



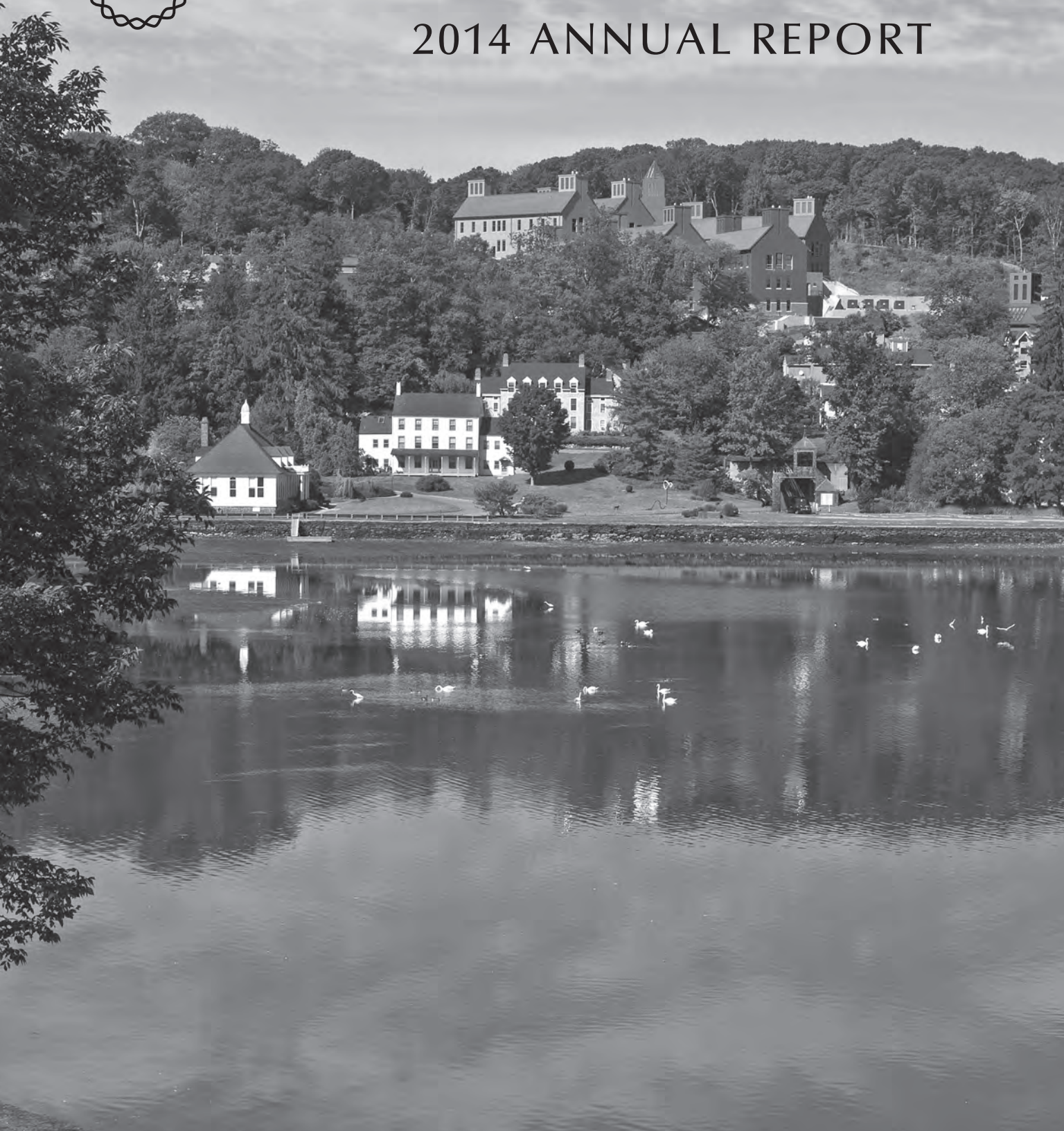
Cold Spring Harbor Laboratory

2014 ANNUAL REPORT



Cold Spring Harbor Laboratory

# 2014 ANNUAL REPORT



**ANNUAL REPORT 2014**

© 2015 by Cold Spring Harbor Laboratory

Cold Spring Harbor Laboratory  
One Bungtown Road  
Cold Spring Harbor, New York 11724  
[www.cshl.edu](http://www.cshl.edu)

**Managing Editors** Philip Renna, Dagnia Zeidlickis  
**Production Editor** Kathleen Bubbeo  
**Science Writing** Peter Tarr  
**Copy Editors** Dorothy Brown, Kathleen Bubbeo  
**Production Manager** Denise Weiss  
**Nonscientific Photography** Constance Brukin, Gina Motisi  
**Cover Designers** Margot Bennett, Denise Weiss

Front cover: *First Leaves of Fall*;  
photography, Philip Renna

# Contents

---

Officers of the Corporation and Board of Trustees	iv–v
CSHL Scientific Advisory Council	vi
Governance	vii
Committees of the Board	viii
David L. Luke III (1923–2014) and Fanny (Curtis) Luke (1923–2014)	ix–x
<b>PRESIDENT’S REPORT</b>	<b>1</b>
Highlights of the Year	5
<b>CHIEF OPERATING OFFICER’S REPORT</b>	<b>19</b>
Long-Term Service	21
<hr/>	
<b>RESEARCH</b>	<b>23</b>
Cancer: Gene Regulation and Cell Proliferation	25
Cancer: Genetics	52
Cancer: Signal Transduction	82
Neuroscience	107
Plant Biology	168
Genomics	192
Quantitative Biology	224
Quantitative Biology Fellow	240
Cold Spring Harbor Laboratory Fellow	244
Author Index	247
<hr/>	
<b>WATSON SCHOOL OF BIOLOGICAL SCIENCES</b>	<b>251</b>
Dean’s Report	253
Spring Curriculum	267
Fall Curriculum	270
Postdoctoral Program	275
Undergraduate Research Program	278
Partners for the Future	281
<hr/>	
<b>MEETINGS &amp; COURSES PROGRAM</b>	<b>283</b>
Academic Affairs	285
Symposium on Quantitative Biology	288
Meetings	291
Postgraduate Courses	350
Seminars	409
<hr/>	
<b>BANBURY CENTER</b>	<b>411</b>
Executive Director’s Report	413
Meetings	419
<hr/>	
<b>DNA LEARNING CENTER</b>	<b>457</b>
Executive Director’s Report	459
Sites of Major Faculty Workshops	474
Workshops, Meetings, Collaborations, and Site Visits	478
<hr/>	
<b>COLD SPRING HARBOR LABORATORY PRESS</b>	<b>485</b>
2014 Press Publications	486
Executive Director’s Report	487
<hr/>	
<b>FINANCE</b>	<b>489</b>
Financial Statements	490
Financial Support of the Laboratory	493
Corporate Sponsor Program for Meetings Support	503
Development	504
<hr/>	
<b>LABORATORY MANAGEMENT</b>	<b>513</b>

# Officers of the Corporation

---

Jamie C. Nicholls, *Chairman*  
Robert D. Lindsay, *Vice-Chairman*  
Marilyn H. Simons, *Vice-Chairman*  
Leo A. Guthart, *Treasurer*

Paul J. Taubman, *Secretary*  
Bruce Stillman, Ph.D., *President*  
W. Dillaway Ayres, Jr., *Chief Operating Officer*

## Board of Trustees

---

### Individual Trustees



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



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



**Jacob Goldfield**  
*New York, New York*



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



**Jeffrey E. Kelter**  
*Senior Partner and CEO, KTR Capital Partners*



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



**Thomas D. Lehrman**  
*CEO, MasterStreet*



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



**Robert W. Lourie, Ph.D.**  
*Head of Future Research, Renaissance Technologies Corp.*



**Howard L. Morgan, Ph.D.**  
*Partner, First Round Capital; Director, Idealab*



**Jamie C. Nicholls**  
*Chairman of the Board, New York, New York*



**Frank O'Keefe**  
*Lloyd Harbor, NY*



**Louise M. Parent**  
*Of Counsel, Cleary Gottlieb Steen & Parent LLP*



**John C. Phelan**  
*Managing Partner, MSD Capital, L.P.*



**Thomas C. Quick**  
*Palm Beach, FL*



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



**George Sard**  
*Chairman and CEO, Sard Verbinnen & Co.*



**Thomas A. Saunders III**  
*CEO, Ivor & Co., LLC*



**Marilyn H. Simons, Ph.D.**  
*President, The Simons Foundation*



**Dinakar Singh**  
*Founding Partner, TPG-Axon Capital*



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



**Paul J. Taubman**  
*PJT Partners, LLC*



**Edward Travagianti**  
*President, TD Bank Long Island*



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

## Individual Trustees Beginning November 2014



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



**David M. Knott**  
Founder,  
Knott Partners



**Elizabeth McCaul**  
Partner-in-Charge, Promontory  
Financial Group, New York;  
CEO, Promontory Europe

## Scientific Trustees



**Tania A. Baker, Ph.D.**  
E.C. Whitehead Professor,  
Massachusetts Institute of  
Technology, Investigator,  
HHMI



**Cornelia Isabella Bargmann**  
Investigator, HHMI; Torsten  
N. Wiesel Professor, The  
Rockefeller University,  
New York



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



**Charles L. Sawyers, M.D.**  
Investigator, HHMI;  
Chairman, Oncology &  
Pathogenesis, Memorial  
Sloan-Kettering Cancer  
Center



**Samuel L. Stanley,  
Jr., M.D.**  
President, Stony Brook  
University



**Bruce Stillman, Ph.D.**  
President, Cold Spring  
Harbor Laboratory

## Honorary Trustees



**Bayard Clarkson, M.D.**  
Memorial Sloan-Kettering  
Cancer Center



**Norris W. Darrell**  
Senior Counsel,  
Sullivan & Cromwell LLP



**Charles F. Dolan**  
Cablevision Systems Corp.



**Helen Ann Dolan**  
Cablevision Systems Corp.



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



**David H. Koch**  
Koch Industries, Inc.



**Mary D. Lindsay**  
Laurel Hollow, New York



**David L. Luke III\***  
New York, New York



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



**Eduardo G. Mestre**  
Vice Chairman,  
Evercore Partners



**William R. Miller**  
New York, New York



**Alan Seligson**  
NAK International



**James D. Watson, Ph.D.**  
Chancellor Emeritus, Cold  
Spring Harbor Laboratory



**Henry Wendt III**  
Friday Harbor, Washington

\*David L. Luke III passed away  
in December of 2014.

# 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 is a nine-member Council, including 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 eight members are world leaders in their respective fields and as such are able to provide advice on the different research areas of the Laboratory.

**Angelika Amon, Ph.D.**

*Massachusetts Institute of Technology*

**Leonid Kruglyak, Ph.D.**

*University of California, Los Angeles*

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

*Harvard University*

**Markus Meister, Ph.D.**

*California Institute of Technology*

**Cornelia Isabella Bargmann, Ph.D.**

*The Rockefeller University*

**Kevan M. Shokat, Ph.D.**

*University of California, San Francisco  
University of California, Berkeley*

**Joanne Chory, Ph.D.**

*The Salk Institute for Biological Studies*

**Max S. Wicha, M.D.**

*University of Michigan*

**Carol W. Greider, Ph.D.**

*Johns Hopkins School of Medicine*

# Governance

---

The Laboratory is governed by a Board of Trustees of up to 35 members that meets three or four times a year. Authority to act for the Board of Trustees between meetings is vested in the Executive Committee of the Board of Trustees. 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 itself is divided between business and community leaders and scientists from major educational and research institutions.

The Laboratory is chartered as an educational and research institution by the Board of Regents 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, Watson School of Biological Sciences" and thereat to confer the degrees of Doctor of Philosophy (Ph.D.), Master of Science (M.S.), and Doctor of Science (Sc.D.), Honorary.

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





# Committees of the Board

---

## Academic Affairs

### James M. Stone, Ph.D., Chair

W. Dillaway Ayres, Jr.  
Tania A. Baker, Ph.D.  
Michael R. Botchan, Ph.D.  
Laurie J. Landeau, V.M.D.  
Thomas D. Lehrman  
Robert Lourie  
Bruce Stillman, Ph.D.

## Audit

### Leo A. Guthart, Chair

Louise M. Parent  
Edward Travagianti

## Commercial Relations

### James M. Stone, Ph.D., Chair

W. Dillaway Ayres, Jr.  
John P. Maroney, Esq.  
Robert Martienssen, Ph.D.  
Bruce Stillman, Ph.D.

## Development

### Marilyn H. Simons, Ph.D., Chair

W. Dillaway Ayres, Jr.  
David Boies  
Charles I. Cogut  
Kristina Perkin Davison  
Jeffrey Kelter  
Laurie J. Landeau, V.M.D.  
Howard Lee Morgan, Ph.D.  
Jamie C. Nicholls  
Louise M. Parent  
George Sard  
Dinakar Singh

Bruce Stillman, Ph.D.

Paul J. Taubman  
Roy J. Zuckerberg

## DNA Learning Center

### Laurie J. Landeau, V.M.D., Chair

W. Dillaway Ayres, Jr.  
Edward A. Chernoff  
Maria de Lesseps  
Suzanne Leeds  
David A. Micklos  
Bruce Stillman, Ph.D.

## Executive

### Jamie C. Nicholls, Chair

W. Dillaway Ayres, Jr.  
Leo A. Guthart  
Louise M. Parent  
Marilyn H. Simons, Ph.D.  
Bruce Stillman, Ph.D.  
James M. Stone, Ph.D.  
Paul J. Taubman  
Edward Travagianti  
Sandra Tytel  
Roy J. Zuckerberg

## Finance

### Leo A. Guthart, Chairman

W. Dillaway Ayres, Jr.  
Jamie C. Nicholls  
John C. Phelan  
William S. Robertson  
Lari C. Russo  
Marilyn H. Simons, Ph.D.  
Edward Travagianti

## Investment

### John C. Phelan, Chair

W. Dillaway Ayres, Jr.  
Thomas D. Lehrman  
Jamie C. Nicholls  
William S. Robertson  
Thomas A. Saunders III  
Dinakar Singh  
Edward Travagianti  
Remy Trafelet

## Nominating

### Louise M. Parent, Chair

W. Dillaway Ayres, Jr.  
Charles I. Cogut  
Leo A. Guthart  
Robert D. Lindsay  
Jamie C. Nicholls  
Thomas A. Saunders III  
Bruce Stillman, Ph.D.  
Paul J. Taubman  
Roy J. Zuckerberg

## Robertson Research Fund

W. Dillaway Ayres, Jr.  
Katherine Ernst  
Walter Goldschmidts  
William Gridley  
Leo A. Guthart  
John H. Linnartz  
Jamie C. Nicholls  
William S. Robertson  
Bruce Stillman, Ph.D.



David L. Luke III  
(1923–2014)

Fanny (Curtis) Luke  
(1923–2014)

In 2014, we lost both David and Fanny Luke. Fanny passed on October 16, and David joined his beloved wife of nearly 60 years soon thereafter on December 13. Their lives began and ended in the same years.

In 1999, the Cold Spring Harbor Laboratory Board of Trustees and then President Jim Watson hosted a “Tribute to the Legacy of David L. Luke III.” David and Fanny had already done so much for the Laboratory. Their support and friendship continued strong into the new millennium. As we mourn their deaths, I write this remembrance of their transformative legacy at this institution.

Fanny graduated from the Westover School in 1941, and in 1943, she joined the U.S. Navy WAVES. She was involved in the final modification of combat aircraft for U.S. Naval Forces in the Pacific Theatre, including the planes her future husband David would fly. She completed her military service in 1946 and then attended Radcliffe College, receiving a B.A. degree in 1951. Fanny worked briefly for Houghton Mifflin and also for Harvard University before moving back to New York. In New York, she took courses at the Bank Street College of Education and then devoted nearly 12 years to teaching young children. From 1972 to 1980 she served on the Board of the Westover School. She was also a member of the Board of the City Gardens Club and served for 2 years as its president.

