprogramming cell identity | Jessica Tollkuhn |
| Shruti Naik, Ph.D., Assistant Professor, Department of Pathology, Department of Medicine, Ronald O. Perelman Department of Dermatology, NYU Langone Health | Immune-mediated mechanisms of adaptive and maladaptive tissue responses | WiSE |
| **February** | | |
| Emma Teeling, Ph.D., Professor, School of Biology and Environmental Science, University College Dublin, Ireland | What can bat genomes teach us and why should we care? | Dick McCombie |
| Luis de Lecea, Ph.D., Professor of Psychiatry and Behavioral Sciences, Stanford | Sleep/wake control during life span | Jeremy Borniger |
| Mike Rosen, Ph.D., Professor and Chair, Department of Biophysics, UT Southwestern; Investigator, HHMI | Cell organization by liquid–liquid phase separation | Leemor Joshua-Tor |
| **March** | | |
| Janelle Ayres, Ph.D., Professor and Laboratory Head, Gene Expression Laboratory, Molecular and Systems Physiology Laboratory, NOMIS Center for Immunobiology and Microbial Pathogenesis, Salk Institute | Host–pathogen interactions: harnessing co-evolution to treat disease | T. Janowitz |
| Hao Wu, Ph.D., Asa and Patricia Springer Professor of Structural Biology, Professor of Biological Chemistry and Molecular Pharmacology, Harvard Medical School | Inflammasomes at the crossroad of health and disease | Hiro Furukawa |
| Beth Stevens, Ph.D., Research Associate in Neurology, Boston Children's Hospital; Associate Professor of Neurology, Harvard Medical School; Investigator, HHMI | Mapping microglia states to function in development and disease | Lucas Cheadle |
| Elizabeth Villa, Ph.D., Associate Professor, Molecular Biology, University of California, San Diego; Investigator, HHMI | Opening windows into the cell: bringing structure into cell biology using cryo-electron tomography | WiSE |
| **April** | | |
| Erica Glasper, Ph.D., Associate Professor, Department of Neuroscience, The Ohio State University, College of Medicine | Breaking bonds: neurobiological consequences of loss across the life span | DiAS |

| Speaker | Title | Host |
|---|---|---|
| Petr Svoboda, Ph.D., Group Leader, Laboratory of Epigenetic Regulations, Institute of Molecular Genetics of the Czech Academy of Sciences | Conserved and divergent features of mammalian RNA silencing | Andrea Schorn |
| Xuemei Chen, Ph.D., Dean of the School of Life Sciences, Peking University, Beijing, China | Noncanonical RNA caps | Rob Martienssen |
| Lisa Monteggia, Ph.D., Professor of Pharmacology, Barlow Family Director of the Vanderbilt Brain Institute | Mechanism of rapid antidepressant action | Linda Van Aelst |
| **October** | | |
| Ted Farmer, Ph.D., Professor, University of Lausanne | Electrical signaling in wounded plants | Dave Jackson |
| Oliver Hobert, Ph.D., Professor, Department of Biochemistry and Molecular Biophysics, Columbia University; Investigator, HHMI | Male and female brains: lessons from *C. elegans* | Chris Hammell |
| Liqun Luo, Ph.D., Professor Stanford University; Investigator, HHMI | Wiring specificity of neural circuits | Linda Van Aelst |
| Mike Summers, Ph.D., Professor of Chemistry and Biochemistry, University of Maryland, Baltimore County; Investigator, HHMI | Mechanism that controls HIV-1 transcript function and genome packaging<br>The Meyerhoff Scholars—a successful and replicable approach for promoting inclusive excellence in STEM | Leemor Joshua-Tor |
| **November** | | |
| Denis Jabaudon, Ph.D., Faculty of Medicine, Department of Basic Neurosciences, Group Leader, Developmental neurobiology and plasticity, Geneva University | Untwining space and time in the developing brain | Postdocs |
| Daniel Schramek, Ph.D., Principal Investigator, LTRI; Assistant Professor, Dept. of Molecular Genetics, U of Toronto; Canadian Research Chair in Functional Cancer Genomics | In vivo veritas—using CRISPR technologies in mice to unravel novel cancer driver alterations and novel drug targets | Camila dos Santos |
| **December** | | |
| Ashani T. Weeraratna, Ph.D., E.V. McCollum Chair of Biochemistry and Molecular Biology, Johns Hopkins Bloomberg School of Public Health; Bloomberg Distinguished Professor of Cancer Biology, Co-Program Leader, Cancer Invasion and Metastasis, Sidney Kimmel Cancer Center | Age against the machine: how the aging microenvironment drives tumor progression | WiSE |
| Norbert Perrimon, Ph.D., Professor of Genetics at Harvard Medical School; Investigator, HHMI | A holistic understanding of inter-organ communication and metabolic regulation in *Drosophila* | Tobias Janowitz |

## CSHL IN-HOUSE SEMINAR PROGRAM

Cold Spring Harbor's In-House Seminars were initiated to provide a semiformal avenue for communication among the various research groups at the Laboratory. The seminars also afford a necessary opportunity for graduate students and postgraduate staff to develop their skills in organizing, presenting, and defending their research.

| Speaker | Title |
|---|---|
| **January** | |
| Rasmani Hazra | Platr4 is an early embryonic lncRNA that exerts its function downstream on cardiogenic mesodermal lineage commitment |
| Arka Banerjee | Neural circuits for vocal communication |
| Michael Wigler and Dan Levy | Counting cancer genomes in a sample |
| **February** | |
| Qian Zhang | Antisense therapy in mouse models of histone H3.3K27M diffuse midline glioma inhibits tumor growth |
| Molly Gale Hammel | Diseases of the nervous system: the contribution of ERVs, LINEs, and other transposable elements to neurodegenerative disease |
| Jose M. Adrover | A vascular-restricted, tumor-induced neutrophil population drives vascular occlusion, pleomorphic necrosis, and metastasis |
| Corina Amor Vegas | Deconstructing aging with senolytic CAR T cells |
| **March** | |
| Peter Westcott | Modeling the evolution of immune dysfunction in cancer |
| Nicholas Tonks | Protein tyrosine phosphatases: mighty oaks from little acorns grow |
| Lucas Cheadle | Sensory experience engages brain-resident immune cells to coordinate neural development and function |
| Lloyd Trotman | Hiding in the shadows, hiding in plain sight: a tumor suppressor comes to light |
| **April** | |
| Semir Beyaz | Defining metabolic switches that boost the fitness of anti-tumor immunity |
| Benjamin Cowley | Compact deep neural network models of visual cortex |
| **October** | |
| Christopher Vakoc | Transcription factor interactions revealed by genetic screens |
| Ledong Wan | Dysregulated spliceosomal components promote pancreatic cancer progression |
| Corina Amor Vegas | Deconstructing aging with senolytic CAR T cells |
| **November** | |
| Xiaosa (Jack) Xu | Single-cell analysis of plant shoot meristems opens a "goldmine" for functional studies |
| Sergey Shuvaev | A normative theory of social conflict |
| **December** | |
| Andrea Schorn | Small RNA regulation of endogenous retroviruses |

# BANBURY CENTER

# BANBURY CENTER EXECUTIVE DIRECTOR'S REPORT

Since 1978, Cold Spring Harbor Laboratory's Banbury Center has convened impactful discussion meetings that allow small groups of experts to debate important issues, inspire fresh thinking, and form new connections. Meetings are organized around issues and challenges in the biosciences that benefit from the Center's unique style of discussion: emerging issues in need of strategy, established fields in need of review, controversial subjects calling for compromise or consensus, and areas in which diverse stakeholders/sectors need to engage or collaborate.

The year 2023 was expected to bring a return to pre-COVID business. However, lingering challenges reminded us that the pandemic's effects would not be effortlessly undone. We continued to need virtual participation for special circumstances; however, the spirit of open discussion and engagement remained.

## Activities

Despite the three-month pause, the estate was the site for more than 30 events in 2023, including traditional Banbury meetings, Meetings & Courses Program workshops, and laboratory retreats.

Banbury meeting themes in 2023 targeted important research and policy issues, with translational topics dominating the schedule. *The Future of Investigational Medicine: Utilizing Science to Optimize the Early Phase Oncology Clinical Trial Effort* convened experts from research and medicine to technology and economics to consider challenges and opportunities related to reliable trial methodology. Cancer remained on the agenda for two retreats that aimed to bring together Cold Spring Harbor Laboratory and Northwell Health colleagues. The first centered on pancreatic cancer, the second on brain tumors. These convenings were part of broader, long-term efforts to foster collaboration between the two institutions.

In addition to translational cancer topics, two Banbury meetings took aim at genetic disorders. In partnership with the FRAXA Research Foundation and Kennedy's Disease Association, the Center hosted *FMRP Restoration: Definitive Therapies in Fragile X* and the *Kennedy's Disease (SBMA) Research Workshop.* In each case, the meeting brought together researchers studying the condition, as well as those who work on related studies. The goal was to inspire new research questions and collaborations in hopes of driving progress for treatment.

Another Banbury meeting targeted an all-too-common occurrence in hospitals—sepsis. Fifty million people develop sepsis every year, with one in five dying when multiple organ systems fail and vital support is withdrawn. Participants in *Enhancing Recovery from Sepsis-Induced Organ Dysfunction* tackled questions such as: Why do vital organs stop functioning in sepsis, and once organs stop functioning, what processes allow them to start functioning again?

Rounding out our translational research topics was a meeting focused on *Integrating Exposomics into the Biomedical Enterprise*. A term describing the totality of environmental contributors to human health and disease, the field of "exposomics" has made slow progress in the biomedical industry.

Participants in this Banbury meeting aimed to develop an operational definition for the concept and strategized ways to better incorporate exposomics across biomedical research and practice.

Two meetings in 2023 centered on basic science research questions, both with translational potential. First, *Persistence, Senescence, and Cell Death* participants discussed research on cell states and fates, especially as related to survival in the face of otherwise fatal therapies. Practical applications in tumor cells may prove promising for the development of new cancer therapies. We moved from animals to plants for the second meeting. *The Future of Plant–Environment Interactions: Challenges and Opportunities in a Changing Climate* envisioned strategies to address the survival of plants (and thus humanity), despite the threats associated with climate change. Although many of the meeting's discussions and recommendations targeted policy needs, the meeting's core work centered on plant biology innovations.

Banbury convened two policy-focused meetings in 2023, on two very different topics. The recent resurgence of research on the therapeutic use of psychedelics, combined with FDA designations of certain psychedelics as breakthrough therapies for PTSD and depression, and the decriminalization of these drugs by several U.S. cities, has highlighted the need to consider ethical concerns with their use. *Developing an Ethical Framework for Psychedelics Research and Use* brought together policymakers, ethics scholars, mental health practitioners, and community representatives to consider ethical issues that have the potential to pose a risk to the responsible and safe uptake of psychedelics as important therapy options. We also convened a truly global group of experts to discuss the United Nations Sustainable Development Goals (SDG), specifically the practice of evaluating whether (and how) global health programs are effective in achieving their aims. Experts representing academic evaluation, country-level practical evaluation, citizen voices, and user perspectives came together to share experiences, discuss challenges, and identify ways to strengthen evaluation in the SDG agenda for health.

G. Waibel, I. Mulvany

The Banbury Center is an ideal location for discussions of scientific challenges, and the bucolic environment is also the perfect site for organizations to develop strategy and review progress. Each year, we welcome the Lustgarten Foundation for its *Scientific Meeting*, an opportunity to discuss cutting-edge biomedical science and progress toward treatments for pancreatic cancer. In 2023, we also hosted the Project Santa Fe Foundation's *CL2.0 Colloquium*, as well as a *Dryad Strategy Retreat*.

## Participants

In 2023, Banbury meeting participants represented the global scientific community, hailing from six continents, 30 countries, and 35 U.S. states. Ninety-six percent participated in person, with only 4% joining virtually. The hybrid arrangement still required significant effort from Audiovisual (AV) and Banbury staff, and created some distractions for all participants during the meeting.

Since 2018, Banbury has instituted efforts to improve gender diversity of our meetings. We have seen steady improvement, and in 2023 we neared balance, with 42% women or nonbinary participants. This ratio varied across meetings, ranging from 28% to 59% women or nonbinary participants. The Center is also committed to improving diversity in other demographic areas, including minoritized races/ethnicities. In 2022, the Center began collecting demographic data in these categories. With 98% of participants reporting in 2023, Banbury participants identified as

follows: 67% white; 17% Asian; 4% Black/African–American or African; 5% Hispanic, Latinx, or Spanish Origin; 3% Middle Eastern or North African; 0.7% Indigenous, American Indian, or Alaska Native; and 3% identifying as part of other minoritized groups. These numbers do not reflect the level of diversity we would like to achieve, and we are instituting efforts to improve in 2024.

## Support

Funding continues to be a major hurdle for Banbury meetings, as topics often fall at new intersections of science and technology or deal with delicate ethical or policy issues. We are ever grateful to the organizations and individuals that provide the financial support enabling the Center to convene global leaders. In 2022, Banbury secured 61% of its financial support from not-for-profit organizations, with our CSHL Corporate Sponsor Program and individual industry sources making up the remaining 39%.

## The Banbury Team

The Center is successful thanks to a team of professionals who ensure the estate and programs are running at a high level. In the main office, meetings organization and delivery are expertly managed by Vanessa Franco as Finance and Development Coordinator, Duncan Yates as Lodging Manager, and Jenna Jacobs as Communications and Special Projects Coordinator. Paulo Krizanovski, Juan Colocho, and Trinidad Velasquez skillfully maintained 55 acres of impeccable grounds, keeping the estate beautiful and accessible. AV staff, particularly Ira Russo, Bill Dickerson, and Robert Eifert, battled technological hiccups without disrupting meeting discussions, and housekeepers ensured our guest rooms and convening spaces are welcoming to visitors. Constance Brukin's gorgeous photography captured important memories and continues to build our historical archive. As always, we are grateful to all of our colleagues and partners across Cold Spring Harbor Laboratory.

**Rebecca Leshan**
*Executive Director*

## 2023 Publications Resulting from Banbury Meetings

Amaral P, Carbonell-Sala S, De La Vega FM, Faial T, Frankish A, Gingeras T, Guigo R, Harrow JL, Hatzigeorgiou AG, Johnson R, et al. 2023. The status of the human gene catalogue. *Nature* **622:** 41–47. doi:10.1038/s41586-023-06490-x

McClarty LM, Becker ML, García PJ, Garnett GP, Dallabetta GA, Ward H, Aral SO, Blanchard JF. 2023. Programme science: a route to effective coverage and population-level impact for HIV and sexually transmitted infection prevention. *Lancet HIV* **10:** e825–e834. doi:10.1016/S2352-3018(23)00224-2

Rahimzadeh V, Fogarty J, Caulfield T, Auñón-Chancellor S, Borry P, Candia J, Cohen IG, Covington M, Lynch HF, Greely HT, et al. 2023. Ethically cleared to launch? *Science* **381:** 1408–1411. doi:10.1126/science.adh9028

Seifert AW, Duncan EM, Zayas RM. 2023. Enduring questions in regenerative biology and the search for answers. *Commun Biol* **6:** 1139. doi:10.1038/s42003-023-05505-7

Williams JJ, Tractenberg RE, Batut B, Becker EA, Brown AM, Burke ML, Busby B, Cooch NK, Dillman AA, Donovan SS, et al. 2023. An international consensus on effective, inclusive, and career-spanning short-format training in the life sciences and beyond. *PLoS ONE* **18:** e0293879. doi:10.1371/journal.pone.0293879.

# BANBURY CENTER MEETINGS

| *Dates* | *Title* | *Organizer(s)* |
|---|---|---|
| January 23–26 | Persistence, Senescence, and Cell Death | Judith Campisi, Anthony G. Letai |
| February 9 | Cold Spring Harbor Laboratory–Northwell Health Affiliation: Pancreatic Cancer Meeting | David Tuveson |
| March 13 | Cold Spring Harbor Laboratory–Northwell Health Affiliation: Brain Tumor Workshop | John Boockvar, Alea Mills |
| March 23–26 | The Future of Investigational Medicine: Utilizing Science to Optimize the Early Phase Oncology Clinical Trial Effort | Tobias Janowitz, Lillian Siu |
| April 16–19 | Enhancing Recovery from Sepsis-Induced Organ Dysfunction | Derek C. Angus, Steven Q. Simpson, Kevin Tracey |
| April 23–26 | Dryad Strategy Retreat | |
| April 30–May 3 | FMRP Restoration: Definitive Therapies in Fragile X | Glenn Cohen, Amy McGuire, Joel D. Richter, |
| June 9–12 | Developing an Ethical Framework for Psychedelics Research and Use | Dominic Sisti, Michael Tranfaglia |
| September 10–12 | Kennedy's Disease (SBMA) Research Workshop | J. Gibson |
| September 27–29 | The 2023 CL2.0 Colloquium | The Board of Directors, Project Santa Fe Foundation LLC |
| October 10–13 | Strengthening the Role of Evaluation in the Sustainable Development Agenda for Health | Kerry Albright, Dugan Fraser, Geoff Garnett, James Hargreaves, Mira Johri, Miriam Sabin, Kabir Sheikh |
| October 23–25 | Lustgarten Foundation Scientific Advisory Board Meeting | Andrew Rakeman, Linda Tantawi, David Tuveson |
| October 29–November 1 | The Future of Plant–Environment Interactions: Challenges and Opportunities in a Changing Climate | Ullas Pedmale, Lucia Strader |
| December 3–6 | Integrating Exposomics into the Biomedical Enterprise | L. Michelle Bennett, Gary W. Miller |

# BANBURY CENTER MEETINGS

## Persistence, Senescence, and Cell Death

**January 23–26**

ARRANGED BY **J. Campisi,** Buck Institute for Research on Aging, Novato, California
**A.G. Letai,** Dana-Farber Cancer Institute, Boston, Massachusetts

FUNDED BY **Genentech and The Cold Spring Harbor Laboratory Corporate Sponsor Program**

Recently, there has been an increased appreciation of the opportunity for some cancer cells to enter a distinct epigenetic state that permits survival in the face of therapies that would otherwise be fatal to the cells. Cells in such states have received names like "persisters" and "resisters." This meeting brought together investigators in the fields of cellular persistence, senescence, and cell death to tackle questions such as: What is the overlap between the persister state and the senescent state? Is persistence a reversible cell fate? Are there therapeutic vulnerabilities specific to persister cells? Understanding mechanisms of persistence and resistance is critical to identifying novel therapies to overcome these states in tumor cells.

**Welcoming Remarks: R. Leshan,** Executive Director, The Banbury Center, Cold Spring Harbor Laboratory

**Welcome and Meeting Objectives: A.G. Letai,** Dana-Farber Cancer Institute, Boston, Massachusetts
**J. Campisi,** Buck Institute for Research on Aging, Novato, California (virtual)

S. Lee, S.W. Lowe

J.A. MacDonald, A. Roe, A.G. Letai

**SESSION 1:** Senescence

**Chairperson: C.A. Schmitt,** Max Delbrück Center for Molecular Medicine, Berlin, Germany

D. Zhou, University of Texas Health, San Antonio: Using PROTAC technology to develop better senolytics.

J. Kirkland, Mayo Clinic, Rochester, Minnesota: Cellular senescence and senolytics: the path to translation.

J. Campisi, Buck Institute for Research on Aging, Novato, California (virtual): Cellular senescence: yin and yang.

M. Xu, University of Connecticut Health, Farmington: The heterogeneity of cellular senescence.

**SESSION 2:** Persistence after Drug Treatment

**Chairperson: C.S. Mitsiades,** Dana-Farber Cancer Institute, Boston, Massachusetts

J. Montero, University of Barcelona, Spain: Rational therapeutic combinations with BH3 mimetics to prevent cancer resistance to treatment.

J.G. Jackson, Tulane University School of Medicine, New Orleans, Louisiana: Mechanisms of tumor cell survival and persistence in chemotherapy treated breast cancer.

**SESSION 3:** Stress and Cell Fate Changes

**Chairperson: M. Hangauer,** University of California, San Diego, La Jolla

J.E. Dick, Princess Margaret Cancer Centre, University Health Network, Toronto, Ontario, Canada: Stem cells, aging and the origins of leukemia in humans.

C. Muñoz-Pinedo, IDIBELL, Barcelona, Spain: Metabolic stress drives a SASP-like response to recruit innate immune cells.

S.W. Lowe, HHMI, Memorial Sloan Kettering Cancer Center, New York, New York: How does p53 suppress cancer? Persistence, senescence, and cell death.

A. Shahbandi, Tulane University School of Medicine, New Orleans, Louisiana: How senescent cancer cells evade the immune system to cause relapse.

S. Lee, Johannes Kepler University, Linz, Austria: Virus-induced senescence and COVID-19.

**SESSION 4:** Cancer Cell Persistence

**Chairperson: C. Amor Vegas,** Cold Spring Harbor Laboratory

J.S. Brugge, Harvard Medical School, Boston, Massachusetts: Phenotypes of cancer therapy resistant persister cells.

T. Ní Chonghaile, Royal College of Surgeons in Ireland, Dublin: Identifying sites of resistance to ABT-199 treatment.

C.A. O'Brien, Toronto General Hospital, Ontario, Canada: GI cancer hibernation and persistence.

**SESSION 5:** Novel Mechanisms of Drug Resistance

**Chairperson: C.A. O'Brien,** Toronto General Hospital, Ontario, Canada

K.A. Sarosiek, Harvard T.H. Chan School of Public Health, Boston, Massachusetts: Broad sensitivity or resistance to cancer therapies is driven by dynamic regulation of cell death pathways in cell origin.

M.K. Classon, Pfizer Center for Therapeutic Innovation, South San Francisco, California: De-regulation of repetitive genomic elements (TEs) as an emerging therapeutic opportunity in cancer.

C.A. Schmitt, Max Delbrück Center for Molecular Medicine, Berlin, Germany: Persisting senescence in cancer—clinical opportunities and challenges.

**SESSION 6:** Persistence and Cell Fate

**Chairperson: C. Muñoz-Pinedo,** IDIBELL, Barcelona, Spain

C. Amor Vegas, Cold Spring Harbor Laboratory: Harnessing senolytic CAR T cells to reverse and prevent aging.

A.G. Letai, Dana-Farber Cancer Institute, Boston, Massachusetts: Apoptosis vulnerabilities in persisters and senescent cells.

S.A. Stewart, Washington University in St. Louis, Missouri: Therapy-induced senescence drives bone loss.

A. Roe, Royal College of Surgeons in Ireland, Dublin: Small molecule screen identifies compound capable of killing apoptotic resistant cells.

C.S. Mitsiades, Dana-Farber Cancer Institute, Boston, Massachusetts: Tumor cell persistence through cytotoxic pharmacological therapies or immune effector cells.

M. Hangauer, University of California, San Diego, La Jolla: Targeting cancer persister cells.

S.W.G. Tait, Cancer Research UK Beatson Institute, Glasgow, United Kingdom: Mitochondrial driven inflammation in cell death and senescence.

J.A. MacDonald, Dana-Farber Cancer Institute, Boston, Massachusetts: Mitochondrial apoptotic priming in senescence.

D.R. Green, St. Jude Children's Research Hospital, Memphis, Tennessee: Persistence, resistance, and the Bcl-2 family.

**SESSION 7:** Mitochondria and Persistence

**Chairperson: K.A. Sarosiek,** Harvard T.H. Chan School of Public Health, Boston, Massachusetts

**SESSION 8:** Meeting Wrap-Up, Next Steps

**Chairperson: A.G. Letai,** Dana-Farber Cancer Institute, Boston, Massachusetts

# Cold Spring Harbor Laboratory–Northwell Health Affiliation: Pancreatic Cancer Meeting

**February 9**

ARRANGED BY **D.A. Tuveson,** Cold Spring Harbor Laboratory

FUNDED BY **Cold Spring Harbor Laboratory–Northwell Health Affiliation**

In 2015, Cold Spring Harbor Laboratory and Northwell Health entered into a strategic affiliation that combines mechanistic biology with the highest level of healthcare to advance diagnosis and treatment. Since that time, the Affiliation has launched new education initiatives, expanded access to clinical trials, and provided funding for new investigations that will have a rapid impact in the clinic. The 2023 Pancreatic Cancer Meeting brought together Cold Spring Harbor Laboratory and Northwell experts working to advance new discoveries and treatments.

D.A. Tuveson, Cold Spring Harbor Laboratory: Welcome, overview, and goals

A. Krasnitz, Cold Spring Harbor Laboratory; N. Chambwe, Feinstein Institutes for Medical Research, Manhasset, New York; and A.N. Habowski, Cold Spring Harbor Laboratory: Disparities and etiology for African Americans with PDAC; genomics; pharmaco-omics.

P. Ferguson, Cold Spring Harbor Laboratory: Obesity and pancreatic cancer—new insights.

S. Nadella and J. Kastan, Cold Spring Harbor Laboratory: Inflammation and pancreatic cancer—new ideas.

D.A. King, Feinstein Institutes for Medical Research, New Hyde Park, New York; A. Rishi, Northwell Health, Greenvale, New York; and M.J. Weiss, Northwell Health, New Hyde Park, New York: Pancreatitis and pancreatic cancer clinical trial—window of opportunity.

K.J. Tracey, Feinstein Institutes for Medical Research, Manhasset, New York, and J. Nigri, Cold Spring Harbor Laboratory: Nerves in pancreatitis and PDAC.

G. Caligiuri, Cold Spring Harbor Laboratory: CAFs in PDAC.
M. Shakiba, Cold Spring Harbor Laboratory: Immune infiltrates in PDAC.
H. Ting, Cold Spring Harbor Laboratory: KRAS as a target in PDAC.
C. Tonelli, Cold Spring Harbor Laboratory: Redox as a targetable pathway in PDAC.
J.E. Moses, Cold Spring Harbor Laboratory: New drugs targeting SOAT1.
A.N. Habowski, Cold Spring Harbor Laboratory: PASS-01.
T. Janowitz, Cold Spring Harbor Laboratory, and D.A. King, Feinstein Institutes for Medical Research, Manhasset, New York: Early phase clinical trial unit for PDAC.
J. Kastan, Cold Spring Harbor Laboratory: New biomarkers for PDAC.
S. Lyons and J. Yeh, Cold Spring Harbor Laboratory: Theranostics with novel antibodies for PDAC.
D.A. Tuveson, Cold Spring Harbor Laboratory: Next steps, opportunities, challenges.

# Cold Spring Harbor Laboratory–Northwell Health Affiliation: Brain Tumor Workshop

**March 13**

ARRANGED BY **J.A. Boockvar,** Lenox Hill Hospital and Zucker School of Medicine at Hofstra/Northwell, New York, New York
**A.A. Mills,** Cold Spring Harbor Laboratory

FUNDED BY **Cold Spring Harbor Laboratory–Northwell Health Affiliation**

In 2015, Cold Spring Harbor Laboratory and Northwell Health entered into a strategic affiliation that combines mechanistic biology with the highest level of healthcare to advance diagnosis and treatment. Since that time, the Affiliation has launched new education initiatives, expanded access to clinical trials, and provided funding for new investigations that will have a rapid impact in the clinic. The 2023 Brain Tumor Workshop brought together Cold Spring Harbor Laboratory and Northwell experts working to advance new discoveries and treatments for brain tumors.

**Welcome and Meeting Objectives:** **A.A. Mills,** Cold Spring Harbor Laboratory
**J.A. Boockvar,** Lenox Hill Hospital, Zucker School of Medicine at Hofstra/Northwell, New York, New York

## SESSION 1

**Chairperson: M. Schulder,** Northwell Health, Lake Success, New York

A.R. Krainer, Cold Spring Harbor Laboratory: Antisense strategies for H3.3K27M diffuse midline glioma.

M. Schulder, Northwell Health, Lake Success, New York: Focused ultrasound for patients with glioblastoma.

## SESSION 2

**Chairperson: A. Goenka,** Northwell Health, New Hyde Park, New York

X. He, Northwell Health, Manhasset, New York: Blood–brain barrier permeability imaging.

N.K. Tonks, Cold Spring Harbor Laboratory: Novel small molecule inhibitors of DYRK1A and their application in models of glioblastoma.

A. Goenka, Northwell Health, New Hyde Park, New York, and A.G. Wernicke, Northwell Health, New York, New York: New techniques in radiotherapy.

## SESSION 3

**Chairperson: D.B. Gruber,** Northwell Health, Lake Success, New York

L.M. Cheadle, HHMI, Cold Spring Harbor Laboratory: Non-neuronal brain cells in health and disease.

S. Singer, Northwell Health, New Hyde Park, New York, and D.B. Gruber, Northwell Health, Lake Success, New York: Update in NeuroOnc clinical trials.

J.C. Borniger, Cold Spring Harbor Laboratory: Glucocorticoids and chronotherapy in cancer.

## SESSION 4

**Chairperson: J.I. Krystal,** Northwell Health, New Hyde Park, New York

J.I. Krystal, Northwell Health, New Hyde Park, New York: Clinical trials in pediatric neuro-oncology.

X. Sun, Cold Spring Harbor Laboratory: Oncogenic chromatin remodeling in glioblastoma.

M. Vojnic and R.S. D'Amico, Northwell Health, New York, New York: CNS metastases.

## SESSION 5

**Chairpersons: A.A. Mills,** Cold Spring Harbor Laboratory; **J.A. Boockvar,** Lenox Hill Hospital, Zucker School of Medicine at Hofstra/Northwell, New York, New York

A.A. Mills, Cold Spring Harbor Laboratory; J.A. Boockvar, Lenox Hill Hospital, Zucker School of Medicine at Hofstra/Northwell, New York, New York: General discussion and closing remarks.

# The Future of Investigational Medicine: Utilizing Science to Optimize the Early Phase Oncology Clinical Trial Effort

**March 23–26**

ARRANGED BY **T. Janowitz,** Cold Spring Harbor Laboratory
**L.L. Siu,** Princess Margaret Cancer Centre, Toronto, Ontario, Canada

FUNDED BY **The Cold Spring Harbor Laboratory–Northwell Health Affiliation**

Improving current clinical trial approaches can improve healthcare for all. This Banbury Center meeting brought together experts to discuss strategies for overcoming challenges related to leveraging advances in scientific discoveries, biomedical research technologies, biomarker development and testing, patient stratification, equitable patient enrollment, data collection and analysis using advanced statistical methods, and converting new strategies into reliable trial methodology together with the regulatory agencies.

**Welcoming Remarks: R. Leshan,** Executive Director, Banbury Center, Cold Spring Harbor Laboratory

**Welcome and Meeting Objectives: T. Janowitz,** Cold Spring Harbor Laboratory
**L.L. Siu,** Princess Margaret Cancer Centre, Toronto, Ontario, Canada

R.G. Maki, Memorial Sloan Kettering Cancer Center, New York, New York: Cancer clinical trials: abject failures, stunning successes and some big ideas.

**SESSION 1A:** Trial Operations: Collaborations across Disciplines, Sectors, and Institutions to Enable Early Phase Clinical Trials

M.D. Goncalves, T. Janowitz

S.J. Mandrekar, R.G. Maki, A. Whiteley, S.M. Rothenberg, E. Fox

**Chairperson: Y. Shyr,** Vanderbilt University Medical Center, Nashville, Tennessee

S.J. Mandrekar, Mayo Clinic, St. Augustine, Florida: Statistical trial optimization for early phase clinical trials.

S.M. Rothenberg, Pfizer Inc., Boulder, Colorado: Industry perspective on innovative, accelerated early phase trials and go/no-go decision-making.

D.S.W. Tan, National Cancer Centre Singapore, Singapore: Local versus regional versus global trial collaborations and delivery.

**SESSION 1B:** Trial Operations: Regulations and Cost-Effective Delivery of Early Phase Clinical Trials

**Chairperson: Victoria Aranda,** *Nature*, New York, New York

A.W. Lo, Massachusetts Institute of Technology, Cambridge, Massachusetts (virtual): Predictions for early phase trial economics.

N.J. Dashdorj, Onom Foundation, Ulaanbaatar, Mongolia: Low-income country clinical trials.

K. Liu, Marengo Therapeutics, Cambridge, Massachusetts: Including the FDA in data-driven trial design.

**SESSION 2:** Incorporating New Technologies and Biomarkers into Clinical Trials

**Chairperson: T. Janowitz,** Cold Spring Harbor Laboratory

M.C. Liu, Natera Inc., San Carlos, California: Incorporating new technologies and biomarkers into clinical trials—the technical perspective.

L.L. Siu, Princess Margaret Cancer Centre, Toronto, Ontario, Canada: Incorporating of new technologies and biomarkers into clinical trials—the clinician's perspective.

C. Yap, Institute of Cancer Research, Sutton, United Kingdom: Utilising novel trial designs to improve efficiency in clinical trials.

**SESSION 3:** What Are the Best Trial Endpoints, Especially If Remote Trial Technology Is Utilized?

**Chairperson: S. Kleeman,** Cold Spring Harbor Laboratory

M.S. Beg, Science 37, Dallas, Texas: Decentralized clinical trials to enable early phase and precision oncology clinical trials.

N.J. Meropol, Flatiron Health, New York, New York: New technologies to enable the conduct of efficient and representative clinical trials.

C.M. Hartshorn, National Institutes of Health (NIH), Bethesda, Maryland: Moving towards more decentralized clinical research and translational science with digital health technologies: efforts at the NIH to accelerate.

**SESSION 4:** Patient Inclusion and Trial Access

**Chairperson: R.G. Maki,** Memorial Sloan Kettering Cancer Center, New York, New York

D.E. Collyar, Patient Advocates In Research (PAIR), Danville, California: Patient wishes for early phase clinical trials.

K.M. Mustian, Wilmot Cancer Institute, Rochester, New York: Recruitment and clinical trial execution in community-based oncology practices.

E. Fox, St. Jude Children's Research Hospital, Memphis, Tennessee: Overcoming barriers for children in trials.

N. Chambwe, Feinstein Institutes for Medical Research, Manhasset, New York: Inclusion of African–American patients in early phase clinical trials.

**SESSION 5:** Combinatorial Strategies in Drug Development—How to Nominate the Most Relevant and Truly Synergistic Combinations for Clinical Testing

**Chairperson: K.A. Schalper,** Yale University, New Haven, Connecticut

H.X. Chen, National Cancer Institute Cancer Treatment Evaluation Program, Bethesda, Maryland: Biomarker-driven combination approaches for cancer immunotherapy.
M.D. Goncalves, Weill Cornell Medicine, New York, New York: Combination of diet and drugs to target cancer.
D.A. King, Feinstein Institutes for Medical Research, Manhasset, New York: Leveraging organoid technology for personalized drug combination decisions.

**SESSION 6:** Biological Validation in Early Phase Trials: Approaches and Implementation Strategies

**Chairpersons: N. Chambwe,** Feinstein Institutes for Medical Research, Manhasset, New York, and **K.T. Flaherty,** Massachusetts General Hospital, Cambridge

K.T. Flaherty, Massachusetts General Hospital, Cambridge, Massachusetts: Precision oncology platform trials: Where is efficiency gained and lost?
K.A. Schalper, Yale University, New Haven, Connecticut: Precision immuno-oncology? Role of next-generation tissue-based molecular studies in clinical research and patient care.
C. Le Tourneau, Institut Curie, Paris, France: Design of next-generation precision medicine trials using AI.
Y. Shyr, Vanderbilt University Medical Center, Nashville, Tennessee: High-dimensional data analysis to inform trial enrollment and analysis.
T. Janowitz, Cold Spring Harbor Laboratory: Combination tracing of biology and patient-reported outcomes in early phase clinical trials.

**SESSION 7A:** Summarizing & Consensus-Building

**Chairpersons: T. Janowitz,** Cold Spring Harbor Laboratory, and **L.L. Siu,** Princess Margaret Cancer Centre, Toronto, Ontario, Canada

**SESSION 7B:** Meeting Output, Next Steps

**Chairpersons: T. Janowitz,** Cold Spring Harbor Laboratory, and **L.L. Siu,** Princess Margaret Cancer Centre, Toronto, Ontario, Canada

# Enhancing Recovery from Sepsis-Induced Organ Dysfunction

**April 16–19**

ARRANGED BY **D.C. Angus,** University of Pittsburgh, Pennsylvania
**S.Q. Simpson,** University of Kansas, Kansas City
**K.J. Tracey,** Feinstein Institutes for Medical Research, Manhasset, New York

IN PARTNERSHIP WITH **Sepsis Alliance**

FUNDED BY **Siemens Healthineers**

Each year, 50 million people develop sepsis, with one in five dying. Death occurs most commonly when it is recognized that patients are simply not recovering from multisystem organ failure, and therefore vital organ support is withdrawn. A huge body of work has uncovered mechanisms by which the body first mounts a response to microbial invasion; far less is understood about why multisystem organ dysfunction persists. No unifying theory successfully describes the underlying pathophysiology and links that pathophysiology to clinical disease expression or therapeutic targets. This meeting brought together experts working in sepsis and those with related expertise to consider critical yet unanswered questions such as: Why do vital organs stop functioning in sepsis? Once organs stop functioning, what processes allow them to start functioning again? How uniform or common are these processes across organ systems? Could the processes that govern dysfunction and subsequent recovery of function represent therapeutic targets to speed recovery and enhance survival?

**Welcoming Remarks: R. Leshan,** Executive Director, The Banbury Center, Cold Spring Harbor Laboratory

A. Wang, H. Bayir, D. Annane, R. Pirracchio

K.J. Tracey, D.C. Angus, A. Bihorac

**Welcome and Meeting Objectives: D.C. Angus,** University of Pittsburgh, Pennsylvania
**S.Q. Simpson,** University of Kansas, Kansas City
**K.J. Tracey,** Feinstein Institutes for Medical Research, Manhasset, New York

**SESSION 1:** Setting the Stage—Current Understanding of "Injury" and Recovery I

**Chairpersons: H. Bayir,** Columbia University, New York, New York, and **A. Wang,** Yale University, New Haven, Connecticut

D.C. Angus, University of Pittsburgh, Pennsylvania: Recovery from sepsis: defining the problem.

M. Singer, University College London, . United Kingdom: Prevailing conceptual models of organ failure and recovery.

A. Bihorac, University of Florida, Gainesville: Acute kidney injury: what is it, what causes it, and what might accelerate recovery?