Having graduated from the Hotchkiss School, David entered Yale University in 1941, but left after Pearl Harbor for naval aviation flight training. He received his naval aviator’s wings and a commission in the Marine Corps in June 1943. He saw combat in the Pacific over the Marianas and Okinawa. After military service, Mr. Luke returned to graduate from Yale in 1948, and then he joined the American Research and Development Corporation. He went to New York City 3 years later to begin a distinguished career with the Westvaco Corporation.

The Luke family established the Westvaco Corporation in 1888, and by the time David joined its corporate financial staff in 1952, the company was thriving not merely as a producer of paper and packaging products but as a developer of new technologies for forestry, tree breeding, and manufacturing. David steadily advanced through the executive ranks to become President in 1962 and also Chief Executive Officer in 1963. He became Chairman in 1980, retiring from that position in 1996.

After purchasing a home in Locust Valley in 1971, David and Fanny almost naturally found themselves at Cold Spring Harbor Laboratory. To David, CSHL’s renowned reputation in plant biology and genetics was practically irresistible. After attending only a few meetings with long-time family friend and soon-to-be CSHL Chairman Barney Clarkson, David joined the Board in 1986 and quickly rose to the position of treasurer. He became Chairman of CSHL’s Board of Trustees in 1992.

When David came to the Board, the Laboratory's plant genetics program was at a turning point as the need for modern infrastructure intersected with advances in genome sequencing technologies. He boldly chaired the Second Century Fund, raising more than \$50 million to meet these challenges at the dawn of the Laboratory's second century in 1990. The Second Century Fund initiative allowed CSHL to build the Beckman Neuroscience Center and powerfully enter the field of neuroscience. These funds also expanded our education programs by allowing for more visiting scientists to attend conferences and courses.

With the success of the Second Century Fund behind him, Chairman Luke was eager to do more. Jim Watson remarked that "David was acutely aware that in order for any organization to be successful in its endeavors—be it corporate or non-profit—it needs to be financially sound." David saw opportunities to refocus the Laboratory's policies on commercial relations and technology transfer. In 1992, he also spearheaded the establishment of a new component of the Laboratory's endowment, the Science Fund. This is a pool into which the proceeds from commercialization of the Laboratory's scientific research would be placed. I know he would be pleased by our current commercialization efforts that are sparking new companies and bringing our research more rapidly to the clinic.

Under David's chairmanship, CSHL acquired 12 acres in Woodbury that would be the home of the Genome Center, the CSHL Press, and a state-of-the-art greenhouse. In 2015, the Preclinical Experimental Therapeutics (PETx) facility will open at Woodbury to support advanced drug testing opportunities as part of the Cancer Therapeutics Initiative.

In 1997, the Board of Trustees made the historic decision to establish a Ph.D.-granting program at CSHL. David and Fanny showed great leadership in this effort, creating a fellowship that fully supports one student each year, in perpetuity. When David stepped down as board chairman in 1998, he became an Honorary Trustee and assumed a new role as chairman of the Capital Campaign for the Watson School of Biological Sciences.

Westvaco Corporation, now MeadWestvaco and led by David's nephew, John A. Luke, Jr., has a significant research program on plant science, and David was always interested in bringing our two research organizations together to think about ways to improve tree growth and yield per acre. David recognized the importance of our basic plant science, and he and Fanny continued to support this research up until their death. David was also fascinated by the telomere research by former CSHL scientist and Nobel laureate Carol Greider, intrigued by the possibility that telomeres would be the secret to longevity. David always took great pride in understanding the research at the Laboratory, asking probing and insightful questions as well as being proud of our accomplishments. David was one of our most active board chairs, and through his tenure on our board, the Laboratory went through significant expansion and a key leadership change.

In recognition of David's outstanding contributions to CSHL, the renovated Power House building that holds administrative offices was named the Luke Building. I enjoy my visits there, always passing by the beautiful portrait of David and Fanny that greets all building visitors.

Since the start of the new millennium, not a year has gone by without notes from David or Fanny accompanied by generous donations to the Laboratory's research and education programs. We were grateful for this continued support that has recognized CSHL events including the Women's Partnership Luncheons, the Double Helix Medals Dinner, and Jim Watson's 80th birthday year, as well as campaigns to support the annual fund and larger initiatives such as the capital campaign for the Hillside Laboratory complex completed in 2009.

David and Fanny made many friends at the Laboratory over the last three decades and we all feel their loss deeply. They left so much of themselves with us personally and to the benefit of our institution that I would like to think they will always be with us. Our deepest gratitude to them and their family who shared them with us.

**Bruce Stillman**

# PRESIDENT'S REPORT

---

In 2015, Cold Spring Harbor Laboratory marks the 125th anniversary of its founding. In 1890, our institutional ancestor, The Biological Laboratory, brought together a small group of biology teachers. A pristine tidal estuary off the Long Island Sound provided these men and women with a living laboratory in which to explore the principles of Darwinian evolution. This was, in retrospect, a wonderful anticipation of the “hands-on” principle of learning by doing that the contemporary DNA Learning Center (DNALC) has spread far and wide. Our institutional DNA also reflects an investment made in the first years of the twentieth century by the Carnegie Institution of Washington. It breathed life into a Field Station for evolutionary studies at Cold Spring Harbor, merging Gregor Mendel’s ideas on inheritable traits with Charles Darwin’s explanation of natural selection. Housed in the structure we now call the Carnegie Building, the Field Station conducted experiments inspired by the re-discovery of Mendel’s research on quantitative traits that he first presented exactly 150 years ago.

I would like to take the occasion of our quasiquicentennial not only to remember our rich history, but to look forward. The contemporary Laboratory is the product of an evolutionary process in which major capabilities have been continuously added—particularly since 1968, the year Jim Watson was named Director. For friends and supporters of the Laboratory, I want to start by reflecting on what I think is most extraordinary about what we are today and then suggest how this informs the path we intend to take in the years ahead, indicating the critical role of the Laboratory’s new Capital Campaign.

I cannot think of a contemporary institution, anywhere in the world, that does what Cold Spring Harbor Laboratory does. It is an institution like no other. We are not only a research institution of the highest caliber, we are also a major innovator in biology education. Dr. Francis Collins, one of the prime movers in the Human Genome Project and for the last 5 years the Director of the National Institutes of Health, had it right when he noted some years ago that “Cold Spring Harbor has placed itself at the center of great science *and* great education.”

The educational part of our mission, a legacy of our earliest days, is indeed singular. Our campus is a crossroads of professional activity in the biomedical and life sciences. That reputation reflects the richness of our Meetings program, which brings nearly 9000 scientists to Bungtown Road each year and another 3000 to our meetings center in Suzhou, China. In the lobby of Grace Auditorium between meeting sessions, you are likely to see distinguished professors mixing with postdoctoral fellows and graduate students, an experience I encountered when, as a second-year graduate student, I was fortunate to speak at the Symposium celebrating the 25th anniversary of the discovery of the double helix structure. There is a kind of intellectual quickening that occurs in these situations that is priceless. The same can be said for what happens among those who participate in the over two dozen professional Courses we offer annually, which collectively serve to advance the entire field’s capabilities in the latest technologies and methods.

The impact of our educational program further reflects the excellence of our “think-tank”-inspired programs at Banbury Center and the wide-ranging efforts of the Cold Spring Harbor Laboratory Press to publish high-impact research journals as well as lab handbooks and monographs on important subjects across the biological sciences. All this activity probably qualifies our educational program as unusual, particularly for a research laboratory. But there are two critical components I have not yet mentioned. One is the DNA Learning Center. Its prodigious efforts over a quarter-century have brought the insight and plain fun of hands-on biological experimentation to hundreds of thousands of Long Island youths, most of them of middle- and high-school age, as well as thousands of their teachers. The DNALC also has educated millions globally, both

at centers established in many countries and through over two dozen educational websites that travel everywhere, instantly and for free, via the Internet.

The other critical component is our bold experiment in graduate education, the Watson School of Biological Sciences. Now in its 16th year, the school has performed as well as we had hoped in the 1990s. Watson School students, along with students from nearby Stony Brook University and nearly 200 postdoctoral fellows, add vibrancy to life on our campus. It is the combined breadth of our educational and basic research activities that makes Cold Spring Harbor Laboratory different. The graduate education program complements the long-standing advanced courses that we teach for scientists who come to learn new techniques and approaches to a field of science.

How do we sustain and expand upon the excellence I have just described? We may be a singular kind of place, but we are facing the same challenges that other institutions face. Never in my 20-year tenure in guiding our science have we been in an environment in which securing operational funds has been more challenging, mostly because of the dramatic reductions in federal funding of science—particularly of fundamental discovery science. In support of both our educational and research missions, therefore, the Laboratory is embarking on a “125th Capital Campaign.” It is one that looks to the future and seeks to secure the means that will enable us to continue playing the wide-ranging, truly unique role that I have described. The fact is that if Cold Spring Harbor is to remain “a Lab like no other,” it will have to significantly add to its endowment. Our goal for the Campaign is to raise at least \$250 million.

Some \$50 million of this sum will enable our extraordinary educational offerings to continue to thrive. About \$13 million will endow student fellowships at the Watson School. These are critical in sustaining one of the school's key innovations. Every student accepted for the program knows that he or she, or their research mentor, is not going to have to worry about securing funds to support their stipend. Our primary goal is to bring the best students here and provide them the freedom to blossom as future leaders in the field. The Meetings and Courses Program will also benefit from an initial endowment because the pressures on federal funding no longer make it obvious that we should receive support for training scientists. As we note each year in this report, the excellence of our program has critically depended on corporate donations. Those funds have declined, in part because of business consolidation among companies but also because of diminished corporate and federal investments in education—hence the need for private support to sustain these highly influential courses. Another portion of the educational endowment will help provide much-needed funds for the DNA Learning Center, including our plans to build a flagship site in Manhattan that will directly serve students in the nation's largest public school system. I can think of no more important investment in America's future.

The bulk of the Capital Campaign—some \$200 million—will endow research at Cold Spring Harbor Laboratory, the excellence of which is the bedrock of our reputation. Two factors make the success of the Campaign essential. First, the cost of doing research has skyrocketed since the dawn of the genome era. Genome science is powerful, but costly to do—and those costs have risen steadily at a rate significantly exceeding that of inflation. As the pie chart on this page indicates, only about 40% of the Laboratory's research operating budget is now supported by federal grants. Even as the government's contribution has declined over the last decade, a second factor has come into play. Today, the amount of money an investigator can expect to receive from a typical “modular” government grant is frozen at \$250,000. If these grants had kept up with inflation in recent years, they would now amount to \$340,000. Our investigators are receiving less money to perform work that costs much more to do. This is seriously eroding the ability of scientists at CSHL and elsewhere to plan ambitious and innovative programs of research. But we are not throwing our hands up in exasperation. Instead, we are seeking help from nongovernmental sources—private individuals, foundations, enlightened corporations—to endow operations critical for Cold Spring Harbor Laboratory's continuing leadership in discovery science.

In research, the Capital Campaign will endow programs that are a critical part of our ongoing Cancer Therapeutics Initiative (CTI), as well as the President's Fund that will support the most basic science and the infrastructure to enable innovation. The CTI represents our effort to identify targets for better cancer drugs. For decades, CSHL has been a world leader in developing tools and methods to characterize the genetic changes that are among cancer's hallmarks. Our understanding of cancer genetics and the interaction between cancer cells and the surrounding tissue cells now provides us with a realistic chance to reduce certain forms of cancer to at least chronic illnesses, manageable over long periods of time, if not attain true cures. Technologies developed at CSHL have made it possible to compare the genomic properties of cancerous versus healthy cells; to model human cancers with remarkable fidelity in mice; and to shut off genes one by one in animal models to dissect genetic and biochemical pathways that drive cancers. We are now at an inflection point. Our next step—and one of the key goals of the Capital Campaign—is to integrate cancer genetics with cancer cell metabolism and whole-organism physiology.

Our new thrust recognizes the importance of thinking about human health in organismal terms. By this I mean understanding disease—cancer, but other diseases as well—not just in terms of what goes wrong in specific cells or tissues, but in terms of the functioning of the entire human system. The biology of the whole affects what diseases we get. One example is the immune system. We are learning, through the work of faculty members including Mikala Egeblad, Douglas Fearon, and David Tuveson, about how stromal tissue—the environment in which cancers develop and thrive—interacts with tumors to affect their course. Nicholas Tonks is exploring the relationship between metabolism and cancer. His recent work provides insights into linkages between the biology of nutrition and disease. Obesity is a state that has broad impacts on the human system; it plays a role in causing diseases, including cancer. Aging affects the entire human system; its processes impact disease susceptibility. Our scientists now have clues that other changes over the life span, for instance, pregnancy, appear to alter disease susceptibility, in this case for breast cancer. What causes ductal-cell precursors in breast tissue to be differentially responsive to oncogenesis—but only if a woman has had a baby before a certain age? How is this effect sustained over many years?

Funds raised through the Capital Campaign will support our exploration of these kinds of questions. We will renovate the Demerec Laboratory, a research building dating from the 1950s where four Nobel laureates have worked, to provide state-of-the-art facilities supporting a Center for Therapeutics Research. As part of this new initiative, we will integrate studies on cancer metabolism and nutrition into CSHL's already strong cancer program. Building on our ability to examine cancer cells and tissue at the single-cell level and employing advanced genomics developed here, the new core facilities will enable our scientists to identify metabolic and oxidative vulnerabilities of cancers that have heretofore remained resistant to therapy.

The Capital Campaign will also add to the capabilities of our Preclinical Experimental Therapeutics (PET<sup>x</sup>) facility, which will open on our Woodbury campus in 2015. We hope that PET<sup>x</sup> will enable Cold Spring Harbor Laboratory to deliver well-validated drug candidates to industry. The PET<sup>x</sup> facility encompasses two core facilities. One enables us to conduct drug testing in animal models for cancer that recapitulate cancer in patients. In particular, we look forward to testing novel drug combinations in these models, which is hard to do in human trials. The other core function in the PET<sup>x</sup> facility is advanced imaging. Our scientists will be able to look in real time at the progression of different cancers, particularly cancers being treated with experimental therapeutics. We want to assess not only whether particular drugs are working in our models, but also how they are working. About 40 years of fundamental cancer research has brought us to the stage where rational approaches to cancer therapy are now possible. I look forward to the day, in the not-too-distant future, when our neuroscientists will likewise apply their knowledge of how the brain works to use the PET<sup>x</sup> facility to advance treatments of the numerous neurological and psychiatric disorders.

Cold Spring Harbor Laboratory has had a very important impact on science over its 125-year history. The mix of basic research and education comes together in a model that is well proven. The Capital Campaign will ensure that the Laboratory continues to be an innovator and a world leader in fundamental discovery science—which is critical to our future, even in times of restricted federal support of such science. I hope you will support it.

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

# Highlights of the Year

---

## Research

Cold Spring Harbor Laboratory is a world-renowned center of research in Cancer, Neuroscience, Plant Genetics, Quantitative Biology, and Genomics. The research environment is highly collaborative. Scientists at the Laboratory work together, frequently across disciplines, to solve biology's most challenging problems. One goal is to apply this research on basic biological mechanisms to improve the diagnosis and treatment of cancer, neurological disorders, and other diseases; another is to provide insights that will speed the development of food and energy supplies of the future. The scope of the faculty's interests is suggested in this summary of several of the past year's important findings.

### *New View of NMDA Receptor Will Guide Drug Developers*

Associate Professor Hiro Furukawa's team this year obtained a full-length view of an important brain cell receptor that previously had been mapped only in segments. The receptor is called the *N*-methyl-D-aspartate (NMDA) receptor, and it is important for several reasons. When the receptor is functioning properly, it is essential in learning and memory. But both overactivity and underactivity of the receptor have been linked with a wide range of disorders, including depression, schizophrenia, and Alzheimer's and Parkinson's diseases. The NMDA receptor is a popular target for candidate drugs to treat these problems. But it is a large and complex multisubunit protein, and, until now, its full-length structure had never been imaged at the atomic level. Early in the year, Furukawa's team provided new detail of how one of the receptor's subunits interacts with various compounds known to inhibit the receptor's activity. These were firsts and will aid drug development. Then later in the year, the team published the first view of all the subunits operating together. It identified numerous interactions among the subunits and offered new insight into how the whole is regulated. In rough terms, the receptor looks like a hot air balloon. The "basket" sits in the membrane of the neuron and provides a passageway for entering and exiting ions. The large "balloon" section sits outside the membrane. One domain in the "balloon" portion called the amino-terminal domain directly regulates the "basket," or ion channel. This function distinguishes the NMDA receptor from other receptors that are found in excitatory neurons throughout the brain. It is an important piece of information that will guide the design of new drugs to specifically and selectively regulate the receptor.



H. Furukawa

### *An Orchestrator of Developmental Timing*

Development has often been likened to a symphony—a complex series of events that unfolds in time and space, culminating in the emergence of a viable organism. This past year, Assistant Professor Chris Hammell and his team identified the gene *lin-42* as a key regulator of developmental timing. Hammell was working in *Caenorhabditis elegans*, the roundworm that is the subject of so many basic science investigations. But *lin-42* is found in creatures of all kinds, throughout the tree of life. Roundworms are very simple creatures, and thus make perfect models for studying not only what genes do at critical points but also when they are active and for how long. Hammell's team likens *lin-42* to the genes in people that control circadian rhythms. In the simple worm, *lin-42* orchestrates key events during development by regulating the activity of literally hundreds of other genes. Levels of the protein encoded by *lin-42* were observed to oscillate in a characteristic manner, suggesting that it is a kind of developmental clock. It provides the organism with a temporal memory that enables it to "remember" that it has completed one developmental step before moving on to the

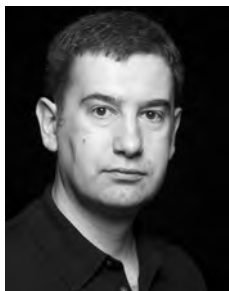


C. Hammell



next. Without such coordination, events would occur out of sequence or overlap, causing serious, irreversible developmental errors.

### *Our Clearest View Yet of Spontaneous Mutations in Autism*



I. Iossifov

CSHL scientists published an important genetic study this year of families affected by autism—2500 “simplex” families in a collection curated by the Simons Foundation. These families have only one child with autism spectrum disorder (ASD); neither siblings nor parents are affected. The study, led by quantitative biologist Ivan Iossifov, a CSHL Assistant Professor, provides our clearest view to date of what is, in its totality, a very complicated picture. The team looked for *de novo* mutations—variations in genes occurring spontaneously in the affected child and not in either parent. Further, they tried to pinpoint which *de novo* mutations might be devastating ones—that is, mutations in genes whose function is so vital in early development that an error might be serious enough to derail normal brain development, giving rise to ASD. Iossifov and colleagues now estimate that at least 30% of all ASD is caused by *de novo* mutations. A related finding was the parsing of this set of mutations into those called “missense” mutations and “likely gene-disrupting” (LGD) mutations. Missense mutations are more common, but less often devastating (1 in 7, on average); LGDs, as their name implies, are often devastating (43%), but much less common. The team, which included Professors Mike Wigler and Dick McCombie as well as faculty members Dan Levy and Michael Schatz, estimated that the total pool of spontaneously mutated genes contributing to ASD numbers around 400, and that these appear to segregate into two risk classes. One portion is found in comparatively high-IQ males; the other in males with low IQ and in nearly all affected females. Females are thought to have a protective factor against the impact of ASD-linked mutations, but when females do develop ASD they tend to have the devastating, gene-disrupting mutations found in this second class.

### *A Way to Objectively Measure Confidence*



A. Kepecs

Associate Professor Adam Kepecs and colleagues carried out a series of experiments this year that demonstrated something extraordinary: that confidence is not just a feeling or emotion, but something that we can measure objectively in the laboratory. The team worked with rats, which were given a difficult choice of discrimination. To obtain a reward, they needed to correctly assess which of two similar paired odors was dominant. Kepecs hypothesized that the amount of time an animal was willing to wait for its reward was directly related to its degree of confidence in the choice it made in a given trial. In fact, the team found that the time a rat was willing to wait enabled them to predict the likelihood of the decision being correct. This is an objective way to measure confidence and should also apply in humans. Kepecs’ team separately used optogenetics to inactivate neurons in a part of the rat brain called the orbitofrontal cortex (OFC), known to be involved in making decisions. They were surprised to find that an inactive OFC did not impair the rats’ ability to make correct decisions. The only thing affected was how long the animals were willing to wait for a reward. Wait time ceased to correlate with confidence; animals often waited a long time even when they had chosen incorrectly. Thus, inactivating the OFC impaired not decision-making but confidence.

### *Elucidating Fear and Depression Mechanisms in the Mouse Brain*

Associate Professor Bo Li and colleagues continued this year to explore pathways in the mammalian brain that are activated in two often disabling conditions: depression and acute fear.

In tracing the fear response, the team identified a group of neurons in the central amygdala that send long-range projections to a part of the brainstem called the midbrain periaqueductal gray (PAG). When they conditioned mice to associate a particular sound with a shock, the team was able to see that activity of these neurons in the central amygdala was enhanced. They had traced a specific pathway used by the brain to translate fear into action. In their work on depression, Li and his team confirmed in mice modeling symptoms of the illness that neurons in the medial prefrontal cortex (mPFC) are abnormally active. Using chemical genetics techniques, they artificially enhanced activity in mPFC neurons in mice that were not only not depressed but better able than most mice to remain unperturbed by stress. These resilient animals quickly became helpless, showing all of the classic signs of depression. This could help account for the often dramatic therapeutic effect of the experimental therapy called deep brain stimulation (DBS), which suppresses the activity of neurons in a small portion of the brain called area 25. Li's work has the potential to make DBS even more targeted and powerful.



B. Li

### *How Gene Expression Patterns Are Carried across Generations*

Cells in our bodies are specialized. Heart cells, blood cells, and liver cells all attain their specificity as they develop out of stem-like precursor cells. Although all our cells carry the full complement of human genes, only certain of these will be expressed as a cell of a given type differentiates. A team led by Professor Rob Martienssen this year showed how even the slightest variations in a class of code-carrying proteins called histones can dramatically alter the way a cell's genes are expressed. Histones come together to form spools around which our DNA is wound, part of the remarkable process through which 6 feet of genetic material is compressed inside the nucleus of each human cell. Histone proteins are marked with chemical tags, such as methyl groups. The histone code consists of the patterns formed by these epigenetic marks across the genome. There are many kinds of histones. One, called H3, comes in two subtypes, H3.1 and H3.3. The former is found only where genes are silent; the latter, only where they are active. Martienssen and colleagues have now figured out why these histones segregate in this manner. Remarkably, they found that a single amino acid difference in the structure of histone H3.3 enables it to serve as a kind of memory device for cells, indicating genes that need to remain active. The amino acid in question is one that cannot be modified with an epigenetic mark. This becomes extremely important when cells divide. Cells need to preserve epigenetic marks that delineate active and inactive genes. Martienssen's work shows how gene-silencing machinery deposits epigenetic marks on histone H3.1 during cell replication but not on H3.3, thus enabling cells to "remember" critical information across generations.



R. Martienssen

### *A Program to Find Disease-Related "Cuts" in the Human Genome*

There are three billion DNA "letters" in the human genome—more than enough to fill a thousand books the length of Melville's *Moby-Dick*. Remarkably, we have learned that adding or removing a single DNA base can have dramatic impacts on biological function. Devastating insertions and deletions—called "indels"—are quite rare, but when they do occur they can affect a person's chances of getting a disease such as autism. Genome sequencing technology has made it possible to spot very large indels spanning thousands of DNA letters, but it has been poor at finding tiny ones—indels of, say, a single letter. Because of the way our DNA encodes proteins—in groups consisting of three letters—single-letter indels can interfere with the production of essential proteins. This year Associate Professor Michael Schatz and colleagues published a computer algorithm



M. Schatz

called SCALPEL that mines existing genome data sets for indel mutations. Schatz compares the effect of small indels with fine cuts in the genome and likens SCALPEL to a means of zeroing in on precisely where these “cuts” occur. With Assistant Professor Gholson Lyon, Schatz used SCALPEL to search for indels in samples from a patient with severe Tourette syndrome and obsessive-compulsive disorder. The team validated more than 1000 indels, demonstrating the accuracy of the method.

### ***Newly Discovered Variant of Tumor-Suppressor Protein May Control Metastasis***



R. Sordella

An intriguing discovery by Associate Professor Raffaella Sordella and colleagues was cited as one of 2014’s “signaling breakthroughs of the year” by the journal *Science Signaling*. The team discovered a new variant of the very well-known protein p53, which they named p53Ψ (the Greek letter “psi”). p53 has often been called “the guardian of the genome.” It is a tumor-suppressing protein that comes to the fore when cells sense damage to their DNA. When the damage is severe, p53 can cause cells to commit suicide, a process called apoptosis. Mutations of p53 are the most common genetic alterations seen in human tumors. Sordella’s team found that p53Ψ, in contrast to normal p53, has pro-growth and even pro-metastatic effects, which it produces by interacting with a protein at the membrane of the mitochondria, the cell’s ubiquitous energy factories. When expressed in cells, p53Ψ reduces the expression of a molecular glue called E-cadherin, which normally keeps cells in contact within epithelial tissue. Such tissue forms the lining of the lung and many other body organs. In fact, Sordella found elevated levels of p53Ψ in samples of early-stage lung tumors with poor prognosis. Normally, p53Ψ appears to play a role in wound healing. It is possible that mutant forms of p53, which play a key role in so many cancers, may have “hijacked” the ability to promote tumor growth and spread from the program used by p53Ψ to promote healing after tissue injury.

### ***Mutant Protein Linked with Intellectual Disability Impairs Synaptic Plasticity***



L. Van Aelst

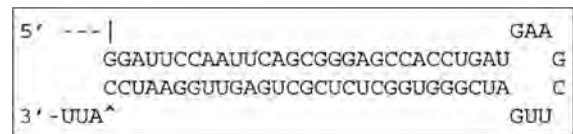
In recent years, Professor Linda Van Aelst has made a series of discoveries revealing the relationship of defects in a gene on the X chromosome called *oligophrenin-1* (*OPHN1*) with various brain dysfunctions contributing to intellectual disability. (It is one form of so-called X-linked intellectual disability.) The OPHN1 protein is understood to help enable excitatory nerve transmission, particularly in cells that have receptors for the neurotransmitter AMPA. This year Van Aelst and colleagues were able to demonstrate how defective forms of the OPHN protein, caused by defects in the gene of the same name, cause impairments in synaptic plasticity. This is the critical process through which our neurons are constantly adjusting the strength of their connections, a key to our ability to both remember and learn. Specifically, Van Aelst found that mutated OPHN1 proteins are not able to properly interact with another protein, called Homer1b/c. When this happens, AMPA “trafficking” is impaired. The receptors, which after they are active normally are brought back inside neurons and pooled for later use, fail to recycle to the surface at synapses when OPHN1 and Homer1b/c are not interacting properly. This impairs neurons’ ability to synapse with one another and may be one of the mechanisms underlying X-linked intellectual disability.

## **Technology Transfer**

Cold Spring Harbor Laboratory (CSHL) is engaged in technology transfer as one way of delivering important discoveries to the public. We continue to ramp up institutional support to

CSHL scientists in partnering with companies, investors, and others to achieve their research-related objectives and add value to the basic research that is the bedrock of science at Cold Spring Harbor. Our success can be measured in terms of dollars, transactions, and relationships, but most importantly how these efforts eventually translate into applications that benefit society. In 2014, CSHL had 561 active patent cases in progress for 133 technologies and 1627 active technology transfer agreements. Net licensing revenue totaled \$3,866,042. The return on commercialization efforts is reinvested in research and innovation at CSHL.

In June, CSHL announced the settlement of a long-running malpractice suit against law firm Ropes & Gray and attorney Matthew P. Vincent just days before the case was to go to a jury in Massachusetts state court. Malpractice impeded the timely issuance of a CSHL patent on short-hairpin RNAs (shRNAs). These are engineered molecules that allow researchers to turn off expression of virtually any target gene or combination of genes in human and other mammalian cells. The shRNA technology at issue was developed by Professor and Howard Hughes Medical Institute Investigator Gregory Hannon, Ph.D. It represents a considerable advance over other technologies, including the RNA interference technology developed by Nobel Prize winners Drs. Andrew Fire and Craig Mello. The Hannon shRNA technology has had a powerful impact across academic and industry medical research, including its use to identify new drug targets and therapies for cancer.



shRNA, named for its hairpin shape

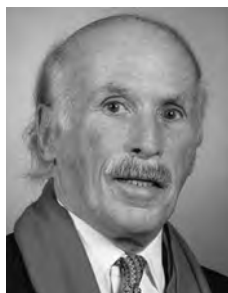
Ultimately patented by CSHL in 2012 while the Laboratory simultaneously pursued the malpractice suit, the proprietary shRNA technology is now a widely used biomedical research tool in cancer research and drug development. Many companies have already signed licenses for this technology with CSHL, and the Laboratory is pursuing additional licensing opportunities.

## Cold Spring Harbor Laboratory Board of Trustees

The CSHL Board of Trustees elects members whose academic and professional accomplishments reach beyond the boundaries of science, providing well-informed governance in an increasingly complex fiscal and regulatory environment. The Board welcomed Lalit Bahl, Ph.D., of Renaissance Technologies Corporation, David M. Knott of Knott Partners, and Elizabeth McCaul of Promontory Financial Group.

The Board congratulated retiring member Thomas Lehrman, who served from 2008 to 2014 and was active on the Planning, Investment, and Academic Affairs committees.

James H. Simons was elected Honorary Trustee. Jim is the founder of Renaissance Technologies and Chairman of the Simons Foundation. He and his wife, Marilyn, who serves as Vice Chairman of the CSHL Board of Trustees, are leaders in private funding of basic research at CSHL and throughout the country.



D. Knott

Trustees, faculty, students, and employees mourned the passing of long-time friends, former CSHL Chairman and Honorary Trustee David Luke and his wife Fanny.

With support from the CSHL Association, the Corporate Advisory Board, foundations, and individuals, CSHL raised more than \$6.6 million for the CSHL Annual Fund, which provides unrestricted support for research and education. The 9th Double Helix Medals Dinner, honoring Matthew Meselson, Marlo Thomas, and former CSHL trustee Andrew Solomon, raised a record \$4 million. Through his award-winning writings and lectures, Andrew has promoted awareness and understanding of mental illness,



L. Bahl



E. McCaul



B. Stillman and Double Helix Medal winners M. Meselson, M. Thomas, and A. Solomon

including a number of disorders that have a genetic basis and are under active investigation at the Laboratory. Dr. Meselson is a legendary geneticist whose landmark experiment in 1958 with Franklin Stahl revealed critical aspects of DNA replication, essentially proving Jim Watson and Francis Crick's hypothesis as to how the double helix of DNA is copied. Marlo Thomas has raised the visibility of research in health care as National Outreach Director for St. Jude Children's Research Hospital, spearheading national campaigns in support of the hospital's mission to advance cures and prevent catastrophic pediatric diseases through research and treatment. In 1962, Marlo's father, Danny Thomas, founded St. Jude Children's Hospital in Memphis, Tennessee.

With generous support from Jim and Marilyn Simons, we established the Simons Center for Quantitative Biology in 2009. The Center brings together experts in applied mathematics, computer science, physics, and engineering to further basic research and investigation into illnesses including cancer, autism, bipolar disorder, and depression. Thanks to Jim and Marilyn, this year we were able to recruit computational biologist Adam Siepel from Cornell University to be the inaugural Chair of the Simons Center.