**SESSION 2:** Setting the Stage—Current Understanding of "Injury" and Recovery II

**Chairpersons: B. Diamond,** Feinstein Institutes for Medical Research, Manhasset, New York, and **A. Ilanges**, Janelia Research Campus, HHMI, Ashburn, Virginia

T.D. Girard, University of Pittsburgh, Pennsylvania: Brain dysfunction in sepsis: What is it, what causes it, and what might accelerate recovery?

S.Q. Simpson, University of Kansas, Kansas City: Acute lung injury: What is it, what causes it, and what might accelerate repair and recovery?

## GENERAL DISCUSSION

**Chairpersons: L.A. O'Neill,** Trinity College Dublin, Dublin, Ireland, and **H. Wunsch**, University of Toronto, Ontario, Canada

**SESSION 3:** Setting the Stage—Current Understanding of "Injury" and Recovery III

**Chairpersons: R. Pirracchio,** University of California, San Francisco, and **H.V. Carey**, University of Wisconsin–Madison

C.W. Seymour, University of Pittsburgh, Pennsylvania: Heterogeneity in sepsis-induced organ dysfunction syndromes.

S.P. Taylor, University of Michigan, Ann Arbor: Long-term recovery trajectories from sepsis.

H. Wunsch, University of Toronto, Ontario, Canada: The problem is the ICU: drugs and other iatrogenic/nosocomial factors.

## GENERAL DISCUSSION

**Chairpersons: H. Bayir,** Columbia University, New York, New York, and **T.G. Buchman**, Emory University, Atlanta, Georgia

**SESSION 4:** Potential Culprits and Theories—Inflammation and Metabolism

**Chairpersons: A. Bihorac,** University of Florida, Gainesville, and **S.Q. Simpson,** University of Kansas, Kansas City

L.A. O'Neill, Trinity College Dublin, Ireland: It's the immune system's metabolism, stupid.

H. Bayir, Columbia University, New York, New York: It's the (parenchymal) mitochondria, stupid.

D.V. Bohórquez, Duke University, Cary, North Carolina: It's the gut, stupid.

**SESSION 5:** Potential Culprits and Theories—Neural Control and Related Top-Down Control Mechanisms

**Chairpersons: H.V. Carey,** University of Wisconsin–Madison, and **T.D. Girard,** University of Pittsburgh, Pennsylvania

R. Pirracchio, University of California, San Francisco: Surely it's the endocrine axis.

P.S. Olofsson, Karolinska Institutet, Stockholm, Sweden: Surely it's the neuroinflammatory axis.

A. Ilanges, Janelia Research Campus, HHMI, Ashburn, Virginia: Surely it's the brainstem axis.

**GENERAL DISCUSSION**

**Chairpersons: M. Singer,** University College London, United Kingdom, and **S.P. Taylor,** University of Michigan, Ann Arbor

**SESSION 6:** Looking for the Reboot Switch—How to Accelerate Return of Organ Function I

**Chairpersons: T.D. Girard,** University of Pittsburgh, Pennsylvania, and **H. Wunsch,** University of Toronto, Ontario, Canada

D. Chan, Massachusetts Institute of Technology, Cambridge: Using gamma frequency sensory stimulation to restore or preserve neurocognition.

H.V. Carey, University of Wisconsin–Madison: Exploring hibernation pathways to boost return of organ function.

D. Annane, Paris-Saclay University, Garches, France: Personalized corticotherapy in sepsis.

**SESSION 7:** Looking for the Reboot Switch—How to Accelerate Return of Organ Function II

**Chairpersons: D.V. Bohórquez,** Duke University, Cary, North Carolina, and **S.P. Taylor,** University of Michigan, Ann Arbor

B. Diamond, Feinstein Institutes for Medical Research, Manhasset, New York: Manipulating neural control of inflammation and somatic function.

A. Wang, Yale University, New Haven, Connecticut: Immunometabolic reprogramming.

**GENERAL DISCUSSION:** The Search for Reboot Switches

**Chairpersons: A. Bihorac,** University of Florida, Gainesville, and **C.W. Seymour,** University of Pittsburgh, Pennsylvania

**SESSION 8:** Toward an Integrated Vision of the Problem and Next Steps

**Chairpersons: D. Annane,** Paris-Saclay University, Garches, France, and **B. Diamond,** Feinstein Institutes for Medical Research, Manhasset, New York

T.G. Buchman, Emory University, Atlanta, Georgia: Thinking in an integrated way about multisystem organ dysfunction and recovery.

K.J. Tracey, Feinstein Institutes for Medical Research, Manhasset, New York: What would the multisystem organ (dys)function "moonshoot" look like?

**GENERAL DISCUSSION 5:** Summary of Key Themes around Problem Definition, Causes, Solutions and Future Research

**Chairpersons: L.A. O'Neill,** Trinity College Dublin, Ireland, and **M. Singer,** University College London, United Kingdom

# Dryad Strategy Retreat

**April 23–26**

ARRANGED BY **J. Gibson,** Dryad, London, United Kingdom

FUNDED BY **Dryad**

Dryad's Strategy Retreat was a convening of the Board of Directors and Executive Leadership, the first in recent memory. It was an opportunity to co-develop their strategic plan for the next three to five years.

**PARTICIPANTS**

S. Edmunds, GigaScience, Hong Kong, China
J. Gibson, Dryad, London, United Kingdom
M. Guerreiro, Dryad, Lisbon, Portugal
B. Hanson, American Geophysical Union, Washington, D.C.
J.L. Herzog, Dryad, Lincoln, Nebraska
M. Kurtz, Dryad, Boston, Massachusetts
S. Lippincott, Dryad, Tours, France
D.P. Madalli, Indian Statistical Institute, Bengaluru, India
I. Mulvany, The BMJ, London, United Kingdom
D. Okubo, Dryad, San Francisco, California
I. Puebla, ASAPbio, Cambridge, United Kingdom
J. Ruttenberg, Association of Research Libraries, Washington, D.C.
R. Scherle, Dryad, Durham, North Carolina
J. Treadway, Great North Wood Consulting, London, United Kingdom
G. Waibel, University of California Office of the President, California Digital Library, San Francisco, California
J. Williams, Cold Spring Harbor Laboratory DNA Learning Center, Cold Spring Harbor, New York

# FMRP Restoration: Definitive Therapies in Fragile X

**April 30–May 3**

ARRANGED BY **J.D. Richter,** University of Massachusetts Chan Medical School, Worcester, Massachusetts
**M.R. Tranfaglia,** FRAXA Research Foundation, Newburyport, Massachusetts

FUNDED BY **FRAXA Research Foundation**

Cloning of the *FMR1* gene in 1991 was a watershed discovery for Fragile X syndrome, allowing researchers to study how this gene is silenced and how it controls neural activity leading to intellectual impairment. Animal models, particularly FMRP knockout mice, have been essential tools to understand FMRP deficiency; however, translation of these findings into human clinical trials has been largely disappointing. More recent technologies have allowed us the opportunity to study FMR1 reactivation or FMRP restoration in a human-based system, and in individuals with Fragile X; the results have been promising, and not observed in mouse models. This Banbury Center meeting convened experts to review the state of Fragile X/FMRP research and propose strategies for progress in the field.

**Welcoming Remarks: R. Leshan,** Executive Director, The Banbury Center, Cold Spring Harbor Laboratory

**Welcome and Meeting Objectives: K. Clapp,** FRAXA Research Foundation, Newburyport, Massachusetts

Z. Wen, P. Jin

J. Richter, M. Shannon, J.R. Fallon, L.E. Maquat

**SESSION 1:** Setting the Stage for Fragile X Therapeutics

**Chairperson: E.M. Berry-Kravis,** Rush University Medical Center, Chicago, Illinois

M.R. Tranfaglia, FRAXA Research Foundation, Newburyport, Massachusetts: Definitive therapies for Fragile X: progress and remaining obstacles.

J. Sommer, Simons Foundation, New York, New York: SFARI: facilitating insights into autism by funding basic biology and resource creation.

**SESSION 2:** ASO Restoration of FMRP

**Chairperson: L.E. Maquat,** University of Rochester Medical Center, New York

J.D. Richter, University of Massachusetts Chan Medical School, Worcester: ASO rescue of CGG expansion-dependent FMR1 mis-splicing in Fragile X syndrome restores FMRP.

S. Shah, University of Massachusetts Chan Medical School, Worcester, Massachusetts: Rescue of an FXS-specific mis-spliced RNA using antisense oligonucleotides restores FMRP levels.

J.K. Watts, University of Massachusetts Chan Medical School, Worcester: Healthy brains and mixed backbones: improving oligonucleotide chemistry to treat neurological diseases.

**SESSION 3:** Fragile X Model Systems

**Chairperson: J.T. Lee,** Harvard Medical School and Massachusetts General Hospital, Boston

N. Benvenisty, Hebrew University of Jerusalem, Israel: Modeling and treating Fragile X syndrome using human pluripotent stem cells.

X. Zhao, University of Wisconsin–Madison: From mice to humans and back: using multimodal approaches to interrogate FMRP functions and Fragile X syndrome.

Z. Wen, Emory University, Atlanta, Georgia: Modeling Fragile X syndrome using human 3-D brain organoids.

**SESSION 4:** Fragile X Clinical Trials

**Chairperson: E.V. Pedapati,** Cincinnati Children's Hospital, Ohio

C.A. Erickson, Cincinnati Children's Hospital, Ohio (virtual): Personalizing medicine in Fragile X syndrome.

E.M. Berry-Kravis, Rush University Medical Center, Cincinnati, Ohio (virtual): Updated trial designs, outcome measures and drug delivery methods for FXS and neurodevelopmental disorders.

**SESSION 5:** Reactivation

**Chairperson: J.R. Fallon,** Brown University, Providence, Rhode Island

J.T. Lee, Harvard Medical School and Massachusetts General Hospital, Boston: A reactivation platform to treat Fragile X syndrome.

D. Kumari, National Institute of Diabetes and Digestive and Kidney Diseases, NIH, Bethesda, Maryland: Therapeutic potential of CRISPR-Cas9-mediated deletion of CGG repeats for *FMR1* gene reactivation in Fragile X syndrome.

K.P. Usdin, National Institute of Diabetes and Digestive and Kidney Diseases, NIH, Bethesda, Maryland: CGG repeat expansion-induced epigenetic changes on the *FMR1* gene in Fragile X syndrome.

**SESSION 6:** RNA-Based Therapeutics

**Chairperson: X. Zhao,** University of Wisconsin–Madison

K.A. Whitehead, Carnegie Mellon University, Pittsburgh, Pennsylvania: Nonviral mRNA delivery to the brain using lipid nanoparticles.

J.R. Fallon, Brown University, Providence, Rhode Island: Promoting adult hippocampal neurogenesis with splice-modifying ASOs targeting the MuSK-BMP pathway.
M.L. Hastings, University of Michigan, Ann Arbor: Splice-switching antisense oligonucleotides strategies to combat disease.

**SESSION 7:** Developing Therapeutics for Neurodevelopmental Disorders

**Chairperson: Mira C. Puri,** The Azrieli Foundation, Toronto, Ontario, Canada

D.C. Lo, European Infrastructure for Translational Medicine (EATRIS), Amsterdam, Netherlands: Transforming the drug repurposing ecosystem for rare diseases.
P. Feliciano, Simons Foundation, New York, New York: SPARK: gene discovery and return of genetic results in 50,000 people with autism.
M.L. Shannon, Prevail Therapeutics, Eli Lilly, New York, New York: Gene therapy for Fragile X.

**SESSION 8:** Human Electrophysiology

**Chairperson: C.A. Erickson,** Cincinnati Children's Hospital, Ohio (virtual)

J.F. Lepage, University of Sherbrooke, Sherbrooke, Québec, Canada: The usefulness of transcranial magnetic stimulation in clinical trials for FXS.
E.V. Pedapati, Cincinnati Children's Hospital, Cincinnati, Ohio: Neural timing in Fragile X syndrome.

**SESSION 9:** Integration of Cellular Mechanisms

**Chairperson: K.P. Usdin,** National Institute of Diabetes and Digestive and Kidney Diseases, NIH, Bethesda, Maryland

L.E. Maquat, University of Rochester Medical Center, Rochester, New York: Nonsense-mediated mRNA decay (NMD) is hyperactivated in FXS.
P. Jin, Emory University School of Medicine, Atlanta, Georgia: Human brain organoids for therapeutic development in Fragile X–associated disorders.
C.L. Sirois, University of Wisconsin–Madison: Use of human stem cell–derived neurons to screen for potential therapeutic compounds for Fragile X syndrome.

**SESSION 10:** Fragile X Syndrome: Where We Are and Where We Are Going

**Chairpersons: E.M. Berry-Kravis,** Rush University Medical Center, Chicago, Illinois, **J.D. Richter,** University of Massachusetts Chan Medical School, Worcester, and **M.R. Tranfaglia,** FRAXA Research Foundation, Newburyport, Massachusetts

# Developing an Ethical Framework for Psychedelics Research and Use

**June 9–12**

ARRANGED BY **I.G. Cohen,** Petrie-Flom Center, Harvard Law School, Cambridge, Massachusetts
**A.L. McGuire,** Baylor College of Medicine, Houston, Texas
**D.A. Sisti,** University of Pennsylvania, Philadelphia

FUNDED BY **Petrie-Flom Center, Harvard Law School; Baylor College of Medicine; The Cold Spring Harbor Laboratory Corporate Sponsor Program**

With the resurgence of research on the therapeutic use of psychedelics in the United States, ethical concerns are mounting regarding potential for off-label use without a corresponding evidence base, unsafe practices by corporations and unlicensed providers, and epistemic and material injustices against indigenous communities that have used psychedelics for millennia. Moreover, the boundaries between medical and nonmedical settings for psychedelic therapies are porous and unclear, leading to confusion about the basic ethical norms and practices that should apply across sectors. Unchecked, these issues pose a clear risk to the responsible and safe uptake of psychedelic therapies that could affect patient safety, have legal implications, and deleteriously impact consumer trust. This workshop brought together practitioners, policy makers, community representatives, and ethicists to develop an ethical framework for psychedelics research and use in the United States.

**Introduction and Meeting Objectives:** **I.G. Cohen,** Petrie-Flom Center, Harvard Law School, Cambridge, Massachusetts
**A.L. McGuire,** Baylor College of Medicine, Houston, Texas
**D.A. Sisti,** University of Pennsylvania, Philadelphia

E.E. Schenberg, C.S. Grob

**SESSION 1:** Setting the Stage I: Where Have We Been, Where Are We Now, and Where Are We Headed?

**Chairperson: A.L. McGuire,** Baylor College of Medicine, Houston, Texas

Y. Celidwen, United Nations and University of California, Berkeley: Indigenous spirit medicine beyond mind medicine: frameworks, possibilities, and perils.

M.C. Mithoefer, MAPS Public Benefit Corporation, Asheville, North Carolina: Psychedelic-assisted therapy.

D.B. Yaden, Johns Hopkins University School of Medicine, Baltimore, Maryland: Psychedelics and psychiatry (and beyond): the state of science.

**SESSION 2:** Setting the Stage II: Where Have We Been, Where Are We Now, and Where Are We Headed?

**Chairperson: I.G. Cohen,** Petrie-Flom Center, Harvard Law School, Cambridge, Massachusetts

A. Pallas, Beckley Retreats, Baltimore, Maryland (virtual): Ethics and psychedelic retreats: the why, the what, and the how.

T.G. Patterson, Usona Institute, Madison, Wisconsin: Therapeutic drug development and clinical trials.

M. Marks, POPLAR and Petrie-Flom Center, Harvard Law School, Cambridge, Massachusetts (virtual): Policy landscape.

**SESSION 3:** Panel—Equity and Access

**Chairperson: I.G. Cohen,** Petrie-Flom Center, Harvard Law School, Cambridge, Massachusetts

M.T. Williams, University of Ottawa, Carleton Place, Ontario, Canada: Psychedelics and racial justice.

S.P. Levine, COMPASS Pathways, Princeton, New Jersey: Broad, equitable, and safe patient access to emerging investigational psychedelic therapies: the regulatory path.

P. Hernandez-Wolfe, Lewis & Clark College, Portland, Oregon (virtual): Relational ethics: including all and hearing everyone's voices.

I. Harvey, People of Color Psychedelic Collective, New York, New York: Policy and regulatory models for psychedelics rooted in equity.

**SESSION 4:** Panel—Research Ethics Issues

**Chairperson: D.A. Sisti,** University of Pennsylvania, Philadelphia

J.J. Sabbagh, National Institute of Mental Health, NIH, Milwaukee, Wisconsin: NIMH considerations for psychedelic research.

C.S. Grob, Harbor-UCLA Medical Center, Los Angeles, California: The use of psychedelics in psychiatry: safety and ethical considerations.

B. Waters, Reason for Hope, New York, New York: Research with and access to treatment for veterans.

**SESSION 5:** Panel—Managing Expectations and Informed Consent

**Chairperson: B.M. Kious,** University of Utah, Salt Lake City

A.H. Peterson, George Mason University, Fairfax, Virginia: Transformative experiences.

P. Summergrad, Tufts University and Tufts Medical Center, Boston, Massachusetts (virtual): The boundaries of psychiatry.

D. Öngür, McLean Hospital and Harvard Medical School, Belmont, Massachusetts: A clinical psychiatry research perspective on psychedelics.

**SESSION 6:** Panel—Therapeutic Ethics

**Chairperson: A.H. Peterson,** George Mason University, Fairfax, Virginia

M.C. Mithoefer, MAPS Public Benefit Corporation, Asheville, North Carolina: Ethical issues in psychedelic-assisted therapy.

N.L. Devenot, University of Cincinnati, Ohio: Therapeutic touch and safety concerns.

M.J. Baggott, Tactogen, Palo Alto, California: Roles and responsibilities in clinical research with MDMA-like drugs.

**SESSION 7:** Panel—Training, Education, and Licensure of Practitioners

**Chairperson: A.L. McGuire,** Baylor College of Medicine, Houston, Texas

J. Sun Cooper, MAPS Public Benefit Corporation, Bentonville, Arkansas: The obligations and limitations of pharmaceutical companies in ensuring appropriate training for psychedelic-assisted therapies.

E.M. Nielson, Fluence and Columbia University, Woodstock, New York: Balancing the psychedelic therapist training burden with patient safety.

B.M. Kious, University of Utah, Salt Lake City, Utah: Personal experience and therapeutic bias.

**SESSION 8:** Panel—Appropriate Role of Gatekeepers in Psychedelic Use

**Chairperson: S. Gracias,** Ortus Foundation, Chicago, Illinois

E.E. Schenberg, Instituto Phaneros, Lisbon, Portugal: Epistemic challenges and opportunities for the ethical development of psychedelic therapies worldwide.

M. Volat, Indigenous Peyote Conservation Initiative & Indigenous Medicine Conservation Fund, Occidental, California (virtual): Indigenous medicine conservation and right relationship during psychedelic field expansion.

M.C. Davis, Usona Institute, Rockville, Maryland: The role of the FDA in regulating off-label medical use of psychedelics

**SESSION 9:** Group Discussion I

**Chairpersons: I.G. Cohen,** Petrie-Flom Center, Harvard Law School, Cambridge, Massachusetts, **A.L. McGuire,** Baylor College of Medicine, Houston, Texas, and **D.A. Sisti,** University of Pennsylvania, Philadelphia

**SESSION 10:** Group Discussion II

**Chairpersons: I.G. Cohen,** Petrie-Flom Center, Harvard Law School, Cambridge, Massachusetts, **A.L. McGuire,** Baylor College of Medicine, Houston, Texas, and **D.A. Sisti,** University of Pennsylvania, Philadelphia

# Kennedy's Disease (SBMA) Research Workshop

**September 10–12**

FUNDED BY **Kennedy's Disease Association and 2023 KD Golf Scramble**

Kennedy's disease, or spinal and bulbar muscular atrophy (SBMA), is a rare X-linked, genetic, slowly progressing neuromuscular disease. The genetic mutation causing the disease was discovered in 1991, and as yet there is no cure or effective treatment. This workshop was convened by the Kennedy's Disease Association to stimulate new research into the mechanisms of SBMA pathology and to identify potential therapeutic approaches. Discussions focused on (1) the pathology of the neuromuscular junction and its role and potential modulation in SBMA, and (2) the use of iPSCs to model and understand the mechanisms of SBMA.

**Welcoming Remarks: R. Leshan,** Executive Director, The Banbury Center, Cold Spring Harbor Laboratory

## SESSION 1

**Chairperson: K.H. Fischbeck,** National Institute of Neurological Disorders and Stroke, NIH, Bethesda, Maryland

**Topics:** Overview of SBMA research, Clinical views, Living with Kennedy's disease

**Speakers**

K.H. Fischbeck, National Institute of Neurological Disorders and Stroke, NIH, Bethesda, Maryland

C. Grunseich, National Institute of Neurological Disorders and Stroke, NIH, Bethesda, Maryland

E. Meyertholen, Kennedy's Disease Association, Alexandria, Virginia

D.E. Merry, A. Lieberman, K.H. Fischbeck

H. Miranda, A. Al Qahtani

J. Parker, Kennedy's Disease Association, Brookfield, Connecticut

**SESSION 2**

**Chairperson: A. Lieberman,** University of Michigan, Ann Arbor, Michigan

Topics: Role of muscle and nerve in SBMA pathology, Role of neuromuscular junction pathology

Speakers

A. Lieberman, University of Michigan, Ann Arbor, Michigan

S.J. Burden, Massachusetts General Hospital & Harvard Medical School, Cambridge, Massachusetts (virtual)

R. Burgess, The Jackson Laboratory, Bar Harbor, Maine

B.G. Burnett, Uniformed Services University of the Health Sciences, Bethesda, Maryland

D.J. Glass, Regeneron Pharmaceuticals, Inc., Tarrytown, New York

E. Molotsky, Johns Hopkins University, Baltimore, Maryland

M.M. Rich, Wright State University, Yellow Springs, Ohio

R. Robitaille, Université de Montréal, Laval, Québec, Canada

**SESSION 3**

**Chairperson: A.R. La Spada,** University of California, Irvine

Topic: Modeling cell biology using IPSCs

Speakers

A.R. La Spada, University of California, Irvine

S. Da Cruz, VIB-KU Leuven, Leuven, Belgium

M. Gouti, Max Delbrück Center for Molecular Medicine, Berlin, Germany (virtual)

C. Grunseich, National Institute of Neurological Disorders and Stroke, NIH, Bethesda, Maryland

H.C. Miranda, Case Western Reserve University, Cleveland, Ohio

**SESSION 4**

**Chairperson: D.E. Merry,** Thomas Jefferson University, Philadelphia, Pennsylvania

**Topics:** Summary of discussions, Next experimental steps, Possibilities for collaboration, Possible therapeutic opportunities

**Speaker**

D.E. Merry, Thomas Jefferson University, Philadelphia, Pennsylvania

# The 2023 CL2.0 Colloquium

**September 27–29**

ARRANGED BY **The Board of Directors,** Project Santa Fe Foundation, LLC

FUNDED BY **Project Santa Fe Foundation, LLC**

The first CL2.0 Colloquium took place in 2016 in Santa Fe, New Mexico, catalyzing the formation of the Project Santa Fe Foundation. The subsequent CL2.0 Colloquium series is an ongoing invitation-only, membership-based think tank that drives the thought leadership and evidence base for CL2.0. Participants contribute to CL2.0 business model creation, guide existing CL2.0 demonstration projects, identify and prioritize future studies and publications, and explore the key elements for a successful CL2.0 model. Discussion at the 2023 CL2.0 Colloquium focused on the business model.

## PARTICIPANTS

B. Bailey, Project Santa Fe Foundation, LLC, Albuquerque, New Mexico
A. Baldwin, Henry Ford Health System, Detroit, Michigan
D.A. Breining, Northwell Health, New Hyde Park, New York
D. Chhieng, COLA, Inc., Columbia, Maryland
J.M. Crawford, Northwell Health, New Hyde Park, New York
B. D'Ambrosio, Northwell Health, New York, New York
C. DeGraff-Murphy, The University of Vermont Health Network, Burlington, Vermont

M.D. Dixon, The Mark Dixon Group, LLC, Minneapolis, Minnesota
K. Donaldson, Geisinger Health System, Hershey, Pennsylvania
L.A. Fleisher, University of Pennsylvania & Rubrum Advising, Philadelphia, Pennsylvania
P.R. Forlenza, Geisinger Health Plan, Dallas, Pennsylvania
M.K. Fung, The University of Vermont Health Network, Burlington, Vermont
A. Hay, Sysmex America, Inc., Barrington, Illinois
D. Ingemansen, Beckman Coulter Diagnostics, Sioux Falls, South Dakota
B.R. Jackson, University of Utah, Salt Lake City, Utah
C. Jasper, Thermo Fisher Scientific, College Station, Texas
V. Joy, Thermo Fisher Scientific, Philadelphia, Pennsylvania
J. Krisa, MVP Health Care, Albany, New York
R.K. Laughman, CareTinuum Advisors, Nashville, Tennessee
A. LeBlanc, Siemens Healthineers, Greenwich, Connecticut
R. McGonnagle, College of American Pathologists (CAP), Winnetka, Illinois
J.V. Mitsios, Siemens Healthineers, Tarrytown, New York
J. Murray, Beckman Coulter Diagnostics, Chicago, Illinois
K. Nemmers, Abbott, Chicago, Illinois
K. Nucifora, COLA, Inc., Arnold, Maryland
S.R. Peskin, SRP Advisors, LLC, Belle Mead, New Jersey
R.M. Salerno, U.S. Centers for Disease Control & Prevention (CDC), Atlanta, Georgia
J. Schulman, Northwell Health, New York, New York
N. Stratton, COLA, Inc., Columbia, Maryland
K. Swanson, Project Santa Fe Foundation, LLC, Albuquerque, New Mexico
M.L. Wilkerson, Geisinger Health System, Danville, Pennsylvania
R.J. Zarbo, Henry Ford Health System, Detroit, Michigan

# Strengthening the Role of Evaluation in the Sustainable Development Agenda for Health

**October 10–13**


ARRANGED BY **K. Albright,** UNICEF, New York, New York
**D.I. Fraser,** Global Evaluation Initiative (GEI), Brussels, Belgium
**G. Garnett,** Bill & Melinda Gates Foundation, Seattle, Washington
**J. Hargreaves,** London School of Hygiene & Tropical Medicine (LSHTM), United Kingdom
**M. Johri,** Université de Montréal, Québec, Canada
**M. Sabin,** *The Lancet,* New York, New York
**K. Sheikh,** University College London, United Kingdom



FUNDED BY **Bill & Melinda Gates Foundation**


Evaluation is a critical component of the United Nations Sustainable Development Goals (SDG) agenda for health. It involves the practice of systematically collecting, analyzing, and interpreting data. It facilitates accountability and oversight, supports organizational learning and course correction, and ensures evidence-based practice. Evaluation in global health diverges based on priorities, and practices have evolved because this type of evaluation involves a range of stakeholders, and standards, norms, guidance and policies for evaluation come in many forms. Global health evaluation is important and can be made stronger. This Banbury meeting convened a diverse


G. Garnett, J. Hargreaves, D.I. Fraser, C. Morkel

R.R. Ved, K. Sheikh

group of participants with the overarching aim of identifying ways to strengthen evaluation within the SDG agenda for health. Participants considered themes including (1) evaluation methods, (2) power imbalances, (3) utility, (4) governance, and (5) investment.

**Welcoming Remarks and Introduction: R. Leshan,** Executive Director, The Banbury Center, Cold Spring Harbor Laboratory,
**G. Garnett,** Bill & Melinda Gates Foundation, Seattle, Washington
**M. Sabin,** *The Lancet*, New York, New York
**M. Johri,** Université de Montréal, Québec, Canada

**SESSION 1A:** The "SDG for Health" Evaluation Landscape

**Chairperson: M. Sabin,** *The Lancet*, New York, New York

J. Hargreaves, London School of Hygiene & Tropical Medicine (LSHTM), United Kingdom: What is "evaluation": a view from academia.

R. McLean, International Development Research Centre (IDRC), Ottawa, Ontario, Canada: The untapped potential for evaluation.

D.I. Fraser, Global Evaluation Initiative (GEI), Brussels, Belgium: "Health SDGs evaluation landscape," Global Evaluation Initiative view.

**SESSION 1B:** Evaluation Approaches

**Chairperson: G. Garnett,** Bill & Melinda Gates Foundation, Seattle, Washington

M. Gaarder, International Initiative for Impact Evaluation (3ie), Oslo, Norway: The state of development effectiveness evidence for health: Where are the gaps, where is there richer evidence, and where are the trends?

K. Sheikh, University College London, United Kingdom: What does a Learning Health Systems approach mean for evaluation?

J. Blanchard, University of Manitoba, Winnipeg, Canada: Metaphors and metrics: applying programme science concepts in the evaluation of complex health programmes.

**SESSION 1C:** Evaluation at Its Best, Evaluation Challenges

**Chairperson: D.I. Fraser,** Global Evaluation Initiative (GEI), Brussels, Belgium

G. Hernández Licona, Multidimensional Poverty Peer Network (MPPN-OPHI) & Global Evaluation Initiative (GEI), Mexico City, Mexico: Monitoring and evaluating health outcomes in Mexico based on rights.

R.R. Ved, Bill & Melinda Gates Foundation, New Delhi, India: The role of evaluation in policy and programme change: learnings from India.

S. Arifeen, International Centre for Diarrhoeal Disease Research, Bangladesh, Dhaka, Bangladesh: Evaluation for change—experiences from Bangladesh.

**SESSION 1D:** How Evaluation Is Conducted: Independence, Governance

**Chairperson: M. Johri,** Université de Montréal, Québec, Canada

C. Morkel, CLEAR-AA, Gqeberha, South Africa: The role of the state in national evaluation systems, with a focus on Africa and the developmental state debate.

P. Hansen, Global Financing Facility for Women, Children and Adolescents, World Bank Group, Washington, D.C.: Evaluation independence: experiences from the Global Financing Facility to balancing independence with country leadership of the evaluation agenda.

N. Schwalbe, Spark Street Advisors, New York, New York: Behavioural independence: pipe dream or possibility in global health?

## VIRTUAL ADDRESS

**P. Gluckman,** International Science Council (ISC), Auckland, New Zealand: Flipping the science model: the International Science Council's roadmap to science missions for sustainability.

**SESSION 2A:** Evaluation, the SDGs, and Beyond

**Chairperson: K. Sheikh,** University College London, United Kingdom

Z.A. Bhutta, Aga Khan University & University of Toronto, Karachi, Pakistan: Addressing SDGs in the reality of complex crises.

M. Johri, Université de Montréal, Québec, Canada: Expanding the focus of evaluation to consider root causes and transformative solutions: challenges and strategies to attain the SDGs.

R. Polastro, World Health Organization (WHO), Geneva, Switzerland: Joint evaluations: strengthening the role of evaluation in the sustainable development agenda for health.

**SESSION 2B:** Maximizing the Benefits of Evaluations

**Chairperson: K. Albright,** UNICEF, New York, New York

J.N. Lavis, Global Commission on Evidence to Address Societal Challenges, Toronto, Ontario, Canada: How can evaluators help to formalize and strengthen domestic evidence-support systems and the global evidence architecture, and to put evidence at the centre of everyday life?

J. Nonvignon, Africa Centres for Disease Control and Prevention, Addis Ababa, Ethiopia: The importance of economic evaluation for decision-making.

T. Lubanga, Office of the Prime Minister, Kampala, Uganda: Mainstreaming climate change in the national evaluation policy and practice in Uganda.

**SESSION 2C:** Innovative Evaluation Partnerships

**Chairperson: D.I. Fraser,** Global Evaluation Initiative (GEI), Brussels, Belgium

J. Grove, The Global Fund, Geneva, Switzerland: Optimising the credibility of evaluations at The Global Fund.

A.M. Ocampo Cobos, IDEAS International, Lima, Peru: Evaluating the SDGs through country-led processes.

R. Wanyenze, Makerere University, Kampala, Uganda: Experiences with evaluation and the SDGs in Uganda.

**SESSION 2D:** Future Directions

**Chairperson: J. Hargreaves,** London School of Hygiene & Tropical Medicine, United Kingdom

K. Albright, UNICEF, New York, New York: The SDG synthesis coalition and future of evaluation.

K. Hanson, London School of Hygiene & Tropical Medicine, United Kingdom: The SDG synthesis coalition and future of evaluation.

**SESSION 3A:** Blue Skies—What Role Can Evaluation Play to Reignite and Accelerate Progress on the Implementation of the 2030 Agenda for Sustainable Development, with a Focus on Global Health?

**Chairperson: M. Johri,** Université de Montréal, Québec, Canada

M. Johri, Université de Montréal, Québec, Canada: Ideas submitted to the Chair.

K. Sheikh, University College London, United Kingdom; D.I. Fraser, Global Evaluation Initiative (GEI), Brussels, Belgium; and K. Albright, UNICEF, New York, New York: Reflections from the steering committee.

**SESSION 3B:** What Next? The Road Ahead

**Chairperson: G. Garnett,** Bill & Melinda Gates Foundation, Seattle, Washington

# Lustgarten Foundation Scientific Advisory Board Meeting

**October 23–25**

ARRANGED BY **A. Rakeman,** Lustgarten Foundation, Woodbury, New York
**L. Tantawi,** Lustgarten Foundation, Woodbury, New York
**D. Tuveson,** Cold Spring Harbor Laboratory

FUNDED BY **The Lustgarten Foundation**

Banbury was pleased to welcome back the Lustgarten Foundation for their 2023 Scientific Meeting, which provided an opportunity for the Scientific Advisory Board and special guests to discuss cutting-edge scientific and clinical approaches in pancreatic cancer research. The meeting served as forum for the open sharing of data and ideas, supporting advances in pancreatic cancer research.

## PARTICIPANTS

J. Adams, Stand Up To Cancer (SU2C), Boston, Massachusetts
P. Allen, Duke University, Durham, North Carolina
S. Chung, Lustgarten Foundation, San Diego, California
L.A. Diaz, Jr., Memorial Sloan Kettering Cancer Center, New York, New York
R.N. DuBois, The Mark Foundation for Cancer Research, Charleston, South Carolina

E.K. Fishman, Johns Hopkins University, Baltimore, Maryland
M.G. Goggins, Johns Hopkins University, Baltimore, Maryland
T.S. Hong, Massachusetts General Hospital, Boston, Massachusetts
P. Hou, Rutgers University, Newark, New Jersey
T. Jacks, Koch Institute at MIT, Cambridge, Massachusetts (virtual)
E.M. Jaffee, Johns Hopkins University, Baltimore, Maryland
A. Lustgarten, Lustgarten Foundation, New York, New York
J. Lyman, Lustgarten Foundation, Woodbury, New York
E.R. Manuel, Beckman Research Institute, City of Hope, Duarte, California
A. Rakeman, Lustgarten Foundation, Woodbury, New York
C. Sander, Dana-Farber Cancer Institute, Boston, Massachusetts
R.J. Shaw, Salk Institute for Biological Studies, La Jolla, California
S. Singh, Lustgarten Foundation, Woodbury, New York
B.W. Stillman, Cold Spring Harbor Laboratory
L. Tantawi, Lustgarten Foundation, Woodbury, New York
D.A. Tuveson, Cold Spring Harbor Laboratory
B. Vogelstein, Johns Hopkins University, Baltimore, Maryland
B.M. Wolpin, Dana-Farber Cancer Institute, Boston, Massachusetts
L.D. Wood, Johns Hopkins University, Baltimore, Maryland
J. Yeh, University of North Carolina at Chapel Hill, North Carolina

# The Future of Plant–Environment Interactions: Challenges and Opportunities in a Changing Climate

**October 29–November 1**

ARRANGED BY **U. Pedmale,** Cold Spring Harbor Laboratory
**L.C. Strader,** Duke University, Durham, North Carolina

FUNDED BY **Cold Spring Harbor Laboratory Corporate Sponsor Program**

Plants sustain humanity by providing food, fuel, fiber, pharmacological compounds, timber, and more. Climate change poses an unprecedented threat to plants despite their crucial role. The quantity, productivity, bioregions, and ecosystems of plants are being significantly impacted by climate change because of their sensitivity to their environment. This Banbury meeting convened experts to discuss strategies for addressing issues at the plant–climate interface. Participants considered questions such as: What are the current challenges in "future-proofing" plants to adapt to climate change? What are the mechanisms by which plants balance growth and cope with multiple stresses and pathogens? What are the implications of climate change for plant carbon sequestration and photosynthetic capacity, and how can we mitigate unanticipated feedback loops? How can different stakeholders such as plant scientists, food producers, funders, and policymakers collaborate to address food security challenges in the face of climate change?

**Welcoming Remarks: R. Leshan,** Executive Director, The Banbury Center, Cold Spring Harbor Laboratory

**Welcome and Meeting Objectives: U. Pedmale,** Cold Spring Harbor Laboratory
**L.C. Strader,** Duke University, Durham, North Carolina

## SESSION 1

**Chairperson: U. Pedmale,** Cold Spring Harbor Laboratory

L.C. Strader, Duke University, Durham, North Carolina: State of agricultural research in the U.S.

E.S. Buckler, USDA-ARS at Cornell University, Ithaca, New York: Eliminating agriculture's greenhouse gas emissions by focusing on the flow of nitrogen.

M. Guerinot, Dartmouth College, Hanover, New Hampshire: Iron: nutritious, noxious and not readily available.

## SESSION 2

**Chairperson: U. Pedmale,** Cold Spring Harbor Laboratory

X. Dong, J. Bailey-Serres

C. Williams, North Carolina State University, Raleigh: Multi-scale modeling approaches for understanding plant and agronomic systems across biological scales.

D. Jackson, Cold Spring Harbor Laboratory: The role of TERMINAL EAR1 in maize development and heat responses.

W. Busch, Salk Institute for Biological Studies, La Jolla, California: Engineering root traits for climate change mitigation.

## SESSION 3

**Chairperson: L.C. Strader,** Duke University, Durham, North Carolina

X. Dong, HHMI, Duke University, Durham, North Carolina: Precision in plant immune responses: from basic research to application.