We are grateful to all major donors who in 2014 supported research and education programs at the Laboratory. The Stanley Medical Research Institute contributed \$10 million to support the Laboratory's Stanley Institute for Cognitive Genomics. The G. Harold and Leila Y. Mathers Foundation continued its support of neuroscience research to map the mouse brain.



Jim and Marilyn Simons

A gift from the Seraph Foundation will help support the Laboratory's new state-of-the-art Genetic Screening Facility, named for John and Edna Davenport. The New York State Empire State Development Corporation awarded CSHL \$1.5 million to equip a Center for Metabolomics, where researchers will assess cell metabolism and oxidative vulnerabilities to develop new small-molecule inhibitors for treating cancers, autism, diabetes, and orphan diseases.

The Lustgarten Foundation contributed \$5 million to support pancreatic cancer researcher Douglas Fearon, M.D., who was recruited from the University of Cambridge, U.K., to continue his work on

harnessing the immune system to fight cancer. This support is in addition to a similar amount that was previously given by the Lustgarten Foundation to support pancreas cancer research in David Tuveson's laboratory.

The year's outstanding additions to our Board, scientific staff, and community of supporters all bode well for a future of discovery and teaching innovation. All help CSHL maintain its leadership position in both research and education in biology and the life sciences.

## Research Faculty

### *Awards*

The National Science Foundation (NSF) awarded Early Concept Grants for Exploratory Research (EAGER) to Assistant Professor Florin Albeanu and Professor Partha Mitra, who are working to develop new technologies that will provide insight into the structure and operation of neural circuits and address how complex behaviors emerge from them. The awards are part of President Barack Obama's multiyear BRAIN Initiative that aims to answer fundamental questions about how the brain works.

The National Science Foundation (NSF) selected Assistant Professor Michael Schatz to receive its prestigious CAREER Award. The Faculty Early Career Development Program is targeted for teacher-scholars who are most effective in integrating research and education. Mike was one of a small group selected from among more than 200 young scientists who were under consideration.

Assistant Professor Molly Hammell was one of seven 2014 Rita Allen Scholars. The award supports promising early-career investigators, and Molly received the Foundation's highest honor: designation as the Milton E. Cassel Scholar, which pays tribute to a long-time president of the Rita Allen Foundation.

Assistant Professor Anne Churchland was named a Pew Scholar in the Biomedical Sciences by the Pew Charitable Trusts. She also received a Klingenstein-Simons Fellowship in the Neurosciences from the Simons Foundation and the Esther A. and Joseph Klingenstein Fund.

Associate Professor Mikala Egeblad won an Era of Hope Scholar Award from the United States Department of Defense (DoD) Breast Cancer Research Program. This prestigious award supports early-career scientists who have demonstrated extraordinary creativity, vision, and leadership potential within the field of breast cancer research. Mikala is one of only two recipients of the award this year.

EMBO, founded 50 years ago as the European Molecular Biology Organization, elected Professor David L. Spector, Ph.D., to its ranks. David, who is CSHL's Director of Research, is one of 50 outstanding researchers in the life sciences newly elected to membership. He was also elected to the American Academy of Arts and Sciences, joining some of the world's most accomplished leaders from academia, business, public affairs, the humanities, and the arts.

The Pershing Square Sohn Cancer Research Alliance honored Lloyd Trotman, Associate Professor, with one of six prizes for Young Investigators in Cancer Research. The prize seeks to fuel innovative scientific discoveries in the search for a cure.

I was honored to be selected as the 2014 recipient of the Herbert Tabor Research Award from the American Society for Biochemistry and Molecular Biology (ASBMB). The award is given for excellence in biological chemistry and molecular biology and for contributions to the community of scientists. It was my pleasure to give the opening lecture at this year's ASBMB Annual meeting in San Diego, CA, in April.

### *New Faculty*

Professor Douglas Fearon, who received his M.D. from Johns Hopkins University School of Medicine in 1968, joined us from the University of Cambridge to further study the interaction



D. Fearon



A. Siepel



J. Lee



J. Tolkuhn

between the immune system and tumors. With support from the Lustgarten Foundation, Doug's goal is to develop an immunotherapy to treat cancer.

Thanks to the Simons Foundation, we welcomed Adam Siepel, Ph.D., Professor and Chair of the Simons Center for Quantitative Biology. Adam was previously Associate Professor at Cornell University, focused on comparative genomics and the development of statistical methods and software tools to identify evolutionarily conserved sequences and the complex processes by which genomes evolve over time. Dr. Siepel attended graduate school in computer science, receiving an M.S. from the University of New Mexico (2001) and a Ph.D. from the University of California, Santa Cruz (2005).

Je H. Lee, with an M.D., Ph.D. from Tufts School of Medicine (2002), is an Assistant Professor, joining CSHL from George Church's laboratory at Harvard Medical School. Je studies how cells sense and remember timing, location, and history and how their surroundings influence their signals with other cells. He also develops imaging and molecular sequencing technologies for tracking genes, molecules, and cells.

A neuroscientist from the University of California, San Francisco, Assistant Professor Jessica Tolkuhn, received her Ph.D. from the University of California, San Diego, in 2006 and is studying how estrogen and testosterone generate sex differences in the brain and behavior.

### **Promotions**

Congratulations to newly promoted Associate Professors Anne Churchland, Mikala Egeblad, and Michael Schatz. Justin Kinney was named Assistant Professor, being promoted from his previous position of CSHL Quantitative Biology Fellow.

### **Education Programs**

#### ***Banbury Center***

More than 500 people took part in 16 by-invitation meetings at the Banbury Center this year. The meetings included two that involved families of people affected by serious illnesses. *Eliminating the Stigma of Mental Illness* was devoted to mobilization of a plan conceived by the foundation BringChange2Mind, spearheaded by the actress Glenn Close, whose sister and nephew have bipolar disorder.

*Rhabdomyosarcoma: A Critical Review* brought together world experts on this common, often fatal, soft-tissue sarcoma that afflicts children. In this meeting, too, urgency to initiate a course of action was spurred by advocacy on the part of affected families. The meeting resulted in the publication of a white paper outlining future research directions and marked the start of the Laboratory's Sarcoma Research Project, funded by local, family-based nonprofit organizations to support targeted research on specific sarcomas.



Eliminating the Stigma of Mental Illness Meeting. Glenn Close models Stigmasucks jacket.



Hands-on biology at the DNALC

Other important Banbury meetings in 2014 included *Connections and Communications in the Brain*; *Defeating Ovarian Cancer*; *The Immune System and Cancer*; *Interpreting Personal Genomes*; *Reactive Oxygen Species in Biology and Cancer*; and *Epigenetics and Agriculture*.

Unusually, in addition to the Rhabdomyosarcoma White Paper, this year's Banbury meetings gave rise to two published papers, one on adolescent mental health (published in *Science*), the other on redefining the privacy of genomic data (published in *PloS Biology*).

### DNA Learning Center

Operations of the DNALC in 2014 brought an additional 19,300 local students on field trips to the Cold Spring Harbor, Lake Success, and Harlem, New York, locations. An additional 9000 students were reached through in-school instruction by DNALC staff. During the summer months, 60 week-long biology and genetics summer camps were held in Cold Spring Harbor and eight other locations in the tristate area, involving 1300 students.

The Urban Barcode Project, in its third year, enjoyed continued success, with 145 students presenting posters and talks at the American Museum of Natural History in the spring. The Urban Barcode Research Program, supported by the Pinkerton Foundation, made possible an extraordinary opportunity for 38 students from the New York area to take part in summer barcoding workshops followed up with 100+ hours of research time spent with mentor-scientists at the Museum, CUNY, Columbia University, and the New York Botanical Garden. A new *Barcode Long Island* program was initiated.

The DNALC's worldwide reach via the Internet was reflected in record-setting numbers: more than 5.2 million visits to 22 DNALC websites; 875,000 views of DNALC YouTube videos; and 668,000 downloads of popular *3D Brain*, *Weed to Wonder*, and *Gene Screen* apps. Total multimedia visits numbered 6.75 million, up 7.7% over the prior year.

### Meetings and Courses Program

In 2014, just less than 7000 scientists attended meetings held on the main campus. Attendance at each meeting ranged from 100 participants to more than 500, and more than half of attendees are either graduate students or



2014 Meetings and Courses



postdoctoral researchers. Nearly one-third of the attendees come from nations other than the United States. The Cold Spring Harbor Asia program, now in its fifth year, included 20 conferences and one summer school in 2014 and attracted more than 3400 participants.

2014 saw the continuation of many successful annual and biennial meetings and the introduction of several new ones: *Avian Model Systems*; *the PARP Family & Friends*; and *Biological Data Science*. The Courses program, meanwhile, drew about 600 trainees, including advanced graduate students, postdoctoral researchers, and faculty. A new offering in 2014 focused on the *Genetics and Neurobiology of Language*, and this meeting will be held biennially.

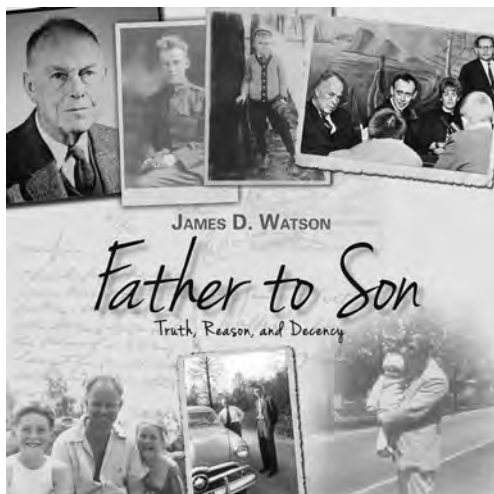
Strong backing of the Howard Hughes Medical Institute continues to sustain the Courses program. Courses also draw essential support from the National Institutes of Health and the National Science Foundation, and they rely on equipment and reagents loaned or donated by many companies. These partnerships enable CSHL to consistently offer training in the latest technologies.

### *Cold Spring Harbor Laboratory Press*

2014 was another successful year for the Press. It maintained an output of high-quality publications, made a substantial financial contribution to the Laboratory, introduced a promising new service, and laid the groundwork for a new research journal. Subscription and advertising revenue streams for its four research journals were stable or higher. *Genes & Development* and *Genome Research* remain two of the world's four most highly cited genetics journals. There was strong growth for the newer journals, *Protocols*, *Perspectives in Biology*, and *Perspectives in Medicine*, each of which represents a strategic online transformation of print manuals and monographs. Total downloads of Press journal articles worldwide reached a record level of 13 million.

Although only a year old, the preprint service bioRxiv made a significant mark. More than 1000 draft research papers in 25 subject categories were posted, and authors received much feedback through on-site comments, social media, and private email. More than 300 of the posted preprints have now been published, in over 100 journals.

The Press also published 21 new book titles in 2014, including Jim Watson's family memoir *Father to Son*.



J. Watson's family memoir.

### *Watson School of Biological Sciences*

The Watson School welcomed its 16th incoming class and graduated its 11th class. The achievements of the graduate program continued to grow. The quality of scientific publications produced by school's students remained highly impressive. Watson School students continued to graduate considerably faster than students in comparable Ph.D.-granting institutions and demonstrated an ability to secure excellent jobs. Nineteen WSBS graduates, or 27% of all graduates, have thus far secured tenure-track faculty positions, an impressive record since the first graduates only left Cold Spring Harbor Laboratory in 2004.

Eleven WSBS students were awarded Ph.D. degrees, bringing the total since the school's inception to 71. We were pleased to award an Honorary Doctor of Science degree to Richard Burgess of the University of Wisconsin, Madison, who taught the protein biochemistry course here for many years and who has made major contributions to understanding gene transcription. During the year, scientific papers published by students of the school appeared in major journals, bringing the cumulative total of publications to nearly 300. Current and former students

won prestigious and highly competitive scholarships and fellowships, as in past years.

In August, the WSBS welcomed nine new students. Members of the Class of 2014 were selected from among 276 applicants and represent the United States, Canada, Germany, India, Lithuania, and the United Kingdom. Other new graduate students entered as visitors from other institutions, including seven from Long Island's Stony Brook University. Current visitors hail from more distant institutions, including Cornell University and several from China, Germany, and France, among others.



WSBS class of 2014

During the summer, undergraduates from across the United States, as well as Austria, Bulgaria, Canada, Greece, Mexico, Singapore, Syria, and the United Kingdom, took up residence at CSHL to take part in the historic Undergraduate Research Program. These “URPs” (chosen from among 819 applicants!) had the remarkable opportunity to perform advanced research in the laboratory of a CSHL faculty member.

### *Library and Archives*

The Library and Archives organized a meeting in Grace Auditorium entitled “The History of mRNA” as part of a series held jointly with the Meetings and Courses division. This meeting brought together pioneers in the field of RNA biology, including Jim Watson and Sydney Brenner—who were co-discoverers of messenger RNA (mRNA)—and those involved in the discovery of long heteronuclear RNA (hnRNA) and RNA splicing, which resulted in a Nobel Prize to former CSHL scientists Richard Roberts and Phillip Sharp. The meeting also included talks by contemporary scientists studying mRNA processing, splicing, and translation. This meeting is part of a series on the history of molecular biology and genetics held here each year.

### **Infrastructure**

#### *Preclinical Experimental Therapeutics Facility (PET<sup>x</sup>)*

Construction on the 8000-square-foot addition to the Woodbury Genome Center that will open in 2015 as the Preclinical Experimental Therapeutics Facility (PET<sup>x</sup>) began in late 2013 and was delayed by a particularly harsh winter. When completed, the facility will house state-of-the-art imaging equipment and a dedicated staff to allow for the discovery and development of novel therapeutics and diagnostics strategies. Equipment will include noninvasive imaging such as PET/CT, ultrasound, and optical imaging equipment. The facility will have two surgical suites, necropsy rooms, a pharmacy, and a diagnostic clinical laboratory.



Preclinical Experimental Therapeutics Facility

### ***Nicholls Biondi Hall***

At the beginning of the year, we broke ground for construction of Nicholls Biondi Hall, a 2000-square-foot poster pavilion space to enhance the Meetings and Courses Program and provide additional configurable space for conferences and seminars. The new building opened in June 2015, looking eastward over a great lawn bounded by the Carnegie Building and the recently constructed



Nicholls Biondi Hall

Hershey Laboratory. The 1904 Carnegie Building that was originally built as a laboratory building and now houses our Library and Archives as a bookend to the new Hall.

The Laboratory this year also dealt with two significant catastrophic system failures. One was the aging, 20-year old McClintock Laboratory chiller plant. Replacement of similarly aging systems have been scheduled beginning in early 2015. Another major event was flooding of the Hershey Laboratory caused by record low temperatures that contributed to the freezing of a fire sprinkler line. Tens of thousands of gallons of water cascaded across the second

floor of the building down to the floor below, causing major damage to the Flow Cytometry and Microscopy facilities and the total loss of a laser confocal microscope. Luckily, the Facilities Department responded immediately and the building was quickly dewatered. Flood damage was repaired and new equipment was procured in less than 60 days.

## **Community Outreach**

### ***Public Lectures***

**March 11**—**Mikala Egeblad, Ph.D.**, Assistant Professor, Cold Spring Harbor Laboratory; **Nicholas K. Tonks, Ph.D., F.R.S.**, Professor, Cold Spring Harbor Laboratory; **Lora R. Weiselberg, M.D.**, Clinical Associate Professor, Hofstra, North Shore–LIJ School of Medicine. Lecture: “Cold Spring Harbor Laboratory Breast Cancer Research Update.”

**March 28**—**Siddhartha Mukherjee**, author, *The Emperor of All Maladies: A Biography of Cancer*, 2014 Roy J. Zuckerberg Lecture.

**April 8**—**Craig Garner, Ph.D.**, Professor, Department of Psychiatry and Behavioral Sciences Stanford University, Co-director Stanford Center for Research and Treatment of Down syndrome. Lecture: “Developing Therapies to Improve Cognitive Abilities of Individuals with Down Syndrome.”

**June 18**—**Michael Schatz, Ph.D.**, Assistant Professor, Simons Center for Quantitative Biology at Cold Spring Harbor Laboratory. Lecture: “Big Data: How Biological Data Science Can Improve Our Health, Foods, and Energy.”



Roy J. Zuckerberg Lecture

**June 24**—**Michael Ronemus, Ph.D.**, Research Assistant Professor Cold Spring Harbor Laboratory; **Rebecca Sachs, Ph.D.**, Staff Psychologist, Fay J. Linder Center for Autism and Developmental Disabilities. Lecture: “Understanding Autism Spectrum Disorder: Focus on the Facts.”

**October 23**—**Bruce Adolphe**, composer and Resident Lecturer and Director of Family Concerts for the Chamber Music Society of Lincoln Center; **Alexis Gambis, Ph.D.**, founder, Science Films. Lecture: “Creativity in Music, Film, and Neuroscience.”



C. Albright

**November 4**—Lynn Pasquerella, Ph.D., President, Mount Holyoke College. Lecture: 2014 Lorraine Grace Lecture on societal issues of biomedical research: “Is Death Un-American?”

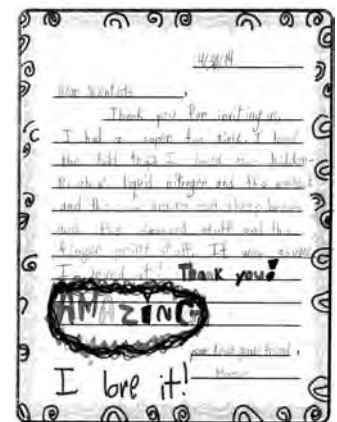
### *Public Concerts*

<b>March 21:</b>	Igor Lovchinsky, piano
<b>April 25:</b>	Cicely Parnas, cello
<b>May 2:</b>	Jerome Lowenthal, piano
<b>May 16:</b>	Ken Noda with Anthony Kalil, piano with tenor
<b>August 15:</b>	Charlie Albright, piano
<b>September 12:</b>	Ji, piano
<b>September 19:</b>	Yun-Chin Zhou, piano
<b>October 2:</b>	Andrey Tchekmazov, cello
<b>October 10:</b>	Einav Yarden, piano
<b>March 27, 2015:</b>	Martin Kasik, Piano
<b>April 17, 2015:</b>	Jiayin Shen and Alan Woo, piano duo
<b>May 1, 2015:</b>	Julia Bullock, soprano
<b>May 15, 2015:</b>	Trio Solisti, violin, cello, and piano

We thank our team of enthusiastic graduate students and postdoctoral fellows who staff CSHL’s public tour program. They conducted 57 campus tours for more than 800 participants during the course of the year.

Local families enjoyed a day exploring the history of Cold Spring Harbor Village and DNA on April 26 as part of the DNA Day Scavenger Hunt. Local institutions including the Cold Spring Harbor Library, the Cold Spring Harbor Whaling Museum, the Firehouse Museum, and CSHL’s DNA Learning Center took part. DNA Day is celebrated across the country, commemorating the completion of the Human Genome Project in April 2003 and the discovery of the DNA double helix by Chancellor Emeritus James D. Watson and Francis Crick.

In April, first graders from Goosehill Primary and Friends Academy participated in a science fair on the CSHL campus. With graduate students and DNA Learning Center instructors manning interactive stations, more than 100 local children, teachers, and parents explored the brain, seaweed, rainbows, and their own fingerprints.



Thank you from student visitor



Hour of Code @ CSHL

In December, Maria Nattestad, a WSBS Graduate Student, led a community event called “Hour of Code.” Part of a global effort to introduce children to computer coding at early ages, the event brought local families to the Laboratory. Maria also leads a “Girls Who Code” club for girls in grades 7–12 together with the DNA Learning Center Assistant Director Amanda McBrien.

Congressman Steve Israel (D-NY) joined CSHL in congratulating Long Island high school students recognized as Intel Science Prize Semifinalists. Two of the 40 local semifinalists were part of CSHL’s Partners for the Future Program, working in the laboratories of Mikala Egeblad and Leemor Joshua-Tor.

CSHL participated in the 11th Long Island 2-Day Walk that raised more than \$22,800 for breast cancer in the lab of Mikala Egeblad, whose research is aimed at developing new strategies to treat breast cancer. Rather than attacking tumor cells directly, Mikala’s group is investigating whether it is possible to target immune cells in the tumor microenvironment that act to promote cancer metastasis and suppress the response to chemotherapy.

The LI2DAY Walk also awards six \$2500 scholarships to Long Island high school seniors who have a parent or guardian with breast or other women’s cancer. CSHL is proud to participate in the scholarship committee.

Interacting with the public through lectures, tours, and special events is important in our efforts to raise visibility about the significance of basic research and science education. The understanding and support of the general public is critical to CSHL and other independent research institutions that rely on taxpayer funding and private philanthropic giving. We continue to expand our efforts to communicate with the public through use of the latest digital technologies. Follow us on Facebook and Twitter!

## Looking Forward

We look forward to our 125th anniversary year and the future beyond it with excitement and optimism, poised for even bigger breakthroughs that will undoubtedly change the world for the better.

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

# CHIEF OPERATING OFFICER'S REPORT

---

In his President's message, Bruce Stillman makes a compelling case for support of Cold Spring Harbor Laboratory's 125th Anniversary Capital Campaign. He cites not only the Laboratory's rich and productive record of achievement in research and education, but the unprecedented promise and opportunity that lies before us in the postgenomic era.

A well-documented decline in federal research funding, combined with dramatic increases in the cost of doing cutting-edge science and recruiting world-class faculty, requires a level of private philanthropic support that we would not have anticipated a decade ago. A primary role of the Chief Operating Officer, in this environment, is to ensure that the institution is worthy of such investment from the perspective of operational and fiscal management. In fact, there are many successes in these areas to which we can point.

Operating expense management has been an ongoing priority for the last decade. Although this has been difficult, at times, for our loyal staff who have endured years of minimal salary increases as well as recent increases in employee contributions to our healthcare plan, such measures have been necessary. These efforts have accomplished their critical objectives of keeping operating expenses relatively flat and conserving cash over the past 5 years, even in the face of absorbing a campus expansion and a substantial rise in the information technology costs associated with data-driven science.

Our balance sheet, with \$800 million in total assets, is strong. We carry a comfortable level of long-term debt (\$97 million), the cost of which is synthetically fixed at a reasonable level for 30 years. The ratings agencies, Moody's and Standard & Poor's, concur, having recently reaffirmed, respectively, our Aa3 and AA ratings. In fact, the Moody's review included a "positive outlook," stating that "it reflects our expectation of continued strong growth in financial resources that outpaces the growth of similarly-rated peers." Moody's added: "Our rating incorporates CSHL's position as a leading not-for-profit research institute, consistently positive operating performance, and healthy and growing financial reserves that provide significant operating flexibility."

In this milestone anniversary year, as we work to meet our financial challenges, we recognize that endowment funds are the key to securing the Laboratory's future as a leading and enduring independent research and education institution. It is imperative that we continue to grow the endowment "organically" through prudent investment management and "inorganically" through vigorous fund-raising. We have had good success in both. Since the original endowment gift of \$8 million was received from the Charles Robertson family in 1975, the total fund has grown to a high-water mark of \$444 million at year-end 2014. The goal of our Investment Committee is to maintain a diversified portfolio designed to achieve healthy risk-adjusted returns while preserving capital and maintaining a reasonable level of liquidity. At the same time, we hold to a disciplined policy of spending no more than 5% of the fund, annually, on a 12-quarter moving average of year-end market value.

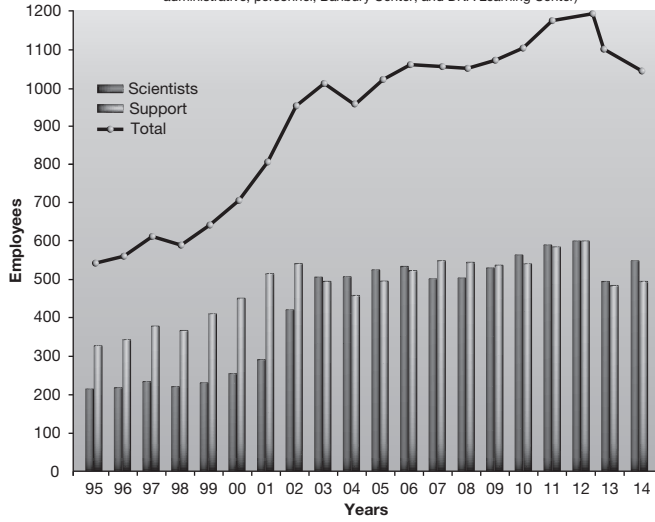
As we rely more heavily, over time, on private philanthropic support to sustain the Laboratory's mission and secure its future, we are profoundly grateful for our generous donors. Cold Spring Harbor Laboratory does important work that we believe is worthy of your investment in us, and we work hard every day to earn and preserve your confidence.

Here's to the next 125 years!

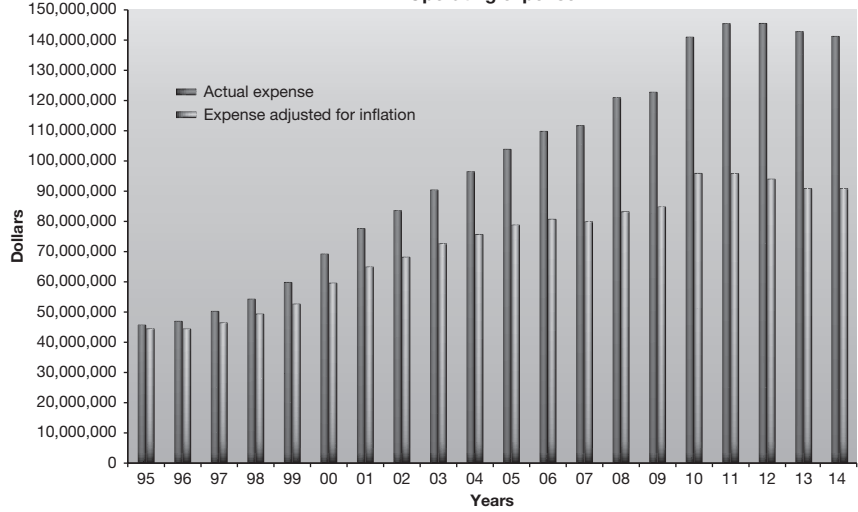
**W. Dillaway Ayres, Jr.**  
*Chief Operating Officer*

### Staff

(Consists of full-time and part-time technical support, core services, publications, meetings, library, public affairs, buildings and grounds, administrative, personnel, Banbury Center, and DNA Learning Center)

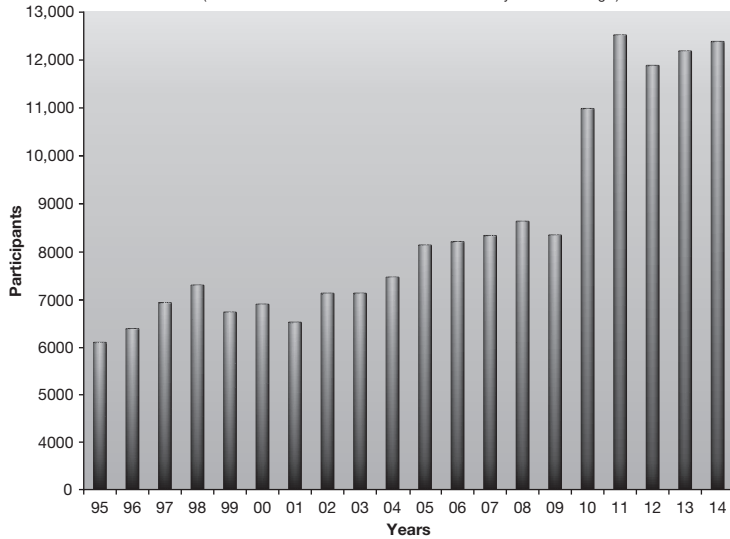


### Operating expense



### Meetings & Courses Participants

(Year 2010 forward includes CSH Asia and Banbury Center meetings.)



# Long-Term Service

---



*Front row (left to right): Susan De Angelo, Sharon Bense, Alison McDermott, Diane Esposito, Carmelita Bautista, Mary Ellen Goldstein, Dessie Carter; back row (left to right): David Spector, Daniel Miller, Philip Lembo, Idee Mallardi, Dill Ayres, James Watson, Bruce Stillman, Jennifer Troge, Spencer Teplin, Andrew Mendelsohn, Janet Argentine, Maureen Berejka, James Hope.*

The following employees celebrated milestone anniversaries in 2014:

- 35 years           Maureen Berejka, James Hope, Bruce Stillman
  
- 30 years           Carmelita Bautista, Dessie Carter, Robert Gensel, Mary Ellen Goldstein,  
Daniel Miller, Steven Tang
  
- 25 years           Sharon Bense, Charlene De Poto, Robert Martiensen, Alison McDermott,  
Spencer Teplin
  
- 20 years           Leslie Allen, Janet Argentine, Susan De Angelo, Carol DuPree, Diane Esposito,  
Philip Lembo, Idee Mallardi, Andrew Mendelsohn, Jennifer Troge



## Long-Term Service: 15 Years

---



*Front row (left to right):* Karine Boyce, Ludmila Pollock, Pierre Jean-Baptiste, Jonathan Parsons; *back row (left to right):* James Watson, Chun-hua Yang, Bruce Stillman, Inessa Hakker, Maureen Bell, Kevin Donohue, Alexander Gann, Dill Ayres.

15 years

Jeannette Amato, Maureen Bell, Karine Boyce, Kevin Donohue, Alexander Gann, Inessa Hakker, Pierre Jean-Baptiste, Frederik Munter, Jonathan Parsons, Ludmila Pollock, John Sundman, David Trimboli, Jenna Williams, Chun-hua Yang, Anthony Zador



RESEARCH

See previous page for photos of the following scientific staff.

*Row 1:* K. Weber (Mitra Lab); C. Hammell, K. Doerfel (C. Hammell Lab); A. Stenlund, S. Schuck (Stenlund Lab); Y. Jin (M. Hammell Lab)

*Row 2:* A. Pommier, D. Fearon (Fearon Lab); Y. Tai (Van Aelst Lab); A. Ambrico (Trotman Lab); K. Rivera (Pappin Lab)

*Row 3:* M. Moissidis (Huang Lab); C. Zepeda-Mendoza, D. Spector (Spector Lab); E. Olund, D. Filippini (Tuveson Lab); P. Mocombe (McCombie Lab)

*Row 4:* M. Wu, J. Tollkuhn (Tollkuhn Lab); A. Scavelli (Gingeras Lab); M. Kaufman (Churchland Lab); T. Forcier (Wigler Lab)

*Row 5:* T. Hige, M. Brill, M. Modi (Turner Lab); M. Regan (Furukawa Lab); B. Lau (Shea Lab); J. Gillis, S. Ballouz, W. Verleyen (Gillis Lab)

# CANCER: GENE REGULATION AND CELL PROLIFERATION

---

Many of the best cancer drugs are highly toxic chemotherapeutics that kill normal and malignant cells alike. In contrast, targeted compounds that recognize cancer-specific pathways represent a kind of silver bullet that would be able to distinguish tumor cells from their healthy counterparts. A few targeted drugs have been identified, such as the BRAF inhibitor vemurafenib for melanoma, and initially these agents were highly promising. But patients rapidly relapsed as their cancers became resistant to treatment. **Molly Hammell** is working to tackle this problem, known as “acquired resistance,” in melanoma. Her lab, in collaboration with the Wistar Institute, combines the power of systems-level, high-throughput data analysis with patient-derived tumor samples. Hammell has developed computational algorithms for the integration of multiple types of high-throughput sequencing data into gene regulatory circuits. She is now applying these methods to explore the global changes in gene regulation that enable melanoma cells to bypass inhibitors of the BRAF signaling pathway, including DNA mutations and epigenetic modifications. Her work will identify the most clinically relevant pathways of interest for additional therapeutic approaches to inhibit tumor growth in melanoma. In addition to her work on melanoma, Hammell is using her expertise in bioinformatics in collaboration with other members of the CSHL community (including Marja Timmermans, Josh Dubnau, and Greg Hannon) to understand gene regulation in diverse systems, from maize to *Drosophila*.