A.M. Hancock, Max Planck Institute for Plant Breeding Research, Cologne, Germany: Adaptations to challenging environments.

R. Jinkerson, University of California, Riverside, Riverside, California: A hybrid inorganic-biological artificial photosynthesis system for energy-efficient food production.

## SESSION 4

**Chairperson: L.C. Strader,** Duke University, Durham, North Carolina

M. Quint, Martin Luther University Halle-Wittenberg, Halle, Germany: What we think we know—thermomorphogenesis in the lab versus real life.
R. Palanivelu, University of Arizona, Tucson: Genomics of reproductive thermotolerance in tomato.

### SESSION 5

**Chairperson: L.C. Strader,** Duke University, Durham, North Carolina

S. Rhee, Plant Resilience Institute, Michigan State University, East Lansing: Understanding and engineering plant resilience from molecular to ecological scales via team science.
U. Pedmale, Cold Spring Harbor Laboratory: Plant response to light limitation: growth and carbon assimilation.
R. Bart, Donald Danforth Plant Science Center, St. Louis, Missouri: Bridging the lab-to-field gap in microbiome research (the search for magic microbes).

### SESSION 6

**Chairperson: L.C. Strader,** Duke University, Durham, North Carolina

J. Bailey-Serres, Center for Plant Cell Biology, University of California, Riverside: Water-wise rice (direct seeded rice to limit methane emissions and water extremes).
J. Casal, University of Buenos Aires, Argentina: Plant growth in a changing climate.
S.P. Long, University of Illinois, Urbana-Champaign: Future-proofing crop photosynthesis.

### SESSION 7

**Chairperson: U. Pedmale,** Cold Spring Harbor Laboratory

R.A. Martienssen, HHMI, Cold Spring Harbor Laboratory: The genomes and epigenomes of aquatic plants (Lemnacae): blueprints for carbon remediation.
R. Sozzani, North Carolina State University, Raleigh: Print better plants: accelerating systematic testing of traits.

### SESSION 8

**Chairperson: U. Pedmale,** Cold Spring Harbor Laboratory

C. Fankhauser, University of Lausanne, Switzerland: Regulation of resource allocation by light cues indicative of crowding.
H. Hirt, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia: PlantACT! Plants for climate action—challenges and opportunities to the climate crisis.

### SESSION 9

**Chairperson: L.C. Strader,** Duke University, Durham, North Carolina

C.F. Weil, National Science Foundation, Alexandria, Virginia: Plants, Environment, Training, Research, and Impact: keeping things growing in the PETRI dish.
N.J. Bate, Bill & Melinda Gates Foundation, Raleigh, North Carolina: Translating science and technology into developing world agriculture.

### SESSION 10

**Chairpersons: L.C. Strader,** Duke University, Durham, North Carolina, and **U. Pedmale,** Cold Spring Harbor Laboratory

Plenary discussion led by: J. Chory, HHMI, Salk Institute for Biological Studies, La Jolla, California

### CLOSING DISCUSSIONS

**Chairpersons: L.C. Strader,** Duke University, Durham, North Carolina, and **U. Pedmale,** Cold Spring Harbor Laboratory

# Integrating Exposomics into the Biomedical Enterprise

**December 3–6**

ARRANGED BY **L.M. Bennett,** Roger Schwarz & Associates, LLC, Potomac, Maryland
**G.W. Miller,** Columbia University, New York, New York

FUNDED BY **Columbia University; Cold Spring Harbor Laboratory Corporate Sponsor Program; and MBX Capital**

The concept of the exposome was introduced in 2005 as a way to identify the environmental contributors to human health and disease. Originally the exposome was defined as "the totality of exposures throughout life," something impossible to scientifically assess in any current or conceivable future setting, and something that does not mesh well with the U.S. biomedical enterprise. Yet, in the age of omic-scale biology, it is essential to study environmental factors of disease in the proper context, with all of its inherent complexity. The goals of this Banbury Center meeting were to develop an operational definition of exposomics for biomedical research; identify the components of the biomedical enterprise that will benefit most from exposomics (research, discovery, clinical trials, diagnosis, etc.); and outline conceptual and technical innovations that will be needed to establish exposomics as a bona fide omics discipline.

S. Li, J. Klánová

G.W. Miller, G.S. Bhutani

**WELCOME: R. Leshan,** Executive Director, The Banbury Center, Cold Spring Harbor Laboratory: Welcoming remarks
**G.W. Miller,** Columbia University, New York, New York, and **L.M. Bennett,** Roger Schwarz & Associates, LLC, Potomac, Maryland: Welcome, overview of meeting, goals

**SESSION 1:** Setting the Stage

**Chairperson: L.M. Bennett,** Roger Schwarz & Associates, LLC, Potomac, Maryland

G.W. Miller, Columbia University, New York, New York: From confusion to clarity in exposomics.

L.M. Bennett, Roger Schwarz & Associates, LLC, Potomac, Maryland: "We are here": how this workshop is situated in the context of exposomics.

**SESSION 2:** The Biomedical Enterprise: Definition for this Workshop

**Chairperson: P.J. Lein,** University of California, Davis

G.W. Miller, Columbia University, New York, New York: Framing the biomedical enterprise from a U.S. perspective.

J. Klánová, Masaryk University, RECETOX, Brno, Czech Republic: EIRENE—European infrastructure for the exposome research.

**SESSION 3:** An Operational Definition for Exposomics Is Needed

**Chairperson: C. Patel,** Harvard University, Boston, Massachusetts

L.M. Bennett, Roger Schwarz & Associates, LLC, Potomac, Maryland: Criteria for an operational definition.

R. Barouki, INSERM, Paris, France: The need for an operational definition.

**SESSION 4:** Developing an Operational Definition for Exposomics I (Breakout Groups)

**Chairpersons: R.O. Wright,** Icahn School of Medicine at Mount Sinai, New York, New York, and **C.K. Ward-Caviness,** U.S. Environmental Protection Agency (EPA), Chapel Hill, North Carolina

## DAY TWO INTRODUCTION

**D.A. Jett,** National Institute of Neurological Disorders and Stroke, NIH, Bethesda, Maryland: Integrating exposomics into an NIH research portfolio—NINDS as an example.

**SESSION 5:** Developing an Operational Definition for Exposomics II (Breakout Groups)

**Chairpersons: C. Patel,** Harvard University, Boston, Massachusetts, and **D.C. Dolinoy,** University of Michigan, Ann Arbor

T.O. Metz, Pacific Northwest National Laboratory, Richland, Washington: Lessons from the "Decoding the Molecular Universe" meeting and initiative.

**SESSION 6:** Identifying the Components of the Biomedical Enterprise That Will Benefit Most from Exposomics (Panel)

**Chairperson: R. Vermeulen,** Utrecht University, Netherlands

### Panelists

K. Pollitt, Yale University, New Haven, Connecticut

D.C. Dolinoy, University of Michigan, Ann Arbor, Michigan

P. Gao, University of Pittsburgh, Pennsylvania

R.O. Wright, Icahn School of Medicine at Mount Sinai, New York, New York

C.K. Ward-Caviness, U.S. Environmental Protection Agency (EPA), Chapel Hill, North Carolina

**SESSION 7:** Identifying the Components of the Biomedical Enterprise That Will Benefit Most from Exposomics (Breakout Groups)

**Chairperson: D. Balshaw,** National Institute of Environmental Health Sciences, NIH, Research Triangle Park, North Carolina

## DAY THREE INTRODUCTION

**A. Rajasekar,** University of North Carolina at Chapel Hill: Outlining conceptual and technical innovations that will be needed to accomplish making exposomics a bona fide omics discipline: infrastructure and community of practice.

**SESSION 8A:** Solutions for Conceptual and Technical Innovations: Developing Impossible Solutions to Solve Challenges in Measuring the Exposome

**Chairperson: L.M. Bennett,** Roger Schwarz & Associates, LLC, Potomac, Maryland

**SESSION 8B:** Solutions for Conceptual and Technical Innovations: Developing Impossible Solutions to Solve Challenges in Measuring the Exposome

**Chairpersons: G.W. Miller,** Columbia University, New York, New York, and **L.M. Bennett,** Roger Schwarz & Associates, LLC, Potomac, Maryland

P.J. Lein, University of California, Davis: Stakeholder engagement and spreading the word about exposomics in the biomedical enterprise.


DNA LEARNING CENTER

# DNA LEARNING CENTER
# 2023 EXECUTIVE DIRECTOR'S REPORT


**ADMINISTRATION**

Elizabeth Asaro
Shreemattie Budhram
Lauren Correri
David Micklos
Donna Smith

**INSTRUCTION**

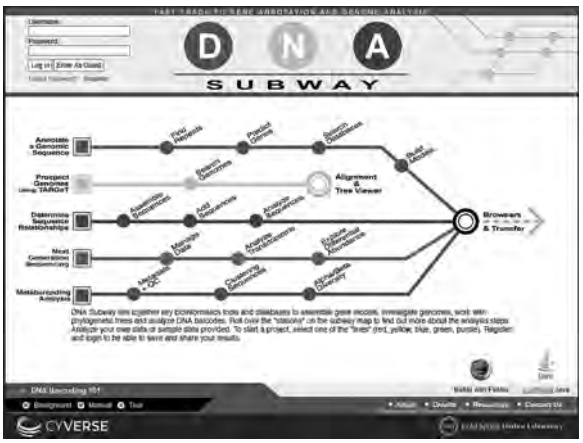
Kelsie Anson
Allison Astudillo
Keith Bannerman
Elna Carrasco-Gottlieb
Kelly Eames
Anna Feitzinger
Cristina Fernandez-Marco
Jennifer Hackett
Carol Henger
Ria Jasuja
Brittany Johnson
Arie Kaz
Allison Mayle
Amanda McBrien
Christina Newkirk
Jeffry Petracca
Tiffani Rushford

**BIOMEDIA**

Daniel Jacobs
Susan Lauter
Jason Williams
Chun-hua Yang


Data science is one of the 10 fastest growing careers in the United States, and bioinformatics is an essential part of many bioscience careers today. Bioinformatics technicians earn a 30% salary premium over other biological technicians. Over the last 25 years, the DNALC has created pathways to these biological jobs of the future. This began in 1998 with the first "personal genetics" experiment that allows students to sequence and analyze their own mitochondrial (mt) DNA—well in advance of *23andMe*, the *Genographic Project*, and *Ancestry DNA*. With support from the National Science Foundation (NSF) Advanced Technological Education (ATE) Program, in 2000 we developed *BioServers* as a simple bioinformatics tool to compare human mtDNA sequences between classmates and with reference data from world populations and ancient human ancestors. This platform has proven remarkably durable, engaging 662,379 users in 2.08 million sessions averaging 15+ minutes. The *BioServers* database currently contains more than 169,000 student mtDNA sequences.

This was followed by the launch in 2010 of *DNA Subway*, an intuitive online interface that makes sophisticated bioinformatics analysis accessible to students without computation experience. Operating under the umbrella of the NSF *CyVerse* cyberinfrastructure, the project coordinated contributions from more than 25 scientists, computer programmers, and bioinformaticians at more than a dozen research institutions. Using the metaphor of a subway, students can "ride" any of five different lines to access and analyze DNA sequences. *DNA Subway* has garnered a dedicated following, with 296,037 total users logging 739,852 sessions averaging 17 minutes from 2010 to 2023.




Screenshot of the current *DNA Subway* website.

More than 260,000 DNA sequences have been uploaded by our partner Azenta (formerly GENEWIZ), and the popular Blue Line, primarily used for DNA barcoding, accounts for ~72% of traffic. Undergraduate students make up 68% of 62,140 registered users. Blue Line users have published 2,391 barcode records on GenBank with 1,594 unique student and faculty authors. Supporting the contention that *DNA Subway* is a research-grade tool, it has become one of the two most widely used infrastructures to support course-based undergraduate research experiences (CUREs), and 41% of nonstudent users are researchers.

In July, we received a $650,000 NSF ATE grant to update the aging *DNA Subway* infrastructure to make it a more flexible, accessible, and capable resource to prepare students for the modern


Mock-up design of *DNA Subway 2.0* scaled for a mobile phone.

bioscience workforce. The project is taking place within the context of the DNALC's role as Genomics Hub of the InnovATE*BIO* National Biotechnology Education Center. *DNA Subway 2.0* will rely on NSF's *Jetstream2* cloud computing infrastructure and systems expertise from the Texas Advanced Computing Center and NSF's Science Gateways Community Institute. By adopting a "mobile-first" approach, we believe that this will be the first set of high-level bioinformatics tools to run smoothly on a smartphone. At the same time, design improvements will prioritize accessibility for users with disabilities, including those with low vision and those who use assistive devices to explore web content. The reimagined mobile and desktop *DNA Subway 2.0* versions will give students complete flexibility to collect data and complete assignments in school, at home, or on the road from their preferred device. In this way, *DNA Subway* will serve the needs of low-income, rural, and limited-sight students—who rely on mobile devices.

## DNA Sequence Analysis Anytime, Anywhere, by Anyone

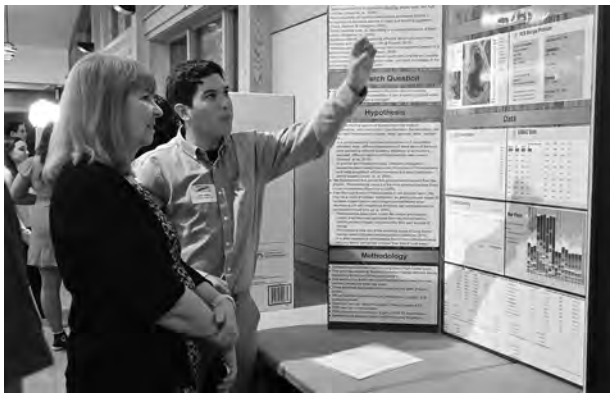
Since its founding in 1988, the DNALC has popularized experiments in molecular genetics for use in high school and college teaching. DNALC experiments and commercial kits based on DNALC technology are used by millions of students per year. These include the most widely used methods for putting DNA into bacteria, examining personal genetics, and producing DNA barcodes.

In summer 2023, we signed a memorandum of understanding (MOU) with Oxford Nanopore to help popularize the "next big thing" in biology education—the ability to analyze individual DNA molecules in real time. In addition to massively parallel DNA sequencers used in the CSHL Genome Center, Oxford also produces the MinION, a "personal" DNA sequencer about half the size of a mobile phone that plugs into a computer USB port. At ~$2,000 for the machine and ~$10 for a gene or small-genome sequence, the MinION makes gene analysis affordable for almost anyone. Under the MOU, the DNALC will help develop improved chemistry, workflows, directions, and packaging/pricing attractive to high school and college faculty.

*DNA Subway 2.0* will include a new line for nanopore sequence analysis. The combination of MinION sequencer and mobile *DNA Subway 2.0* will be the first integrated system to support DNA sequencing and analysis any time, any place, by anyone. We envision a day when a MinION sequencer and *DNA Subway*–powered analysis joins the polymerase chain reaction (PCR) machine as indispensable equipment in every bioscience teaching lab.

## High School DNA Barcoding Research Programs

The DNALC continued efforts to enable high school students to conduct authentic biodiversity research using DNA barcoding. *Barcode Long Island* (*BLI*) involves students in "campaigns" to compare biodiversity across Long Island. The *Urban Barcode Project* (*UBP*), funded by the Thompson Family Foundation, and *Urban Barcode Research Program* (*UBRP*), funded by the Pinkerton Foundation, involve students in research of biodiversity in New York City (NYC).



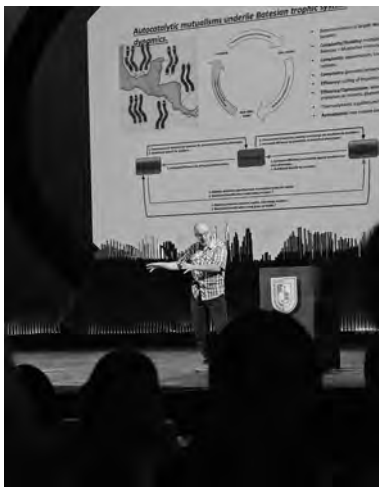
(*Left*) Dr. James Lendemer described lichen diversity during a beautifully illustrated keynote talk. (*Right*) Sam Adler, a student from Long Beach High School, presents research results to Jennifer Newitt, a mentor from Friends Academy.

Science teachers are mentors for *BLI* and *UBP* students, whereas scientists from NYC institutions mentor *UBRP* students.

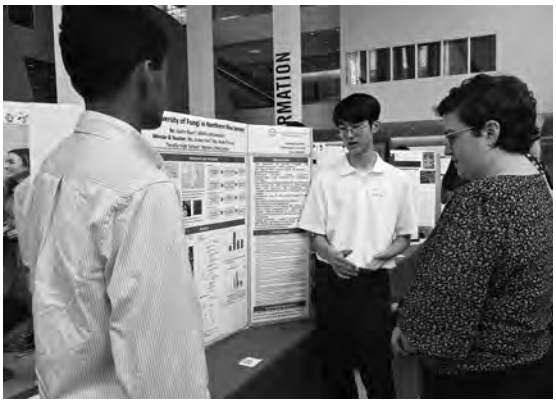
After the completion of funding from the NIH National Institute of General Medical Sciences (NIGMS) Science Education Partnership Award (SEPA), *BLI* continued the transition to a sustainable support model in which schools covered the costs of materials and sequencing for their teams, participated in DNALC memberships, or received scholarships based on financial need. One hundred and seventy-three students across 60 teams completed projects, with 15 teams utilizing Open Lab sessions at the Dolan DNALC and Brookhaven National Lab, and 39 teams borrowing equipment. A significant portion of student-authored sequences (82) were published in GenBank, including four new barcode records and 22 sequences with variable DNA. Additionally, 11 teachers attended a five-day summer (6) or one-day fall (5) training workshop, with more than half submitting student project proposals by December.

The annual *Barcode Long Island* Student Symposium, held June 6, 2023, on the Cold Spring Harbor Laboratory campus, featured keynote speaker Dr. James Lendemer, Curator of Botany, Research & Collections at the New York State Museum. Dr. Lendemer presented his work on deciphering the dimensions of lichen diversity from cities to wilderness.

The annual *Urban Barcode Project* and *Urban Barcode Research Program* Symposium featured keynote speaker Dr. David Kizirian of the American Museum of Natural History, who presented on the evolution of biological systems. Three teams were recognized with outstanding poster awards: in the *UBP,* a team from Tenafly High School for "Biodiversity of Fungi in Northern New Jersey," and in the *UBRP*, a team mentored at the New York Botanical Gardens for "Effect of Specimen Age on DNA Barcoding Success" and a team mentored at Pace University for "Antibiotic Prospecting from Plant Microbial Endophytes."



Dr. David Kizirian gives an engaging presentation on the importance of biological processes like autocatalysis.



Urban barcoding student teams present project results to symposium attendees. CSHL Board of Trustees Chair Marilyn Simons learns about ant barcoding (at *left* in *left* photo) and (*right*) one of the outstanding poster teams that researched fungi in New Jersey speaks with Kelsie Anson, DNALC NYC educator.

One hundred and twenty-four students on 45 teams completed projects in the *UBP* and 39 students on 15 teams completed projects in the *UBRP.* These students made ample use of DNALC resources: 16 teams attended open-lab sessions at the *Harlem DNA Lab* or *DNALC NYC*, and 23 teams borrowed equipment. Thirty-eight *UBP* teams and all 15 *UBRP* teams presented posters to peers and science professionals at the annual research symposium on May 31 at the New York City College of Technology. Eight *UBRP* teams also presented posters to peers and science professionals at the annual Science Research Mentoring Program (SRMP) Colloquium on June 9.

*UBRP* students emerged from the program with stronger interest and confidence in pursuing a path in science. Fifty-five percent of students from Cohort 10 were more interested in studying biology following their *UBRP* projects. Seventy-five percent felt more capable of going further in science than they did prior to participation in *UBRP*. Importantly, most (82%) felt the approach to problem-solving they learned in *UBRP* would help them succeed in future science courses and their future career.

## Citizen Science DNA Barcoding

The *Citizen DNA Barcode Network* (*CDBN*) entered its fourth year of NIGMS SEPA funding. The project mobilizes citizen scientists around the United States, working under the guidance of trained staff at science and nature centers, to collect and barcode insects including ants, beetles, and mosquitos. Citizen science–derived barcodes can then be used to better inform range maps, identify new species, and contribute new barcode records to sequence databases.

California Academy of Sciences (CAS) fulfilled their role as a Year 3 *CDBN* Hub, which included working with a cohort of Careers in Science high school interns to learn, execute, and communicate DNA barcoding programming among themselves and with the general public at CAS. The DNALC also supported citizen science DNA barcoding programming at a number of local institutions, including the New York Hall of Science, Jones Beach Energy and Nature Center, Sweetbriar Nature Center, South Fork Natural History Museum, and the Long Island Aquarium.

*CDBN* supported the Natural History Museum of Utah (NHMU) as a Year 4 Hub, successfully engaging 972 citizen scientists through a variety of DNA barcoding activities, including tabletop events, one- or two-day wet labs, and internships. One highlight was the inclusion of DNA barcoding at NHMU's BugFest 2023, where they set up stations focusing on each part of the wet lab and invited NHMU's Mycology lab to discuss how they used DNA barcoding in their research. This tabletop-style event engaged 340 members of the public. Beyond NHMU, a new *CDBN* collaboration was initiated with Discover Life in America (DLIA) and the National

Park Service in Great Smokey Mountains National Park (GSMP). DLIA regularly conducted DNA barcoding programming with park staff, volunteers, and naturalists; notably, a firefly beetle collected during a DLIA *CDBN* event, *Photuris quadrifulgens*, made local news as it was the first time that the species was recorded in GSMNP.

Educators at *CDBN* training at the Dolan DNA Learning Center.

During the fourth program year *CDBN* supported 1,355 participants who spanned age ranges and skill levels. In total, 1,212 specimens with photos and metadata were entered into our Barcode Sample Database, and 1,161 of those specimens were processed to generate DNA barcodes. Two hundred and fifty DNA barcodes were published to GenBank with citizen scientists as authors, including 16 previously unpublished barcode sequences and 51 new variants. *CDBN* continues to engage citizen scientists into its fifth program year, bringing the Missouri Botanical Garden on as a new barcoding Hub.

## Suzhou Barcoding Project

During the summer, DNALC educator Jeffry Petracca traveled to the DNALC's licensed center in Suzhou, China, *Cold Spring Harbor Asia DNA Learning Center*, to instruct a two-week DNA barcoding summer course. The course engaged local secondary school students and undergraduate *DNALC Asia* interns on the use of DNA barcoding in research projects. The group was asked to design a project over the course of the two-week session, and they produced and presented a scientific poster to the class. In addition, a GenBank BioProject was established for the *DNALC Asia* team to help track and monitor GenBank sequence submissions. Ninety-four DNA barcodes were published to GenBank with Chinese students as authors, including 13 previously unpublished barcode sequences and 22 new variants.

## Lounsbery eDNA

The Richard Lounsbery Foundation funded creation of a website illustrating and explaining key topics of environmental DNA (eDNA) research. Because of the departure of Bruce Nash, the grant transitioned to the *BioMedia* Group and is now being spearheaded by Jason Williams and Carol Henger, whose expertise in metabarcoding is well suited for this project. The website will be geared toward students, teachers, and researchers entering the field of eDNA research who need help learning the basics of best practices and protocols. In late 2023, we conducted a small survey of scientists actively involved in eDNA research to ask them which topics they believe would benefit from an explanatory educational aid. With the responses gathered, we are in the process of developing new and contributed content to create an "eDNA Primer."

## Carolina Biological

DNALC collaborated to update more than 20 DNALC laboratory kits distributed by Carolina Biological Supply Company, including redesign of manuals, incorporation of Next Generation Science Standards, and testing of different DNA polymerases to reduce costs for PCR laboratories. In addition, the DNALC finalized the testing and development of teaching materials for a new kit to be released in 2024, "*Taq* Polymerase Production and Validation."

## National Center for Biotechnology Education

The DNALC continued its work as Genomics Hub of the InnovATE*BIO* National Biotechnology Education Center. This project is funded through the National Science Foundation's Advanced Technological Education (ATE) Program, which supports training for America's workforce. The MOU with Oxford Nanopore and receipt of funding for *DNA Subway 2.0* supported our new hub goal of advancing nanopore sequencing in workforce development. We returned to an active summer of teacher training, reaching 58 faculty at workshops conducted in three states. We collaborated with co-PI Jim Hewlett to introduce metabarcoding and nanopore sequencing at the Great Bay Undergraduate Skills Workshop, May 13–15. Workshops at Maricopa Community College (Phoenix, June 5–7) and Santiago Canyon College (LA, June 7–9) introduced Methods in Personal Genomics and DNA Barcoding, including PCR, DNA sequencing, and *Taq* polymerase production. These workshops were organized with Pushpa Ramakrisna, formerly at NSF, and co-PI Terry Quenzer.

## NSF CyVerse and *DNA Subway*

As we reach the final year of the CyVerse award—the DNALC's largest single NSF grant—there is reason to look back at major accomplishments and ongoing impact. The kickoff meeting for what was then known as the iPlant Collaborative was held at CSHL in 2008. For 15 years, the DNALC has played the primary role in the project's education and outreach activities, ultimately reaching more than 3,700 educators and faculty at more than 100 multiday workshops and training events, in addition to thousands more reached in person and online. Although it is impossible to quantify all the outcomes from such an immense and collaborative project, several of the most important programs in the current DNALC portfolio resulted from our connection to CyVerse. DNA barcoding, for example, began as an exercise for the iPlant *Genomics in Education* Workshops. Just prior to that, in 2010, *DNA Subway* was the first major software platform for the project and served as an "educational discovery environment" for gene annotation projects and transposon detection.

When iPlant started, there were fewer than 10 fully sequenced plant genomes available, and the cost of genome sequencing was ~$10,000 per megabase. Today, there are hundreds of released plant genomes, and generating a megabase of sequence costs less than ten cents! Next- and third-generation sequencing platforms have enabled these cost reductions, including long-read sequencing made possible by Oxford Nanopore. iPlant steeped the DNALC in the world of plant genomics, and expertise and connections to researchers in this space also drew us to Oxford Nanopore early on. *DNA Subway 2.0*'s update will enable the use of Oxford Nanopore reads both for Blue Line DNA barcoding projects and include a new dedicated *DNA Subway* line for sequence assembly (e.g., plasmids, phages, plastids). As the DNALC's formal collaboration with CyVerse draws to a close, the new and expanding DNA barcoding, *DNA Subway*, and Oxford Nanopore sequencing programs are important legacies from our work with iPlant and CyVerse. This collaboration, and the many educators and colleagues with whom we have worked, are sources of immense pride and gratitude for what we have accomplished together.

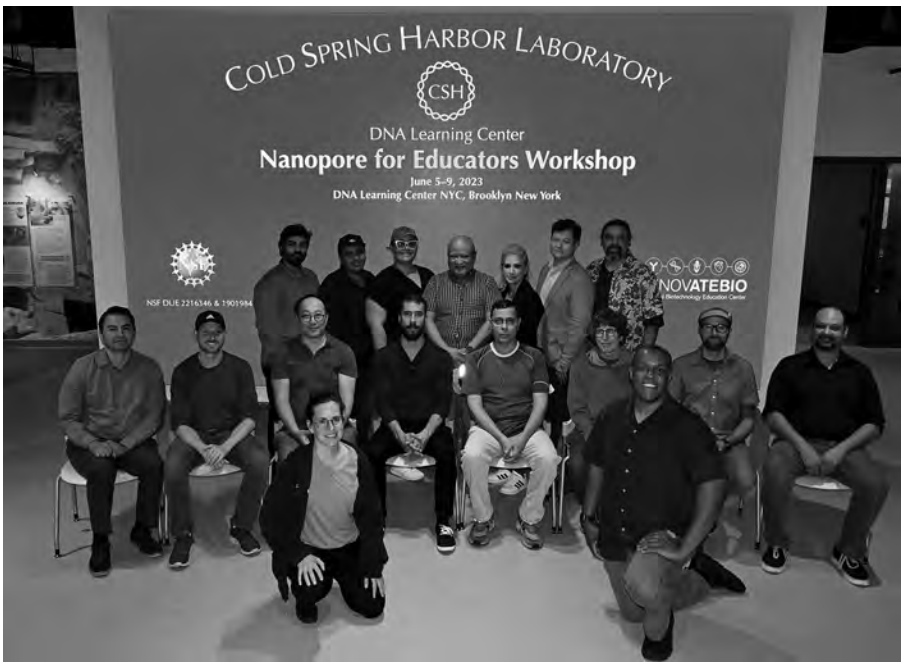
## Phage Biomanufacturing

In 2022, the DNALC was awarded a Future Manufacturing Research Grant (FMRG) titled *"Enabling Cell-Free Engineering and Biomanufacturing of Bacteriophages as a Universal Platform for Tailorable Bioactive Materials."* The project is a collaboration with synthetic biologist Vincent Noireaux and microbiologist Steve Bowden at the University of Minnesota. In year one, there was a necessary change in co-PI at CSHL's DNALC from Bruce Nash to David Micklos. In July of 2023, DNALC staff were trained by PI Vincent Noireaux at a Synthetic Biology Course

held at CSHL. The intensive four-day course included training on experimental design, reagent preparation, and execution of experiments using the cell-free transcription-translation (TXTL) system originally developed in the Noireaux laboratory. The training showcased the flexibility, modularity, and ease of doing molecular biology manipulations in a cell-free system. The experiments focused on three main areas: gene expression from plasmids, modulation of gene expression using CRISPR-Cas9, and the cell-free assembly of phages. In the fall, DNALC staff began regular planning meetings with Noireaux and Bowden to prepare for educator workshops scheduled for the following year.

## NSF IUSE Nanopore

In the second year of our NSF Improving Undergraduate STEM Education (IUSE) pilot grant, *Developing Foundations for Nanopore DNA Sequencing Course-Based Undergraduate Research Experiences at Minority-Serving Institutions*, we have built a small but growing network of educators across the country interested in bringing Oxford Nanopore sequencing into the classroom. More than 60 educators have registered on our QUBES-hosted website, and more than half of the registrants have participated in one or more of our 23 semimonthly faculty mentoring network online sessions. These 1.5- to 2-hour sessions are interactive discussions at which faculty members work together to solve challenges related to teaching. This summer, we hosted our first week-long faculty workshop for nanopore sequencing in Brooklyn at *DNALC NYC*. Sixteen educators from New York to California gathered to work on all aspects of nanopore sequencing from sample preparation to data analysis. Mentoring sessions continue as we reach the last quarter of the project, and we are working on consolidating the teaching materials and experience the group has developed into a collection of resources. The resources will be used in our planned summer workshops at *DNALC NYC* in collaboration with Jeremy Seto from City Tech, as well as with our collaborators in Atlanta at Spelman College, and in collaboration with the University of Puerto Rico, Río Piedras team, hosted at the new Arecibo C3 Center.




Anna Feitzinger and Jason Williams (kneeling) with attendees at the first nanopore sequencing faculty workshop at *DNALC NYC*.

(*Left*) STARS campers enjoy a lunchtime discussion with researcher David Jackson (standing), who spoke about his study of genes that regulate plant growth. (*Right*) A STARS student examines a tiny ear of Teosinte, a wild ancestor of maize that Jackson shared.

## Diversity, Equity, and Inclusion

### *STARS*

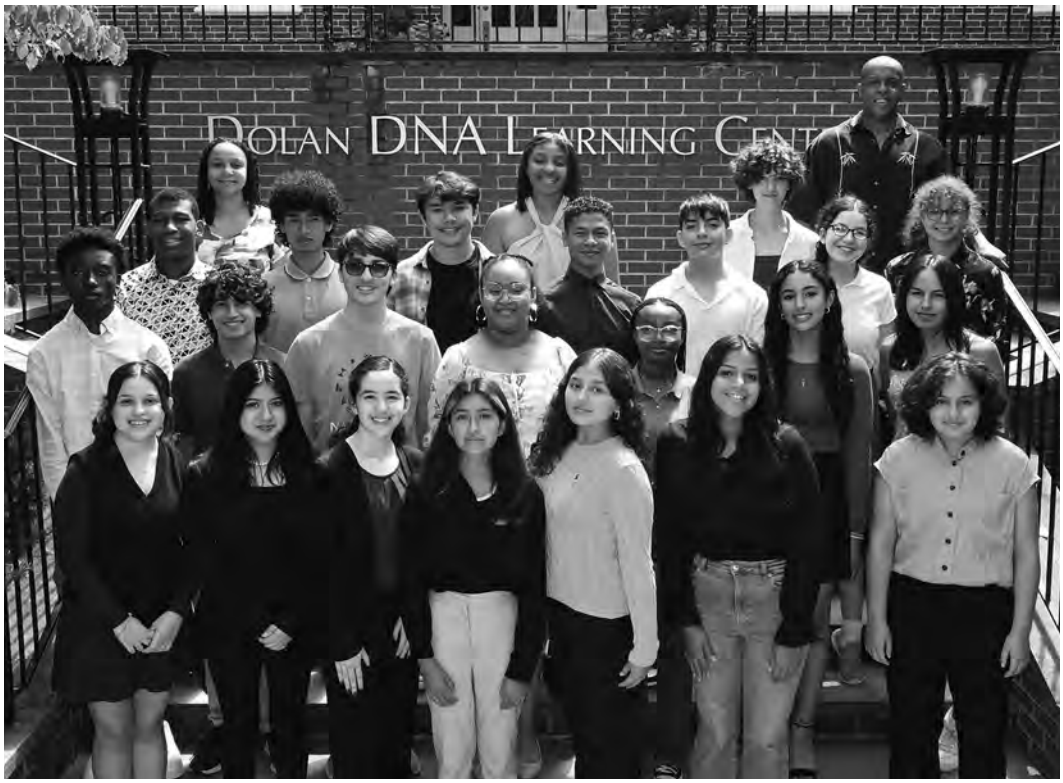
This year, the *Science, Technology, and Research Scholars (STARS)* program passed its 100th student milestone as we welcomed 23 students from 21 school districts, including six private or homeschooled students. We continue to build our network of alumni, engaging students year-round through virtual and in-person mentoring and meetups. STARS physician shadowing grew from one student in the prior year to five student participants this year. Funds gifted by CSHL Trustee Laurie Landeau were used to provide bussing for nine students who would not otherwise be able to attend.

Alumni also integrated into other science research opportunities at CSHL; several students have participated in the Partners for the Future program and other CSHL internships. Carlos Diaz Sanchez (2019) continued study as a volunteer in the Lloyd Trotman laboratory and was also selected for the prestigious MIT Research Science Institute. We also had several alumni participate in a college panel during our family orientation and alumni meetup, including Madison Krug (STARS 2021), a Partners for the Future alumna attending Harvard; Ellis Eisenberg (STARS 2020), a Regeneron ISEF 4th Place Award Winner in Physics and Astronomy currently at Yale; and Jenifer Martinez (STARS 2019), attending Quinnipiac University.

We have linked STARS with partner DEI-focused STEM programs, including a new collaboration with Gina Granger of Hofstra School of Medicine. Ms. Granger is head of their Pipeline Programs, and through shared interests and work with the NYC Gateway Schools, we have already instituted student exchanges where STARS students participate in Hofstra-led programs and several of their students participate in DNALC programs. As we also make and strengthen connections with Brookhaven National Laboratory and Northwell Health, we have new opportunities to strengthen diverse introductory student STEM programs across the region.

### *STEP-UP*

This year, we completed a full cycle as a Coordinating Center for the NIH's Short-Term Research Experiences to Unlock Potential (STEP-UP) program. STEP-UP supports high school students from groups historically excluded from science by offering a paid eight-week summer internship at a research laboratory near their homes. Students work with scientist mentors on a research project to be presented at an end-of-summer symposium. In our second year, we placed 20 students in colleges from Michigan to New York. STEP-UP is highly complementary to our STARS program, and we had our first STARS student, Gabriella Williams (2022), successfully apply for




The STARS 2023 cohort, including workshop organizers/instructors Brittany Johnson (*back left*) and Jason Williams (*back right*).

the program. With the worst of the pandemic behind us, this was the first year we could participate in the in-person symposium held at NIH in Bethesda, Maryland. In addition to organizing travel and lodging for our own students, we managed 20 chaperones who supervised the group of more than 75 total students—from New York to Palau—who traveled for the event. Previously managed by Michelle Juarez, Brittany Johnson has come on board to assist with all aspects of STEP-UP student mentoring and support. As with STARS, we provide regular virtual check-ins with students during the school year to assist high school seniors who continue after the summer with college prep and applications, as well as other research skill-building activities. We are also working to promote the program by building regional relationships with educators who serve historically excluded students.

### *Research Ready*

The year 2023 was the first full calendar year for the Richard Lounsbery Foundation–supported *Research Ready* program in New York City. As part of our commitment to level the science playing field for underrepresented minority and disadvantaged students, two public high schools have been receiving customizable, in-school instruction and opportunities equivalent to our Partner Member program.

- Manhattan Center for Science and Mathematics (MCSM) in Harlem has used their program membership to offer forensics lab activities in school and free field trips to the *Harlem DNA Lab*. Many MCSM students also participated in the *UBRP* and *UBP*.
- Satellite Academy in midtown Manhattan is a transfer high school that admits students who need a fresh start. They participated in hands-on in-school science instruction on a variety of

topics including bacterial transformation, human mitochondrial DNA sequencing, and forensic DNA analysis. Access to lab equipment at Satellite Academy is very limited, so the students really enjoyed the opportunity to feel like real scientists and use lab equipment.

We continue to grow and develop the Research Ready program to adapt to the needs of public schools in New York City and to expand the number of schools we are able to serve.

## DNALC around the World

### *NSF Arecibo C3*

In September, the National Science Foundation awarded a $5.5 million grant to a DNALC-led collaboration to redevelop Puerto Rico's Arecibo Observatory into a new center for science outreach and research. The Arecibo Center for Culturally Relevant and Inclusive Science Education, Computational Skills, and Community Engagement has the mission to integrate Ciencia (Science) across the breadth of STEM disciplines, empower learners with Computación (Computing), and foster Comunidad (Community) through culturally relevant and inclusive values—hence the acronym AC3. Previously managed by NSF's Astronomical Sciences Division, the current award is under the purview of NSF's Directorate for STEM Education. This shift in support opens new avenues for the site's mission, with a primary focus on education.