In **Leemor Joshua-Tor**'s lab, 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 critical 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. More recently, members of the Joshua-Tor lab 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 lab also studies the generation of PIWI-interacting RNAs (piRNAs), which serve to protect the genome of germ cells. With colleagues in the Hannon lab, Joshua-Tor's team also determined the structure and function of Zucchini, a key nuclease in the initial generation of piRNAs in fruit flies. In other work, the lab 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 lab studies the mechanisms of RNA splicing, ways in which they go awry in disease, and the means by which faulty splicing can be corrected. In particular, 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 lab is able to correct this defect using a potentially powerful therapeutic approach. It is possible to stimulate protein production by altering mRNA splicing through the introduction of chemically modified pieces of RNA called antisense oligonucleotides (ASOs) into the spinal cords of mice. A drug based on this technology is now in late-stage human trials. Using

ASOs in mice carrying a transgene of human SMN2, Krainer's team has developed a model for SMA using a technique they called TSUNAMI (shorthand for targeting-splicing using negative ASOs to model illness). They have developed a mouse model for adult-onset SMA and they are working to develop models for the study of other diseases caused by splicing defects, including familial dysautonomia. The Krainer lab has also worked to shed light on the role of splicing proteins in cancer. They have found that the splicing factor SRSF1 functions as an oncogene stimulating the proliferation of immortal cells. They have discovered that SRSF1 can actually stop cell growth by stabilizing a powerful tumor suppressor protein, called p53, suggesting that the cell is responding to the aberrant SRSF1 activity. This discovery offers insight into how tumors arise and the pathways that lead to transformation.

**David L. Spector's** laboratory studies the spatial organization and regulation of gene expression. Their recent studies have demonstrated an increase in random monoallelic gene expression upon the differentiation of mouse embryonic stem cells (ESCs) to neural progenitor cells (NPCs). These data support a model where 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. In addition, the Spector lab is characterizing long nuclear retained noncoding RNAs (lncRNAs) that exhibit altered levels of expression as mESCs transition from the pluripotent state to neural progenitor cells, and they are studying lncRNAs that are misregulated in cancer. Their efforts have focused on *Malat1* lncRNA, one of the most abundant noncoding RNAs. The Spector lab previously identified a novel mechanism of 3'-end processing of this RNA. Current studies have revealed that altered levels of *Malat1* lncRNA impact breast cancer initiation and progression. Studies are currently under way to elucidate the mechanism of action of this abundant nuclear retained RNA.

**Arne Stenlund** and colleagues have obtained a detailed understanding of processes required for initiation of DNA replication from the papillomavirus, using this system to gain a general biochemical understanding applicable in other systems. Papillomaviruses are a large viral family that induces cell proliferation at the site of infection, usually giving rise to benign tumors. But certain types of human papillomaviruses (HPVs) generate tumors that progress toward malignancy. Among these are HPVs that cause most cervical cancers. Members of the Stenlund lab also pursue studies aimed at developing an effective small-molecule inhibitor of HPVs that might someday be used by women who do not receive the preventive anti-HPV vaccine now available or those already infected with HPV who would not be helped by the vaccine.

**Bruce Stillman's** lab 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 lab 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.

Acute myeloid leukemia (AML) is a particularly devastating and aggressive blood cancer that is currently incurable in 70% of patients. Research in **Chris Vakoc's** lab seeks to understand this disease as well as others, such as lymphoid leukemias and epithelial tumors, by studying them at the level of genomic regulation. He is particularly interested in the proteins that regulate chromatin in the nucleus of the cell. To identify proteins involved in AML, which may also be targets for drug therapy, he deploys large-scale genetic screens using RNA interference (RNAi) as well as genetically engineered mouse models that display the hallmarks of human cancer. In collaboration with Jay Bradner at the Dana-Farber Cancer Institute, Vakoc has shown that the small-molecule drug candidate JQ1 has potent anti-AML activity. It works by suppressing the protein BRD4, which is a critical regulator of the potent oncogene *c-Myc*. JQ1 is currently in clinical trials as a therapeutic strategy for AML. Vakoc's team identified other chromatin regulators, including PRC2 and RNF20, that if blocked with small-molecule inhibitors may halt AML. Vakoc's lab has gained new insights into the mechanism behind these drug targets, finding that Brd4 and other proteins required for AML bind to enhancers, short stretches of DNA more than a million bases away from the oncogene *c-Myc*. When Brd4 binds to the enhancer, it bends the DNA within the nucleus so that it touches the *c-Myc* region of the genome and activates it, causing cells to proliferate without restraint. Understanding how these proteins function in AML will allow researchers to design more effective and safer therapies to treat this intractable cancer.

# INTEGRATING GENOMIC DATA SETS TO UNDERSTAND GENE REGULATION IN DEVELOPMENT AND DISEASE

M. Hammell   Y. Hao   Y. Jin   E. Paniagua   J. Regalado Perez  
Y.-J. Ho   D. Molik   A. Patel   O. Tam

Human development requires the regulated expression of thousands of genes in hundreds of distinct cell types throughout life. One requirement for this process is that each cell must contain an intact, functional genome. Furthermore, there is a set of rules for how the genes in our genomes coordinate their activities, and these rules are encoded within gene regulatory networks. Many human diseases occur when these basic processes are altered, either through altering the genome itself (as in the mutations seen in cancers) or through altering the way in which genes interact within gene networks. The focus of our lab is to understand how mutations in our genomes lead to alterations in the function of the mutated gene itself and also the repercussion of these alterations on the hundreds of other neighboring genes within the network. To this end, we use computational algorithms to integrate multiple types of genomic and transcriptomic profiling data into models of regulatory rewiring events in human disease. This includes an emphasis on developing novel tools for the statistical analysis of high-throughput data, developing novel algorithms for modeling the flow of signals through genetic pathways, and importantly, testing these models using the tools of molecular genetics. The ultimate goal is to understand how human diseases like cancer take advantage of the cell's ability to rewire its own regulatory networks into programs that serve the needs of the cancer cells.

Much of this work involves tight collaborations with other experimental groups at CSHL. One example involves a collaboration with Marja Timmermans' lab to understand the diversity of small RNA regulators in the maize genome and how these small RNAs dictate leaf patterning in development. Another collaboration, with Josh Dubnau's lab, is exploring the contribution of transposons to human neurodegenerative diseases such as amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD). A collaboration with Greg Hannon's lab probes the molecular mechanisms of

transposon control more generally. A separate project with Greg Hannon's lab aims to understand the changes that occur during early breast cancer development so that we can better predict which patients are most likely to progress to invasive malignant cancers. An independent project in the M. Hammell lab combines both computational and experimental techniques to understand how melanoma cells develop resistance to targeted inhibitors of the melanoma cell growth pathways. Although the experimental systems for each of these projects vary greatly, each project seeks to understand how gene regulation contributes to maintaining cellular function.

## Mechanisms of Acquired Drug Resistance in Melanoma

A. Patel, J. Regalado Perez

---

The genetic basis of melanoma development is fairly well understood, with activating mutations in the oncogene *BRAF* occurring in a majority of melanoma patients (Hodis et al. 2012). Specific inhibitors that target activated *BRAF* as well as the downstream MAP/ERK (mitogen-activated protein kinase/extracellular signal-related kinase) signaling pathway have been developed that dramatically reduce the growth of melanoma cells in patients. However, the effects of these drugs typically extend patient life span for 6 mo or less, as the tumors rapidly develop resistance to these targeted therapies (Villanueva et al. 2011). Although some tumors resistant to *BRAF* inhibitors acquire additional genetic lesions that elevate ERK or AKT signaling (Alcala et al. 2012), an astounding number of therapy-resistant cell lines establish resistance without significant alteration of the genome (M. Herlyn, pers. comm.). Furthermore, serial passage of some of these cell lines in drug-free media results in the re-acquisition of sensitivity to the targeted therapeutics, suggesting that the resistance mechanism in these cases is

likely to be due to changes in gene regulation rather than secondary acquired mutations. Many factors could be mediating this reversible acquired resistance phenotype, the most likely of which would involve alterations in the expression of small RNAs, alternative splicing of key transcripts in the MEK/ERK or AKT signaling pathways, or chromatin-modifying enzymes that alter DNA methylation or histone modification. Members of our lab are using genomic and transcriptomic profiling studies to identify the gene regulatory factors that accompany acquired resistance to BRAF inhibitors. These experimental data sets are fed into custom algorithms designed to integrate these disparate data types into a systems-level view of the cellular signaling pathways that underlie melanoma growth and BRAF inhibitor resistance.

### The Progression from Ductal Carcinoma In Situ to Invasive Ductal Carcinoma in Early Breast Cancer

Y.-J. Ho, O. Tam [in collaboration with G. Hannon, Cold Spring Harbor Laboratory]

Ductal carcinoma in situ (DCIS) is currently believed to be the precursor of invasive ductal carcinoma (IDC), the most common form of breast cancer, accounting for 80% of all breast cancers (DeSantis et al. 2014). It has been estimated that more than 50% of all patients diagnosed with DCIS would never progress to invasive cancer in the patient's lifetime (Cowell et al. 2013), yet nearly all patients opt for treatment to avoid that risk. This reflects an urgent need for understanding the pathway for progression from relatively benign DCIS to the mortality-causing IDC. Adequately characterizing DCIS lesions, however, requires new experimental and computational tools that can handle the extremely small cell counts typically seen—on the order of 10–100 cells per biopsy section. We seek to develop tools that are suitable for single-cell and low-input (1–100 cells) differential expression analysis by improving gene abundance normalization methods, better formulating the relationship between abundance and variance in high-variance samples, and adjusting the standard statistical models currently used to model the sequence counts distribution. Application of these tools to DCIS transcriptome samples will enable us to probe the small populations of cells within early DCIS lesions, understand

the heterogeneity that exists both within and between individual DCIS lesions, understand the changes that occur as benign DCIS lesions develop aggressive and invasive characteristics, and find molecular markers that can be used to predict which DCIS lesions will eventually progress to IDC tumors.

This project will be aided by thousands of tissue samples available from the Duke Breast SPORE bank. Fresh frozen DCIS biopsy samples are available for 159 lesions from 85 patients, with many patient samples exhibiting co-occurring DCIS and IDC tumors. In addition, 3000 DCIS samples are available as formalin-fixed paraffin-embedded (FFPE) tissue blocks from the Duke surgical pathology archives. Although these FFPE samples are much more degraded than the freshly frozen samples, this archive includes a crucial set of biopsies from 25 patients originally diagnosed with DCIS who later developed IDC tumors in the same breast. These cases can be compared to a set of 50 samples with extended clinical follow-up, where the patients are known not to have progressed to IDC within 10 years of the initial DCIS diagnosis. Taken together, these samples present an excellent resource for surveying the range of variance in DCIS transcriptomes, for asking questions about the progression from DCIS to IDC, and for validating any prognostic markers determined from the larger case-control using the smaller longitudinal cohort.

### Small RNA Pathways in Maize

O. Tam [in collaboration with M. Timmermans, Cold Spring Harbor Laboratory]

The maize genome has remained relatively unexplored despite the importance of maize as one of the early genetic model organisms. Preliminary characterizations of the maize genome and transcriptome suggest that substantial differences exist, between both maize ecotypes and the well-characterized plant model *Ara-bidopsis* (Springer et al. 2009; Eveland et al. 2010). Furthermore, these maize inbreds exhibit dramatic phenotypic differences in response to mutations in small RNA biogenesis factors, suggesting extensive evolution in both small RNA function and activity. In collaboration with the Timmermans lab, my group has undertaken a project to characterize the dynamics of small RNA expression and activity in a variety of maize tissues and inbreds. The focus of this project



is to determine the small RNA-mediated regulatory circuits that give rise to inbred-specific phenotypic diversity both within the maize species and within the larger group of grasses. Our analysis of maize *trans*-acting small interfering RNAs (tasiRNAs), a group of mobile endogenous 21-nucleotide siRNAs, has revealed that in maize shoots, tasiRNAs are generated through just a single genetic pathway and function solely in leaf development. In contrast, the many tasiRNAs found in *Arabidopsis* are processed via several distinct biogenesis pathways and function in development as well as pathogen defense. This same analysis also identified a new class of 22-nucleotide phased siRNAs, produced by an unidentified pathway that does not overlap with the tasiRNA biogenesis pathway. This class of small RNAs is entirely novel, with no homologs in other sequenced plants and no defined mechanism for biogenesis or activity. Finally, only 25% of the known microRNAs (miRNAs) from *Arabidopsis* are conserved in maize, leaving a wealth of novel miRNAs to explore. Members of our lab are collaborating to provide computational analysis of the profiles of small RNAs and their targets across maize ecotypes. Our lab has developed novel algorithms for the identification of phased small RNAs, for miRNA gene identification, and for miRNA target interaction analysis. These algorithms are being applied to deep-sequencing data sets of small RNAs and mRNAs isolated from developing maize tissues.

### Transposon Control Systems in Animals

Y. Jin, O. Tam [in collaboration with G. Hannon and J. Dubnau, Cold Spring Harbor Laboratory]

TDP-43 is an RNA-binding protein that is known to control proper splicing and translation of many RNA targets in neurons. Mutation of TDP-43 has been associated with a variety of neurodegenerative diseases including amyotrophic lateral sclerosis (ALS), frontotemporal lobar degeneration (FTLD), and Alzheimer's disease (AD) (Cohen et al. 2011). However, the normal function of TDP-43 in neuronal development and maintenance has not been fully characterized, and few of its mRNA targets have been definitively associated with the neurodegenerative diseases that result from loss of TDP-43 function (Da Cruz et al. 2011). In collaboration with the Dubnau lab, my group has explored the novel hypothesis that TDP-43 normally

plays a large and hitherto uncharacterized role in regulating the expression of transposable elements (TEs), mobile genetic elements whose unregulated expression leads to genetic instability as well as cellular toxicity. Members of our lab have shown that TE transcripts are elevated upon expression of mutant, aggregate-prone forms of human TDP-43 in the fly brain and that neurodegeneration results from expression of this human TDP-43 protein. Members of my group have shown that TDP-43 binds widely to TE transcripts in mammals and that TDP-43 binding to TEs is lost in human patients diagnosed with FTLD (Li et al. 2012), a disease characterized by TDP-43 proteinopathy. Although these studies support a role for TDP-43 in regulating TE expression, our future goals are centered on defining a causal role for TDP-43-mediated regulation of TEs in neurodegenerative disease. One important element of this project will be to identify how TDP-43 interacts with the small RNA regulators of TE expression known to be involved in controlling TE mobility. We are currently exploring the molecular mechanism behind TDP-43 function in the fly and mouse model organisms, where many genetic mutants are available, to carefully describe the factors that mediate TDP-43 involvement in TE regulation, biogenesis, and activity. My group is pursuing a parallel approach in close collaboration with the Dubnau lab to quantify the extent to which TDP-43 contributes to TE regulation as compared to its other mRNA targets and the potential contribution of each function to clinical outcomes.

In a separate collaboration, we work closely with members of the G. Hannon lab to better understand the biological mechanisms of transposon control in animals. This basic research into the factors that regulate transposon activity sets the foundation for understanding disease processes affected by transposon misregulation. Most families of TEs are highly distinct at the primary sequence level and use different strategies for their propagation, challenging the ability of host genomes to discriminate transposons from endogenous genes (Levin and Moran 2011). Several recent studies have indicated that a small RNA-based innate immune system, the Piwi-interacting RNA (piRNA) pathway, addresses the challenge of transposon recognition and silencing. In germ cells, primary piRNAs are loaded into Piwi and Aubergine to control transposons through a posttranscriptional mechanism. However, several factors, including Piwi's nuclear localization, have suggested that Piwi might

also regulate its targets by controlling their transcription. We therefore probed the contributions of Piwi to transposon silencing in germ cells and in their somatic support cells using tissue-specific piwi knockdowns.