Before its collapse on December 3, 2020, the Arecibo Observatory boasted the world's largest radio telescope with a 305-meter dish. Although the remaining buildings and spaces within the 118-acre site were operational, the collapse of the telescope due to structural damage cast doubt on the site's future. Fortunately, NSF's commitment to repurpose the site gained support from the 2022 CHIPS and Science Act, which included a provision "encouraging the NSF to consult with other federal agencies to enhance and broaden the Arecibo Observatory's role in Puerto Rico through education, outreach, diversity programs, and future research and technology capabilities."

AC3 builds upon DNALC's growing commitment to Diversity, Equity, and Inclusion (DEI). As the first DNA Learning Center directly supported by a federal agency, AC3 also extends our approach to licensed DNALC centers. The proposal was spearheaded by DNALC Assistant Director for Diversity, Jason Williams, who grew up in New York with his grandfather from Puerto Rico. The proposal leverages existing partnerships, including those developed by Williams for nanopore sequencing, with collaborators at the University of Puerto Rico in Río Piedras. It also brings in new collaborators from the Universidad del Sagrado Corazón and the University of Maryland, Baltimore County. In addition to deploying DNALC field trips and summer camps in new biology labs planned for the space, the project will leverage the observatory's on-site lodging and meeting spaces, fostering integration with the CSHL Meetings and Courses Department. Over the program's five-year duration, efforts from the DNALC and collaborating institutions will emphasize inclusion, accessibility, and cultural relevance. The program aims to enrich undergraduate research as well as student and educator programs in astronomy, data science, and biotechnology.

### *Passaic*

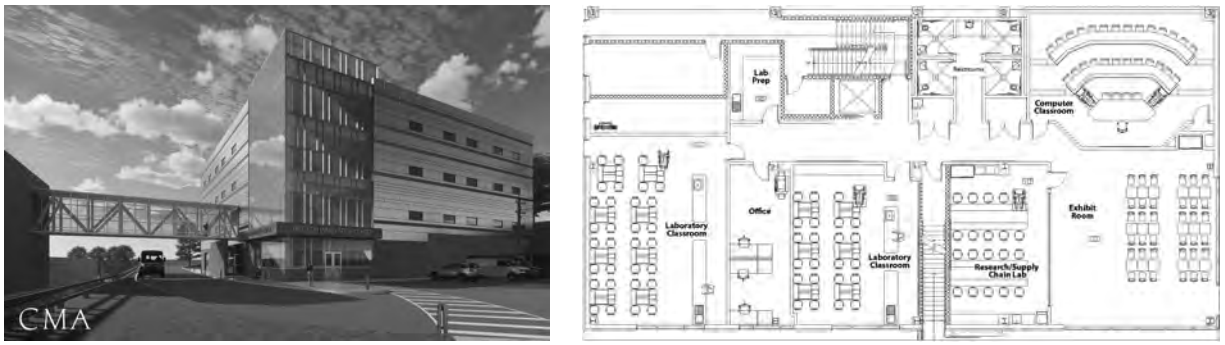
On May 24, the Board of Education of the Passaic County Technical Institute (PCTI) approved a license agreement for a DNA Learning Center (DNALC) in Passaic County, New Jersey. At $210,000 per year, this license is second in size only to the Cold Spring Harbor Asia DNALC in Suzhou. The Passaic center will include two teaching labs, a bioinformatics lab, a research/supply chain lab, a prep lab, and staff offices. Although the 6,500-square-foot facility was originally


NSF | Arecibo C3 Ciencia Computación Comunidad

The Arecibo C3 Kickoff Meeting was held in October. While on site, we toured all facilities and gathered photos and measurements to begin planning for classrooms and exhibition spaces. The Visitor Center (*top left*) includes exhibition space and an auditorium. (*Bottom left*) The rear of the center overlooks the damaged telescope dish that remains after the collapse.

planned to open in winter 2023–2024, delays in construction have pushed opening to later in 2024. Like other locations in the New York metropolitan area, the Passaic center will be staffed and directly run as part of the core DNALC enterprise—so we can ensure that students throughout the metro area will have the same great lab experience at any of our centers.

The entire Passaic project is in synch with the DNALC's role as the Genomic Hub of InnovATE*BIO*, the National Biotechnology Education Center—whose goal is to promote workforce development at two-year colleges. The Biotechnology Innovation Center, in which the Passaic DNALC is located, literally bridges the campuses of a 5,000-student vocational–technical high school and a community college. High school students enrolled in the PCTI biotech program will receive dual enrollment credits at Passaic County Community College and graduate with a 60-credit AS degree in biotechnology. Construction of the Innovation Center, including the DNALC, was funded by $25 million in state and local bonds, and each student's AS degree is entirely funded by state and federal grants. About 150 students have been recruited for the




Architectural rendering and floor plans for the future DNA Learning Center in New Jersey.

inaugural biotech class, immediately making it one of the largest dual-enrollment programs in the county. Similar cohorts will be added over the next three years to create a full complement of freshman through senior students.

The Passaic DNALC will support students in the dual-degree program with specialized lab and research experiences, including a supply chain lab in which students will run a "company" that produces biotech reagents for use in high schools throughout Passaic County. This will include producing *Taq* polymerase, the key reagent of the widely used PCR. This project is based on a specialized plasmid, which constitutively produces *Taq* polymerase, and simplified extraction methods developed by the DNALC. Our facility will also serve approximately 1,000 students enrolled at a science magnet school on the PCTI campus. Middle and high school students from throughout Passaic County, including a large proportion of underrepresented minority and disadvantaged students, will have access to the DNALC's full menu of academic-year lab field trips and summer camps.

The Passaic County venture will further bolster the proposition that we can potentially expand the "run by the DNALC" model to sites across the country. We hope to test this soon, as we are currently discussing developing a DNALC with the Morehouse School of Medicine and Atlanta Public Schools.

### *Regeneron DNALC*

*Regeneron DNALC* continued to see an increase in field trips and summer camp attendees. A total of 1,816 students visited for field trips and 339 campers attended 18 camps during the summer. We partnered with Regeneron for their annual *Day for Doing Good* event, at which DNALC-trained Regeneron employees volunteered to lead a series of student lab activities to help increase STEM exposure in the local community. Regeneron volunteers taught a total of 236 students from Sleepy Hollow Middle School and Intellectus Preparatory Charter School. We continue to nurture relationships with schools that visit *Regeneron DNALC* for field trips and we perform outreach to form new relationships by hosting teacher training and *Saturday DNA!* sessions (122 students) as well as attending STEM fair events. *Regeneron DNALC* is on pace to reach full occupancy in 2024–2025.

### *Notre Dame*

The Notre Dame DNA Learning Center (ND DNALC) continues to rebound from COVID-19 shutdown challenges. Although visitation has not yet returned to pre-2020 totals, the growth trajectory remains strong with 425 student lab exposures—60% of whom were from low-income schools—a 45% increase over 2022. Summer programs brought campers from all over the United States, and one international student from Barbados. The high school research program continued to thrive, drawing participants from the pool of 2022 summer campers, and one student advanced

to the Regeneron International Science and Engineering Fair for the second consecutive year. Additionally, the ND DNALC continued to participate in broad College of Science and Notre Dame activities throughout the year.

The Advisory Board was reformed and diversified to include community, regional, and other members who will bring valuable perspectives to ensure continued growth and success. In November 2023, the primary donors of the ND DNALC, Dr. John and Heidi Passarelli, were celebrated for their cumulative support and dedication of the endowed Directorship, Dr. John and Heidi Passarelli Family Director of the DNA Learning Center.

## International Partnerships

### *Suzhou*

*DNALC Asia* welcomed 12,383 visitors in 2023, including 10,014 course participants, 1,229 visitors, and 1,140 outreach participants. The center hosted projects at both district and municipal levels, with activities spanning 12 cities and six provinces, and actively collaborates with various institutions, including government bodies, media outlets, and renowned scientific research institutions in China and worldwide.

In July, first-time application of third-generation sequencing (Oxford Nanopore technology) was introduced to Chinese high school students by DNALC Executive Director Dave Micklos and Educator Jeffry Petracca in *Genome Science* and *DNA Barcoding Research* summer workshops. Suzhou students' experiment results were published by a professional scientific research database (NCBI) for the first time.

首次面向中国高中生授课！全国37名高中生尝鲜第三代DNA测序技术

Dave Micklos and Jeff Petracca introduced nanopore sequencing to *Genome Science* and *DNA Barcoding* students in July. The article title translates to "First class for Chinese high school students! 37 high school students from across the country tried the third generation DNA sequencing technology." Source: https://app.suzhou-news.cn/news/300576610

The center organized several successful public outreach events in 2023. The *Colorful Wetlands* attracted 51 schools and 1,555 student submissions, with their work displayed at the center. They collaborated with Illumina to deliver the Life Science Enters Campus project—providing DNA science classes, online and offline, to more than 3,448 elementary and middle school students in six provinces. Additionally, the 7th Annual Suzhou Young Life Scientist Cultivation Program expanded its reach by involving teachers, fostering scientific literacy in both students and educators. The program's success is reflected in the increased participation (173 schools and 1,632 students) and the development of project manuals to guide future iterations.

The center received news of college admissions from 42 students who participated in one or more programs at *DNALC Asia*. From 2017 to 2022, 30 out of 76 international high school students who reported their college admissions (out of 131 total students who conducted 47 DNA Barcoding Research Projects) were accepted into top universities abroad, resulting in a 39.5% acceptance rate.

### *Beijing No. 166 Middle/High School*

In the wake of consistent travel restrictions early in the year, we continued to work with Beijing No. 166 faculty to implement modified *On-Demand* Camps. Using comprehensive prerecorded video instruction, 19 high school students performed *Genome Science* labs using real equipment in their classroom over winter break. Additional recordings of class data analysis were filmed and shared to help with DNA sequence analysis during the course. Fortunately, travel restrictions were lifted by summer, and meetings with faculty from the school resulted in a plan to resume in-person instruction. To start, an educator will visit Beijing for a *Green Genes* workshop in 2024, moving us one step closer to fulfilling our prepandemic lab instruction contract.

### *Beyond Beijing and Suzhou*

In April, Peking University Affiliated High School in Beijing contacted us about conducting a summer study abroad program in New York. Because of limited time for planning and U.S. visa applications, the program was postponed until February 2024. Twenty-four students and two biology teachers will participate in a two-week high school workshop at *DNALC NYC*. We have also received inquiries from several Chinese education consultant agencies regarding their summer travel programs. With relaxed travel restrictions and the growing reputation of DNALC programs, we look forward to new partnerships in China, with the goal of hosting international students for winter and summer camps at *DNALC NYC* every year.

### *DNA Learning Center Nigeria*

We continued to provide financial and reagent support for *DNALC Nigeria*, located in Enugu State—a region that has never fully recovered since its failed attempt to secede from Nigeria in 1967 resulted in civil war. The 2023 election in Nigeria brought increased inflation and economic instability, including a tripling of fuel prices, which severely strained the center's already tight budget. Political unrest, insecurity, and safety concerns in Southeast Nigeria limited engagement with local communities and the ability of schools to travel to the center. Even so, the center conducted six extended workshops, reaching 345 students, teachers, and industry professionals. In September, Michael Okoro and George Ude supported a workshop in DNA barcoding conducted at the SENA Institute of Technology, in Ghana.

## Lab Instruction and Outreach

With visitation building at New York DNALC facilities—Dolan DNALC, *Harlem DNA Lab*, *DNALC NYC,* and *Regeneron DNALC*—income from local school programs and camps reached


Students teach their families all that they learned during the week at a *Fun with DNA* Parent Day at Regeneron DNALC.

a record high of $1.5 million, 26% more than pre-COVID-19 income. Focused email, social media, and Google Ads marketing to new audiences in Westchester and NYC paid off with a 30% increase in field trip visitation over 2022. Combined, a total of 21,230 students attended lab field trips, and 6,698 received in-school instruction. An additional 963 used DNALC footlocker kits, 186 of whom were conducting independent research through *UBP*, *UBRP*, or *BLI*. Tuition assistance for field trips to Dolan and *Regeneron DNALC* was provided for 1,704 students from the Amityville, Brentwood, Floral Park, Freeport, Hempstead, Yonkers, Malverne, Nyack, NYC, Ossining, Uniondale, and Westbury school districts. We offered a full schedule of in-person summer camps at three locations. We hosted 662 campers in Cold Spring Harbor and 339 in Sleepy Hollow for nearly full occupancy in its second summer season! In Brooklyn we hosted 489 campers, including 168 who attended subsidized *UBRP* prep courses.

Our operating contract with the City University of New York requires that our Brooklyn center provide scholarships to 50% of New York City public school students who participate in lab field trips. We exceeded this statutory challenge in 2023, providing academic year scholarships to 56% of students at our NYC centers. Scholarships also were provided to 29% of summer campers in NYC. In all, 4,124 NYC students received scholarships valued at $161,760. Across all DNALC locations, 5,909 students received scholarships totaling $254,600. Endowment provided $171,328 of this amount. Grants from the William Townsend Porter Foundation, National Grid, STEM Matters NYC, and the City Tech STEP Program provided $83,272.

In partnership with CSHL Women in Science and Engineering (WiSE), we hosted the seventh annual WiSE *Fun with DNA* summer camp. Twenty-one young female science enthusiasts, including one who received a full scholarship from WiSE, had the opportunity to meet engaging role models with careers in science. After completing the core *Fun with DNA* curriculum each day, campers participated in WiSE activities on advanced topics like cancer research, neuroscience, and gene expression. In addition to the Parent Participation showcase on the final day of camp, parents and campers enjoyed guided tours of the CSHL campus.

*Sustaining Membership* enrollment in our School Membership Program increased to 21 schools, with the addition of Archbishop Stepinac High School at *Regeneron DNALC* and Magen David Yeshivah High School at *DNALC NYC*. Dolan *Associate Partner* Friends Academy continued development of their research program with implementation of DNA barcoding projects for

entry-level students and metabarcoding projects for returning researchers, whereas the Glen Cove City School District focused on consistent lab exposures for every Living Environment and AP Biology class and continued support for DNA barcoding teams.

*Dolan Partner Members* Long Beach City School District and Massapequa School District entered year three of membership with plans to continue implementation of DNALC research supports and uniform exposures for specific student cohorts. In the Massapequa research program, every ninth grader completed a DNA barcoding project, whereas every eighth grader had three structured exposures to learn lab techniques. Additionally, all Living Environment classes completed electrophoresis labs, and all fifth grade classes performed microscopy investigations. Long Beach continued lab field trips for all of the eighth grade *Living Environment* and life science classes, and added fifth grade field trips to the mix. A new research teacher was trained, so, with our help, all ninth grade researchers will participate in *BLI*. In the summer, 40 students attended *Fun with DNA* and *World of Enzymes* camps taught at Long Beach Middle School.

As part of our ongoing partnership with St. Dominic High School, 11 students participated in the *Molecular and Genomic Biology Research* course, where they received daily hands-on instruction in *DNA Barcoding*, *DNA Science*, and *Genome Science* curricula from DNALC educators. After a year-long hiatus, the Cold Spring Harbor High School *Molecular and Genomic Biology* course returned. Participants spent two periods every other day this fall immersed in experimentation and independent projects at the DNALC. All students in these courses develop research projects and present posters at the annual *Barcode Long Island* Symposium.

Six independent schools benefitted from custom instructional sequences and advanced electives as *DNALC NYC Partner Members* or *Associate Partners*.

- Dwight School integrated DNALC laboratories within middle school science electives. Grade seven students explored bacterial transformation through inquiry, whereas grade eight students analyzed DNA evidence in a forensic mystery.
- Portfolio School grade five and six students used experiments and models to learn the fundamentals of chemistry through inquiry. In a culminating research experience, they used DNA barcoding to identify plant and insect species.
- Lycée Français de New York implemented genetics programs in their AP Biology courses. Grade seven students solved a mystery using DNA evidence. Grade 10 students explored Genetically Modified Organisms, and grade 11 studied human mitochondrial DNA and what it can tell us about evolution.
- At Marymount School of New York, genetics programs were incorporated as key parts of the biology curriculum at multiple grade levels, including Advanced Molecular Biology. Grade six students explored the roles of gene mutation and natural selection in evolution.
- St. David's School integrated basic genetics with existing curricula in grade five. Grade eight used DNA barcoding to survey the ants of Central Park.
- The Chapin School coordinated genetics programs at several grade levels, including the advanced *Molecular Genetics* elective. Grade nine students used PCR and restriction enzyme analysis to determine their genotypes for a bitter taste allele, which they correlated with their phenotypes.

During the year, our Ötzi the Iceman exhibition drew 3,637 visitors, most of whom were students on field trips. Fifty visitors were members of the general public taking self-guided tours. Sixteen *Saturday DNA!* sessions drew 417 participants at the Dolan DNALC and *Regeneron DNALC*. Participants learned about the true story behind the mystery of Anastasia Romanov, crime scene analysis, what Ötzi the Iceman can tell us about ancient human life, and the curious way that milk has shaped human evolution. Some created artwork by "painting" with genetically engineered

Cris Fernandez-Marco leads a DNA extraction activity during the *Day in the Lab* event.

bacteria, whereas others learned how to construct custom plasmids, such as those used to manufacture human insulin. An additional 140 students attended *School Break Bio* classes over winter and spring breaks, ranging from full-day sessions on GMOs, forensics, and Ötzi the Iceman, to two-hour, field-trip style lessons on infectious disease and electrophoresis. On Veteran's Day we hosted the second annual CSHL *Day in the Lab* event for local youngsters. With the help of the CSHL Association Directors, we enrolled 106 children and their parents for fun hands-on activities designed to expose a young audience to STEAM (Science, Technology, Engineering, Art, and Math), including observing biochemical reactions, using stereomicroscopes, and modeling DNA.

Math for America teachers attended a mini-course in Brooklyn where they learned about human mitochondrial sequencing. Teachers who attended mini-courses became eligible to borrow DNALC footlockers to implement these labs at school, and many decided to bring students on a field trip after seeing our new space. STANYS (New York State Science Teachers Association) teachers attended a professional development session in Brooklyn where they learned about DNA fingerprinting and how to teach about gel electrophoresis. The *DNALC NYC* hosted STEM Teachers NYC for a Molecular Modeling workshop when Daniel Fried from Biochemistry Literacy for Kids taught teachers and DNALC staff about molecular modeling for the classroom, and teachers learned about the Ötzi the Iceman exhibit and field trips.

*DNALC NYC* staff participated in Brooklyn community events to meet our neighbors and do hands-on science. Families attending the Brooklyn "Atlantic Antic Festival" and SUBMERGE Marine Science Festival learned about eDNA or did "pipette painting" with our team.

The virtual "Meet a Scientist" series continued to facilitate connections between public audiences and Regeneron Pharmaceuticals researchers. From January through May, four presenters were invited to share their career journeys and research endeavors. Research specialist Zachary Oberholtzer discussed his work on biochemical and biophysical characterization in drug development, and Senior Research Associate Terrence Turner talked to us about the use of DNA sequencing in the modern age. Dr. Turner also highlighted Regeneron's commitment to DEI initiatives, as well as their efforts in STEM outreach programs. Assistant Investigator Dr. Joel N.H. Stern shared his career path, emphasizing his approach to balancing teaching and research responsibilities, and veterinarian and Associate Director of Veterinary Services and Vivarium Operations

in VelociGene, Dr. Stephanie E. Woods, provided her perspective on the ethical treatment of animals in research settings, underscoring the importance of safeguarding animal welfare while conducting scientific experiments. These "Meet a Scientist" presentations are archived on our webpage and have received 1,782 views.

Annual training with graduate students from the CSHL School of Biological Sciences continued with the 2023 cohort of first-year students. Over a series of 12 half-day sessions, students worked with DNALC instructors to develop skills for communicating science to a variety of audiences. Beginning with observing field trips, then progressing to co-teaching and independently teaching lab classes, each grad student interacted with both middle and high school students, and rounded out the semester with elective teaching or lab development projects.

## *BioMedia* Visitation and Projects

In 2023, 3.5 million visitors accessed our suite of multimedia resources, a decrease from the previous year. Our YouTube channel had 1,236,206 views with 44,306 hours of watch time and added 3,941 subscribers. Google Analytics counted two million user sessions on DNALC websites, 85.3% of the prior year. These visitation statistics include a transition beginning in July to an updated version of Google Analytics, the tool we use to track visitation data on our websites. Although the old and new versions are not quite an "apples to apples" comparison, the metric reported, user sessions, is fairly similar in months where we collected data from both versions. Overall, decreased visitation is likely attributed to shifted focus of the BioMedia team away from developing online educational content as well as aging of our existing web resources. *3D Brain* and *Gene Screen* smartphone/tablet apps were downloaded 235,801 times, including 4,615 *3D Brain HQ* in-app purchases earning $3,136.


*All the World in New York City* features some faces of the many races, cultures, and ethnicities represented by people living today in New York. Photographs and subject biographies were judged "blind," without submitter's name. The photographers of the selected images received a $500 award.

We benefitted from an ongoing nonprofit Google Ads grant that funds ads for our websites in the Google search engine; ads generated 126,720 impressions and 13,763 clicks equaling a 10.86% click-through rate; the equivalent of $18,840 in advertisement spending.

We continue to develop the *DNALC NYC* exhibition. In consultation with CSHL trustee Jeanne Moutoussamy-Ashe, the *All the World in New York City* gallery space was completed in April. We also initiated a unique collaboration with City Tech's Communication Design Department (COMD), which is located on the floor below the *DNALC NYC*. COMD freshman students sequenced their own DNA and then integrated this personal perspective into assigned exhibition designs to complement the DNA maps to be displayed on the video wall. In the spring semester, advanced Design Team Students proposed designs for DNA Future, which will feature work by CSHL scientists. In the summer session, students created designs for interior and exterior signage; DNALC staff voted on the submissions and plan to install a lobby mural and exterior sign in spring 2024. The hope is that these students will continue to develop exhibit components in advanced classes or as interns.

We also started planning a display on Eugenics and immigration to acknowledge 100 years since the Immigration Act of 1924 imposed stringent quotas on people coming to the United States. The *BioMedia* team continued to provide additional support for DNALC programs through print and web design, photography, and videography.

## Staff and Interns

2023 saw a number of significant staff changes, with poignant partings and promising engagements. From our Dolan location, we said goodbye to our Assistant Director of Science, Bruce Nash, Ph.D., and our Manager of Student and Public Research, Sharon Pepenella, Ph.D.

Bruce Nash was recruited in 2005 to spearhead our initiative to popularize *Caenorhabditis elegans* and RNAi in biology instruction. He was instrumental in developing curriculum for high school labs and summer workshops, including an NSF-funded integrated experiment- and bioinformatics-based curriculum on RNAi in *C. elegans*, which later became the *Silencing Genomes* Workshop with a supporting website. In 2007 he was promoted to Assistant Director of Science and in 2012 co-authored the textbook *Genome Science: A Practical and Conceptual Introduction to Molecular Genetic Analysis in Eukaryotes*. Bruce fostered a collaboration in 2013 with the first regional Doshi Science, Technology, Engineering, and Math (STEM) high school, which prepared students for the evolving high-tech science arena. In 2017, he received a supplemental grant from the NIH Big Data to Knowledge Program to adapt microbiome research for high school students. In 2019, he directed the start-up for InnovATE*BIO* (for high schools and community colleges) and iUSE (for four-year colleges/universities), programs that provide teacher and student support, troubleshooting, and provision of reagents and footlockers. Bruce left the DNALC after 18 years of service.

Sharon Pepenella joined the teaching staff in June 2015 and plunged into teaching and initiating new projects. Her carefully constructed lesson plans are an indispensable resource for training new DNALC teaching staff. She took on the program manager role for the NIH SEPA *BLI* and *CDBN* projects and managed *Partner Member* schools. In April 2021, she advanced to the position Manager of Student and Public Research. After organizing her final *BLI* Symposium in June 2023, she transitioned to Brookhaven National Laboratory as a Senior Research Programs Representative through the Office of Educational Programs, where she will focus on developing and evolving formal and informal science education research projects. Although we will miss Sharon's unwavering commitment to science and education, along with her festive holiday hats, we look forward to continued collaboration with her as a BNL representative.

We also said goodbye to Michelle Juarez, Ph.D., Arden Feil, and Chaunna Henry from *DNALC NYC*, and Jack Kellogg from the *Regeneron DNALC*.

Michelle Juarez joined *DNALC NYC* in March 2022 as the Assistant Director of Diversity and Research Readiness. She oversaw collaborations and partnerships to prepare students for research careers, expanded DNALC's outreach to underrepresented STEM students, worked with research-ready schools and CUNY faculty, and recruited scholarship students for the summer. In August 2023, she accepted a position as a Diversity Outreach Coordinator at Stony Brook University.

Arden Feil began her journey with us in June of 2021. She immediately helped with the start-up of the Brooklyn facility during our inaugural summer. Arden managed *UBP* and *UBRP*, including recruiting participants, conducting training workshops, providing scientific support, and organizing the annual symposium. She also worked with *Associate Partner* the Portfolio School and managed purchasing supplies, reagents, and equipment for the *DNALC NYC*. In April she accepted the Science Research Mentoring Consortium Manager position at the American Museum of Natural History (AMNH), where she continues to collaborate with the *UBRP*.

Chaunna Henry joined the Brooklyn team in November 2022 as the Administrative Manager at *DNALC NYC*. She left in April for a position with a remote work option. Her extensive background in a variety of roles and her adaptability, creativity, and problem-solving skills were invaluable during her short time with the DNALC.

In November of 2021, Jack Kellogg joined the DNALC in preparation for reopening the *Regeneron DNALC*. He administered classes and laboratory experiments; prepared, tested, cataloged, and ordered supplies and reagents for all programs; and hired and managed high school interns. Because of family relocation, Jack joined the staff at Harvard Medical School in April as a research technician, where he prepares ancient human samples for next-generation sequencing and maintains lab safety and clean room protocols to uphold the integrity of the ancient DNA.

2023 brought new staff and staff changes to all DNALC locations.

Carol Henger started with the *Regeneron DNALC* team in March. Originally from Dallas, Texas, she has a Ph.D. in biological sciences from Fordham University, an M.A. in animal behavior and conservation from Hunter College, and a B.S. in environmental science from Texas Christian University. She has taught genomics at the university level, and as a postdoctoral fellow she developed a protocol for detecting big cats in the wild using DNA analysis. Carol has mentored students on wildlife research projects and has authored multiple scientific papers on coyotes and bird behavior.

In May, we welcomed Allison Astudillo, lab manager, and Keith Bannerman, educator, to the Dolan DNA Learning Center team.

Born and raised on Long Island, Allison Astudillo's strong interest in STEM with a focus on health can be seen throughout her academic and professional career. She is currently pursuing a Ph.D. in Health Sciences from Liberty University, where she received an M.A. in medical sciences, and she also has a B.S. in health studies from Monmouth University. Her previous work experience includes serving as a quality control–focused microbiologist for AKORN, and a medical scribe for City MD. In addition to filling the crucial role of lab manager and managing the interns, Allison is an Adjunct Assistant Professor for the Biological Sciences Department at New York Institute of Technology.

Keith Bannerman, who started in May, earned a B.S. in biochemistry and a B.A. in philosophy from Stony Brook University Honors College, then went on to complete an M.A. in bioethics, medical administration, and compassionate care. While pursuing his graduate degree, Keith served as a genomics researcher at Allied Microbiota, where he studied bacterial remediation techniques.

In May, *Regeneron DNALC* found a new educator in Arie Kaz. Originally from Scranton, Pennsylvania, Arie has an M.S. in secondary science education from the University of Scranton, and a B.S. in life science from Penn State University. Arie was a high school teacher in North Carolina, where he taught several courses including AP Biology, Honors Research Methods and Techniques, and Honors Biology. Arie brings with him a strong background in NGSS standards and a passion for teaching.

New staff in 2023 (*left* to *right*) in order of arrival: Carol Henger, Allison Astudillo, Keith Bannerman, Arie Kaz, Christina Newkirk, Ria Jasuja, and Shreemattie Budhram.

Two staff joined the *DNALC NYC* educator team in October: Christina Newkirk and Ria Jasuja.

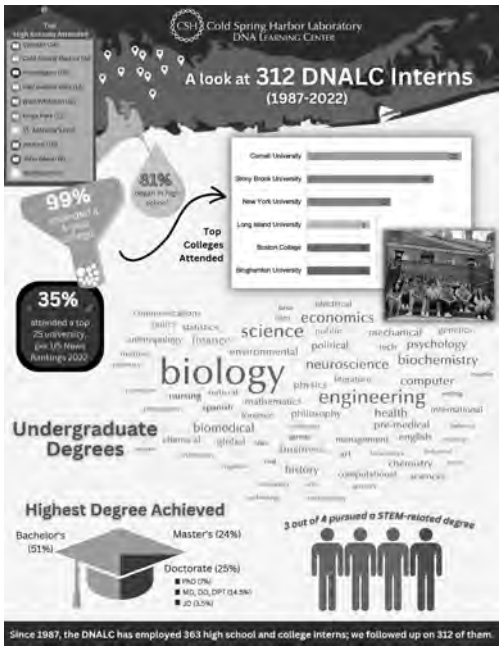
Christina Newkirk has an M.A. in environmental conservation education from New York University, and a B.A. in English from the University of Iowa. After graduating, Christina realized she wanted to share her love of science with diverse audiences. She has worked at the New York Academy of Sciences, Christadora, the American Museum of Natural History, and the Intrepid Museum in roles focused on science education and outreach.

Ria Jasuja holds a B.A. in biology, with a concentration in molecular biology/biochemistry, from Washington University in St. Louis. As an undergraduate, she was a research technician in the Goldfarb Laboratory for the Department of Cell Biology and Physiology at Washington University School of Medicine. After graduating, Ria stayed on at Wash U as a laboratory technician. Ria co-led a summer program, "Astronaut Training Camp—Mars," and volunteered at the Challenger Learning Center in St. Louis.

Shreemattie "Sarika" Budhram joined the Brooklyn team in December as the Administrative Manager at *DNALC NYC*. Sarika has a B.S. from the University of Phoenix and an A.A. in business from Western International University and brings experience in business administration and start-up companies. She served as a management consultant and was director of operations of Azimuth Resources for the Australian Consulate in Guyana. She volunteered in Guyana for more than 15 years, raising funds for several charitable initiatives.

January saw the first promotional title change for staff; Chun-hua Yang, now Associate Design Director, continues to be a selfless coworker adept at anticipating department needs and working collaboratively. In addition, she has stepped in to facilitate communication with collaborators in China and ensures annual flow of intellectual property, including field trip, summer camp, and website materials.

The departures of Bruce and Sharon provided opportunities for Anna Feitzinger, Cristina Fernandez-Marco, and Jeffry Petracca to take on new roles with increased responsibilities. Anna




(*Top*) Neil and Mina at work on the (*bottom*) intern survey.

stepped in to be DNALC's Assistant Director, Science, and picked up where Bruce left off on two NSF projects. Cris is now Manager, External Collaborations, and lead of the *BLI* program. Jeff, previously a part-time educator, is now full-time Manager, Student and Public Research, and leads the NIH *CDBN* project.

Kelly Eames was elevated to Partnership Manager in August after taking on middle school field trip reservations for our CSH location and serving as the DNALC Relationship Manager with School Membership Program districts. Kelly helps to shape customized curricula of field trips, in-school instruction, and research mentorship for their students, as well as arranging focused training and follow-up consultation for Partner faculty.

Since the inception of the STARS program, Brittany Johnson has assumed multiple responsibilities that advanced the DNALC's mentoring and inclusion initiatives, including teaching the high school–level STARS curriculum and providing support for monthly student mentoring sessions. She has collaborated with partners to facilitate student participation in medical shadowing programs and student placement in additional research initiatives. Following Michelle Juarez's departure, Brittany seamlessly took over NIH STEP-UP program responsibilities and was promoted to Manager, Diversity and Research Readiness.

Since the DNALC opened, we have relied on high school and college interns to support our day-to-day operations. An internship offers students the unique opportunity to gain real laboratory or design experience in an educational environment. We have often wondered how DNALC interns fare after they leave our employment. So, we were pleased when Brittany Johnson worked with interns Mina Sarmas (Brown University) and Neal Mehta (Boston College) to conduct a survey of DNALC interns going back to 1987! The following infographic summarizes the impressive results.

This year an amazing group of interns helped out, and we said farewell as others left for college:

**High School Interns**

*Dolan DNALC*

Alexander Gottlieb, St. John the Baptist High School
Charlotte Gordon, Cold Spring Harbor High School
Daniel Galvin Gusmano, Portledge High School
Francesca Mango, St. Dominic's High School
Genevieve Decker, St. Dominic's High School
Ian Smith, Cold Spring Harbor High School
Jakob Rechtweg, Friends Academy
Jordyn Shafran, Pine Crest High School
Alexis Panebianco, Mineola High School
Andres Orellana, Portledge High School
Lauren Belkin, Syosset High School
Lauren Graziosi, Syosset High School
Lily Rodgers, Syosset High School
Ryan Koenigsberger, Cold Spring Harbor High School
Ryan Lee, Plainview Old Bethpage JFK High School

*Regeneron DNALC*

Ashley Alexander, Yonkers High School
Katelyn Battacharia, Ossining High School
Jacob Leobandung, John Jay Senior High School
Faye Luneau, Pelham Memorial High School
Jackson Meyercord, Bronxville High School
Naflah Mohammed, John Jay Senior High School
Maya Shetty, Scarsdale High School
Brady Wang, Horace Greeley High School

**High School Interns Departing for College**

*Dolan DNALC*

Croi Spillane, Quinnipiac University
Min Hur, University of California Los Angeles

**College Interns**

*Dolan DNALC*

Charlie Whitman, The Taft School
Elena Gnilitskaya, SUNY Stony Brook
Holly Schadler, Fairfield University
Ian Quinn, Tulane University
Jason Long, SUNY Stony Brook
Jillian Hooey, University of Chicago
Julia Salatti, Cornell University
Juliana Dovi, SUNY Stony Brook
Kimberly Cardinale, Oxford College of Emory University
Maggie Wang, SUNY Stony Brook
Michael Stabile, Cornell University
Nicholas Liotta, Nassau Community College
Nicholas Stabile, University of Notre Dame
Raquel Belkin, SUNY Stony Brook
Rosemarie Russo, Siena College
Ryan Cui, SUNY Stony Brook
Sandhya LoGalbo, Hofstra University
Timmeree Koepele, University of New Mexico

*DNALC NYC*

Ashley LaSalle, New York City College of Technology
Cristofer Hernandez, Hunter College
Sebastian Maurice, The City College of New York
Jamie Price, Lehman College
Morgan Serbagi, Hunter College
Derbie Desir, New York City College of Technology
Faith Tsentner, Hunter College
Tess Fleming, Brooklyn College
Jens Salva, The City College of New York
Ella Cervi, Princeton University
Rosalie Ye, Stonybrook University
Sheilaya Gresham, Hunter College
Milany Bruno, Hunter College
Marc Louis, John Jay College of Criminal Justice
Harold Miller, Borough of Manhattan Community College

*Regeneron DNALC*

Madelyn Meehan, Penn State
Kelly Tito, University of Massachusetts, Amherst
Mikayla Tucci, Sacred Heart University
Lily Wong, Northeastern University

## Sites of Major Faculty Workshops

Program Key: *Middle School* High School **College** Informal Education

| Location/State | Institution |
|---|---|
| ARIZONA | **Maricopa Community College, Phoenix** |
| CALIFORNIA | **Santiago Canyon College, Orange** |
| NEW HAMPSHIRE | **Great Bay Community College, Portsmouth** |
| NEW YORK | Dolan DNA Learning Center |
| | ***DNA Learning Center NYC*** |
| TENNESSEE | Discover Life in America, Gatlinburg |
| UTAH | **Natural History Museum of Utah, Salt Lake City** |

## Workshops and Visitors

| | |
|---|---|
| January 13–18 | "A Moderated Discussion on the Missing Components in Genomics: Justice, Equity, Diversity, Inclusion," "Advanced Computational Methods—CyVerse for Machine Learning, Containers and Clouds," "Understanding Barriers to Bioinformatics Education—Network for Integrating Bioinformatics into Life Sciences Education," International Plant and Animal Genome Conference (PAG30), San Diego, California |
| January 19 | "Meet a Scientist: Zachary Oberholtzer," Virtual Webinar, *Regeneron DNALC* |
| January 20 | Pop-Up Training for CUNY Biology Educators: Getting Started with Barcoding, Metabarcoding, and Nanopore DNA Sequencing, *DNALC NYC* |
| January 21 | *Saturday DNA!* "Pathogens, Plasmids, and Petri Dishes—Oh My!" DNALC<br>*Urban Barcode Project* Open Lab, *Harlem DNA Lab* |
| January 23 | Pop-Up Training for CUNY Biology Educators: Getting Started with Barcoding, Metabarcoding, and Nanopore DNA Sequencing, *DNALC NYC* |
| January 25–26 | Site visit and tour by Kwasi Agbleke, Sena Institute of Technology, DNALC |
| January 26 | Site visit and tour by Joe Leniczak and Kissaou Tchendre, Austin Community College, *DNALC NYC* |
| January 28–31 | Site visit and tour by Mrs. Lorena Martínez Rodríguez, Mrs. Verónica González de Alba, Dr. Armando Barriguete Meléndez, and Dra. Liliana Hernández, Aguascalientes DNA Learning Center Project Delegation, *DNALC NYC*, DNALC & *Regeneron DNALC* |

| | |
|---|---|
| January 30 | RockEDU STEM Research Teacher Summit, Rockefeller University, New York, New York |
| February 3 | The Young Women's Leadership School's Summer Enrichment Fair, Harlem, New York |
| February 11 | *Saturday DNA!* "Pollen Tells a Story," *DNALC NYC* |
| | *Saturday DNA!* "A Royal Ruse," DNALC |
| | *Saturday DNA!* "Agar Art," *Regeneron DNALC* |
| February 14 | The Young Women's Leadership School's Summer Enrichment Fair, Brooklyn, New York |
| February 22 | *School Break Bio!* "Science Detectives: Use Forensic Techniques to Solve a Crime!" *DNALC NYC* |
| | *School Break Bio!* "GMO Analysis: Is My Snack Genetically Modified?" *DNALC NYC* |
| February 24 | *School Break Bio!* "Infectious Diseases," DNALC |
| March 2–4 | "Proteins, Biology and Artificial Intelligence," National Academy of Sciences Kavli Frontiers of Science 33rd U.S. Symposium, Arnold and Mabel Beckman Center, Irvine, California |
| March 10–11 | Site visit & tour by Dr. Santiago March, Jack Landsmanas, and Dr. Armando Barriguete Meléndez, EDOMEX DNA Learning Center Project, Mexico, *DNALC NYC* and DNALC |
| March 11 | *Urban Barcode Research Program* Open Lab, *DNALC NYC* |
| March 15 | *Urban Barcode Research Program* Open Lab, *DNALC NYC* |
| March 16 | "Meet a Scientist: Terrence Turner," Virtual Webinar, *Regeneron DNALC* |
| March 18 | *Saturday DNA!* "A Day in the Life of the Iceman," DNALC |
| | *Saturday DNA!* "Dust Away Crime," *Regeneron DNALC* |
| March 25 | *Barcode Long Island* Open Lab, DNALC |
| | *Urban Barcode Project* Open Lab, *Harlem DNA Lab* |
| March 27–30 | CyVerse/Bioinformatics Conference, Tucson, Arizona |
| March 29–31 | NIH *Citizen DNA Barcoding Network* Collaborator Training, Discover Life in America, Gatlinburg, Tennessee |
| April 1 | *Urban Barcode Research Program* Open Lab, *DNALC NYC* |
| April 10 | *School Break Bio!* "Fresh Fruit DNA Extraction," DNALC |
| | *School Break Bio!* "Forensic DNA Fingerprint," *Regeneron DNALC* |
| April 11–12 | "It's Impossible to Keep Up—Career-Spanning Learning in the Life Sciences," Seminar, HudsonAlpha Institute for Biotechnology, Huntsville, Alabama |
| April 12 | *School Break Bio!* "Dust Away Crime," DNALC |
| | *School Break Bio!* "Ötzi the Iceman: A Museum Tour and Interactive Laboratory Experience," *DNALC NYC* |
| | *School Break Bio!* "DNA Barcoding: Using DNA to Study Species Biodiversity and Ecology," *DNALC NYC* |
| April 13 | *School Break Bio!* "Forensic DNA Fingerprint," DNALC |
| April 13–18 | NIH *Citizen DNA Barcoding Network* Hub Training, Natural History Museum of Utah, Salt Lake City, Utah |
| April 17–21 | Chan Zuckerberg Initiative Open Science Workshop, Buenos Aires, Argentina |
| April 19 | Math for America Teacher Training Workshop, *DNALC NYC* |
| April 20 | "Meet a Scientist: Dr. Joel N.H. Stern," Virtual Webinar, DNALC |
| April 22 | *Saturday DNA!* "The Mystery of Anastasia," DNALC |
| April 27 | Lehman College Biology Club Agar Art, Lehman College, Bronx, New York |
| | STANYS/SCONYC Teacher Training Workshop, *DNALC NYC* |
| April 29 | *Urban Barcode Research Program* Open Lab, *DNALC NYC* |
| May 2 | "It's Impossible to Keep Up—Career-Spanning Learning in the Life Sciences," AgBioData Conference/ CyVerse, Chicago, Illinois |
| May 3 | Math for America Teacher Training Workshop, *DNALC NYC* |
| May 13–14 | InnovATE*BIO* Undergraduate Skills Workshop, Great Bay Community College, Portsmouth, New Hampshire |
| May 15–17 | "Building Communities of Practice," Global Bioinformatics Education Summit, Hinxton, United Kingdom |
| May 15 | "Meet a Scientist: Dr. Stephanie E. Woods, D.V.M., M.S. LAM, DACLAM," Virtual Webinar, *Regeneron DNALC* |
| May 17 | Math for America Teacher Training Workshop, *DNALC NYC* |
| May 20 | *Saturday DNA!* "The Magic of Microscopes," DNALC |
| | *Saturday DNA!* "Plasmid Manipulation," *Regeneron DNALC* |
| May 25 | Office of Community Schools Professional Learning Series, Fordham University, Bronx, New York |
| May 31 | *Urban Barcode Project/Urban Barcode Research Program* Student Symposium at The Theater at City Tech, hosted by *DNALC NYC* |
| June 1–2 | NIH SciEd Conference, Durham, North Carolina |
| June 5–7 | InnovATE*BIO* National Biotechnology Education Center: Methods in Personal Genetics and DNA Barcoding, Sequencing, PCR, and *Taq* Production Teacher Workshop, Maricopa Community College, Phoenix, Arizona |
| June 5–9 | Nanopore Sequencing Foundations for Course-Based Research; Teacher Workshop, *DNALC NYC* |

| | |
|---|---|
| June 6 | *Barcode Long Island* Student Symposium, CSHL |
| June 7–9 | InnovATE*BIO* National Biotechnology Education Center: Methods in Personal Genetics and DNA Barcoding, Sequencing, PCR, and *Taq* Production Teacher Workshop, Santiago Canyon College, Orange, California |
| June 10 | *Saturday DNA!* "Got Lactase?" *Regeneron DNALC* |
| June 12–16 | *BioCoding* Workshop, Lycée Français de New York, New York |
| | *Green Genes* Workshop, Lycée Français de New York, New York |
| June 20–21 | Site visit & tour by Dr. Adrian Tyndall, Dr. Sandra Harris-Hooker, Dr. Rita Finley, Dr. Lisa Herring, Dr. Matt Smith, Dr. Selena Florence, and Dr. Dwionne Freeman, Morehouse School of Medicine and Atlanta Public Schools Delegation, DNALC and *DNALC NYC* |
| June 26–30 | *Fun with DNA* Workshop, DNALC |
| | *World of Enzymes* Workshop, DNALC |
| | *Green Genes* Workshop, DNALC |
| | *DNA Science* Workshop, DNALC |
| | *Green Genes* Workshop, *DNALC NYC* |
| | *DNA Science* Workshop, *DNALC NYC* |
| | *Fun with DNA* Workshop, *Regeneron DNALC* |
| | *DNA Science* Workshop, *Regeneron DNALC* |
| June 26–30 | *"Making Life Science CURES Inclusive and Accessible,"* Gordon Research Conference: Undergraduate Biology Education Research, Bates College, Lewiston, Maine |
| July 3 | *DNALC NYC* Teacher Training Workshop, "PCR and Human DNA Variation: Part 1: Human Mitochondrial Sequencing," *DNALC NYC* |
| July 3–7 | *World of Enzymes* Workshop, DNALC |
| | *Green Genes* Workshop, DNALC |
| | *Forensic Detectives* Workshop, DNALC |
| | *Fun with DNA* Workshop, *DNALC NYC* |
| | *Forensic Detectives* Workshop, *DNALC NYC* |
| | *World of Enzymes* Workshop, *Regeneron DNALC* |
| July 5 | *DNALC NYC* Teacher Training Workshop, "DNA Structure, Isolation, and Mutation," *DNALC NYC* |
| July 6 | *DNALC NYC* Teacher Training Workshop, "DNA Restriction Analysis and Bacterial Transformation," *DNALC NYC* |
| July 7 | *DNALC NYC* Teacher Training Workshop, "PCR and Human DNA Variation: Part 2: Detecting a Jumping Gene (*Alu*)," *DNALC NYC* |
| July 10–14 | *Fun with DNA* Workshop, DNALC |
| | *Forensic Detectives* Workshop, DNALC |
| | *DNA Science* Workshop, DNALC |
| | *Genome Science* Workshop, DNALC |
| | *Forensic Detectives* Workshop, *DNALC NYC* |
| | *DNA Science* Workshop, *DNALC NYC* |
| | *Green Genes* Workshop, *Regeneron DNALC* |
| | *DNA Science* Workshop, *Regeneron DNALC* |
| | *Urban Barcode Research Program Conservation Genetics* Workshop, *DNALC NYC* |
| July 11 | NIH *Citizen DNA Barcoding Network* DNA Collection Event, Sweet Briar Nature Center, Smithtown, New York |
| July 17–21 | *Fun with DNA* Workshop, DNALC |
| | *World of Enzymes* Workshop, DNALC |
| | *Forensic Detectives* Workshop, DNALC |
| | *DNA Science* Workshop, DNALC |
| | *BioCoding* Workshop, DNALC |
| | *Fun with DNA* Workshop, *DNALC NYC* |
| | *World of Enzymes* Workshop, *DNALC NYC* |
| | *Green Genes* Workshop, *DNALC NYC* |
| | *Fun with DNA* Workshop, *Regeneron DNALC* |
| | *DNA Barcoding* Workshop, *Regeneron DNALC* |
| | *Urban Barcode Research Program DNA Barcoding* Workshop, *DNALC NYC* |
| | *Urban Barcode Research Program Conservation Genetics* Workshop, *Harlem DNA Lab* |
| | *DNA Barcoding* Workshop, *DNALC Asia* |
| July 23–27 | Intelligent Systems for Molecular Biology and the European Conference on Computational Biology, Centre de Congrès de Lyon, Lyon, France |

| | |
|---|---|
| July 24–28 | *World of Enzymes* Workshop, DNALC |
| | *Green Genes* Workshop, DNALC |
| | *DNA Science* Workshop, DNALC |
| | *DNA Barcoding* Workshop, DNALC |
| | *Forensic Detectives* Workshop, *DNALC NYC* |
| | *DNA Science* Workshop, *DNALC NYC* |
| | *DNA Science* Workshop, *Regeneron DNALC* |
| | *World of Enzymes* Workshop, *Regeneron DNALC* |
| | *Urban Barcode Research Program Conservation Genetics* Workshop, *DNALC NYC* |
| | *Urban Barcode Research Program DNA Barcoding* Workshop, *Harlem DNA Lab* |
| | *DNA Barcoding* Workshop, *DNALC Asia* |
| | *Genome Science* Workshop, *DNALC Asia* |
| July 31–August 4 | *Fun with DNA* Workshop, DNALC |
| | *Green Genes* Workshop, DNALC |
| | *Genome Science* Workshop, DNALC |
| | NIH *Citizen DNA Barcode Network* Educator Workshop, DNALC |
| | *Fun with DNA* Workshop, *DNALC NYC* |
| | *Green Genes* Workshop, *DNALC NYC* |
| | *BioCoding* Workshop, *DNALC NYC* |
| | *STEM Matters Forensic Detectives* Workshop, *DNALC NYC* |
| | *Forensic Detectives* Workshop, *Regeneron DNALC* |
| | *Genome Science* Workshop, *Regeneron DNALC* |
| | *Urban Barcode Research Program DNA Barcoding* Workshop, *DNALC NYC* |
| | *Fun with DNA* Workshop, Long Beach High School, Long Beach, New York |
| August 3 | Site visit and tour by Dr. Garrett Dunlap, British Consulate General, *DNALC NYC* |
| August 7–11 | *Forensic Detectives* Workshop, DNALC |
| | *DNA Science* Workshop, DNALC |
| | *STARS DNA Barcoding* Workshop, DNALC |
| | *Forensic Detectives* Workshop, *DNALC NYC* |
| | *DNA Science* Workshop, *DNALC NYC* |
| | *DNA Barcoding* Workshop, *DNALC NYC* |
| | *Urban Barcode Research Program Conservation Genetics* Workshop, *DNALC NYC* |
| | *Fun with DNA* Workshop, *Regeneron DNALC* |
| | *DNA Science* Workshop, *Regeneron DNALC* |
| | *World of Enzymes* Workshop, Long Beach High School, Long Beach, New York |
| August 14 | NIH *Citizen DNA Barcoding Network* DNA Collection Event, Jones Beach Energy & Nature Center, Wantagh, New York |
| August 14–18 | *Fun with DNA* Workshop, DNALC |
| | *World of Enzymes* Workshop, DNALC |
| | *Green Genes* Workshop, DNALC |
| | *DNA Science* Workshop, DNALC |
| | *STARS BioCoding* Workshop, DNALC |
| | *Fun with DNA* Workshop, *DNALC NYC* |
| | *World of Enzymes* Workshop, *DNALC NYC* |
| | *Genome Science* Workshop, *DNALC NYC* |
| | *Urban Barcode Research Program DNA Barcoding* Workshop, *DNALC NYC* |
| | *World of Enzymes* Workshop, *Regeneron DNALC* |
| August 21–25 | *WiSE Fun with DNA* Workshop, DNALC |
| | *World of Enzymes* Workshop, DNALC |
| | *Forensic Detectives* Workshop, DNALC |
| | *Sequence a Genome!* Workshop, DNALC |
| | *Fun with DNA* Workshop, *DNALC NYC* |
| | *Green Genes* Workshop, *DNALC NYC* |
| | *DNA Science* Workshop, *DNALC NYC* |
| | *Urban Barcode Program* Teacher Training, *DNALC NYC* |
| | *DNA Barcoding* Workshop, *Regeneron DNALC* |
| | *Green Genes* Workshop, *Regeneron DNALC* |
| August 28–September 1 | *Fun with DNA* Workshop, DNALC |
| | *Green Genes* Workshop, DNALC |
| | *Forensic Detectives* Workshop, DNALC |
| | *DNA Science* Workshop, DNALC |

| | |
|---|---|
| | *World of Enzymes* Workshop, *DNALC NYC* |
| | *Forensic Detectives* Workshop, *DNALC NYC* |
| | *Fun with DNA* Workshop, *Regeneron DNALC* |
| | *Forensic Detectives* Workshop, *Regeneron DNALC* |
| September 11–12 | Hudson Valley Community College Biotech Meeting, Troy, New York |
| September 20 | Site visit & tour by Professor Yaping Zhang, Mr. Zhenyu Wang, and Ms. Ting Tong, Delegation from the Chinese Academy of Science with Bruce Stillman, DNALC |
| September 23 | NIH *Citizen DNA Barcoding Network* DNA Collection and Wet-Lab Event, South Fork Natural History Museum, Bridgehampton, New York |
| | Science Saturday 2023 Festival, Rockefeller University, New York, New York |
| September 30 | SUBMERGE Marine Science Festival, Hudson River Park, New York, New York |
| | Mercy University STEM Education Conference, Dobbs Ferry, New York |
| October 14 | *Saturday DNA!* "Jack-O'-Lantern DNA!" *Regeneron DNALC* |
| October 17 | Site visit by Carissa Jordan, CSHL Association Board Director and friend Rachel Sitman, with Karen Orzel, *DNALC NYC* |
| October 20 | Regeneron Day for Doing Good, Regeneron Pharmaceuticals, Sleepy Hollow, New York |
| October 21 | *Saturday DNA!* "Beholding Blood," DNALC |
| October 24 | Regeneron Day for Doing Good, Regeneron Pharmaceuticals, Sleepy Hollow, New York |
| November 3 | Site visit and tour by John Buck, Jessica Raba, and Erik Weinstein, Delegation from Long Island Lutheran Middle and High School with Charlie Prizzi, DNALC |
| November 7 | *Barcode Long Island* Teacher Training Workshop, "DNA Barcoding," DNALC |
| | *Regeneron DNALC* Teacher Training Workshop, "Human Mitochondrial Sequencing," *Regeneron DNALC* |
| November 10 | CSHL Association "A Day in the Lab," DNALC |
| November 11 | *Saturday DNA!* "Protein Purification," DNALC |
| | *Saturday DNA!* "Forensic DNA Fingerprint," *Regeneron DNALC* |
| December 4 | "It's Impossible to Keep Up—Career-Spanning Learning in the Life Sciences," Genetics and Genomics Academy, North Carolina State University, Raleigh, North Carolina |
| December 6 | Site visit by Delegation from Meharry Medical College, *Regeneron DNALC* |
| December 9 | *Saturday DNA!* "Got Lactase?" *Regeneron DNALC* |
| December 16 | *Saturday DNA!* "Christmas Cactus Cloning," DNALC |


COLD SPRING HARBOR
LABORATORY PRESS

# PRESS PUBLICATIONS

## Serials

*Genes & Development,* Vol. 37 (www.genesdev.org)

*Genome Research,* Vol. 33 (www.genome.org)

*Learning & Memory,* Vol. 30 (www.learnmem.org)

*RNA,* Vol. 29 (www.rnajournal.org)

*Cold Spring Harbor Molecular Case Studies,* Vol. 9 (www.molecularcasestudies.org)

*Life Science Alliance,* Vol. 6 (www.life-science-alliance.org)

*Cold Spring Harbor Perspectives in Biology,* Vol. 15 (www.cshperspectives.cshlp.org)

*Cold Spring Harbor Perspectives in Medicine,* Vol. 13 (www.perspectivesinmedicine.org)

*Cold Spring Harbor Protocols* (www.cshprotocols.org)

## Monographs (Topic Collections from Perspectives in Biology and Perspectives in Medicine)

*Wound Healing: From Bench to Bedside,* edited by Xing Dai, Sabine Werner, Cheng-Ming Chuong, and Maksim Plikus

## Other

*Mouse Phenotypes: Generation and Analysis of Mutants: Second Edition: A Laboratory Manual,* by Virginia E. Papaioannou and Richard R. Behringer

*African Turquoise Killifish* (Nothobranchius furzeri)*: A Laboratory Manual,* edited by Anne Brunet

*Experiments in Bacterial Genetics: A Laboratory Manual,* edited by Lionello Bossi, Andrew Camilli, and Angelika Gründling

*Inside Science: Revolution in Biology and Its Impact,* by Benjamin Lewin

*The Medical Revolution of Messenger RNA,* by Fabrice Delaye

*Xenopus: A Laboratory Manual,* by Hazel L. Sive

CSHL Annual Report 2020, Yearbook Edition

## E-books

*Retinal Disorders: Approaches to Diagnosis and Treatment,* Second Edition, edited by Eyan Banin, Jean Bennett, Jacque L. Duncan, Botond Roska, and José-Alain Sahel

*Synthetic Biology and Greenhouse Gases,* edited by Daniel Drell, L. Val Giddings, Aristides Patrinos, Richard J. Roberts, and Charles DeLisi

*Aging: Geroscience as the New Public Health Frontier,* Second Edition, edited by James L. Kirkland, Jay Olshansky, and George M. Martin

*Breast Cancer: From Fundamental Biology to Therapeutic Strategies,* edited by Jane E. Visvader, Jeffrey M. Rosen, and Samuel Aparicio

*Mouse Phenotypes: Generation and Analysis of Mutants,* Second Edition, by Virginia E. Papaioannou and Richard R. Behringer

*Experiments in Bacterial Genetics: A Laboratory Manual,* edited by Lionello Bossi, Andrew Camilli, and Angelika Gründling

*African Turquoise Killifish* (Nothobranchius furzeri)*: A Laboratory Manual,* by Anne Brunet

*The Medical Revolution of Messenger RNA,* by Fabrice Delaye

*Inside Science: Revolution in Biology and Its Impact,* by Benjamin Lewin

*Wound Healing: From Bench to Bedside,* edited by Xing Dai, Sabine Werner, Cheng-Ming Chuong, and Maksim Plikus

*Xenopus: A Laboratory Manual,* by Hazel L. Sive

## Websites

Cold Spring Harbor Monograph Archive Online (www.cshmonographs.org)

Cold Spring Harbor Symposia on Quantitative Biology Archive (symposium.cshlp.org)

# COLD SPRING HARBOR LABORATORY PRESS EXECUTIVE DIRECTOR'S REPORT 2023

Cold Spring Harbor Laboratory Press creates resources that help scientists advance their research and careers. They include books, laboratory manuals, review journals, and peer-reviewed research journals.

The Press publishes nine journals and offers more than 250 books in print and electronic form. In 2023, its staff published more than 100 journal issues, seven new print books, and 11 new e-books—outputs that collectively enabled the Press to make another substantial financial contribution to the Laboratory that exceeded projection.

## Journals

The research journals *Genes & Development* and *Genome Research* continued to publish the high-quality work the community has come to expect from these well-regarded, broad-scope, selective journals, but since submission rates declined, they published fewer papers than in previous years. Both remained high in the competitive rankings of journals in their burgeoning fields. *RNA* and *Learning & Memory* continued to serve more specialized research communities.

Subscription renewals to all the journals were close to or ahead of projection in 2023. Like many established research journals, our four titles are currently "hybrid," deriving revenue in part from subscriptions and in part from fees paid by authors to make their papers immediately free to read ("open access"). CSHL Press has been working toward phasing out subscriptions to make its research journals entirely open-access, aligning them with the requirements of the Howard Hughes Medical Institute (HHMI), the Wellcome Trust, and other science funders. The National Institutes of Health (NIH) and other U.S. government agencies may have similar requirements by 2026.

Many STEM publishers are transitioning to open access through so-called "Read and Publish (R&P) Agreements" with individual institutions and consortia. These vary in nature, depending on the institution and the publisher. CSHL Press is leveraging the value of its three highly regarded review journals (*Perspectives in Biology*, *Perspectives in Medicine*, and *Protocols*) by offering institutions R&P agreements that provide access to all seven research and review journals and the opportunity for their faculty to publish open access in the research journals without additional charges.

In 2023, the negotiation of R&P agreements was the Press marketing team's highest priority and negotiations were initiated with a large number of institutions. However, there is much turmoil among institutions about the open-access transition and our process uncovered a preference within some institutions for alternatives to R&P agreements with less administrative complexity. One currently popular alternative is known as "subscribe to open" (S2O). In this approach, institutions agree to continue their current level of payment for a journal, using their established subscription payment management systems, and the publisher agrees to make the journal content open as long as sufficient financial support for the journal continues. The chief risk is that institutions eventually stop providing support because the content is already open and paid for by others. Should our R&P agreement fail to become widely accepted and S2O proves too risky, the CSHL journals will continue to be hybrid and funders like HHMI will be satisfied by our allowing their grantees to post an accepted manuscript in an open-access repository such as PubMed Central.

What the publishing ecosystem will look like beyond 2025 is currently unpredictable. But it is clear that our journals will have to compete vigorously for authors, against commercial companies with vast resources who are developing silos of outlets to attract and retain manuscript submissions. Independent, not-for-profit journals like ours continue to strive editorially for high-quality papers, but they have diminished submission rates, and because fewer papers are meeting required standards, publication volume has decreased. In response, new promotional programs intended to encourage submissions and promote the special character of the CSHL journals are being launched.

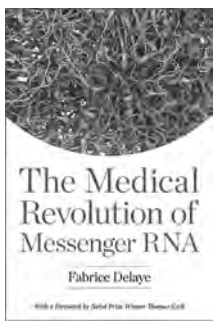
When the editor of *Genes & Development*, Terri Grodzicker, stepped down in December 2022 after 35 years at her post, a new editorial structure was established involving 12 distinguished scientists as Consulting Editors to assist the Executive Editor with decision-making on submitted papers. The six-month pilot period was successful and was continued for another 12 months.

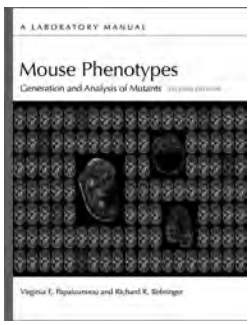
The founding editor of *Learning & Memory*, Jack Byrne, also announced his intention to step down after 26 years in the position, and conversations began about the future scope and direction of the journal.

The Laboratory's two most recently launched journals were born open-access. *Life Science Alliance* (LSA) is owned and published jointly by Cold Spring Harbor Laboratory, the European Molecular Biology Organization (EMBO), and Rockefeller University and offers an efficient route to publication with reduced delay for papers declined by the partners' nine frontline journals. Five years on, the journal is already profitable and the first distribution of financial surplus, repaying each partner's initial investment, has been made. *Molecular Case Studies* (MCS), however, has not prospered after a promising beginning, failing to publish enough papers to cover its annual costs. The decline in submissions was likely due to the impact of the pandemic on clinical science as well as invasion of the journal's editorial niche from commercial journals. The journal was therefore discontinued in December 2023 but *Genes & Development* expanded in scope to embrace content typical of MCS and co-opted the distinguished Editor-in-Chief of MCS, Elaine Mardis, to assist.

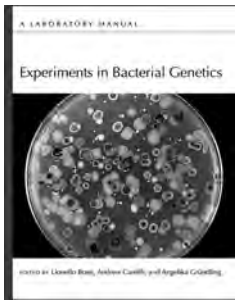
## Books

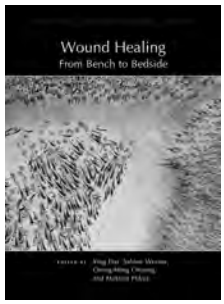
Two books were published in September for which extensive promotional campaigns were orchestrated for the first time by an experienced freelance publicist. *Inside Science: Revolution in Biology and Its Impact,* is by Benjamin Lewin, founder of Cell Press, a longtime observer of the scientific enterprise. He argues that research science has changed with the rise of team-based, multidisciplinary investigations and that generative artificial intelligence technologies will accelerate change still more, transforming science as we know it. In *The Medical Revolution of Messenger RNA*, Fabrice Delaye tells the decades-long story of how RNA-based technologies were eventually harnessed to create new therapeutics such as COVID-19 vaccines. The book extensively features the scientists who were awarded the 2023 Nobel Prize in Physiology or Medicine and was published just a week before the prize was announced.

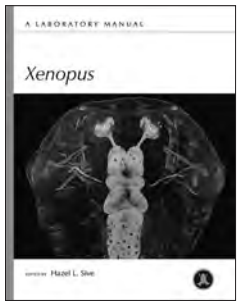




Book sales for the year matched expectation, a consequence of effective direct-to-customer marketing and direct sales through the Press website, which accounted for >25% of book revenue. Our book distributor, Oxford University Press, decided with little notice that it would no longer fulfill


orders from individuals, so our e-commerce operation had to be remade from scratch. Our e-book manufacturer had to be replaced, also at short notice—an important matter, because one-third of book sales include an electronic edition bundled with a print edition, a combination not available elsewhere, including commercial online retailers like Amazon.

Book printing continued to be affected by supply chain delays that made certain types of paper scarce and more expensive and by the closure of plants and a shortage of trained staff in those that remain.

## Staff

During the year, we welcomed four new colleagues. Christin Munkittrick was appointed to the new position of Assistant Editor of *Cold Spring Harbor Protocols*. Danett Gil became an Editorial Assistant, Jacqueline Picone a Marketing Associate, and Jennifer Quereau the Executive Assistant in the Office of the Publisher. All these recruits brought valuable skills and experience and made our organization stronger.

CSHL Press provides the research community with information in many forms that is authoritative, accessible, and reasonably priced. In 2023, the staff maintained the high quality of our books and journals and delivered them with efficiency and financial success. I thank all the staff for their dedication to our enterprise and in particular those who provide outstanding leadership in our diverse activities: Assistant Director Richard Sever, journal editors Laureen Connell, Hillary Sussman, Alejandro Montenegro Montero, and Eric Sawey and department heads K.J. Black, Wayne Manos, Stephen Nussbaum, Ted Roeder, Marcie Siconolfi, and Denise Weiss. And it was a personal pleasure to have Jenn Quereau become a welcoming and efficient presence at the heart of the organization.

**John R. Inglis, Ph.D.**
*Executive Director and Publisher*

# PREPRINT SERVERS

# PREPRINT SERVERS

A preprint is a research manuscript made freely available by its authors without peer review, on a website known as a preprint server. The Laboratory's preprint servers, bioRxiv and medRxiv, were founded in 2013 and 2019, respectively. In 2023, they continued to grow, delivering around 4500 new manuscripts each month. These preprints are free to post and free to read, creating awareness of the latest science being done by hundreds of thousands of authors, in tens of thousands of institutions, in 190 countries. bioRxiv and medRxiv are widely recognized as the most influential preprint servers in biomedicine.

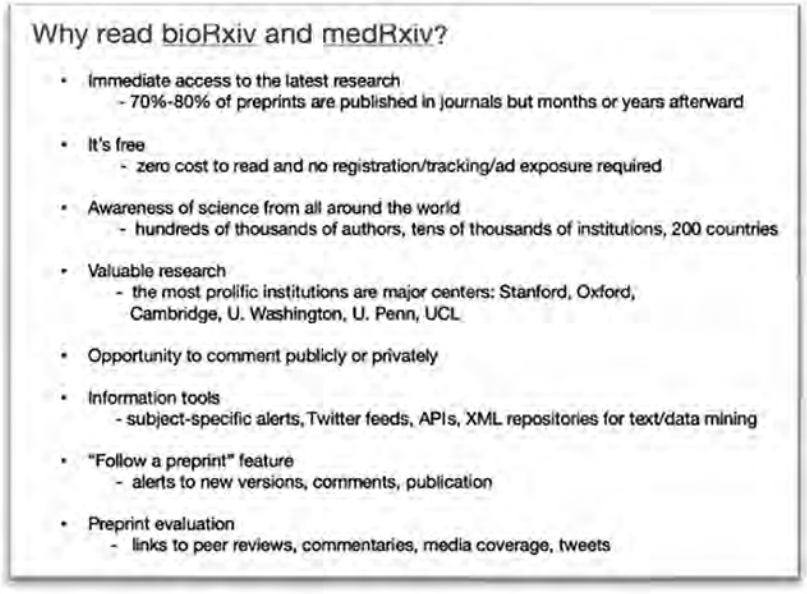
By the end of 2023, bioRxiv was hosting more than 219,300 preprints, having added more than 39,000 new manuscripts during the year. Twenty-six percent of those preprints had been revised more than once. The largest of the 26 subject categories are neuroscience, microbiology, and bioinformatics. bioRxiv's readership exceeded six million page views each month.

During 2023, medRxiv preprints grew to a total of 49,300, adding 11,000 new manuscripts during the year. Reflecting the vital role medRxiv played in distributing new information during the peak pandemic years, the largest subject categories are infectious disease, epidemiology, and public health. The proportion of pandemic-related medRxiv preprints declined during the year, with growth coming for the first time from fields like cardiology and neurology. Usage of the server remained high, at two million page views each month.

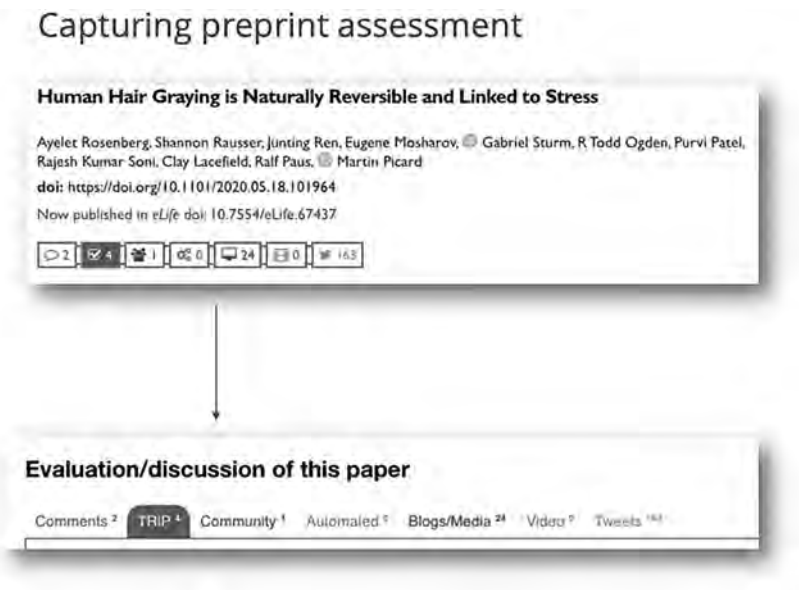
The most prolific sources of preprints on both servers are research-intensive institutions in the United States and United Kingdom such as Stanford University, University of Oxford, University of Cambridge, University of Washington, University of Pennsylvania, University of Michigan, University College London, Imperial College London, Columbia University, Johns Hopkins University, and Icahn School of Medicine at Mount Sinai. Scientists at these prominent institutions often post first-class science months or even years before it appears in peer-reviewed journals. As a consequence, papers on the servers are frequently rapidly cited: By the end of 2023 there were more than 480,000 bioRxiv citations and 230,000 medRxiv citations.

The servers are independent of journals but closely integrated with them. More than 50 journals permit authors to post a submitted manuscript simultaneously on both servers, and preprint authors have the opportunity to submit their paper directly to more than 300 journals from multiple publishers. Seventy to eighty percent of manuscripts posted to the servers are published in a journal within two years, and more than 4,500 journals have published papers that were first posted to the servers.

After initial periods in which growth was organic, an active campaign to promote use of the servers began in 2023 that argued for posting and reading preprints. The campaign centered on two questions:

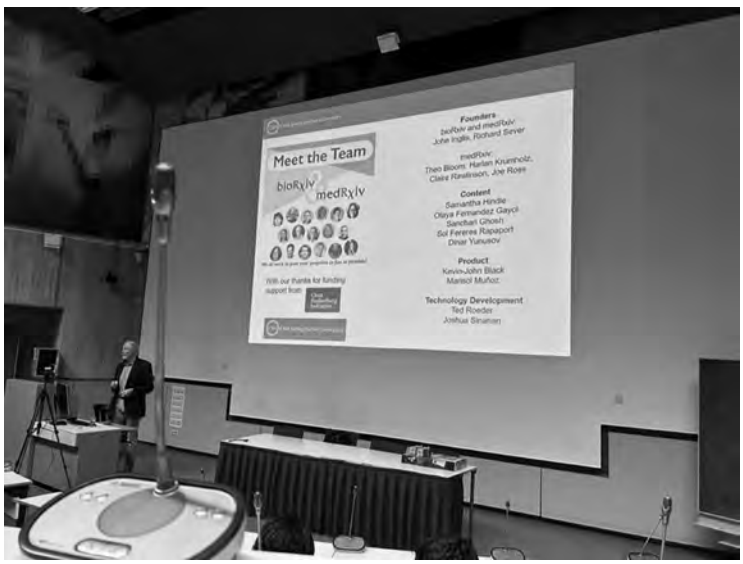

The growing prominence of the Laboratory's servers has prompted interest in the provision of reviews and commentary on preprints that help readers assess the work reported. A tabbed dashboard reveals the varied forms of manuscript assessment available around an individual preprint. The TRIP (Transparent Reviews in Preprint) tab shows formal peer reviews.




By the end of 2023, there were 7,000 such reviews on the servers, commissioned by 10 different organizations. They are part of a movement to decouple the assessment of science from its publication in journals that too often provide a false proxy of importance.

The Chan Zuckerberg Initiative (CZI) provided five years of generous funding for bioRxiv in June 2017 and two years of support for the newly launched medRxiv in June 2020. In November 2022, CZI committed a further $4.9 million for two years. Three years of further funding of $2.0 million/year will be available from CZI—but only if additional funding is obtained to support a sustainability plan for these unique and invaluable resources. We have commissioned a consulting organization to help develop a strategy for sustainability.

The servers' co-founders continue to be invited to give national and international talks, panels, webinars, podcasts, and interviews about the servers' progress and the implications of preprint adoption in biology and medicine.




John Inglis speaking at VIB KU Center for Cancer Biology, Leuven, Belgium, October 26, 2023.


(*Top row*) R. Sever, D. Yunusov, J. McFadden, S. Ghosh, O. Fernandez Gayol, J. Inglis; (*bottom row*) S. Hindle, S. Fereres Rapoport, K-J. Black

bioRxiv's 10th anniversary was celebrated in November in several ways. Awareness campaigns were mounted on the social media platforms LinkedIn, Facebook, Bluesky, Mastodon, and X (formerly Twitter). A questionnaire was sent to more than 100,000 scientists throughout the world, whose answers told us much about the attitudes, behaviors, and wishes of the servers' users. The familiar logo was replaced until the turn of the year with a birthday variation, which online had a subtly flickering candle flame.

bioRxiv could not have come this far without its association with—and support from—the Laboratory, with its a long history as a place for sharing science. The project began with no budget or new hires, just the enthusiasm of Press staff excited about creating something potentially transformational and the willingness of the Laboratory's leadership to take a risk. Ten years on, the bioRxiv project has expanded to embrace clinical medicine, has acquired millions of ardent users, and has attracted millions of dollars in grants to support an extraordinary team of staff dedicated to content management and product and technology development. There is a limitless pool of goodwill for the servers that extends to the Laboratory that provides them. bioRxiv and medRxiv are making science faster, more equitable, more open, and more interactive, and their strongly growing brands are augmenting the Laboratory's reputation for innovation among scientists worldwide.