Integrated analyses of steady-state RNA levels by RNA sequencing (RNA-Seq), and measurements of nascent RNA synthesis using global run-on sequencing (GRO-Seq), chromatin marks by chromatin immunoprecipitation combined with deep sequencing (ChIP-Seq), and small RNA levels indicated multiple roles for Piwi in the piRNA pathway in both the germline and the soma. In the soma, our data support a model where Piwi silences the transcription of targeted elements, such that transcriptional gene silencing (TGS) is the dominant form of Piwi activity. In germ cells, our data revealed that Piwi both acts via TGS and exerts unexpected impacts on piRNAs that occupy Aubergine and Ago3, indirectly reducing the capacity of the entire piRNA pathway to silence transposons via posttranscriptional gene silencing (PTGS). These data sets add new complexity to our understanding of Piwi function, which has primarily been studied for its direct role in PTGS. Ongoing experiments will seek to understand the mechanisms

by which Piwi contributes to TGS and how Piwi indirectly affects other moderators of PTGS.

## PUBLICATIONS

- Dotto MC, Petsch KA, Aukerman MJ, Beatty M, Hammell M, Timmermans MC. 2014. Genome-wide analysis of leafbladeless1-regulated and phased small RNAs underscores the importance of the TAS3 ta-siRNA pathway to maize development. *PLoS Genet* **10**: e1004826.
- Feigin ME, Xue B, Hammell MC, Muthuswamy SK. 2014. G-protein-coupled receptor GPR161 is overexpressed in breast cancer and is a promoter of cell proliferation and invasion. *Proc Natl Acad Sci* **111**: 4191–4196.
- Meikar O, Vagin VV, Chalmel F, Sostar K, Lardenois A, Hammell M, Jin Y, Da Ros M, Wasik KA, Toppari J, Hannon GJ, Kotaja N. 2014. An atlas of chromatoid body components. *RNA* **20**: 483–495.
- In Press*
- Boj SF, Hwang CI, Baker LA, Chio II, Engle DD, Corbo V, Jager M, Ponz-Sarvisé M, Tiriác H, Spector MS, et al. 2015. Organoid models of human and mouse ductal pancreatic cancer. *Cell* **160**: 324–338.
- Yu JR, Tai Y, Jin Y, Hammell MC, Wilkinson JE, Roe JS, Vakoc CR, Van Aelst L. 2015. TGF- $\beta$ /Smad signaling through DOCK4 facilitates lung adenocarcinoma metastasis. *Genes Dev* **29**: 250–261.



O. Tam, D. Molik

# STRUCTURAL BIOLOGY OF NUCLEIC ACID REGULATORY PROCESSES

L. Joshua-Tor   E. Elkayam   C. Faehnle   J. Ipsaro   S.-J. Lee   A. Tocilj  
A. Epstein   S. Goldsmith   C. Kuhn   K. On   J. Walleshauser

We study the molecular basis of nucleic acid regulatory processes, RNA interference (RNAi) and DNA replication in particular. We use the tools of structural biology, biochemistry, and biophysics to study proteins and protein complexes associated with these processes to elucidate how they work. 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 with their function and biology.

## Mechanisms of RNAi and Noncoding RNAs

---

RNAi has made an enormous impact on biology in a very short period of time. Not only are we still discovering new cellular pathways for the regulation of gene expression that use these pathways, but RNAi has also become an extraordinarily useful and simple tool for gene silencing. Almost from its beginnings, people have used genetics, biochemistry, molecular biology, and bioinformatics to study the mechanism of RNAi and related pathways. We argued, however, that in order to get a true mechanistic understanding of these pathways, we must understand how the components of the RNAi machinery work at a molecular level.

### The let-7 Pathway

C. Faehnle, J. Walleshauser

---

let-7 is among the first microRNAs (miRNAs) discovered, and it is involved in controlling cell fate during late-stage larval development in *Caenorhabditis elegans*. Functionally conserved in humans, the let-7 family is broadly expressed in somatic cells and regulates cellular proliferation and differentiation, as well as the repression of several oncogenes and key regulators of mitogenic pathways. Humans have 12 let-7

variants expressed from eight independent chromosomal loci, and many tumors are associated with the coordinated down-regulation of multiple let-7 family members. In contrast to differentiated cells, mature let-7 levels are undetectable in stem cells, but primary and precursor let-7s remain high. Similarly, human tumors generally express the primary transcripts but not the mature form, which was the first indication that a mechanism of posttranscriptional regulation of let-7 biogenesis is at play. The Lin28 protein was discovered to be the posttranscriptional regulator of let-7 biogenesis. Lin28 is highly expressed in embryonic stem cells and is naturally active in early development, where it suppresses the expression of mature let-7 to maintain the pluripotent state. Lin28 has a number of important biological roles, including the maintenance of pluripotency, the reprogramming of cellular metabolism, and the promotion of tissue regeneration. The Lin28–let-7 pathway has been described as a bistable switch. Either Lin28 or mature let-7 is expressed, but not both. Each molecule represses expression of the other, and once the cell changes its state, the result is differentiation or if the switch is reversed, cancer.

In differentiated cells, TUT4 and TUT7 uridylylate certain let-7 precursors (pre-let-7) that carry a single 3′-nucleotide overhang, rather than the more Dicer-ready two-nucleotide overhang pre-miRNAs. In this case, in the absence of Lin28, only a single uridine is added to convert the pre-let-7s to better substrates for Dicer and promote the production of mature let-7. In stem cells, Lin28 binds to pre-let-7, triggering its 3′ oligouridylation by TUT4/7. The oligoU tail serves as a decay signal, as it is rapidly degraded by Dis3L2, a homolog of the catalytic subunit of the exosome. We determined the structure of mouse Dis3L2 in complex with an oligoU RNA to mimic the uridylylated tail of pre-let-7 (Fig. 1). The RNA-binding domains form an open funnel on one face of the catalytic domain that allows RNA to navigate a path to the active site. An extensive network of uracil-specific interactions spans



**Figure 1.** The structure of mouse Dis3L2 in complex with oligoU RNA.

the first 12 nucleotides of the oligoU-tailed RNA, explaining how Dis3L2 recognizes, binds, and processes uridylylated pre-let-7 in the final step of the Lin28-let-7 pathway.

### Small RNA and tRNA Surveillance by the CCA-Adding Enzyme

C. Kuhn [in collaboration with J. Wilusz, University of Pennsylvania; P. Beal, University of California, Davis]

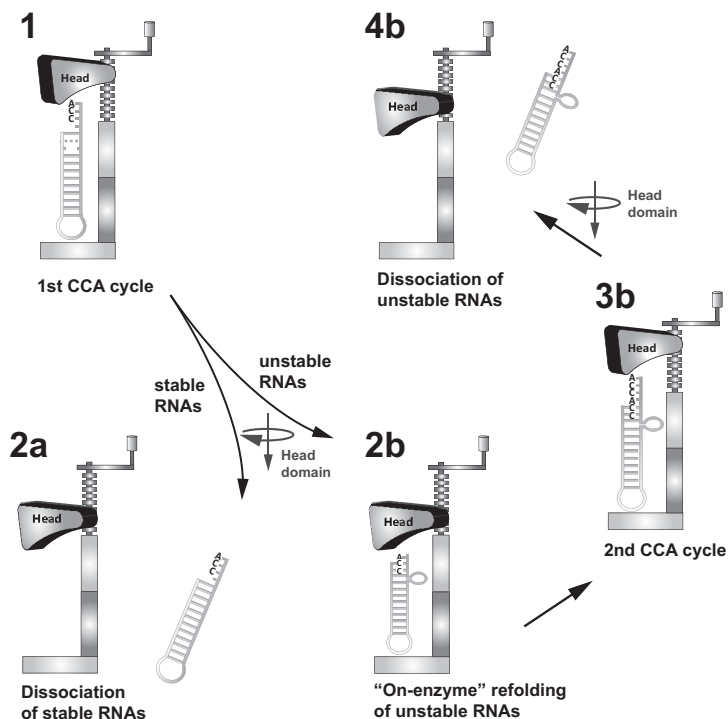
Transcription in eukaryotes produces a number of long noncoding RNAs (lncRNAs). Two of these, MALAT1 and Men $\beta$ , which are well-characterized long nuclear-retained noncoding transcripts that are involved in cancer progression and paraspeckle formation, generate a tRNA-like small RNA in addition to the mature lncRNA. The stability of these tRNA-like small RNAs as well as that of bona fide tRNAs are monitored by the CCA-adding enzyme. Just as for bona fide tRNAs, the trinucleotide CCA is post-transcriptionally added to the 3'-end of mascRNA (MALAT1-associated small cytoplasmic RNA), the tRNA-like transcript originating from the 3'-end of MALAT1. Unlike the acceptor stems of canonical tRNAs and mascRNA, the Men $\beta$  acceptor stem is destabilized through mismatches or wobble base pairs,

somehow prompting the CCA-adding enzyme to repeat its catalytic cycle and thereby add tandem CCA motifs. The CCACCA tail then serves as a degradation signal for the cellular RNA decay machinery. It appears now that many bona fide tRNAs that are destabilized through mutations or the lack of proper modifications are subjected to a similar CCACCA addition. Thus, the CCA-adding enzyme is not only critical for generating functional tRNAs, but also likely has a universal and central role in tRNA and tRNA-like small RNA surveillance and quality control.

We investigated how the CCA-adding enzyme distinguishes structurally stable RNAs from unstable RNAs so as to only mark unstable RNAs with CCACCA. We unraveled how a second CCA cycle can occur despite the enzyme possessing a unique mechanism that normally ensures a single cycle of CCA synthesis. We found that after the first CCA cycle, nucleotide binding to the active site induces the enzyme to apply torque on the RNA. A clockwise screw motion of the enzyme's catalytic domain leads to RNA compression and overwinding. This causes unstable RNAs to extrude a distinctively positioned bulge from the acceptor stem while still bound to the enzyme (Fig. 2). The bulge does not perturb the double-helical nature of the substrate and all other



**Figure 2.** Structure of the CCA-adding enzyme from *Archaeoglobus fulgidus* in complex with a tRNA minihelix (containing the acceptor stem and the TΨC stem-loop) with added CCAC, and CTP.



**Figure 3.** RNA surveillance by the CCA-adding enzyme. The CCA-adding enzyme acts as a molecular vise. The clockwise screw motion of the enzyme's head domain applies torque on the RNA, leading to compression. The RNAs thus essentially proofread themselves through differential responses to their interrogation between stable and unstable substrates.

structural determinants near the active site remain in place. The catalytic mechanism is thus preserved between the first and second cycles of CCA addition. In total, we found that tandem CCA addition is not the result of a modified enzymatic activity that is particular to unstable RNAs. Rather, it is a consequence of the natural activity of the CCA-adding enzyme on a substrate with increased conformational flexibility. By exploiting the versatility of RNA structure, the CCA-adding enzyme is able to trigger the degradation of potentially detrimental small RNAs and tRNAs (Fig. 3).

### RNAi and Heterochromatin Formation

C. Kescu, T. Schalch [in collaboration with R.A. Martienssen, Cold Spring Harbor Laboratory; J. Partridge, St. Jude Children's Research Hospital]

RNAi pathways are also used to repress genes at the transcriptional level by guiding heterochromatin formation. Although we are just beginning to discover how metazoans use this type of mechanism for

transcriptional silencing, these pathways are much better characterized in plants and fungi, in particular in *Schizosaccharomyces pombe*, where heterochromatin is essential for kinetochore assembly required for proper segregation of chromosomes during cell division. The complexes involved in heterochromatinization in *S. pombe* are fairly well defined, and in most cases, only one version of each component exists. However, our understanding of this pathway has not gone much beyond the "blobology" stage either. Assembly of centromeric heterochromatin requires the RITS (RNA-induced transcriptional silencing) complex, a specialized RISC (RNA-induced silencing complex), which physically anchors small RNAs to chromatin. It consists of the Argonaute protein, Ago1, the chromodomain protein Chp1, a GW-like protein Tas3, and small interfering RNAs (siRNAs) derived from centromeric repeats. We previously characterized Chp1's high-affinity binding to chromatin and showed it to be critical for the establishment of centromeric heterochromatin. In addition, apart from the chromodomain, Chp1 had no recognizable

domains. Our structure of the carboxy-terminal half of Chp1 with the Tas3 amino-terminal domain revealed the presence of a PIN domain in Chp1, which contributes to posttranscriptional gene silencing of subtelomeric transcripts independent of RNAi. We also realized that Chp1 and Tas3 form a tight complex and that Argonaute appears to bind more loosely. We suggested that Chp1–Tas3 provides a solid and versatile platform to recruit both RNAi-dependent and RNAi-independent gene-silencing pathways for locus-specific regulation of heterochromatin and that the reach of long and flexible regions in these two proteins might be important for contacting other nucleosomes at these loci.

The RITS complex recruits the H3K9 methylase, Clr4, which is part of a large complex called CLRC. CLRC also contains the cullin Cul4 and its interacting protein Pip1, the adaptor Rik1, a WD repeat protein Dos1, and Dos2. Curiously, CLRC is an active E3 ligase *in vitro*, and this activity is necessary for heterochromatin assembly *in vivo*. Yet the role of this E3 ligase activity is still unknown. We performed a pairwise interaction screen of the CLRC components revealing a subunit arrangement that is reminiscent of CRL4, a different cullin ring ligase complex for which a crystal structure is available. Both complexes contain the ubiquitin ligase Cul4, and sequence similarities between two other subunits were already noted. We have begun delineating structural and functional similarities. In particular the placement of Dos1 indicated that it might have a role similar to that of DDB2 in CRL4, as a target-binding factor for the E3 ligase. We solved the structure of Dos1 and found it to be an eight-bladed  $\beta$ -propeller with one face mediating interaction with the adaptor protein Rik1. We made a series of mutations on the opposite surface and found several regions that are required for heterochromatin silencing, underscoring the notion that Dos1 is the specificity factor for the ligase.

## The Different Faces of E1: A Replicative Hexameric Helicase

S.-J. Lee [in collaboration with A. Stenlund, Cold Spring Harbor Laboratory; T. Ha, University of Illinois, Urbana-Champaign]

Precise replication of the genome is essential for maintaining the integrity of genomic information. As a prerequisite for DNA replication, two complemen-

tary DNA strands are separated and each becomes a template for the synthesis of a new complementary strand. Strand separation is mediated by a helicase enzyme, a molecular machine that uses the energy derived from ATP hydrolysis while moving along the DNA. Our crystal structures of the double-stranded DNA (dsDNA)-binding domain (DBD) of the replicative helicase E1 from papillomavirus in various stages of assembly led us to propose a model in which the transition from the dimer to the ultimate double hexamer results in strand separation. The loading and assembly of this protein separate the double helix, such that each hexameric helicase encircles one strand of DNA. Once assembled, the helicase uses its ATP-driven motor to translocate on the DNA or “pump” the single-stranded DNA (ssDNA) through the hexameric ring. Several competing mechanisms for helicase unwinding were proposed. Having determined a structure of hexameric E1 with ssDNA discretely bound in the central channel and nucleotides at the subunit interfaces, we showed that only one DNA strand passes through the hexamer channel and that the DNA-binding hairpins of each subunit form a spiral staircase that sequentially tracks the DNA backbone. The nucleotide configurations at the subunit interfaces indicate that each subunit sequentially progresses through ATP, ADP, and apo states, whereas its associated DNA-binding hairpin travels from the top to the bottom of the staircase, each escorting one nucleotide of ssDNA through the channel, as if six hands grab the DNA and upon ATP hydrolysis and ADP release pull it through the channel. With this unique look into the mechanism of translocation of this molecular machine along DNA, we have focused on mechanistic aspects of the enzyme in solution.