**John R. Inglis, Ph.D.**
*Co-Founder and Principal Investigator*
*bioRxiv and medRxiv*

# FINANCE

# FINANCIAL STATEMENTS

## Consolidated Balance Sheets
December 31, 2023

| | 2023 | 2022 |
|---|---|---|
| Assets: | | |
| Cash and cash equivalents | | |
| Board designated for capital expansion | $ 19,547,834 | $ 33,415,601 |
| Available for working capital | 17,364,567 | 43,752,374 |
| Grants receivable | 9,278,903 | 9,660,138 |
| Contributions receivable, net | 139,232,929 | 107,020,461 |
| Investments in certificates of deposit | | |
| Board designated for capital expansion | 58,330,665 | 39,302,433 |
| Available for working capital | 32,232,315 | – |
| Investments | 766,313,926 | 683,077,580 |
| Investment in employee residences | 6,809,903 | 6,588,394 |
| Restricted use assets | 545,643 | 3,015,801 |
| Other assets | 9,425,337 | 7,145,053 |
| Land, buildings, and equipment, net | 318,368,175 | 307,777,062 |
| Total assets | $ 1,377,450,197 | $ 1,240,754,897 |
| Liabilities and net assets: | | |
| Liabilities: | | |
| Accounts payable and accrued expenses | $ 11,285,519 | $ 11,471,782 |
| Deferred revenue | 22,376,812 | 34,208,363 |
| Fair value of interest rate swap | 12,189,072 | 12,586,718 |
| Bonds payable | 96,139,041 | 96,072,771 |
| Total liabilities | 141,990,444 | 154,339,634 |
| Commitments and contingencies | | |
| Net assets: | | |
| Without donor restrictions | | |
| Undesignated | 327,644,303 | 286,853,843 |
| Board-designated | 463,799,743 | 410,037,211 |
| With donor restrictions | 444,015,707 | 389,524,209 |
| Total net assets | 1,235,459,753 | 1,086,415,263 |
| Total liabilities and net assets | $ 1,377,450,197 | $ 1,240,754,897 |

# Consolidated Statement of Activities
## Year ended December 31, 2023
(with summarized financial information for the year ended December 31, 2022)

| | Without Donor Restrictions | With Donor Restrictions | 2023 Total | 2022 Total |
|---|---|---|---|---|
| Revenue and other support: | | | | |
| Public support—contributions and nonfederal grant awards | $ 45,688,095 | $ 68,766,741 | $ 114,454,836 | $ 82,017,715 |
| Federal grant awards | 35,774,511 | – | 35,774,511 | 39,188,764 |
| Indirect cost allowances | 33,311,621 | – | 33,311,621 | 30,876,743 |
| Investment return utilized | 39,546,774 | – | 39,546,774 | 34,493,868 |
| Royalty and license revenue | 18,445,260 | – | 18,445,260 | 10,165,173 |
| Program fees | 7,567,933 | – | 7,567,933 | 8,583,126 |
| Publications sales | 9,442,510 | – | 9,442,510 | 9,389,318 |
| Dining services | 5,335,060 | – | 5,335,060 | 4,895,902 |
| Housing services | 5,173,763 | – | 5,173,763 | 4,240,693 |
| Miscellaneous | 1,065,921 | – | 1,065,921 | 1,016,985 |
| Net assets released from restrictions | 43,040,442 | (43,040,442) | – | – |
| Total revenue and other support | 244,391,890 | 25,726,299 | 270,118,189 | 224,868,287 |
| Expenses: | | | | |
| Research | 114,218,434 | – | 114,218,434 | 111,828,147 |
| Educational programs | 20,906,758 | – | 20,906,758 | 18,789,514 |
| Publications | 10,571,538 | – | 10,571,538 | 10,202,241 |
| Banbury Center conferences | 2,513,128 | – | 2,513,128 | 2,323,893 |
| DNA Learning Center programs | 6,319,508 | – | 6,319,508 | 5,819,821 |
| School of Biological Sciences programs | 3,468,247 | – | 3,468,247 | 3,274,739 |
| General and administrative | 31,779,496 | – | 31,779,496 | 29,975,299 |
| Total expenses | 189,777,109 | – | 189,777,109 | 182,213,654 |
| Excess of revenue and other support over expenses | 54,614,781 | 25,726,299 | 80,341,080 | 42,654,633 |
| Other changes in net assets: | | | | |
| Investment income (loss) excluding amount utilized | 39,540,565 | 28,765,199 | 68,305,764 | (112,327,765) |
| Change in fair value of interest rate swap | 397,646 | – | 397,646 | 21,327,803 |
| Increase (decrease) in net assets | 94,552,992 | 54,491,498 | 149,044,490 | (48,345,329) |
| Net assets at beginning of year | 696,891,054 | 389,524,209 | 1,086,415,263 | 1,134,760,592 |
| Net assets at end of year | $ 791,444,046 | $ 444,015,707 | $ 1,235,459,753 | $ 1,086,415,263 |

# Consolidated Statements of Cash Flows
## Year ended December 31, 2023

| | 2023 | 2022 |
|---|---|---|
| Cash flows from operating activities: | | |
| Increase (decrease) in net assets | $ 149,044,490 | $ (48,345,329) |
| Adjustments to reconcile increase (decrease) in net assets to net cash provided by operating activities: | | |
| Change in fair value of interest rate swap | (397,646) | (21,327,803) |
| Depreciation and amortization | 16,042,189 | 16,456,459 |
| Amortization of deferred bond costs | 66,270 | 66,269 |
| Net (appreciation) depreciation in fair value of investments | (93,096,715) | 77,377,459 |
| Contributions restricted for long-term investment or capital | (850,661) | (539,385) |
| Changes in assets and liabilities: | | |
| Grants receivable | 381,235 | 1,363,832 |
| Contributions receivable, net | (29,126,241) | (751,688) |
| Restricted use assets | 2,470,158 | 375,390 |
| Other assets | (2,280,284) | 440,262 |
| Accounts payable and accrued expenses | 1,892,217 | (3,052,624) |
| Deferred revenue | (11,831,551) | (8,609,235) |
| Net cash provided by operating activities | 32,313,461 | 13,453,607 |
| Cash flows from investing activities: | | |
| Capital expenditures | (28,711,782) | (34,995,691) |
| Proceeds from sales and maturities of investments | 185,800,136 | 214,969,326 |
| Purchases of investments | (175,939,767) | (190,267,845) |
| Purchases of investments in certificate of deposits | (51,260,547) | (39,302,433) |
| Net change in investment in employee residences | (221,509) | (222,073) |
| Net cash used in investing activities | (70,333,469) | (49,818,716) |
| Cash flows from financing activities: | | |
| Contributions restricted for long-term investment | 260,500 | 35,168 |
| Contributions restricted for investment in capital | 590,161 | 504,217 |
| (Increase) decrease in contributions receivable | (3,086,227) | 2,954,662 |
| Net cash (used in) provided by financing activities | (2,235,566) | 3,494,047 |
| Net decrease in cash and cash equivalents | (40,255,574) | (32,871,062) |
| Cash and cash equivalents at beginning of year | 77,167,975 | 110,039,037 |
| Cash and cash equivalents at end of year | $ 36,912,401 | $ 77,167,975 |
| Supplemental disclosure: | | |
| Interest paid | $ 3,581,394 | $ 3,675,221 |
| Purchases of capital expenditures in accounts payable | $ 2,322,335 | $ 4,400,815 |
| Operating cash flows from operating leases | $ 337,095 | $ 327,294 |

# FINANCIAL SUPPORT OF THE LABORATORY

Cold Spring Harbor Laboratory, the DNA Learning Center, and the Banbury Center receive a substantial portion of funding through grants from the federal government and through grants, capital gifts, and annual contributions from New York state, private foundations, corporations, and individuals. The following section summarizes funding that occurred during 2023.

## GRANTS January 1–December 31, 2023

### COLD SPRING HARBOR LABORATORY GRANTS

| *Grantor* | *Program/Principal Investigator* | *Project Start Date* | *Project End Date* | *2023 Funding*[1] | |
|---|---|---|---|---|---|
| **FEDERAL GRANTS** | | | | | |
| NATIONAL INSTITUTES OF HEALTH | | | | | |
| *Program Project and Center Support* | Drs. Vakoc/Egeblad/Dobin/dos Santos/ Krainer/Lyons/McCombie/Spector | 02/01/23 | 01/31/28 | 2,880,000 | * |
| | Dr. Tuveson - Cancer Center Core | 08/01/21 | 07/31/26 | 4,495,455 | |
| *Cooperative Research Agreement Support*[2] | Dr. Vakoc | 08/01/19 | 07/31/24 | 430,416 | |
| *Equipment* | Dr. Spector | 07/15/23 | 07/14/24 | 600,000 | * |
| *Research Support* | | | | | |
| | Dr. Albeanu | 09/22/22 | 07/31/27 | 551,762 | |
| | Drs. Albeanu/Koulakov | 06/15/19 | 03/31/25 | 904,662 | |
| | Dr. Amor Vegas | 09/15/22 | 08/31/27 | 480,000 | |
| | Dr. Amor Vegas | 09/15/23 | 05/31/28 | 533,952 | * |
| | Drs. Bandyopadhyay/Zador | 05/03/23 | 04/30/26 | 642,129 | * |
| | Dr. Cheadle | 09/01/22 | 08/31/25 | 531,037 | |
| | Dr. Cheadle | 03/15/23 | 02/29/28 | 636,837 | |
| | Dr. dos Santos | 03/01/20 | 02/28/25 | 430,416 | |
| | Drs. dos Santos/Koo | 09/01/23 | 08/31/28 | 509,988 | * |
| | Drs. dos Santos/Siepel | 09/30/20 | 05/31/25 | 357,683 | |
| | Dr. Egeblad | 03/01/20 | 02/28/25 | 518,311 | |
| | Dr. Furukawa | 03/01/22 | 02/28/27 | 840,682 | |
| | Dr. Furukawa | 04/15/19 | 03/31/24 | 557,574 | |
| | Dr. Goodwin | 09/11/19 | 08/31/24 | 122,709 | |
| | Dr. C. Hammell | 06/23/22 | 05/31/26 | 415,296 | |
| | Dr. Janowitz | 06/01/22 | 05/31/27 | 386,428 | |
| | Dr. Kinney | 09/01/19 | 08/31/24 | 480,000 | |
| | Drs. Kinney/McCandlish | 06/15/22 | 03/31/27 | 805,485 | |
| | Dr. Koo | 08/01/23 | 04/30/27 | 384,000 | |
| | Drs. Koo/Kinney/McCandlish | 09/07/22 | 06/30/27 | 432,000 | |
| | Drs. Koulakov/Li | 09/30/19 | 07/31/24 | 501,159 | |
| | Dr. Krainer | 04/01/23 | 02/28/27 | 806,400 | |
| | Drs. Li/Tollkuhn | 09/28/21 | 12/31/26 | 769,217 | |
| | Drs. Li/Van Aelst | 03/01/20 | 12/31/24 | 709,626 | |
| | Drs. Lukey/Moses | 03/15/23 | 01/31/27 | 393,600 | * |
| | Dr. Martienssen | 08/05/22 | 07/31/27 | 435,072 | |
| | Dr. McCandlish | 09/01/19 | 07/31/24 | 480,000 | |

[1]Awarded, including direct and indirect costs
[2]Funding amounts include only CSHL's portion of the award
*New or competing renewals or supplements awarded in 2023

| *Grantor* | *Program/Principal Investigator* | *Project Start Date* | *Project End Date* | *2023 Funding*[1] | |
|---|---|---|---|---|---|
| | Dr. Meyer | 04/01/22 | 03/31/27 | 575,055 | |
| | Dr. Park | 09/01/21 | 08/31/26 | 248,045 | |
| | Dr. Schorn | 08/17/20 | 06/30/25 | 403,200 | |
| | Dr. Shea | 06/01/19 | 03/31/24 | 480,000 | |
| | Dr. Siepel | 03/01/23 | 02/29/28 | 576,000 | |
| | Dr. D. Spector | 08/20/19 | 07/31/24 | 768,000 | |
| | Dr. Stillman | 07/25/22 | 05/31/26 | 751,488 | |
| | Drs. Tollkuhn/Li | 02/16/23 | 01/31/28 | 621,062 | |
| | Dr. Tonks | 07/01/21 | 06/30/26 | 768,908 | |
| | Dr. Tonks | 03/17/20 | 01/31/24 | 480,000 | |
| | Dr. Trotman | 09/01/22 | 08/31/27 | 625,745 | |
| | Drs. Tuveson/Dobin/Preall | 08/01/21 | 07/31/26 | 591,587 | |
| | Drs. Vakoc/Tuveson | 07/02/19 | 06/30/24 | 404,792 | |
| | Dr. Vakoc | 08/01/23 | 07/31/28 | 742,538 | * |
| | Dr. Van Aelst | 04/01/19 | 01/31/25 | 646,141 | |
| | Dr. Van Aelst | 04/01/20 | 01/31/25 | 586,028 | |
| | Dr. Zador | 08/01/22 | 07/31/23 | 985,210 | |
| | Dr. Zhang | 01/05/23 | 12/31/27 | 439,200 | * |
| *Active Research Awards Continuing without Additional Support* | | | | | |
| | Drs. Albeanu/Koulakov | 09/11/18 | 08/31/24 | | |
| | Drs. Dobin/Gingeras | 08/18/17 | 05/31/23 | | |
| | Dr. Furukawa | 08/01/19 | 03/31/24 | | |
| | Dr. Mitra | 09/15/21 | 09/13/24 | | |
| | Dr. Mitra | 09/01/23 | 09/13/24 | | |
| | Dr. Vakoc | 12/01/22 | 11/30/24 | | |
| | Dr. Zador | 07/13/17 | 05/31/24 | | |
| *Research Subcontracts* | | | | | |
| NIH/Baylor College of Medicine Consortium Agreement | Dr. Zador | 05/01/22 | 04/30/25 | 608,795 | |
| NIH/Baylor College of Medicine Consortium Agreement | Dr. Zador | 09/01/23 | 07/31/26 | 276,835 | * |
| NIH/Columbia University Consortium Agreement | Dr. Zador | 08/15/21 | 07/31/26 | 652,340 | |
| NIH/Cornell University Consortium Agreement | Dr. Siepel | 04/01/23 | 03/31/28 | 94,748 | |
| NIH/Duke University Consortium Agreement | Dr. Koo | 05/01/21 | 04/30/24 | 46,840 | |
| NIH/Memorial Sloan Kettering Cancer Center Consortium Agreement | Dr. Egeblad | 03/15/20 | 12/31/24 | 14,597 | |
| NIH/MIT Consortium Agreement | Dr. Mitra | 09/07/23 | 08/31/24 | 1,017,125 | * |
| NIH/MIT Consortium Agreement | Dr. Mitra | 09/01/21 | 06/30/24 | 292,708 | |
| NIH/New York University Consortium Agreement | Dr. Koulakov | 09/01/19 | 05/31/24 | 374,637 | |
| NIH/The Salk Institute for Biological Studies Consortium Agreement | Dr. Zador | 07/15/19 | 03/31/24 | 156,600 | |
| NIH/The Broad Institute Consortium Agreement | Dr. Dobin | 09/01/22 | 06/30/27 | 32,417 | |
| NIH/The Research Foundation for the State of New York–Stony Brook Consortium Agreement | Dr. Koulakov | 07/01/22 | 03/31/27 | 87,936 | |

[1]Awarded, including direct and indirect costs

*New or competing renewals or supplements awarded in 2022

| *Grantor* | *Program/Principal Investigator* | *Project Start Date* | *Project End Date* | *2023 Funding*[1] | |
|---|---|---|---|---|---|
| NIH/The Research Foundation for the State of New York–Stony Brook Consortium Agreement | Dr. Trotman | 02/01/20 | 01/31/25 | 220,783 | |
| NIH/Thomas Jefferson University Consortium Agreement | Dr. Preall | 09/01/23 | 08/31/25 | 88,502 | * |
| NIH/University of Minnesota Consortium Agreement | Dr. dos Santos | 12/01/19 | 11/30/24 | 44,365 | |
| NIH/University of Minnesota Consortium Agreement | Dr. Mitra | 09/15/23 | 08/31/28 | 74,668 | * |
| NIH/University of Nebraska Medical Center Consortium Agreement | Drs. Tuveson/Lyons/Yeh | 09/08/22 | 08/31/27 | 150,530 | |
| *Fellowship/Career Development Support* | P. Cunniff | 05/16/23 | 02/15/26 | 47,694 | |
| | Y. Gao | 07/01/23 | 06/30/25 | 135,845 | |
| | S. Lewis | 06/01/23 | 05/31/27 | 4,550 | |
| | N. Sivetz | 07/16/23 | 07/15/24 | 47,694 | |
| | P. Westcott | 09/01/23 | 08/31/26 | 201,042 | |
| | L. Yang | 06/01/23 | 05/31/27 | 4,550 | |
| | S. Zebell | 04/01/23 | 03/31/25 | 118,907 | |
| *Institutional Training/Education Program Support* | | | | | |
| | Dr. Gann/CSHL School of Biological Sciences | 02/01/22 | 01/31/27 | 226,317 | |
| *Course Support* | | | | | |
| | Advanced Sequencing Technologies and Applications | 09/21/21 | 06/30/26 | 80,019 | |
| | Advanced Techniques in Molecular Neuroscience | 04/15/20 | 02/28/25 | 128,639 | |
| | Cell and Developmental Biology of Xenopus | 06/01/20 | 05/31/25 | 106,456 | |
| | Computational Genomics | 09/08/20 | 06/30/25 | 72,821 | |
| | Cancer Workshops: Chromatin, Epigenetics and Gene Expression Course | 07/18/23 | 06/30/28 | 161,887 | |
| | Quantitative Imaging: From Acquisition to Analysis | 07/01/98 | 06/30/27 | 99,377 | |
| | Imaging Structure and Function in the Nervous System | 04/15/20 | 02/28/25 | 158,630 | |
| | Linking Circuits, Perception and Behavior Course | 04/01/19 | 03/31/25 | 29,000 | |
| | Cancer Workshops: Molecular Embryology of the Mouse | 07/18/23 | 06/30/28 | 162,113 | |
| | Programming for Biology | 09/01/20 | 07/31/26 | 91,092 | |
| | Single Cell Analysis | 05/24/23 | 04/30/28 | 161,579 | |
| | Statistical Methods for Functional Genomics | 09/07/21 | 06/30/25 | 92,149 | |
| *Meeting Support* | | | | | |
| | Biology of Cancer: Microenvironment and Metastasis | 08/01/23 | 07/31/24 | 15,000 | |
| | Brain Barriers | 03/25/23 | 02/29/24 | 15,000 | |
| | DNA Replication and Genome Maintenance | 09/12/23 | 08/31/24 | 16,626 | |

[1]Awarded, including direct and indirect costs

*New or competing renewals or supplements awarded in 2022

| *Grantor* | *Program/Principal Investigator* | *Project Start Date* | *Project End Date* | *2023 Funding*[1] | |
|---|---|---|---|---|---|
| | Eukaryotic mRNA Processing | 06/01/23 | 03/31/24 | 15,000 | |
| | Global Regulation of Gene Expression | 09/10/19 | 07/31/24 | 29,774 | |
| | Mechanisms of Metabolic Signaling | 05/15/23 | 04/30/24 | 20,000 | |
| | Microbial Pathogenesis and Host Response | 07/13/23 | 06/30/24 | 7,500 | |
| | Network of Biology | 02/01/21 | 01/31/26 | 32,465 | |
| | Genome Informatics | 12/01/23 | 11/30/24 | 34,980 | |
| | Neurobiology of *Drosophila* | 06/01/23 | 05/31/24 | 25,000 | |
| | Retroviruses | 04/01/21 | 03/31/25 | 35,000 | |
| | Telomeres & Telomerase | 04/01/23 | 03/31/24 | 30,985 | |
| | The Biology of Genomes | 05/01/23 | 04/30/28 | 60,185 | |
| **NATIONAL SCIENCE FOUNDATION** | | | | | |
| *Research Support* | | | | | |
| | Dr. C Hammell | 07/01/22 | 06/30/26 | 250,042 | |
| | Dr. Jackson | 08/01/22 | 07/31/25 | 224,987 | |
| | Dr. Jackson | 01/15/23 | 12/31/25 | 267,865 | |
| | Drs. Jackson/Lippman | 09/01/21 | 08/31/25 | 787,993 | |
| | Dr. Lippman | 08/15/22 | 07/31/27 | 496,965 | |
| | Dr. Navlakha | 11/01/19 | 03/31/24 | 129,777 | |
| | Dr. Ware | 10/01/21 | 09/30/25 | 12,874 | |
| *Fellowship Support* | E. Isko | 09/01/23 | 08/31/24 | 49,000 | |
| *Course Support* | *Drosophila* Neurobiology: Genes, Circuits and Behavior | 05/01/20 | 04/30/24 | 107,033 | |
| | Synthetic Biology | 08/01/22 | 07/31/25 | 172,698 | |
| | Advance Courses for Model Genetic Systems: Advance Bacterial Genetics | 06/01/23 | 05/31/26 | 92,350 | * |
| | Advance Courses for Model Genetic Systems: Frontiers & Techniques in Plant Science | 06/01/23 | 05/31/26 | 126,980 | * |
| | Advance Courses for Model Genetic Systems: Yeast Genetics and Genomics | 06/01/23 | 05/31/26 | 114,002 | * |
| **HEALTH RESOURCES AND SERVICES ADMINISTRATION** | | | | | |
| *Equipment* | | | | | |
| | Dr. Zador | 09/30/23 | 09/29/26 | 2,000,000 | |
| **UNITED STATES DEPARTMENT OF AGRICULTURE** | | | | | |
| *Research Support* | | | | | |
| | Dr. McCombie | 09/15/22 | 09/14/24 | 1,912,990 | |
| **UNITED STATES DEPARTMENT OF THE ARMY** | | | | | |
| *Research Support* | | | | | |
| | Dr. Borniger | 07/01/23 | 07/31/24 | 550,035 | |
| | Dr. Borniger | 03/01/22 | 02/29/24 | 192,572 | |
| | Dr. Lukey | 01/15/21 | 10/14/24 | 306,545 | |

[1]Awarded, including direct and indirect costs
*New or competing renewals or supplements awarded in 2022

| *Grantor* | *Program/Principal Investigator* | *Project Start Date* | *Project End Date* | *2023 Funding*[1] | |
|---|---|---|---|---|---|
| | Drs. Trotman/Borniger | 08/15/22 | 08/14/25 | 476,291 | |
| | Dr. Zhang | 07/01/22 | 06/30/25 | 288,000 | |
| | Dr. Doherty | 08/01/22 | 07/31/24 | 296,272 | |
| **UNITED STATES DEPARTMENT OF ENERGY** | | | | | |
| *Research Support* | | | | | |
| | Drs. Lyons/Tuveson | 01/01/23 | 12/31/24 | 141,531 | * |
| *Research Subcontracts* | | | | | |
| DOE/Brookhaven National Laboratory Consortium Agreement | Dr. Ware | 02/07/23 | 12/31/23 | 105,000 | * |
| **NEW YORK STATE** | | | | | |
| NYS Department of Economic Development | Drs. Vakoc/Dobin/dos Santos/Lyons | 02/01/23 | 01/31/28 | 200,000 | * |
| Empire State Development | Dr. Moses | 08/02/22 | 08/01/25 | 106,887 | |
| **NON-FEDERAL** | | | | | |
| **MISCELLANEOUS SOURCES OF FUNDING** | | | | | |
| *Equipment Support* | | | | | |
| F. M. Kirby Foundation, Inc. | Dr. Moses | 06/30/20 | 05/31/24 | 120,000 | |
| *Program Project Support* | | | | | |
| The Simons Foundation/Autism | Dr. Wigler | 01/01/21 | 12/31/24 | 1,286,395 | |
| The Simons Foundation/Cancer | Dr. Wigler | 01/01/21 | 12/31/24 | 3,315,402 | |
| *Research Support* | | | | | |
| Ajces Trust | Dr. Westcott | 01/01/23 | 02/28/24 | 200,000 | * |
| Autobahn Labs | Dr. Fearon | 03/23/23 | 03/22/24 | 550,796 | * |
| Ms. Tasia Ballas | Dr. Vakoc | 09/16/18 | 09/15/24 | 10,000 | |
| Bayer Research & Development Services, LLC | Dr. Martienssen | 10/25/22 | 10/24/24 | 300,572 | |
| Breast Cancer Alliance | Dr. dos Santos | 03/01/23 | 02/28/24 | 100,000 | * |
| Breast Cancer Awareness Day in Memory of Elizabeth McFarland | Dr. Wigler | 01/01/23 | 12/31/23 | 5,950 | |
| Brain & Behavior Research Foundation | Dr. Cheadle | 01/15/22 | 01/14/24 | 35,000 | |
| | Dr. Hou | 07/15/23 | 07/14/24 | 17,500 | |
| The Breast Cancer Research Foundation | Drs. dos Santos/Yeh | 10/01/23 | 09/30/25 | 250,000 | * |
| | Dr. Wigler | 10/01/23 | 09/30/24 | 225,000 | |
| Cancer Research UK (CRUK) | Dr. Janowitz | 06/01/22 | 05/31/27 | 246,050 | |
| Caper Labs | Dr. Beyaz | 01/01/23 | 12/31/23 | 700,000 | * |
| Cedar Hill Foundation | Dr. Fearon | 11/16/22 | 11/15/24 | 65,000 | |
| Mr. and Mrs. Lawrence Cantwell | Dr. Vakoc | 09/16/18 | 09/15/24 | 5,000 | |
| Christina Renna Foundation, Inc. | Dr. Vakoc | 09/16/18 | 09/15/24 | 30,000 | |
| *CSHL/Northwell Clinical Cancer Research Fund* | Dr. Beyaz | 05/01/23 | 04/30/27 | 87,775 | |
| | Dr. dos Santos | 02/01/22 | 07/31/24 | 271,474 | |
| | Dr. dos Santos | 02/01/22 | 07/31/24 | 8,536 | |
| | Dr. Egeblad | 02/01/23 | 06/30/23 | 56,012 | |
| | Dr. Fearon | 07/15/22 | 07/14/24 | 222,238 | |
| | Dr. Fearon | 07/15/22 | 07/14/24 | 162,425 | |
| | Dr. Fearon | 08/22/23 | 08/21/24 | 54,982 | |
| | Dr. Koo | 07/01/23 | 06/30/24 | 96,745 | |

[1]Awarded, including direct and indirect costs

*New or competing renewals or supplements awarded in 2022

| *Grantor* | *Program/Principal Investigator* | *Project Start Date* | *Project End Date* | *2023 Funding*[1] | |
|---|---|---|---|---|---|
| | Dr. Krainer | 07/01/23 | 06/30/24 | 135,123 | |
| | Dr. Krasnitz | 05/17/19 | 05/16/24 | 14,468 | |
| | Dr. Krasnitz | 07/01/23 | 06/30/27 | 62,640 | |
| | Dr. Levy | 05/17/19 | 05/16/24 | 14,895 | |
| | Dr. Levy | 07/01/23 | 06/30/27 | 23,556 | |
| | Dr. Li | 02/01/22 | 01/31/26 | 227,178 | |
| | Dr. Lukey | 10/01/22 | 09/30/25 | 8,772 | |
| | Dr. Lyons | 10/01/22 | 09/30/25 | 28,858 | |
| | Dr. Lyons | 07/01/23 | 06/30/24 | 19,012 | |
| | Dr. Meyer | 05/01/21 | 04/30/24 | 102,630 | |
| | Dr. Meyer | 05/01/21 | 04/30/24 | 24,691 | |
| | Dr. Mills | 04/15/22 | 04/15/24 | 15,469 | |
| | Dr. Mills | 02/01/23 | 01/31/24 | 832,109 | |
| | Dr. Moses | 10/01/22 | 09/30/25 | 77,804 | |
| | Dr. Moses | 10/01/22 | 09/30/25 | 24,170 | |
| | Dr. Stillman | 01/01/23 | 12/31/23 | 4,562,238 | |
| | Dr. D. Spector | 05/01/23 | 04/30/27 | 445,348 | |
| | Dr. Tuveson | 07/01/23 | 06/30/26 | 174,053 | |
| | Dr. Vakoc | 07/01/23 | 06/30/26 | 71,473 | |
| | Dr. Westcott | 02/01/23 | 01/31/26 | 92,479 | |
| | Dr. Wigler | 05/17/19 | 05/16/24 | 486,271 | |
| | Dr. Wigler | 07/01/23 | 06/30/27 | 355,091 | |
| | Dr. Yeh | 10/01/22 | 09/30/25 | 3,708 | |
| | Dr. Zhao | 08/01/23 | 07/31/24 | 24,960 | |
| The Don Monti Memorial Research Foundation | Dr. Westcott | 01/01/23 | 12/31/24 | 50,000 | |
| The Oliver S. and Jennie R. Donaldson Charitable Trust | Dr. Beyaz | 12/07/21 | 12/06/24 | 97,260 | |
| Foundation for Food & Agriculture Research | Dr. Martienssen | 08/01/22 | 07/31/25 | 338,534 | |
| The Feinstein Institutes for Medical Research | Dr. Wigler | 01/01/23 | 03/31/24 | 40,000 | * |
| Douglas and Christine Fox | Dr. Furukawa | 06/01/22 | 05/31/23 | 60,000 | |
| Friends of TJ Foundation | Dr. Vakoc | 09/16/18 | 09/15/24 | 50,000 | |
| Bernard F. and Alva B. Gimbel Foundation, Inc. | Dr. Krainer | 02/01/20 | 01/31/24 | 5,778 | |
| Glen Cove C.A.R.E.S. | Dr. dos Santos | 01/01/23 | 12/31/23 | 5,000 | * |
| Mr. Michael J. Graziano | Dr. Vakoc | 09/16/18 | 09/15/24 | 10,000 | |
| Indian Institute of Technology—Madras | Dr. Mitra | 07/01/20 | 06/30/23 | 34,154 | |
| U.S.–Israel Binational Science Foundation | Dr. Krainer | 10/01/22 | 09/30/26 | 22,712 | |
| Ester A. & Joseph Klingenstein Fund | Dr. Bandyopadhyay | 07/01/23 | 06/30/26 | 100,000 | * |
| Linsey Family Foundation | Dr. Li | 05/01/22 | 04/30/27 | 357,834 | |
| The Lustgarten Foundation | Drs. Fearon/Lyons/Yeh | 07/01/21 | 06/30/24 | 200,000 | |
| | Dr. Tuveson | 09/01/23 | 08/31/26 | 1,154,131 | * |
| | Dr. Tuveson | 01/01/23 | 12/31/27 | 1,000,000 | |
| | Dr. Tuveson | 07/01/23 | 06/30/24 | 70,000 | |
| Maddie's Promise | Dr. Vakoc | 09/16/18 | 09/15/24 | 50,000 | |
| Maggie's Mission, Inc. | Dr. Vakoc | 09/16/18 | 09/15/24 | 50,000 | |
| The Manhasset Women's Coalition | Dr. Amor Vegas | 10/20/22 | 10/20/24 | 36,638 | |
| The Mark Foundation for Cancer Research | Drs. Janowitz/Beyaz | 10/15/20 | 01/31/25 | 750,000 | |
| The Mary Ruchalski Foundation, Inc. | Dr. Vakoc | 09/16/18 | 09/15/24 | 150,000 | |
| The G. Harold and Leila Y. Mathers Foundation | Dr. Beyaz | 06/01/21 | 05/31/24 | 238,336 | |
| | Drs. Koulakov/Zador | 10/01/21 | 09/30/24 | 215,603 | |

[1]Awarded, including direct and indirect costs

*New or competing renewals or supplements awarded in 2022

| *Grantor* | *Program/Principal Investigator* | *Project Start Date* | *Project End Date* | *2023 Funding*[1] | |
|---|---|---|---|---|---|
| Dr. Lee MacCormick Edwards Charitable Foundation | Dr. dos Santos | 10/01/23 | 09/30/24 | 15,000 | |
| The McKnight Endowment Fund for Neuroscience | Dr. Cheadle | 07/01/21 | 06/30/24 | 75,000 | |
| Michelle Paternoster Foundation | Dr. Vakoc | 09/16/18 | 09/15/24 | 50,000 | |
| Mr. and Mrs. Thomas Milana, Jr. | Dr. Trotman | 11/01/23 | 10/31/24 | 10,000 | |
| New York Genome Center | Drs. Tuveson/Krasnitz | 03/02/23 | 03/01/26 | 180,000 | * |
| The Nicholls Biondi Foundation | Dr. Tuveson | 01/01/23 | 12/31/23 | 25,000 | |
| Norn Group | Dr. Amor Vegas | 02/01/22 | 01/31/25 | 72,595 | |
| Ono Pharmaceutical Co., Ltd. | Drs. Tuveson/Chang/Dobin | 01/01/22 | 01/31/24 | 397,270 | |
| Penny's Flight Foundation | Drs. Van Aelst/Borniger/Lukey | 12/01/23 | 11/30/25 | 545,567 | * |
| The Pershing Square Foundation | Dr. Tuveson | 04/30/23 | 04/29/24 | 180,649 | |
| | Dr. Tollkuhn | 06/01/23 | 05/31/26 | 250,000 | * |
| Rita Allen Foundation | Dr. Cheadle | 09/01/21 | 08/31/26 | 100,000 | |
| Searle Scholars Program | Dr. Bandyopadhyay | 07/01/22 | 06/30/25 | 100,000 | |
| The Simons Foundation | Dr. Tobias | 01/01/22 | 12/31/24 | 480,000 | |
| | Dr. Tollkuhn | 03/01/22 | 02/28/26 | 90,000 | |
| | Dr. Tuveson | 05/01/23 | 04/30/28 | 2,000,000 | * |
| Starr Cancer Consortium | Drs. Egeblad/Van Aelst | 01/01/23 | 12/31/24 | 120,000 | * |
| | Drs. Janowitz/Borniger/Engel/Lukey | 01/01/22 | 12/31/24 | 278,400 | |
| | Dr. Joshua-Tor | 01/01/22 | 12/31/24 | 127,400 | |
| | Dr. Siepel | 01/01/23 | 12/31/24 | 74,320 | * |
| Swim Across America Nassau/Suffolk | Drs. Beyaz/Amor Vegas | 12/21/23 | 12/20/24 | 180,000 | |
| The Thompson Family Foundation | Drs. Tuveson/Moses/Lyons/Yeh | 07/01/23 | 06/30/26 | 198,000 | * |
| Treeline Biosciences, Inc. | Dr. Vakoc | 06/08/21 | 05/31/26 | 2,499,431 | |
| American Cancer Society/University of Miami | Dr. dos Santos | 01/01/23 | 12/31/27 | 31,374 | * |
| V Foundation | Dr. Krainer | 11/01/21 | 11/01/24 | 199,611 | |
| | Dr. Vakoc | 03/01/23 | 02/29/28 | 200,000 | * |
| | Dr. Westcott | 09/01/23 | 08/31/26 | 200,000 | * |
| The Wasily Family Foundation, Inc. | Dr. Westcott | 06/01/23 | 05/31/24 | 75,000 | * |
| Wings for Ewing Sarcoma | Dr. Vakoc | 09/16/18 | 09/15/24 | 20,000 | |
| Joan & Sanford I. Weill Medical College | Dr. Fearon | 07/01/22 | 06/30/24 | 104,244 | |
| The Bradley Zankel Foundation, Inc. | Dr. Mills | 11/15/23 | 11/14/24 | 15,000 | |
| Chan Zuckerberg Initiative DAF, an advised fund of Silicon Valley Community Foundation | Dr. Beyaz | 12/01/21 | 11/30/24 | 299,801 | |
| | Dr. Subhash | 11/01/23 | 10/31/24 | 20,000 | |
| *Fellowship Support* | | | | | |
| American Cancer Society | A. Habowski | 09/01/23 | 08/31/26 | 70,000 | |
| Annete Kade Fund | CSHL School of Biological Sciences | 01/01/23 | 12/31/23 | 65,254 | |
| Clare College | CSHL School of Biological Sciences | 08/16/23 | 08/15/24 | 43,508 | |
| Ester A. & Joseph Klingenstein Fund | Dr. Cheadle | 07/01/21 | 06/30/24 | 75,000 | |
| Fraza Research Foundation | D. Dumontier | 06/16/23 | 06/15/24 | 50,000 | |
| Lola A. Goldring | Dr. Stillman | 10/01/21 | 09/30/24 | 100,000 | |
| Howard Hughes Medical Institute | S. Sun | 10/01/23 | 09/30/27 | 100,000 | |
| | D. Adams | 09/01/21 | 08/31/24 | 43,690 | |
| | J. Bauer | 09/01/23 | 08/31/24 | 43,690 | |
| | L. Braviner | 08/16/23 | 08/15/24 | 55,128 | |
| The Meier and Linnartz Family Foundation | Dr. Cheadle | 12/01/23 | 11/30/24 | 50,000 | |

[1]Awarded, including direct and indirect costs

*New or competing renewals or supplements awarded in 2022

| *Grantor* | *Program/Principal Investigator* | *Project Start Date* | *Project End Date* | *2023 Funding*[1] | |
|---|---|---|---|---|---|
| The Research Foundation for State University of New York–Stony Brook | C. Cleary | 03/16/23 | 05/29/23 | 4,550 | |
| | P. Moresco | 07/16/23 | 07/15/24 | 4,550 | |
| | D. Voss | 06/01/23 | 05/31/24 | 4,550 | |
| The Simons Foundation | Gann, Alexander A | 12/01/23 | 11/30/24 | 500,000 | |
| | Dr. Matho | 06/01/23 | 08/31/24 | 10,000 | |
| | Dr. Matho | 06/01/23 | 08/31/24 | 223,975 | * |
| The Swartz Foundation | Drs. Koulakov/Pashakhanloo | 01/01/23 | 12/31/23 | 132,394 | |
| *Course Support* | | | | | |
| BioImaging North America | Course Scholarship Program | 04/01/23 | 03/31/24 | 5,500 | |
| Foundation Fighting Blindness, Inc. | Vision: A Platform for Linking Circuits, Behavior and Perception | 04/01/19 | 12/31/23 | 5,000 | |
| Howard Hughes Medical Institute | Course Program | 09/01/22 | 08/31/25 | 500,000 | |
| Regeneron Pharmaceuticals, Inc. | Regeneron Scholars Account Funder: Regeneron Pharmaceuticals | 01/01/19 | 12/31/23 | 100,000 | |
| Society for Neuroscience/International Brain Research Organization | Summer Neuroscience Course | 09/01/23 | 08/31/24 | 30,050 | |
| *Meeting Support* | | | | | |
| CSHL Translational Cancer Support | Biology of Cancer: Microenvironment and Metastasis | 05/01/23 | 04/30/24 | 105,134 | |
| Illumina, Inc. | Bio History Series | 06/01/23 | 05/31/24 | 6,000 | |
| Inari Agriculture | Plant Genome & Biotechnology: From Genes to Networks | 09/01/23 | 08/31/24 | 5,000 | |
| JXTX Foundation | Biological Data Science Meeting | 01/01/23 | 12/31/23 | 15,380 | |
| Life Science Alliance | Single Cell Analysis; Biology of Cancer | 01/01/23 | 12/31/23 | 3,500 | |
| New England Biolabs, Inc. | Bio History Series | 06/01/23 | 05/31/24 | 25,000 | |
| Pairwise Plants | Plant Genome & Biotechnology: From Genes to Networks | 09/01/23 | 08/31/24 | 1,000 | |
| Pioneer Hi-Bred International, Inc. | Plant Genome & Biotechnology: From Genes to Networks | 09/01/23 | 08/31/24 | 10,000 | |
| Promega Corporation | Bio History Series | 06/01/23 | 05/31/24 | 12,500 | |
| Qiagen LLC–USA | Bio History Series | 06/01/23 | 05/31/24 | 6,000 | |
| ViiV Healthcare | Retrovirus | 05/01/23 | 12/31/24 | 25,000 | |
| *Library Support* | | | | | |
| The New York State Education Department | | 07/01/23 | 06/30/24 | 4,291 | |
| Goelet, LLC | | 03/13/20 | 07/31/25 | 150,000 | |
| Goelet, LLC | | 10/27/21 | 10/26/24 | 10,000 | |
| The River Foundation/Celia and Wally Gilbert | | 10/27/21 | 10/26/24 | 45,000 | |
| Dr. Andrew Pressley | | 10/27/21 | 10/26/24 | 10,000 | |
| The Robert David Lion Gardiner Foundation, Inc. | | 01/01/22 | 12/31/24 | 128,123 | |
| *Preprint Server for Biology* | | | | | |
| This project has been made possible in part by grant number CZIF2022-007484 from the Chan Zuckerberg Initiative Foundation | Dr. Inglis | 07/01/22 | 06/30/24 | 2,000,000 | |
| This program has been made possible in part by a grant from the Chan Zuckerberg Initiative DAF, an advised fund of Silicon Valley Foundation | Dr. Inglis | 07/01/22 | 06/30/23 | 499,790 | |

[1]Awarded, including direct and indirect costs

*New or competing renewals or supplements awarded in 2022

## DNA LEARNING CENTER GRANTS

| *Grantor* | *Program* | *Duration of Grant* | *2023 Funding*[+] |
|---|---|---|---|
| **FEDERAL GRANTS** | | | |
| National Institutes of Health | *Citizen DNA Barcode Network* | 6/20–3/25 | $176,540 |
| National Institutes of Health | *Genomics Step-Up High School* | 5/22–3/27 | 180,429 |
| National Science Foundation | *Implementing DNA Barcoding for Course-Based Undergraduate Research Experiences* | 10/18–9/24 | 281,261 |
| National Science Foundation (University of Arizona) | *CyVerse: Cyberinfrastructure for the Life Sciences* | 8/18–7/24 | 82,301 |
| National Science Foundation | *Enhancing DNA Subway 2.0 as a Shared Resource for Bioscience Workforce Development* | 7/23–6/26 | 69,788 |
| National Science Foundation | *Collaborative Research: Arecibo C3—Center for Culturally Relevant and Inclusive Science Education, Computational Skills, and Community Engagement* | 10/23–9/24 | 64,703 |
| National Science Foundation (Austin Community College) | InnovATEBIO National Biotechnology Education Center | 10/19–9/24 | 213,554 |
| National Science Foundation (Pierce College) | *Advanced Student-Focused Projects: Internship, Research, and Education* (ASPIRE) | 9/21–8/24 | 9,048 |
| National Science Foundation (University of Minnesota) | *FMRG:BIO: Enabling Cell-Free Engineering* | 10/22–9/26 | 28,864 |
| National Science Foundation | *What Works in Workshops—Evolving Short Format Training to Serve Life Sciences STEM Professionals in the 21st Century* | 3/21–2/23 | 4,587 |
| National Science Foundation | *Nanopore DNA Sequence Course-Based Undergraduate Research* | 6/22–5/24 | 79,371 |
| **NONFEDERAL GRANTS** | | | |
| Beijing No. 166 High School | Chinese Collaboration Agreement | 7/19–6/23 | $25,593 |
| Breakthrough Prize Foundation | Laboratory Design and Teacher Training for Breakthrough Junior Challenge Prize Winners | 12/15–12/25 | 69,421 |
| Health Park | Health Park Agreement | 12/15–12/23 | 5,112 |
| Pinkerton Foundation | *Urban Barcode Research Program* | 1/21–5/23 | 100,550 |
| Richard Lounsbery Foundation | Paul Taubman support for *DNALC NYC* Exhibit Development | 6/21–6/24 | 125,703 |
| Richard Lounsbery Foundation | Research Ready Partnerships for NYC Public Schools | 6/21–5/23 | 16,500 |
| Richard Lounsbery Foundation | Videos and Animations to Explain Environmental DNA to a Broad Audience | 2/22–7/24 | 28,869 |
| The Simons Foundation | *Urban Barcode Research Program* | 12/17–8/23 | 26,807 |
| William Townsend Porter Foundation | *Harlem DNA Lab for Underprivileged Students* | 1/20–1/23 | 24,428 |
| DNALC Unrestricted Institutional Grant | DNALC Asia Royalties | 9/15–12/23 | 50,000 |
| NY Harbor Foundation | Billion Oyster Project | 6/20–12/21 | 296 |
| Laurie Landeau Foundation | Laurie Landeau Seed Program | 1/21–1/24 | 931 |

[+] Includes direct and indirect costs.

**The following schools and school districts participated in these membership programs of the DNALC:**

| | |
|---|---|
| ***Sustaining Memberships*** | |
| Bellmore–Merrick Central High School District | $3,150 |
| Elwood UFSD | $3,150 |
| Great Neck | $3,150 |
| Herricks Union Free School District | $3,150 |
| Huntington | $3,150 |
| Island Trees | $3,150 |
| Jericho High School | $3,150 |
| Levittown Union Free School District | $3,150 |
| North Shore Central School District | $3,150 |
| Oceanside Union Free School District | $3,150 |
| Oyster Bay–East Norwich Central School District | $3,150 |
| Plainview–Old Bethpage Central School District | $3,150 |
| Portledge School | $3,150 |
| Port Washington Union Free School District | $3,150 |
| Roslyn Union Free School District | $3,150 |
| Syosset Central School District | $3,150 |
| Yeshiva University High School for Girls | $3,150 |
| ***Associate Memberships*** | |
| Glen Cove Central School District | $17,000 |
| Friends Academy | $17,000 |
| St. Dominic High School | $17,000 |
| ***Partner Memberships*** | |
| Cold Spring Harbor Central School District | $34,000 |
| Long Beach Central School District | $34,000 |
| Massapequa Union Free School District | $34,000 |

**The following schools participated in these membership programs of the *DNALC NYC* at City Tech:**

| | |
|---|---|
| ***Sustaining Membership*** | |
| Magen David Yeshiva High School | $3,150 |
| Stuyvesant High School | $3,150 |
| ***Associate Membership*** | |
| Dwight School | $17,000 |
| ***Partner Memberships*** | |
| The Chapin School | $34,000 |
| Lycée Français de NY | $34,000 |
| Marymount School of NY | $34,000 |
| St. David's School | $34,000 |

**The following school participated in this membership program of the *Regeneron DNALC:***

| | |
|---|---|
| ***Sustaining Membership*** | |
| Archbishop Stepinac High School | $3,150 |

## BANBURY CENTER GRANTS

| *Grantor* | *Program* | *2023 Funding* |
|---|---|---|
| **NONFEDERAL SUPPORT** | | |
| Baylor College of Medicine | Developing an Ethical Framework for Psychedelics Research and Use | $10,000 |
| Bill & Melinda Gates Foundation | Optimizing Coverage of HIV/STI Prevention and Care Programs: A Program Science Approach | 75,156 |
| Cold Spring Harbor Laboratory Corporate Sponsor Program | Persistence, Senescence, and Cell Death | 30,304 |
| Cold Spring Harbor Laboratory Corporate Sponsor Program | Developing an Ethical Framework for Psychedelics Research and Use | 22,828 |
| Cold Spring Harbor Laboratory Corporate Sponsor Program | Integrating Exposomics into the Biomedical Enterprise | 20,000 |
| Cold Spring Harbor Laboratory–Northwell Health Affiliation | CSHL–Northwell Pancreatic Cancer Workshop | 8,287 |
| Cold Spring Harbor Laboratory–Northwell Health Affiliation | CSHL–Northwell Brain Tumor Workshop | 7,117 |
| Cold Spring Harbor Laboratory–Northwell Health Affiliation | The Future of Investigational Medicine: Utilizing Science to Optimize the Early Phase Oncology Clinical Trial Effort | 57,841 |
| Columbia University | Integrating Exposomics into the Biomedical Enterprise | 23,700 |
| Dryad | Dryad Strategy Retreat | 38,579 |
| FRAXA Research Foundation | FMRP Restoration: Definitive Therapies in Fragile X | 55,630 |
| Kennedy's Disease Association | Kennedy's Disease (SBMA) Research Workshop | 30,403 |
| Lustgarten Foundation | The Lustgarten Foundation Scientific Meeting | 37,350 |
| Petrie-Flom Center/POPLAR | Developing an Ethical Framework for Psychedelics Research and Use | 24,000 |
| Project Santa Fe Foundation, LLC | The 2023 CL2.0 Colloquium | 37,010 |

# MAJOR PROGRAM FUNDING FOR MEETINGS & COURSES

We appreciate the ongoing major financial support for our courses from the following: Helmsley Charitable Trust, Howard Hughes Medical Institute, National Institutes of Health, National Science Foundation, and Regeneron. The course program is also supported by equipment and reagent companies that provide in-kind support for the program.

Contributions from the following companies provide core support for the Cold Spring Harbor meetings through the Corporate Sponsor Program:

**CORPORATE BENEFACTORS**

Estée Lauder Companies
Regeneron
Genentech
Merck & Co., Inc.
New England BioLabs
Novartis

**CORPORATE SPONSORS**

Agilent Technologies
Biogen
Bristol-Myers Squibb
Calico Life Sciences LLC

**CORPORATE PARTNERS**

Alexandria Launch Labs

The National Institutes of Health provided multiple grants for individual meetings. Additional companies and foundations provided individual grants and sponsorships in support of individual meetings.

**David Stewart**
*Executive Director, Meetings and Courses Program*
*President, Cold Spring Harbor Asia*

**Terri Grodzicker**
*Academic Guidance*
*Dean of Academic Affairs*

# ADVANCEMENT

The Advancement Department continues to raise funds in support of the Foundations for the Future Campaign—the largest expansion project in Cold Spring Harbor Laboratory's history. At year's end, $322 million was raised toward the $500 million campaign goal. The project will add four new research buildings, expanding into the areas of Neurodegenerative Disease, Cancer-Neuroscience, and Neuroscience/Artificial Intelligence. Additionally, there will be a Conference Center and a Collaborative Science Center to service the thousands of scientists who travel to our campus each year for scientific meetings and courses. We are grateful to Campaign Chair Robert Lourie and Development Committee Chair Howard Morgan for their leadership and generosity.

In 2023, a record $9.2 million was raised for unrestricted support, which included proceeds from fund-raising events. Beginning in June, Eric Carlstrom was honored at the Golf Tournament. Then in October, longtime donor Cynthia Stebbins was our honoree at the Women's Partnership for Science lunch. Finally, science philanthropists Neri Oxman and Bill Ackman and renowned scientist Dr. Jim Allison were honored in November at the Double Helix Medals for their dedication to science and humanity.

The scientific discoveries and advances being made at Cold Spring Harbor Laboratory could not happen without the support of our donors. As always, we deeply appreciate all that they do to propel our science forward.

**Charles V. Prizzi**
*Senior Vice President for Advancement*
*Special Advisor to the President*

29th Annual Golf Tournament honoree Eric Carlstrom with Bruce Stillman

# Cold Spring Harbor Laboratory Corporate Advisory Board

Established in 1992, the Corporate Advisory Board (CAB) serves as a liaison to the corporate community and assists in securing unrestricted dollars for annual support of Cold Spring Harbor Laboratory. Comprised of influential business leaders from large and small companies on Long Island and Manhattan, the CAB is a necessary and vital contributor to the continued success and mission of Cold Spring Harbor Laboratory.

The goals of the CAB are to:

- act as an ambassador for Cold Spring Harbor Laboratory;
- offer a level of expertise in strategic planning and marketing that will ensure success;
- educate business leaders about the importance and values of science on Long Island.

The CAB meets two times per year. The annual financial commitment to CSHL by Board members is $5,000.

**Corporate Advisory Board, 2023**

Edward A. Chernoff, MARS, *Chair*

*Matthew Aboff, Aboff's Inc.
David Altman, Brown & Altman, LLP
Paul Amoruso, Oxford & Simpson Realty
Todd Andrews, Centerbrook Architects and Planners
*Michael Asta, Asta Law
Stephen R. Barrese, Dilworth & Barrese
Eric Belfi, Labaton Sucharow LLP
*Jennifer Bitterman
Edward Blaskey, Sterling National Bank
Ryan J. Bohn, Tempus Financial Group
Jean Cacciabaudo, M.D.
Thomas J. Calabrese, Daniel Gale Sotheby's International Realty
John T. Callaghan, Callaghan LLP
John D. Catalano, Catalano Enterprises, LLC
Richard A. Catalano, KPMG, LLP
Jonathan Connors, Wells Fargo
Marian Conway, Ph.D., NYCB Foundation
Denis Cullen
Alessandro Delfino, McKinsey and Company
Tracy Dellomo, UBS
Gregory DeRosa, Roanoke Holdings
Lauren Emr, Esq.
Jim Ford, Eppendorf North America
*Bob Fox, Fox's
Brian Fox, Klick
*Keith Friedlander
Amit Gandhi, M&R Hotel Group
Tom Giarraputo, Executive Cleaning Services, LLC
Thomas Gsell, R2DConsulting, LLC
*Donald Hehir, Donald Hehir & Associates
Mary Jane Helenek
Richard W. Humann, H2M architects + engineers
Nancy Israeli, M.D.
Alan L. Jakimo, Sidley Austin LLP
Patricia Janco-Tupper, Capital Group
John C. Kean III, Kean Development Company
Michael Keenan, Wells Fargo Bank, N.A.
Laurie J. Landeau, V.M.D.
David Lessing, Lessing's
Patricia Marcin, Rivkin & Radler
Jeffrey L. Martin, M.D., Sight MD
Chris McIntosh, JP Morgan
Stephen F. Melore, Farrell Fritz, P.C.
Richard Nattis, M.D.
Robert Palatnick, DTTC
John G. Passarelli, M.D.
*Marc Perez, Bank of America
Sean Perrotta, Point72
*Jeremy Pole, Boilermatic Welding Industries
Nicole Prizzi, Harbor Digital Properties
*Roberto Rappa, M.D., Ally Anesthesia
Erin Rechler
Joseph Roberto, BankUnited, N.A.
Keith Rooney, National Grid
Stephen Ross, Nikon
Don Saladino, Drive 495
Raju Sarwal, M.D.
*Robert Scoskie, Summit Health—CityMD
Carmella L. Stephens, Ph.D., Carter, DeLuca & Farrell
Edward Strohm, Three Strohm Sisters Family
*Kristin Thomas, Marble Collective
John Topolovec, TD Bank
Ed Travaglianti Jr., The Cullinan & Travaglianti Group
Craig A. Weiss, T. Weiss Realty Corp.
Dave Zuklie, The Swoondle Society

*New 2023 Members.

# Cold Spring Harbor Laboratory Association

Under the leadership of CSHL Association Board President Mark Hamer, in 2023 more than $9.2 million in unrestricted funding was raised through fund-raising events, the annual fund campaign, and community outreach. The sold-out 29th Annual Golf Tournament, chaired by Director Eddie Chernoff, was held June 13 and honored Director Eric Carlstrom. Honoring Cynthia Stebbins, on October 1, the 22nd annual Women's Partnership for Science luncheon featured CSHL Professor and Director of Research Dr. Leemor Joshua-Tor. The Double Helix Medals Dinner was held November 15 at the American Museum of Natural History, where we honored Nobel Laureate Jim Allison and science philanthropists Neri Oxman and Bill Ackman. The CSHLA Directors continue to represent CSHL as community ambassadors, and we are grateful for their service.

## Officers

Mark Hamer, *President*
Ron Gottlieb, *Treasurer*
Elizabeth Ainslie, *Secretary*

## Directors

Mary Auersperg
Hans E.R. Bosch
Christina Bucci-Rechtweg
Eric Carlstrom
Edward A. Chernoff
Steve Chestler
Debra Del Vecchio
Frank DellaFera
Nelson DeMille
Quentin Dolan
David Einbinder
Kelly Gaudreau
Stephanie Gibbons
Brad Glick
David Goldring
Candace D. Hammonds
Carissa Jordan
Ashley Jostrom
Jack Kelly
Terri Keogh
Errol Kitt
Peter Klein
Nick Leopard
Emma Liu
Madelyn Lombardi
Christine Masata
Michael Maturo
Marcia Kramer Mayer
Michèle Bahnik Mercier
Tom Milana, Jr.
Nina Monell Morton
Eileen Otto
Patricia Petersen
George Petrocheilos
Rita Ranieri
Alicia Zarou Scanlon
Frank Sciame, Jr.
Linda Silver
Hope Geier Smith
Kristin Olson Smith
Mary Striano
Peter Tilles
Lara Trafelet

## Honorary Directors

Pien Bosch
Trudy Calabrese
Robert Gay
Ginny Knott
Anne R. Meier
Cathy Cyphers Soref
Cynthia R. Stebbins
Pat Woods

The 2023 CSHL Association Directors at Grace Auditorium

# Honor Roll of Donors

### Lifetime Contributions of $5 million+

Dr. and Mrs. Lalit R. Bahl
The Arnold and Mabel Beckman Foundation
BGI
Jamie C. Nicholls and O. Francis Biondi
Mr. and Mrs. David Boies, Boies, Schiller & Flexner LLP
The Breast Cancer Research Foundation
Chan Zuckerberg Initiative DAF, an advised fund of Silicon Valley Community Foundation
The Dana Foundation
Mrs. Norris W. Darrell, Jr.
Barbara J. Amonson and Vincent J. Della Pietra
Pamela Hurst-Della Pietra and Stephen Della Pietra
DeMatteis Family Foundation
Charles F. Dolan, Dolan Family Foundation
The William Stamps Farish Fund
Charitable Lead Annuity Trust, Will of Louis Feil
Jacob Goldfield
Kate Medina Guthart and Leo A. Guthart
The Leona M. and Harry B. Helmsley Charitable Trust
Mr. and Mrs. Jeffrey E. Kelter
David H. Koch*
Laurie Landeau Foundation
Mr. and Mrs. Robert D. Lindsay and Family
Mr. and Mrs. Thomas H. Lister
Ivana Stolnik-Lourie and Robert W. Lourie
The Lustgarten Foundation
Nancy Abeles Marks
The G. Harold and Leila Y. Mathers Charitable Foundation
Mercer Family Foundation
The Don Monti Memorial Research Foundation
Dr. and Mrs. Howard L. Morgan
New York Empire State Development Corporation
Pershing Square Foundation
Cynthia Hazen Polsky and Leon Polsky
The Quick Family
Charles and Marie Robertson Family
Edith Seligson
Drs. Marilyn and James Simons
The Simons Foundation
Alfred P. Sloan Foundation
St. Giles Foundation
The Starr Foundation
Dr. and Mrs. James M. Stone
The Swartz Foundation
Mr. and Mrs. Paul J. Taubman
The Thompson Family Foundation
Dr. and Mrs. James D. Watson
Holly and Henry Wendt III*
Roy J. Zuckerberg Family Foundation

### Lifetime Contributions of $1 million+ or $100,000+ in 2023

Elizabeth and Lee Ainslie
Betty Ajces Trust under the Last Will and Testament of Leon Ajces
Rita Allen Foundation, Inc.
The Bahnik Foundation
Bloomberg Philanthropies
Mary K. Chapman Foundation
Mr. and Mrs. Edward A. Chernoff, MARS
Ms. Verena F. Cushman
Davenport Family Foundation
Mr. Michel David-Weill*
Laura and John Desmarais
Madelyn and Rick DiBella
Oliver S. and Jennie R. Donaldson Charitable Trust
Edward P. Evans Foundation
Elizabeth Cogan Fascitelli and Michael D. Fascitelli
Robert D.L. Gardiner Foundation
Mr. and Mrs. Robert A. Gay
Mr. and Mrs. Ilan Gluzman
Francis Goelet Charitable Trust
Miriam and Alan Goldberg
Estate of Fred Goldberg
Lola Goldring
The Oliver R. Grace Family
Jon and Mindy Gray Family Foundation
Michael J. Griffin
Guru Krupa Foundation, Inc.
Mr. and Mrs. Mark W. Hamer
Janet Strauss and Jeff Hawkins
Jo-Ellen and Ira Hazan
William Randolph Hearst Foundation
Heartfelt Wings Foundation
Elizabeth McCaul and Francis Ingrassia
Jorge Family Foundation
F.M. Kirby Foundation, Inc.
Kissinger Family Foundation
Peter Klein, Claire Friedlander Family Foundation
Knott Family Foundation
Vesna and Tomislav Kundic
Betsy and Bryan H. Lawrence
The Lehrman Institute
Mr. and Mrs. Stephen M. Lessing
Jerome Levy Foundation
Mary D. Lindsay*
Long Island Real Estate Group
Louis Morin Charitable Trust
The Mark Foundation for Cancer Research
Dr. Marcia Kramer Mayer
Magaro Family Foundation
Estates of Florence and Harold and Ethel McNeill
Gillian and Eduardo Mestre
William R. Miller*
William C. and Joyce C. O'Neil Charitable Trust
Jeff Parsigian
Mr. and Mrs. Douglas S. Partrick
Penny's Flight Foundation
The Pfizer Foundation
Amy and John Phelan
Lyon Polk, Morgan Stanley Private Wealth Management
Linda Johnson and Bruce Ratner
Estate of Vincent Rinando
The Mary Ruchalski Foundation
The Saunders Family
Judy Gibbons and Francesco Scattone
Alison Holtzschue and Douglas Schloss
Wendy and Eric Schmidt, Schmidt Futures
Eleanor Schwartz Charitable Foundation
The Seraph Foundation
Pearl F. Staller*
Starr Companies
Cynthia Rossbach Stebbins
Swim Across America Nassau–Suffolk
Waclaw Szybalski, D.Sc.*
TD Bank
Dr. and Mrs. Stuart T. Weisbrod
Anne Wojcicki Foundation
Dr. George Yancopoulos, Regeneron Pharmaceuticals

### Contributions of $30,000+

Aboff's, Inc.
Autobahn Labs
Caroline E. Bassett
Mr. and Mrs. Hans E. Bosch
Capital Group
Cedar Hill Foundation
Centerbrook Architects and Planners, LLC
Clare College
Ellen and Casey Cogut
Estate of Marilyn Cunniff
DelVecchio Family Foundation
Mr. and Mrs. Gregory DeRosa, G2D Development Corp.
Ike, Molly and Steven Elias Foundation
Andrew Farkas & Island Capital
Douglas and Christine Fox
Friends of TJ Foundation
Mr. and Mrs. Ronald A. Gottlieb
The Marc Haas Foundation
Irving Hansen Foundation
Douglas and Robin Horn
Tom and Mara Hutton, Geoffrey Beene Foundation
Marc and Jennifer Lipschultz Family Foundation
Maddie's Promise

*Deceased

Maggie's Mission
Michael Maturo, RXR Co Property Management LLC
The Meier and Linnartz Family Foundation
The Ambrose Monell Foundation
Northwell Health
Parker Foundation
Michelle Paternoster Foundation
Patricia J. Petersen, Daniel Gale Sotheby's International Realty
George Petrocheilos, Catalio Capital Management, LP
Christina Renna Foundation
The River Foundation
Geoff Robertson, Robertson Family Foundation
Mr. and Mrs. Frank Sciame, Jr.
Christine Anderson and Jake Siewert
Mr. and Mrs. Harry Slatkin
Mr. and Mrs. Dustin Smith
The Wasily Family Foundation

### Contributions of $10,000+

Chang Ge Aaronson
Alpha Omega Charitable Foundation
Paul and Mary Auersperg
Mr. and Mrs. W. Dillaway Ayres, Jr.
Tess and Demetri Ballas
Fred and Janet Baron
Deborah C. Benjamin
Benjamin Moore
Dr. Michael Botchan
Broad Hollow Bioscience Park
Dr. Christina Bucci-Rechtweg and Mr. Jay Rechtweg
The Darlene Carbone Brain Tumor Foundation
CardWorks Servicing, LLC
Mr. and Mrs. Eric Carlstrom
Mr. and Mrs. James J. Celestino
Steven Chestler
Frederic R. Coudert Foundation
Denise R. Coyle
Davies Ward Phillips & Vineberg
Frank and Janet DellaFera Family Foundation
Nelson R. DeMille
James L. Dolan
Dr. Lee MacCormick Edwards Charitable Foundation
Liete and Mark Eichorn
Adam Flatto
Mr. and Mrs. Russell Gaudreau
GEI Consulting Engineers & Scientists
Stephanie and Gregory Gibbons
Brad Glick
Dr. and Mrs. Philip Goelet
Mr. and Mrs. Austen T. Gray, Jr.
Michael and Nicole Graziano
Candace and Brad Hammonds
Susan T. Harris
Josefin and Paul Hilal
Carissa and James Jordan
Ashley and Gabe Jostrom
John Kean, Kean Development Company
Carmela and John Kelly
Terri and Peter Keogh
Ginny and Errol Kitt
Beth and Seth Klarman
Andrea B. and Peter D. Klein
Michael Kramer
Gordon Lamb, Jefferson Family Charitable Foundation
Melissa and Nick Leopard
Cynthia and David Lippe
Emma Liu
Madelyn and Carl Lombardi
John and Bridget Macaskill
Christine and Daniel Masata
Mr. and Mrs. Thomas Milana, Jr., Man Cave Health
Tom Moore and Judy Livingston Moore
Sharmin Mossavar-Rahmani
National Grid Foundation
Ashley and Frank O'Keefe
Susanne C. Olin
O'Neill Family Charitable Trust
Mr. and Mrs. Jonathan Otto
Lauryl and Robert Palatnick
David P. Pearson
Dr. Andrew Pressley
Cheryl and John Pufahl
Marina and Thomas Purcell
Rita and Lewis Ranieri
Erin Rechler, The Morton and Beverly Rechler Family Foundation
John R. Reese
Maggie Lovett and Tim Rotolo
The Pamela & Richard Rubenstein Family Fund
Jonathan Shugar, Goldman Sachs & Co.
Jesse and Linda Silver
Skanska USA Building, Inc.
John Sobolewski
Cary and Marisela Staller
Joseph S. and Diane H. Steinberg 1992 Charitable Trust
Dr. and Mrs. Bruce W. Stillman
The Strohm Foundation
Mr. and Mrs. Harold J. Thompson
Dori and Peter Tilles
Lauren and Bobby Turner
Turner & Townson
United Bank
Robin and Paul Vermylen
Webster Bank
T. Weiss Realty Corp.
Mr. and Mrs. W. Fifield Whitman
William Townsend Porter Foundation
Wings for Ewing Sarcoma
The Bradley Zankel Foundation

Bruce Stillman, Leemor Joshua-Tor, Marilyn Simons, and Cynthia Stebbins, 2023 honoree, at the Women's Partnership for Science lunch.

### Contributions of $5,000+

20th Century Fox Film Corporation
Anthony Acerra, ALC Steel Corporation
David N. Altman, Brown Altman & DiLeo, LLP
Paul Amoruso, Oxford & Simpson Realty Services, Inc.
Anron Heating & Air Condition
Michael Asta, Asta Law, L.P.
Atkinson Koven Feinberg Engineers, LLP
B&G Electric
BankUnited, N.A.
Andre Barnowski, Tri State Dismantling Corp.
Eric Belfi
Dr. and Mrs. Adam Bitterman
Hadley and Ryan Bohn

Ian Mohr, Ilan Gluzman, Terri Grodzicker, Josefa Gluzman, and Mike Botchan at a dedication of the Terri Grodzicker Library and Yakov Gluzman Lounge.

Mike Botto, Botto Mechanical Corporation
Dr. Steven Brenner
Laura Louise Breyer
Mr. and Mrs. Thomas J. Calabrese, Jr.
Callaghan LLP
Lawrence and Teresa Cantwell
Devon Carroll
Certified Interiors, Inc.
Rita Cleary
Susan Cohen
Mr. and Mrs. Jonathan Connors
Mr. and Mrs. James C. Cook
Joe Cooney, Squan
Cullen Family Fund
Leslie and Alessandro Delfino
Steven Dubner Landscaping
Sarah Edwards
Mr. and Mrs. David Einbinder
E-J Electric Installation Co.
Michael L. Focazio, Carissa Maringo Fund
Jim Ford, Eppendorf North America, Inc.
Mr. and Mrs. Robert Fox, Fox's
Carmen and Keith Friedlander
Ellen V. Futter
Anneke Young Gaber
Gil-Bar Health and Life Science
Bernard F. and Alva B. Gimbel Foundation
Glen Cove C.A.R.E.S., Inc.
Raquel Flatow Haas and Michael Haas
Robin Hadley
The Hastings Foundation, Inc.

Jim Allison, 2023 Double Helix honoree, with Bruce Stillman.

Donald Hehir & Associates, LLC
Robin and Jim Herrnstein
Rich Humann, H2M architects + engineers
Drs. Nancy and Ron Israeli
Alan L. Jakimo
Steven B. Klinsky
Lessing's
Rita and Art Levinson
Susan Lucci
Dr. Jean Cacciabaudo and Mr. William Maiorino
Patricia Marcin - Rivkin Radler LLP
Masthead Cove Yacht Club
Edward E. Matthews
Chris McIntosh, JP Morgan Private Wealth Management
Stephen Melore, Farrell Fritz P.C.
Michele M. Miroff
Drs. Ian Mohr and Michelle Pacht
National Grid
Dr. Stephen Ross, Nikon Instruments, Inc.
James J. Norman
Hugh and Arianne O'Kane
Mr. and Mrs. Jonathan Otto
Oxford Nanopore Technologies, Inc.
Anne Pace
Par Plumbing Co., Inc.
Dr. and Mrs. John G. Passarelli
Corine and Sean Perrotta
Jeremy Pole, Boilermatic Welding Industries, Inc.
Nicole and Charlie Prizzi
Dr. Roberto Rappa
Theadora Richmond
Linda Rodgers
Tami and Scott Schneider
Mr. and Mrs. Gordon L. Seaman
Bernadette Casey Smith
Spionkop Charitable Trust
Striano Electric Co., Inc.
Kristin Thomas, Marble Collective
Zara and David Tisch
The Tonna Family
Mr. and Mrs. Edward Travaglianti, Jr.
Sandy Tytel
Drs. Marjorie J. Van de Stouwe and Scott J. Ratner
John B. Vermylen
Wells Fargo Private Bank
Dr. and Mrs. Michael Wigler
Ann Eden Woodward Foundation
Mr. and Mrs. David Zuklie

### Contributions of $1,000+

A.K.S. International, Inc.
Hope and Marc Altheim
Nelle and Todd Andrews
Debra and Scott Arenare
Dorothy T. Baldwin
Shahnaz Batmanghelidj and Radford W. Klotz

Kathy and Gene Bernstein
Vincent and Patricia Breitenbach
Mr. and Mrs. Timothy S. Broadbent
Irina and Sergey Butkevich
Carol and Stephen Canter
Louise Parent and John Casaly
John D. Catalano
Kate Calabrese Chapman
Daniel and Rhonda Chestler
Mr. and Mrs. Elliot Conway, Dau Family Foundation
John and Maggie Cooley
Dr. Angela N. Creager
Henry P. and Kristina Davison II
Estate of Meleanor Deming
Todd A. DiScala
Nancy R. Douzinas
Daniel F. Doyle
Dorothy Engel
Emily T. Frick
Olivia Tiernan Geary
EM Geddes, Jr.
Artie Godsell
Dagnia and Walter Goldschmidts
Eric Goldsmith
Gondelman Foundation
Pierre and Paula Gonthier Family Foundation
Tom Gsell
Lynn and Frank Gundersen
Vivian Ha
Jeffrey Hager
Virginia Hanson
Margaret Hargraves
Gwen N. Harris, M.D.
Mr. and Mrs. Robert Heathwood
Drs. Nouria Hernandez and Winship Herr
William A. Herzog
James Hicks
Hirschhorn and Klebanoff Family Fund
James B. Hoover
Douglas and Robin Horn
International Society of Neuroethology
Walter B. James Fund No. 2
Barclay and Jean Jones
Dr. Leemor Joshua-Tor
Drs. Kathleen and Victor Klein
Mr. and Mrs. Ragnar M. Knutsen
John Kubacki
James and Carol Large
Kevin Leichter
Maureen G. Leness
John Leonard, M.D.
The Jeremy and Robin Lewallen Foundation
George N. Lindsay, Jr. and Nancy Metz
Jane and Philip Mallinson
W. Corby May
Victor K. McElheny
Neil and Amy McGoldrick
Elizabeth Menges
Jeanne Moutoussamy-Ashe
Mullarkey Family Fund
Nassau County Department of Health
Mr. and Mrs. Justin Nelson
Norwood Foundation, Inc.
Claudia Overstrom
Stacey Paci
Jeffrey Pash
Nicholas Paumgarten
Petro's Peers
Ira Platt
Lou Ann Montana and JW Pflugrath
Porter Braden Fund
Mr. and Mrs. Thomas L. Pulling
Bernard and Anne Reynolds
Ritter Family Foundation
Joe Roberto
Rich Roberts
Mr. and Mrs. William Roche
Dr. and Mrs. Gerald M. Rubin
Dr. Marilyn Moffat Salant
Dr. and Mrs. Raju Sarwal
Peter Scalamandre & Sons, Inc.
Robin and Enrique Senior
Mr. and Mrs. Douglas S. Soref
Pat and Jim Stewart
Mr. and Mrs. Keith Stillman
Pamela M. Thye
The Louise Giffuni Tiernan Foundation
Mary Ann Tighe
Patricia W. Timpson
Lynn and Pat Tone
John and Carol-Ann Treiber
Susie and Stanley Trotman
Sarah Trust
Universe Kogaku (America), Inc.
Mr. and Mrs. Charles Vallone
Robert F. and Joan M. Vizza Foundation
Jim Vogel
Gerald I. White
Sandy and Jennifer Williams
Mr. and Mrs. Steven J. Winick
Richard Zoller

Neri Oxman and Bill Ackman, 2023 Double Helix honorees.

### In-Kind

1212
Aboff's Paints
Americana Manhasset
Kris Amplo
Michael Asta
Mary Auersperg
Veronica Beard
Besito
Jennifer Bitterman
Jean Cacciabaudo
Eric Carlstrom
Rita Castagna
Kate Chapman
Eddie Chernoff
Steve Chestler
Cook's Studio
Chris Croken
CSHL WiSE
Daniel Gale Sotheby's International Realty
Tracy Dellomo
Nelson DeMille
Matt Dougherty
Harry Dunn
Estée Lauder
The Farm Italy
Fiorello Dolce
Fox's
Fuentes Cosmetic Surgery
Amit Gandhi
Garden City Hotel
Brandon and Stephanie Garrett
Stephanie and Gregory Gibbons
Dagnia Zeidlickis and Walter Goldschmidts
Ron Gottlieb
Mary Jane Helenek
Annabel Romero Hernandez
Holly Logan Art
Elizabeth McCaul and Frank Ingrassia


The Mary Ruchalski Foundation presented a check to the Vakoc laboratory for their sarcoma research.

Nancy Israeli
John W. Engeman Theater
Jonathan's Ristorante
Carissa Jordan
Natalie Rakowski Kammer
Karmic Grind
Carmela and Jack Kelly
Jeffrey Kelter
Kerber's Farm
Sarah Kitt
Lessings, Inc.
Living
Maddy and Carl Lombardi
Michael Maturo
Mitchells
Nassau Flyers
Nest
Nikon
Sarah Kelly Noderer
Oheka Castle
Eileen Otto
The Paramount
John Passarelli
Patricia Petersen
Jeremy Pole
R&S Meats
Erin Rechler
Red
Rex Burger and Lobster
Joe Roberto
Stephen Ross
Sanctuary Home and Patio
Sandbar
Sapsuckers
Raju Sarwal
Alicia Scanlon
Tom Schaudel
Sedoni Gallery
The Shed
Linda and Jesse Silver
Laura Slatkin
Southdown Coffee
Priya Sridevi
Cynthia Stebbins
Stellina
Bruce Stillman
TD Bank
Terra
Tilles Center
Ed Travaglianti, Jr.
John Tunney
Marjorie Van de Stouwe
Wempe
The Wine Line
Youngs Farm
Dave Zuklie

### Contributions in honor of

Neri Oxman and Bill Ackman
Dill Ayres
Dr. Jillian Berkman
David Boies
Hans and Pien Bosch
Canadian Police Association
Casey Cogut
Susan Cohen
Kristina Perkin Davison
Camila dos Santos
Thomas Gingeras
Toby Goldberg
Lola Goldring
Jessica and Thomas Heckel
Doug and Robin Horn
Ron Israeli
Jenny and Jeff Kelter
Laureen Knutsen
Laurie Landeau
Brian Lee
Robert D. Lindsay
Peggy Milonas
Ontario Provincial Police Association
Lewis Ranieri
John and Doreen Reali
Douglas Schloss
Ilgin Seidner
Adam Siepel and Laura Woodson
Jim Simons
Hope Geier Smith
Cynthia Stebbins
Diana Taylor
Daniel and Justyna Torres
Dave Tuveson
Sallie and Alex Van Rensselaer
James D. Watson

### Contributions in memory of

Shirley R. Aprison
Sandy Aranoff
Bea Aron
Karen Bahar
Roger Bahnik
Jerry Ballas
The Bedells
The Boudreaus
Bill Brenner
Ellen Brenner
Rose Brenner
Adrienne Cammarata
Joe Cerami
James Cleary
John Cleary
Marie Cochran
Helen Dolan
George Durning
Michael Eisler
Claire Finkel
Rory Friedland
Helen Gallagher
Yakov Gluzman
Dr. Christopher G. Goff
Teresa Haire
John W. Hanson
Dr. Ruth M. Patrick and Dr. Charles Hodge IV
Ezra Joshua
Phillip Charles Karda
Brian Kenny
Jerry Willy Kiser
Amar JS Klar
David Knott
John Ligas
Barbara Lobovsky
Fred Lobovsky
Susan Maurice
John J. McGowan
Frank Morelli

2023 Helix Society lunch.

Anthony Petrocelli
Diane Edmin Sachs
Roger Samet
Barbara and James Schubauer
Alan Seligson
David Smith
John Smith
Sandy Sparber
Freddie and Erwin Staller
Jason Stewart
Rajani Lakshmi Velivela
Margaret Wallace
Irv Weinstein
Amy Wessan
Norton Zinder
Dr. Mark Zoller

### Helix Society

Mr. and Mrs. W. Dillaway Ayres, Jr.
Caroline E. Bassett
Elise Best
Marjorie Bhavnani
Mrs. Willliam M. Blair, Jr.
John Broven
Mr. and Mrs. Thomas J. Calabrese, Jr.
Vicki Gruber Callahan
Kate Calabrese Chapman
Edward A. Chernoff
Dr. Bayard D. Clarkson
Verena F. Cushman
Henriette Darrell
Jane Duggan
Jan Eisenman
Mary Epstein
Michael L. Focazio
Mr. and Mrs. Douglas B. Fox
John H. Friedman
Jean G. Gardiner
Drs. Joan E. Brooks and James I. Garrels
Robert A. Gay
Eleanor J. Greenan
Michael J. Griffin
Michael Gurtowski
Robin Hadley
Mr. and Mrs. Mark W. Hamer
Margaret Hargraves
Mr. and Mrs. Herman M. Heinemann
Mrs. Valdemar F. Jacobsen
Mr. and Mrs. Peter D. Klein
Dr. Daniel F. Klessig
Laurie J. Landeau, V.M.D.
Mrs. Leslie S. Learned
Mrs. Henry Lewis III
Lisa M. Manche
Dr. Marcia Kramer Mayer
Michele M. Miroff
Susan Morgan
Richard S. Overton
Ann Parkinson
John S. Popeleski
Whitney F. Posillico
Dr. Gregory Prelich
Nicole and Charlie Prizzi
John R. Reese
Edith Seligson
Jeffrey Shellan, Ph.D.
Barbara Sicherman
Suzanne Slocum
Mr. and Mrs. James L. Spingarn
Cynthia R. Stebbins
Mr. and Mrs. Harold J. Thompson
Dr. Robert Tjian
Dr. and Mrs. Stuart T. Weisbrod
Holly Wendt

# LABORATORY MANAGEMENT

Bruce W. Stillman, Ph.D., President & CEO
John P. Tuke, Chief Operating Officer

**RESEARCH**

David L. Spector, Ph.D.
*Director of Research*

Walter L. Goldschmidts, Ph.D.
*Vice President, Executive Director of Sponsored Programs*

**Research Operations**

Sydney Gary, Ph.D.
*Director, Research Operations*

**Professors**

Jeff Boyd, Ph.D.
Douglas Fearon, M.D.
Hiroyasu Furukawa, Ph.D.
Thomas Gingeras, Ph.D.
*Professor and Head, Functional Genomics*
Christopher Hammell, Ph.D.
Ivan Iossifov, Ph.D.
David Jackson, Ph.D.
Leemor Joshua-Tor*, Ph.D.
*Chair, Cancer and Molecular Biology*
Alexei Koulakov, Ph.D.
Adrian Krainer, Ph.D.
*Deputy Director, CSHL Cancer Center*
Bo Li, Ph.D.
Zachary Lippman*, Ph.D.
*Professor & Director of Graduate Studies*
Robert Martienssen*, Ph.D.
*Chair, Genomics and Plant Biology*
W. Richard McCombie, Ph.D.
Alea Mills, Ph.D.
Partha Mitra, Ph.D.
John Moses, Ph.D.
Stephen Shea, Ph.D.
Adam Siepel, Ph.D.
*Chair, Simons Center Quantitative Biology*
Nicholas Tonks, Ph.D.
*Deputy Director, CSHL Cancer Center*
Lloyd Trotman, Ph.D.
David Tuveson, M.D., Ph.D.
*Director, CSHL Cancer Center*
Christopher Vakoc, M.D., Ph.D.
Linda Van Aelst, Ph.D.
*Chair, Neuroscience*
Doreen Ware**, Ph.D.
Michael Wigler, Ph.D.
Anthony Zador, M.D., Ph.D.

**Associate Professors**

Dinu Albeanu, Ph.D.
Camila dos Santos, Ph.D.
Tobias Janowitz, M.D., Ph.D.
Justin Kinney, Ph.D.
Dan Levy, Ph.D.
David McCandlish, Ph.D.
Saket Navlakha, Ph.D.
Vincent Pedmale, Ph.D.
Jessica Tollkuhn, Ph.D.

**Assistant Professors**

Arkarup Bandyopadhyay, Ph.D.
Semir Beyaz, Ph.D.
Jeremy Borniger, Ph.D.
Lucas Cheadle*, Ph.D.
Benjamin Cowley, Ph.D.
Alexander Dobin, Ph.D.
Xun Hou, Ph.D.
Peter Koo, Ph.D.
Michael Lukey, Ph.D.
Hannah Meyer, Ph.D.
Gabrielle Pouchelon, Ph.D.
Andrea Schorn, Ph.D.
Peter Wescott, Ph.D.
Lingbo Zhang, Ph.D.

**Research Professor**

Alexander Krasnitz, Ph.D.

**Research Associate Professors**

Kenneth Chang, Ph.D.
Scott Lyons, Ph.D.
Jonathan Preall, Ph.D.
*Research Associate Professor and Head of Genomics Tech Development*
Johannes Yeh, D.D.S., Ph.D.

**Research Assistant Professors**

Paolo Cifani, Ph.D.
*Research Assistant Professor and Director, CSHL Mass Spectrometry Shared Resource*
Sara Goodwin, Ph.D.
Siran Li, Ph.D.
Michael Ronemus, Ph.D.
Zihua Wang, Ph.D.
Seungtai Yoon, Ph.D.

**CSH Fellow**

Corina Amor Vegas, M.D., Ph.D.

**Clinical Fellows**

Sandeep Nadella, M.D.
Jung-In Yang, M.D.
Kenneth Yu, M.D.

**Adjunct Professors**

Richard Barakat, M.D., M.B.A.
Mikala Egeblad, Ph.D.
Larry Norton, M.D.
Kevin Tracey, M.D.

**Adjunct Associate Professors**

Tatiana Engel, Ph.D.
Jesse Gillis, Ph.D.

**Adjunct Assistant Professors**

Arnon Arazi, Ph.D.
Nyasha Chambwe, Ph.D.
Daniel King, Ph.D.
Andrea Moffit, Ph.D.
Zhen Zhao, M.D., Ph.D.

**QB Visiting Scholar**

Bhubaneswar Mishra, Ph.D.

**Computer Scientists**

Osama El Demerdash, Ph.D.
Vivek Kumar, Ph.D.

**Senior Research Investigators**

Youngkyu Park, Ph.D.
Damianos Skopelitis, Ph.D.

**Research Investigators**

Joan Alexander, M.D.
Pascal Belleau, Ph.D.
Qing Gao, M.D., Ph.D.
Nicholas Gladman, Ph.D.
Gilbert Henry, Jr., Ph.D.
Joshua Homer, Ph.D.
Manzar Hossain, Ph.D.
Yoon-Ha Lee, Ph.D.
Katherine Matho, Ph.D.
Michael Regulski, Ph.D.
Yi-Jun Sheu, Ph.D.
Asya Stepansky, Ph.D.
Claudia Tonelli, Ph.D.
Yixin Zhao, Ph.D.

**Deputy Director of Administration for the Cancer Center**

Lindsey Baker, Ph.D.

**Director, Clinical & Translational Collaborations**

Soma Prum

**Visiting Scientists/ Collaborative Scientists**

Jose Adrover Montemayor, Ph.D.
Lijuan Sun, Ph.D.
Ledong Wan, Ph.D.
Kenny Ye, Ph.D.

**Visiting Professor**

John Boockvar, M.D.

**Visiting Clinical Researchers**

Mali Barbi, M.D.
Emma Gazzara, M.D.
Divya Gowthaman, M.D.
Shruti Koti, M.D.

*Employee of the Howard Hughes Medical Institute
**Employee of the United States Department of Agriculture

**Director, MAPseq Core Facility**

Huiqing Zhan, Ph.D.

**NeuroAI Scholars**

Kyle Daruwalla, Ph.D.
Christian Gerno Pehle, Ph.D.

**Postdoctoral Fellows**

Nolwenn Adam, Ph.D.
Hudson Alakonya, Ph.D.
Armend Axhemi, Ph.D.
Balasooriya Balasooriya, Ph.D.
Amitava Banerjee, Ph.D.
Walter Bast, Ph.D.
Alexandria Battison, Ph.D.
Ari Benjamin, Ph.D.
Debmalya Bhunia Ph.D.
Francesco Boato, Ph.D.
Janeen Braynen, M.D.
Jonathan Cahn*, Ph.D.
Giuseppina Caligiuri, Ph.D.
Mackenzie Callaway, Ph.D.
Deeptiman Chatterjee, Ph.D.
Tsung Han Chou, Ph.D.
Alessandro Crnjar, Ph.D.
Katherine Day, Ph.D.
Benjamin de la Cruz Thea, Ph.D.
Cristiane de Santis Alves Rosa*, Ph.D.
John Desmarais, Ph.D.
Mary Doherty, Ph.D.
Durgesh Dubey, Ph.D.
Dimitri Dumontier, Ph.D.
Max Epstein, Ph.D.
Lesley Ferguson, Ph.D.
Miriam Ferrer Gonzalez, Ph.D.
Austin Ferro, Ph.D.
Matthew Fisher, Ph.D.
Minakshi Gandhi, Ph.D.
Yuan Gao, Ph.D.
Ankur Garg*, Ph.D.
Aybuke Garipcan, Ph.D.
Shanu George, Ph.D.
Bryan Gitschlag, Ph.D.
Wuqiang Guan, Ph.D.
Guangran Guo, Ph.D.
Priyanka Gupta, Ph.D.
Amber Habowski, Ph.D.
Clifford Harpole, Ph.D.
Alexander Harris, Ph.D.
Jia He*, Ph.D.
Anat Hendelman*, Ph.D.
Diego Hernandez Trejo, Ph.D.
GaRam Hwang, Ph.D.
Yuma Ishigami, Ph.D.
Nandhini Kalavathi Palanisamy, Ph.D.
Rishvanth Kaliappan Prabakar, Ph.D.
Hyunook Kang, Ph.D.
Der-Shyang Kao, Ph.D.
Jonathan Kastan, Ph.D.
Stan Kerstjens, Ph.D.
Olaf Klingbeil, Ph.D.
Shan Kuang, Ph.D.
Wenjun Lan, Ph.D.
Amy Lanctot*, Ph.D.
Viet Hang Le, Ph.D.
Eva Lentsch, Ph.D.
Bo Li, Ph.D.
Yujia Li, Ph.D.
Qianyu Lin, Ph.D.
Yong Lin, Ph.D.
Penelope Lindsay, Ph.D.
Bodu Liu, Ph.D.
Mingzhe Liu, Ph.D.
Kaiser Loell, Ph.D.
Jason Lynn*, Ph.D.
Vinay Mandati, Ph.D.
Carlos Marti Gomez Aldaravi, Ph.D.
Cyrille Mascart, Ph.D.
Cristian Mateo Elizalde, Ph.D.
Kevin Michalski, Ph.D.
Sushanta Mishra, Ph.D.
Sanjay Naik, Ph.D.
Erik Eidy Nakagaki Silva, Ph.D.
Jeremy Nigri, Ph.D.
Kin On*, Ph.D.
Huiwu Ouyang, Ph.D.
Sujay Pal, Ph.D.
Shirsa Palit, Ph.D.
Soumyaranjan Pati, Ph.D.
Om Prakash Chouhan*, Ph.D.
Yijian Qui, Ph.D.
Nilesh Rai, Ph.D.
Samrat Rakshit, Ph.D.
Nissim Ranade, Ph.D.
Ashlan Reid, Ph.D.
Yunus Sahin, Ph.D.
Irene Sanchez Martin, Ph.D.
Miguel Santo Domingo Martinez, Ph.D.
Anirban Sarkar, Ph.D.
James Satterlee, Ph.D.
Armin Scheben, Ph.D.
Evan Seitz, Ph.D.
Helene Sertznig, Ph.D.
Mojdeh Shakiba, Ph.D.
Nandhakumar Shanmugaraj, Ph.D.
Yang Shen, Ph.D.
Hagai Shohat, Ph.D.
Sergey Shuvaev, Ph.D.
Cristian Soitu, Ph.D.
Ruben Steigerwald, Ph.D.
Santhilal Subhash, Ph.D.
Mengyi Sun, Ph.D.
Qingtao Sun, Ph.D.
Shoujun Sun, Ph.D.
Simon Sun, Ph.D.
Xueqin Sun, Ph.D.
Kyle Swentowsky, Ph.D.
Johanna Syrjanen, Ph.D.
Debora Tenenbaum, Ph.D.
Hsiu-Chi Ting, Ph.D.
Ankit Tiwari, Ph.D.
Thu Tran, Ph.D.
Vijina Varapparambath, Ph.D.
Prabhadevi Venkataramani, Ph.D.
Filip Vercruysse, Ph.D.
Dharmendra Vishwakarma, Ph.D.
Dominic Vita, Ph.D.
Athanasios Vouzas, Ph.D.
Yuanting Wang, Ph.D.
Zifei Wang, Ph.D.
Peipei Wu, Ph.D.
Yue Wu, Ph.D.
Alexander Xue, Ph.D.
Tao Yang, Ph.D.
Wen-Hsuan Yang, Ph.D.
Toyoki Yoshimoto, Ph.D., M.D.
Li Yuan, Ph.D.
Sophia Zebell, Ph.D.
Xiang Zhao, Ph.D.
Jessica Zhou, Ph.D.
Imanol Zubiete Franco, Ph.D.
Chao Zhang, Ph.D.

**Graduate Students**

Dexter Adams
Disha Aggarwal
Mia Lin Amato
Dhivyaa Anandan
Javier Anduaga
Batuhan Baserdem, Ph.D.
Jack Bauer
Leah Braviner
Paul Bunk
Salome Carcy
Yuxin Cen
Shirley Chan
King Hei Cheng
Charlie Chung
Danielle Ciren
Emma Courtney
Patrick Cunniff
Todor Cvetanovic
Zarmeena Dawood
Jed de Ruiter-Swain
Shivani Deshpande
Marie Dussauze
Fatima Ejaz
Onur Eskiocak
Santiago Espinosa
Connor Fitzpatrick
Fernanda Renee Garcia Flores
Iacopo Gentile
Matias Gleason
Lijie Han, M.D.
Xiao Han
Kimberly Hane
Samantha Henry
Diego Hernandez
Nikolas Holland
Qingting Hu
Qiyao Huang
Sessen Daniel Iohannes
Emily Isko
Shruti Iyer, Ph.D.
Hoda Jaber Ansari
Yujia Ji
Lu Jia
Robert Johnson
Jakub Kaczmarzyk
Jessica Kahng
Yijie Kang
Vahag Kechejian
Seung Tea Kim
Sun Kim
Sam Kleeman
Alexander Kral
Tan-Chun Kuo
Nemanja Kutlesic
Steven Lewis
Heng Liang
Lingjie Liu
Luiz Carlos Machado de Oliviera, Jr.
Dennis Maharjan
Timothy Maher
Ziyi Mo, Ph.D.
Philip Moresco
Manojit Mosur Swamynathan
Masayuki Nagai
Yuriko Nishino
Emmanuella Nnuji-John
Irene Nozal Martin
Satwik Vasant Pasani, M.D.
Michael Passalacqua
Matthew Peacey
Jordan Pearson
Rachel Polfer
Zhe Qian

*Employee of the Howard Hughes Medical Institute

Yihan Qin
Chandana Rajesh
Leonardo Jared Ramirez Sanchez
Kaeli Rizzo
Khristina Samoilova
Danilo Segovia
Liam Shanley
Kristina Shaw
Longling Shui
Nicole Sivetz
Stephen Staklinski
Joshua Steinberg
Maha Syed
Katie Tam
Kamil Taneja
Samantha Tang
Ziqi Tang
Lucia Téllez Pérez
Tuba Thakir
Yaman Thapa
Shushan Toneyan
Khue Tran
Isabella Valentino
Daniella van de Lisdonk
Dillon Voss
Alice Wang
Patrick Wehrle
Yuqianxun Wu
Yunxin Xie
Wenbo Xu
Lucia Yang
Zhezhen Yu
Narges Zali
Hao Zhang
Jialin Zhang
Yuhan Zhang
Chris Zhao
Xiaoyue Zheng
Xingyu Zheng
Zeru Zhu

### PREP Scholars

Theresa Clark
Daniel DiMartino
Pretty Garcia
Melissa Lozada
Laura Lynn
Germaine Smart-Marshall

### Computational Scientists

Samik Banerjee
Sarah Chapin
Kapeel Chougule
Astrid Deschenes
Evan Ernst*
Rabia Göndür
Rebecca Hassett
Juliene Hinds
John Hover
Vasilisa Kovaleva
Melissa Kramer
Sunita Kumari, Ph.D.
Xu Li
Zhenyuan Lu
Steven Marks
Christopher Mezias, Ph.D.
Terezija Miskic, Ph.D.
Ahmed Mohamed
Andrew Olson
Claire Regan
Jonathan Skaza
Marcela Tello-Ruiz, Ph.D.
Marygrace Trousdell
Raditya Utama, Ph.D.
Huaien Wang, M.D.
Xuehong Wei

### Biostatistician

Taehoon Ha

### Chemistry Data Analyst

Adam Moorhouse

### Research Associates

Kenneth Addison
Maria Antonelli, Ph.D.
Carmelita Bautista
Yon Chang
Hrishikesh Deshpande
Jorg Drenkow
Adrian Gomez, Ph.D.
Song Han
Sanjeev Kaushalya, Ph.D.
Hyun Soo Kim*, Ph.D.
Junyi Luo
Lisa Manche
Joseph Merrill
Joseph Milazzo
Diana Ravens
Suzanne Russo
Caizhi Wu, M.D.
Lifang Zhang, Ph.D.

### Scientific Project Managers

Katherine Brenner
Deepthi Poornim Budagavi, Ph.D.

### Laboratory Facility Managers

Sabrina Boettcher
Amy Brady
Julie Cheong
Robert Eifert
Elena Ghiban
Stephanie Goldsmith*
Kelly Ann Holfester
Pamela Moody
Timothy Mulligan
Eleni Noutsos, Ph.D.
Hardik Patel
Jeremy Patino
Gina Robitalle*
Joseph Simorowski*
Dennis Thomas, Ph.D.
Ming Wang
Tse-Luen Wee, Ph.D.

### Scientific IT Support

John Pane
David Trimboli
Peter Van Buren

### Scientific Administrators

Susan Anderson
Lorraine Baldwin
Christine Bedell
Patricia Bird
Samantha Bowen
Alexandra Boyle
Barbara Cascone
Melissa Daley
Becky Dong
Suzan Downey
Tricia Forys
Susan Fredericks
Dawn Kightley-Sutter
Lisa Kimoto
Caryn Koza
Kelly Lewis
Antonia Little
Theresa Morales
Jessica Peluso
Lauren Richter
Catherine Scalise
Kara Spardy Sweeney
Judith Weigold
Madeline Wisnewski

### Research Scientist

Meng Ouyang, M.D., Ph.D.

### Research Technicians

Maxwell Abramson
Aisha Ademola
Zakeria Aminzada
Kristin Anderson
Andalus Ayaz
Samantha Blau
Regina Borello
Leah Boyd
Daniel Bradford
Denise Cahn
Jingyang Cai
Darren Chen
Eileen Cheng
Lisa Christensen
Michael Ciccone
Anthony Cocozzelli
Devon Cowan
Emma Davidson
Martin Davis
Ethan Ertel
Audrey Fahey
Brendan Farhi*
Brian Farrell
Phoebe Fechtmeyer
Christy Felice
Michael Fiore, Jr.
Patrick Flannery
Victoria Gaeth
Renae Galluccio
Pau Garcia Baucells
Joseph Gerwolb
Diya Ghosh
Alicia Gonzalez
Camila Gonzalez
Antoine Gruet
Inessa Hakker
Kaarina Hanington
Amanda Jensen
Chris Kang
Suyun Kim
Rebecca Koelin
Fatim Kouassi
Veronika Kulik
Breanna Leasure
Samuel Liebman
Ying Hsiu Liu
Brianna Lodato
Maulik Masaliya
Senem Mavruk Eskipehlivan
Kristin Milicich
Francesca Minicozzi
Patrick Morris
Stephanie Muller
Diana Naglic
Payal Naik
Cassandra Nash
Larissa Nogueira de Almeida
Joseph O'Rourke
Justin Park*
Nikita Persaud
Sadia Rahman
Umamaheswari Ramu
Max Richman
Michael Riggs

*Employee of the Howard Hughes Medical Institute

Julie Rosenbaum
Samantha Sanchez
Ava Sann
Adrian Santiago
Stephen Savoia
Vanessa Schoen
Brooke Seman*
Marina Sherman
Noriko Simorowski
Tara Skopelitis
Olivia Stamatatos
Danielle Stauder
Joshua Torres
Caitlin Tsang
Uma Vrudhula
Jing Wang
Robert Wappel
Jenna Wilken
Yi-Chen Wu
Shih Ting Yang
Evan Zhang

### Laboratory Facility Staff

Rachel Rubino, D.V.M.
*Director of Lab Animal Resources/Attending Veterinarian*
Doris Abrego
Alexander Aguilera
Katherine Atehortua
Valerie Bagan
Carlos Ballon
Lisa Bianco
Amanda Bjertnes
Oral Bryan
Amy Burrows
Manuel Canas
Steven Collins
Charlene De Poto
Wisline Delphin
Maria Demott
John Donnelly
Rorke Dunne
Eileen Earl
Sara Fredriksen
Gloria Garcia
Pamela Guerra
Susan Guterman
Jill Habel
Oscar Hernandez Ulloa
Juan Jaramillo
Joseph Kelly
Michael Labarbera
Magdalena Lima Canizalez
Dylan Marcus
Maria Mosquera
Gustavo Munoz
Angel Oliveira
Bryan Rose
Sherry Salgado
Matthew Soethout
Kaarina Stearns
Francisco Velasquez
Walter Vigil Garcia

### Media Makers/Farm Assistants

Nancy Bolanos
Terrance Chisum
Blaine Fitzgerald*
Peter Henneberry
Adriana Hincapie
Shouyun Qiao
Amy Qiu
Kyle Schlecht

## EDUCATION

### Banbury Center

Rebecca Leshan, Ph.D.
*Director*
Alicia Franco
Jenna Jacobs
Duncan Yates

### CSHL Press

John R. Inglis, Ph.D.
*Executive Director*
Barbara Acosta
Kevin John Black
Cynthia Blaut
Tara Bonet-Black
Carol Brown
Kathleen Bubbeo
Kathleen Cirone
Laureen Connell, Ph.D.
Marie Cotter
Samantha Cyrill, Ph.D.
Jennifer DeLeon, Ph.D.
Ahmet Denli, Ph.D.
Corrisa Farmer
Olaya Fernandez Gayol, Ph.D.
Bibiane Garite
Sanchari Ghosh, Ph.D.
Danett Gil
Michael Henigman
Samantha Hindle, Ph.D.
Deborah Jarski
Brandon Kelly
Katherine Kelly
Tara Kulesa
Wayne Manos
Joanne McFadden
Alejandro Montenegro Montero, Ph.D.
Mary Mulligan
Christin Munkittrick, Ph.D.
Marisol Munoz
Stephen Nussbaum
Dorothy Oddo
Jacqueline Picone
Jennifer Quereau
Robert Redmond
Theodore Roeder
Mary Sabala
Eric Sawey, Ph.D.
Diane Schubach
Richard Sever, Ph.D.
Marcie Siconolfi
Joshua Sinanan
Gerard Sputo
Hillary Sussman, Ph.D.
Denise Weiss
Dinar Yunusov, Ph.D.

### DNA Learning Center

David A. Micklos
*Executive Director*
Genesis Acevedo
Kelsie Anson, Ph.D.
Elizabeth Asaro
Allison Astudillo
Keith Bannerman
Milany Bruno
Shreemattie Budhram
Elna Carrasco
Kian Charkhian
Lauren Correri
Genevieve Decker
Derbie Desir
Julianna Dovi
Kelly Eames
Anna Feitzinger
Cristina Fernandez Marco, Ph.D
Tess Fleming
Daniel Galvin Gusmano
Charlotte Gordon
Alexander Gottlieb
Sheilaya Gresham
Jennifer Hackett, Ph.D.
Carol Henger
Cristofer Hernandez
Min Hur
Daniel Jacobs
Ria Jasuja
Brittany Johnson
Arie Kazmierczak
Ryan Koenigsberger
Vanshika Kohli
Taylor Korn
Ashley LaSalle
Susan Lauter
Nicholas Liotta
Kate Lin
Sandhya LoGalbo
Marc Louis
Francesca Mango
Annie Marcus
Aaron Mathew
Fernanda Martinez
Sebastian Maurice
Ella Mayers
Allison Mayle, Ph.D.
Amanda McBrien
Lola Milanese
Harold Miller
Christina Newkirk
Ava Orellana
Jeffry Petracca
Jamie Price
Jakob Rechtweg
Tifanny Rushford
Jesus Salas
Julia Salatti
Jens Regleigh Salva
Morgan Serbagi
Donna Smith
Ian Smith
Croi Spillane
Michael Stabile
Faith Tsentner
Jason Williams
Chun-hua Yang

### Library and Archives

Ludmila T. Pollock
*Executive Director*
Paula Abisognio
Thomas Adams
Jannette D'Esposito
Jacqueline Gunther, Ph.D.
Tricia Loria
Kathleen McGuire
Elizabeth Pessala, JD
Katharine Pigliacelli
Stephanie Satalino
Gail Sherman
Antoinette Sutto, Ph.D.
Clare Tonks, Ph.D.

### Meetings and Courses

David J. Stewart, Ph.D.
*Executive Director*
Alfred Baman
Edward Campodonico

*Employee of the Howard Hughes Medical Institute

Meredith Cassuto
Katiria Ceballos
Maoyen Chi, Ph.D.
Wendy Crowley
Matthew Denninger
William Dickerson
Robert Eifert III
Bradley Frey
Constance Hallaran
Coleen Jackman
Patricia Maroney
Kelley McGrath
Andrew Mendelsohn
Kimberly Mistretta
Olivia Mulligan
Kenneth Orff
Valerie Pakaluk
Jessica Palma
Jonathan Parsons
Natasha Reyes
Ira Russo
Mary Smith
Erica Stang
Caitlin Stellin
Christine Vocson
James Whitaker
Barbara Zane

### Office of the President

Terri I. Grodzicker, Ph.D.
*Professor of Biological Sciences/ Dean of Academic Affairs*
Jan A. Witkowski, Ph.D.
*SBS Professor*

### Cold Spring Harbor Laboratory School of Biological Sciences

Zachary Lippman, Ph.D.
*Professor & Director of Graduate Studies*
Brianna Campmier, Ph.D.
Kimberly Creteur
Alexander Gann, Ph.D.
Alyson Kass-Eisler, Ph.D.
Monn Monn Myat, Ph.D.
Catherine Perez

## ADMINISTRATIVE DEPARTMENTS

### Advancement

Charles V. Prizzi
*Senior Vice President, Advancement & Special Advisor to President*
Dominique D'Anna-Stanley
Stephanie Garrett
Sarah Kitt
Karen Orzel
Ilgin Seidner, Ph.D.
Jill Stone

### Business Development & Technology Transfer

Andrew Whiteley
*Vice President, Business Development/Technology Transfer*
Vladimir Drozdoff, JD, Ph.D.
Eric Greenbaum
Elizabeth Hand
Cynthia Miller
Radhakrishnan Narayanan, Ph.D.
Elizabeth Panagot

### Central Administration

Kimberly James
Carol Link
Irina Miretskiy

### Communications

Dagnia Zeidlickis
*Vice President, Communications*
Margot Bennett
Caroline Cosgrove
Samuel Diamond
Nicholas Fiore
Sara Giarnieri
Marc Persad
Philip Renna
Susan Runkowski
Luis Sandoval, Ph.D.
Brianne Seviroli
Michael Skuthan
Nicholas Wurm
Susan Weil-Kazzaz

### Culinary Services

James Hope
*Director*
Alan Aguirre
Jeannette Amato
Ingrid Amaya
Jessica Avila Hernandez
Francis Bowdren
Maryam Brinkwart
Elmer Canales
Maria Del Pilar Carmona
William Carmona
Francis Cayamcela
Marvin Chavez
Melissa Davis
Erica Douglas
Saira Estrella
Bartola Garcia
Luis Garcia
Martir Garcia Moreno
Andres Gonzalez
Silvia Gonzalez
David Graber
Bacilio Hernandez
Yesenia Herrera
Bartholomew Isaac
Rebecca Lavelli
Juan Magana
Jose Maradiaga Gonzalez
Ann May
Dung Nguyen
Jose Reyes
Susana Roman
Marlenis Romero
Carol Rupp
Katherine Schidlovsky
Lisa Spero
Jose Suarez
Lauren Terilli
Ryan Tojza
Selma Turk
Maria Vilorio
Alexandros Vogiatzis
Dennis Zelaya
Nelson Zepeda
Lidia Zuleta

### Diversity, Equity & Inclusion

Charla Lambert, Ph.D.
*CSHL Diversity, Equity, & Inclusion Officer*
Stephanie Franco

### Facilities

Stephen Monez
*Vice President, Chief Facilities Officer*
Lindsey Ahl
Eliseo Amaya
Noel Arias Garcia
Ariad Beazer
Esther Benitez Chavez
Richard Bryant
Keith Byrne
Joseph Carrieri
Dessie Carter
Reginald Cayo
Daniel Chapman
Rene Chavez Contreras
Delmy Chicas de Zavala
Anthony Ciccarella
Christopher Clark
Matthew Coleman
Robert Collins
Juan Colocho
Drew Comer
Peter Dale
Steven Dale
John Damico
Amanda DeDonato
Michael Dileo
Glen DiMaria
Wesley Dreusike
Paul Edwards
Rafael Estrella
Nicholas Favata
Jose Flores
Cesia Madai Flores De Chavez
Maria Garay
Vito Gassi
Robert Gensel
Michael Gorman
Leo Greene
Philip Grella, Jr.
Philip Grella III
Wilman Gutierrez
Victor Hernandez
Jameel Hodges
Matthew Homburger
Joseph Houser
Matthew Hoyer
Christopher Hubert
Christopher Hughes
Merrill Hughes, Jr.
Louis Hunter
Michael Hutchinson
Jeffrey Klaverweiden
Brandon King
Trevor Knight
Robert Kopetic, Jr.
Paulo Krizanovski
Oscar Lastra
James Lawshe
Joseph Leonard
Zaharia Likoka
Jian Lin
Charles Mandell, Jr.
Gregory Marotta
Patricia McAdams
James McCann
Dennis McGuire
Kevin McHugh
Riley McKenna
Marina Millan
Daniel Miller
Mirian Morales Benitez
Oscar Moreira
Guido Mosquera
Mary Muno
Michael Murphy
Michael Napoleon
Kenneth Nastri
Herbert Neazer
John O'Connor
Theresa Parietti
Wayne Pav

Jose Pena Corvera
Miriam Peralta
Elsy Perez De Lopez
Santos Perez Torres
John Pisciotta
Wilson Ramones
Brandon Reynolds
Jaqueline Reynolds
Jesus Rivera
Alfredo Rivera, Jr.
Justin Romero
Jeffrey Sanchez
Steven Sardone
Bruce Schadler, Jr.
John F. Schindler
John L. Schindler
Teri Schindler
Claudia Schmid
Raymond Stewart
Laura Stonebridge
Michael Szlepcsik
Tracey Anne Taylor
Hock Yew Teo
Kevin Titus
Joseph Uhl
James Van Wie
Fredy Vasquez
Jose Velasquez
Trinidad Velasquez
Wilfredo Velasquez
Benjamin Veneable III
Christopher Venegas
Mirna Vigil
Thomas Walters
Graham Wildt
Daniel Zimmer

### Finance and Accounting

Nicholas B. Milowski
*Chief Financial Officer*
Elise Best
Caitlin Carroll
Damian Desiderio
Maryanne Detweiler
Colleen Eiermann
Stacey Farina
Karen Filasky
Danielle Gamrat
Christine Gentile
Melissa Hicks
Emily Janson
Himani Kapoor
Eileen Lee
Alicia Londono
Mary Mastropaolo
Linda Maurer
Alison McDermott
Denise Miller
Kristine Murphy
Ryan Puglia
Cassandra Schneider
Nicole Solomon

### Human Resources

Katherine G. Raftery
*Vice President, Chief Human Resources Officer*
Andres Alarcon
Wendy Alexander
Jaclyn Cappello
Camille Cava-Belluscio
Patricia Crawford
Jamie Distilli
Evelyn Guillermo
Jessica Heckel
Lesley Inglis
Laura Magri
Laura Moran
Victoria Panebianco
Douglas Ricciardi
Bridget Shanley
Teresa Zagar

### Information Technology

MaryJo Zaborowski
*Vice President, Chief Information Officer*
Alex Barton
Walter Benka III
Paul Braun
David Jonathan Castillo Torres
Jack Demm
Paul Dunn
Frank Falco
Leonid Flaks
Thomas Karow
Thomas Keller
Tyrene LeSane
Matthew Lindsey
Michael Malave
Barbara Peters
Robert Petkus
Vincent Reale
Salvatore Serafino
John Themelis
Oybek Toirov
Kauris Zephyr

### Legal

Debra Arenare, J.D.
*Vice President, General Counsel*
Julie Block-Rosen, J.D.

### Procurement

Michael Marchesiello
*Vice President, Chief Procurement Officer*
Robert Albinson
Peter Anderson
Deborah Aufiero
Christine Bernius
Victoria Bond
Kimberly Bronson
Timothy Cotone
Craig Darcy
Susan De Angelo
Kevin Donohue
Jeffrey DuPree
Wayne Hamilton
Elizabeth Janow
Anne Knoch
Christopher Oravitz
Jorge Ramirez
Daniel Stahl
Maria Villatoro
Leslie Wenzel

### Research Operations

Diane Esposito, Ph.D.
*Director of Research Compliance/ Research Investigator*
Priya Chembakasseril
Shanique Edwards, Ph.D.
Robert Gerdes
Laura Lynn
Joanie O'Connor
Denise Roberts, Ph.D.
Priya Sridevi, Ph.D.
Julie Sutherland

### Sponsored Programs

Walter L. Goldschmidts, Ph.D.
*Vice President, Executive Director of Sponsored Programs*
Renee Cercone-Rilo
Carol DuPree
Samantha Fagone
Kevin Galati
Jill Hemish, Ph.D.
Joanne Janelli
Andrea McCarthy
Cynthia McCormack
Andres Mercado
Dave Neale
Catherine Perdikoylis
Jaclyn Sammut
Jacob Schuster
Caroline Tran
Anthony Walsh
Dawn Winkoff

## RESEARCH STAFF DEPARTURES DURING 2023

### Adjunct Assistant Professor

Jason Sheltzer, Ph.D.

### Adjunct Associate Professor

Michael Schatz, Ph.D.

### Adjunct Professors

Gregory Hannon, Ph.D.
Josh Huang, Ph.D.
Scott Lowe, Ph.D.

### Associate Professor

Molly Hammell, Ph.D.

### Clinical Fellow

Hassal Lee, Ph.D.

### Computational Science

Oliver Tam, Ph.D.

### Postdoctoral Fellows

Cina Aghamohammadi, Ph.D.
Yinan Dong, Ph.D.
Talitha Forcier, Ph.D.
Bahruz Jabiyev, Ph.D.
Christopher Langdon, Ph.D.
Ruchi Lohia, Ph.D.
Sukalp Muzumdar, Ph.D.
Farhad Pashakhanloo, Ph.D.
Yanliang Shi, Ph.D.
Pavel Tolmachev, Ph.D.
Aktan Alpsoy, Ph.D.
Florencia Campetella Mayoral, Ph.D.
Chen Chen, Ph.D.
Taemoon Chung, Ph.D.
Jaya Balan Devasahayam, Ph.D.
Nikita Francis, Ph.D.
Rasmani Hazra, Ph.D.
Xueyan He, Ph.D.
Aleksander Kaplan, Ph.D.
Asad Aziz Lakhani, Ph.D.
Andre Machado Xavier, Ph.D.
Evdokia Michalopoulou, Ph.D.
Rodrigo Munoz Castaneda, Ph.D.
Karthick Natarajan, Ph.D.
Daniele Silva Pereira Rosado, Ph.D.
Dongyan Song, Ph.D.
Laura Taylor, Ph.D.
Anand Vasudevan, Ph.D.
Juanjuan Xie, Ph.D.
Xiaosa Xu, Ph.D.
Wenqiang Zheng, Ph.D.

### Research Assistant

Divyansha
Ece Kilic, Ph.D.

### Research Associates

Charlotte Lee
Regina Shaw
Cristina Veresmortean, Ph.D.
Melody Wu, Ph.D.

**Research Investigators**

Jonathan Ipsaro, Ph.D.
Pramod Kumar, Ph.D.
Apurva Tandon, Ph.D.
Kai Yu, Ph.D.

**Research Professor**

Darryl Pappin, Ph.D.

**Research Scientists**

Martyna Sroka, Ph.D.
Xiaoli Wu, Ph.D.

**Visiting Scientists/ Collaborative Scientists**

Taimour Baslan, Ph.D.
Dmitry Biba
Fernanda Duque Mendoza, Ph.D.
Shujing Li, Ph.D.
Kuan-Ting Lin, Ph.D.
Carlos Rodriguez-Saltos, Ph.D.
Gwen Swinnen, Ph.D.

**Visiting Clinical**

Megan Gorman, M.D.
Alexandra Peyser, M.D.
Oliver Standring, M.D.
Lauren Ursillo, M.D.
Jung-In Yang, M.D.

**Visiting Professor**

Stuart Weisbrod, Ph.D.