By taking a multifaceted approach including single-molecule and ensemble FRET (Förster resonance energy transfer) methods, we have found that E1 is oriented with the amino-terminal side of helicase facing the replication fork, consistent with the crystal structure. We also showed that E1 generates strikingly heterogeneous unwinding patterns stemming from varying degrees of repetitive movements that are modulated by the DNA-binding domain. Furthermore, our studies found that the DNA-binding domain promotes the assembly of E1 helicase onto a forked DNA substrate, acting as an allosteric effector of the helicase. Taken together, our studies

reveal previously unrecognized dynamic facets of replicative helicase unwinding mechanisms, adding another layer of complexity in the workings and regulation of DNA replication.

#### PUBLICATIONS

Faehnle CR, Walleshauser J, Joshua-Tor L. 2014. Mechanism of Dis3L2 substrate recognition in the Lin28/let-pathway. *Nature* **514**: 252–256.

Kuscu C, Zaratiegui M, Kim HS, Wah DA, Martienssen RA, Schalch T, Joshua-Tor L. 2014. The CRL4-like Clr4 Complex

in *S. pombe* depends on an exposed surface of Dos1 for heterochromatin silencing. *Proc Natl Acad Sci* **111**: 1795–1800.

Lee SJ, Syed S, Enemark EJ, Schuck S, Stenlund A, Ha T, Joshua-Tor L. 2014. Dynamic look at DNA unwinding by a replicative helicase. *Proc Natl Acad Sci* **111**: E827–E835.

#### *In Press*

Ipsaro JJ, Joshua-Tor L. 2015. From guide to target: Molecular insights into eukaryotic RNAi machinery. *Nat Struct Mol Biol* **22**: 20–28.

Kuhn CD, Wilusz JE, Zheng Y, Beal PA, Joshua-Tor L. 2015. On-enzyme refolding permits small RNA and tRNA surveillance by the CCA-adding enzyme. *Cell* **160**: 644–658.

# RNA SPLICING

A.R. Krainer	M. Akerman	S. Das	T. Nomakuchi	W. Tan
	O. Anczuków	Y. Hua	J. Novatt	D. Vargas Landín
	I. Aznarez	H.-Y. Jeon	E. Panagot	R. Wappel
	F. Bezirci	Y.-H. Liu	J. Scharner	S. Wong
	W.-J. Chen	L. Manche	M. Sivaramakrishnan	

## 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. The work in our lab 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 misregulation in cancer; and the development

of effective methods to correct defective splicing or modulate alternative splicing, especially in a disease context. A summary of some of our recently published studies is provided below.

## Alternative Splicing and Cancer

---

Alternative splicing has an important role in cancer, partly by modulating the expression of many oncogenes and tumor suppressors, and also because inactivating mutations that affect alternative splicing of various tumor suppressor genes account for some of the inherited and sporadic susceptibility to cancer. In addition, alternative splicing controls a metabolic switch characteristic of all cancer cells, known as the Warburg effect. We have continued to analyze the roles of individual members of the SR protein family of splicing factors in cancer.

We studied the SR protein SRSF6 and identified it as a regulator of wound healing and tissue homeostasis in skin (in collaboration with John Erby Wilkinson, University of Michigan, Ann Arbor). We found that SRSF6 is a proto-oncogene frequently overexpressed in human skin cancer, including malignant melanomas and basal-cell and squamous-cell carcinomas. Overexpressing doxycycline-inducible SRSF6 cDNA in transgenic mice induced epithelial hyperplasia of sensitized skin (e.g., by shaving or hair plucking) and promoted aberrant alternative splicing. We observed severe epidermal thickening and hair-follicle lesions, which were reversible upon withdrawal of the doxycycline inducer. Skin transplantation experiments to nontransgenic donor mice showed that the inducible SRSF6 effects are cell-autonomous.

Microarray analysis identified 139 target genes with SRSF6-induced alternative-splicing changes in skin,



including many genes associated with wound healing. Focusing on the *Tnc* gene coding for the extracellular-matrix protein tenascin C, we demonstrated that SRSF6 binds to alternative exons in its pre-mRNA and promotes the expression of isoforms characteristic of invasive and metastatic cancer, although this particular regulatory event is independent of cell type. SRSF6 overexpression additionally resulted in depletion of LGR6+ stem cells and excessive keratinocyte proliferation and response to injury. Furthermore, we showed that the effects of SRSF6 on wound healing assayed *in vitro* depend on the TNC isoforms. Thus, abnormal expression of this SR protein splicing factor can perturb tissue homeostasis in skin.

### Targeted Antisense Modulation of Alternative Splicing for Therapy and Investigation of Pathogenesis Mechanisms

---

Spinal muscular atrophy (SMA) is a common, autosomal-recessive motor-neuron degeneration disorder caused by homozygous deletion or mutation of the survival-of-motor-neuron gene, *SMN1*. A closely related *SMN1* paralog, *SMN2*, is present in all patients and differs from *SMN1* by a C to T transition in exon 7 that causes substantial skipping of this exon, such that *SMN2* expresses only low levels of functional, full-length SMN protein. *SMN2* decreases the severity of SMA in a copy-number-dependent manner. We previously developed an antisense oligonucleotide (ASO) that efficiently increases the extent of exon 7 inclusion during splicing of *SMN2* transcripts, for therapeutic use in SMA. This translational research was done in collaboration with Isis Pharmaceuticals. After extensive preclinical testing, and phase I and phase II open-label clinical trials, phase III randomized, double-blind, sham-procedure-controlled clinical trials with the ASO compound, ISIS-SMN<sub>Rx</sub>, are well under way. The drug, which is administered to SMA infants and children by lumbar puncture, has been well tolerated at all doses tested, and dose- and time-dependent improvements in motor function have already been observed in the open-label studies.

Although it is clear that SMN deficiency causes SMA, the pathogenesis mechanisms remain elusive. SMN is ubiquitously expressed and is an essential protein, but its reduction appears to selectively affect motor neurons, for unknown reasons. However, selectively restoring SMN in motor neurons only partially

rescues SMA in mouse models. We addressed the relative importance of SMN restoration in the central nervous system (CNS) versus peripheral tissues in mouse models using a systemically administered therapeutic splice-switching ASO to broadly restore SMN, and an intracerebroventricularly administered complementary decoy oligonucleotide to neutralize the effects of the therapeutic ASO within the CNS. This procedure effectively cancels out the effect of ASO that crosses the immature blood-brain barrier in neonate mice. Increasing SMN exclusively in peripheral tissues completely rescued necrosis in mild SMA mice and robustly extended survival in severe SMA mice, with significant improvements in vulnerable tissues and motor function along with preservation of spinal motor neuron counts and neuromuscular-junction integrity. Our data demonstrate a critical role of peripheral pathology in the mortality of SMA mice and indicate that peripheral SMN restoration compensates for its deficiency in the CNS and preserves motor neurons. Thus, SMA is not a cell-autonomous defect of motor neurons in SMA mice.

### Spliceosomal Protein-Interaction Network

---

During spliceosome assembly, protein–protein interactions (PPI) are sequentially formed and disrupted to accommodate the spatial requirements of pre-mRNA substrate recognition and catalysis. Splicing activators and repressors, such as SR proteins and heterogeneous nuclear RNPs (hnRNPs), modulate spliceosome assembly and regulate alternative splicing. However, it remains unclear how they differentially interact with the core spliceosome to perform their functions.

We investigated the protein connectivity of SR and hnRNP proteins to the core spliceosome using probabilistic network reconstruction based on the integration of interactome and gene-expression data. We validated our model by immunoprecipitation and mass spectrometry of the prototypical splicing factors SRSF1 and hnRNPA1. Network analysis revealed that a splicing factor's properties as an activator or repressor can be predicted from its overall connectivity to the rest of the spliceosome. In addition, we discovered and experimentally validated PPIs between the oncoprotein SRSF1 and members of the SF3 subcomplex of U2 snRNP that are targeted by spliceostatin, pladienolide, and related antitumor drugs. Our findings suggest that activators promote

the formation of PPIs between spliceosomal subcomplexes, whereas repressors mostly operate through protein-RNA interactions.

## PUBLICATIONS

- Chen HH, Welling M, Bloch DB, Muñoz J, Mientjes E, Chen X, Tramp C, Wu J, Yabuuchi A, Chou YF, Buecker C, Krainer A, Willemsen R, Heck AJ, Geijsen N. 2014. DAZL limits pluripotency, differentiation, and apoptosis in developing primordial germ cells. *Stem Cell Reports* **3**: 892–904.
- Das S, Krainer AR. 2014. Emerging functions of SRSF1, splicing factor and oncoprotein, in RNA metabolism and cancer. *Mol Cancer Res* **12**: 1195–1204.
- Jensen MA, Wilkinson JE, Krainer AR. 2014. Splicing factor SRSF6 promotes hyperplasia of sensitized skin. *Nat Struct Mol Biol* **21**: 189–197.
- Paz S, Krainer AR, Caputi M. 2014. HIV-1 transcription is regulated by splicing factor SRSF1. *Nucleic Acids Res* **42**: 13812–13823.
- Rigo F, Chun, SJ, Norris DA, Hung G, Lee S, Matson J, Fey RA, Gaus H, Hua Y, Grundy JS, Krainer AR, Henry SP, Bennett CF. 2014. Pharmacology of a central nervous system delivered 2'-O-methoxyethyl-modified survival of motor neuron splicing oligonucleotide in mice and nonhuman primates. *J Pharmacol Exp Ther* **350**: 46–55.
- Senturk S, Yao Z, Camiolo M, Stiles B, Rathod T, Walsh AM, Nemerova A, Lazzara MJ, Altorki NK, Krainer A, Moll UM, Lowe SW, Cartegni L, Sordella R. 2014. p53 $\Psi$  is a transcriptionally inactive p53 isoform able to reprogram cells toward a metastatic-like state. *Proc Natl Acad Sci* **111**: E3287–E3296.
- Weyn-Vanhenenryck SM, Mele A, Yan Q, Sun S, Farny N, Zhang Z, Xue C, Herre M, Silver PA, Zhang MQ, Krainer AR, Darnell RB, Zhang C. 2014. HITS-CLIP and integrative modeling define the Rbfox splicing-regulatory network linked to brain development and autism. *Cell Rep* **6**: 1139–1152.

## In Press

- Akerman M, Fregoso OI, Das S, Ruse C, Jensen MA, Pappin DJ, Zhang MQ, Krainer AR. 2015. Differential connectivity of splicing activators and repressors to the human spliceosome. *Genome Biol* **16**: 119.
- Hua Y, Liu YH, Sahashi K, Rigo F, Bennett CF, Krainer AR. 2015. Motor neuron cell-nonautonomous rescue of spinal muscular atrophy phenotypes in mild and severe transgenic mouse models. *Genes Dev* **29**: 288–297.
- Krainer AR. 2015. Splicing: Still so much to learn. *RNA* **21**: 500–501.
- Staropoli JF, Li H, Chun SJ, Allaire N, Cullen P, Thai A, Fleet CM, Hua Y, Bennett CF, Krainer AR, Kerr D, McCampbell A, Rigo F, Carulli JP. 2015. Rescue of gene-expression changes in an induced mouse model of spinal muscular atrophy by an antisense oligonucleotide that promotes inclusion of SMN2 exon 7. *Genomics* **105**: 220–228.
- Xiong HY, Alipanahi B, Lee LJ, Bretschneider H, Merico D, Yuen RK, Hua Y, Gueroussov S, Najafabadi HS, Hughes TR, Morris Q, Barash Y, Krainer AR, Jovic N, Scherer SW, Blencowe BJ, Frey BJ. 2015. RNA splicing: The human splicing code reveals new insights into the genetic determinants of disease. *Science* **347**: 1254806.

# CELL BIOLOGY OF THE NUCLEUS

D.L. Spector   G. Arun   S. Diermeier   R.I. Kumaran  
J. Bergmann   M. Eckersley-Maslin   J. Li  
M.S. Bodnar   S. Hearn   A. Yu  
K.-C. Chang   M. Hübner   C. Zepeda-Mendoza

Most cellular processes can trace their beginnings to the nucleus, where gene activation results in the production of an RNA molecule, some of which encode for proteins, whereas others are functional in the form of RNA. 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 non-coding RNAs in regulating this multifaceted process are just beginning to be elucidated. During the past year, our research has continued to focus on elucidating various mechanisms of regulating gene expression and DNA repair and the role of long nuclear retained noncoding RNAs in development and cancer progression. The following is an overview of some of our major accomplishments over the past year.

## Expression of lncRNAs in Mouse Embryonic Stem Cells and Neural Progenitor Cells

J. Bergmann, A. Yu, M. Eckersley-Maslin, J. Li  
[in collaboration with F. Rigo and S. Freier,  
Isis Pharmaceuticals]

Long noncoding RNAs (lncRNAs) (>200 nucleotides in length) represent a relatively recently studied class of RNAs that have been implicated in numerous cellular functions, including modulating transcriptional patterns, regulating protein activities, serving structural or organizational roles, altering RNA processing events, and serving as precursors to small RNAs. The majority of lncRNAs are expressed at very low levels, and these RNAs generally exhibit poor primary sequence conservation over evolution.

We performed high-depth poly(A)<sup>+</sup> RNA sequencing across multiple clonal populations of mouse embryonic stem cells (ESCs) and neural progenitor cells (NPCs) to comprehensively identify differentially regulated intergenic lncRNAs. Of the

3506 genes annotated in Gencode M3 as “lincRNA” or “processed transcript” (Ensembl 76; referred to as “lncRNA” for the remainder of this study), we detected 958 lncRNAs expressed at  $\geq 0.1$  FPKM (fragments per kilobases of exons for per million mapped reads) in the union of our ESC and NPC data sets. This set of lncRNAs contained well over 800 that are largely uncharacterized on the molecular level. Of the 800 lncRNAs identified, 508 (64%) displayed a significant change in expression level upon differentiation of ESCs to NPCs (DESeq2, FDR < 0.01), with 439 lncRNAs being exclusively detected in either cell type. We also used *ab initio* transcript assembly using Cufflinks to assess the completeness of the Gencode annotation. From our analyses, we prioritized 50 lncRNA candidates that exhibited expression in ESCs, thus potentially implicating them in the coregulation of self-renewal or pluripotency-related networks. Because we are interested in identifying lncRNAs whose expression is strongly associated with the ESC state, we also analyzed raw poly(A)<sup>+</sup> RNA-Seq data from 22 mouse tissues (including fetal liver and neural tissues of embryonic origin) recently released by the ENCODE project. Gene-level analysis of this matrix demonstrated that the majority of the 50 lncRNAs most specifically expressed in ESCs displayed exclusive expression in this cell type, rendering them viable candidates as pluripotency biomarkers.

We have initiated studies to assess the functional role of these lncRNA candidates. Among our top candidates, we have been able to demonstrate that acute depletion of PAT-14 using antisense oligonucleotide technology impacts the differentiation- and development-associated gene expression program of ESCs. Furthermore, we demonstrate that Firre, an lncRNA highly enriched in the nucleoplasm and previously reported to mediate chromosomal contacts in ESCs, controls a network of genes related to RNA processing. Ongoing studies will further analyze

the functional associations of these and other ESC-specific lncRNAs.

### Probing the Role of the lncRNA *Malat1* in Breast Cancer Progression

G. Arun, S. Diermeier, K.-C. Chang [in collaboration with M. Egeblad, M. Akerman, S. Hearn, and A. Krainer, Cold Spring Harbor Laboratory; J.E. Wilkinson, University of Michigan Medical School; Y. Kim and A.R. MacLeod, Isis Pharmaceuticals; L. Norton and E. Brogi, Memorial Sloan-Kettering Cancer Center]

*MALAT1* (metastasis-associated lung adenocarcinoma transcript 1) is among the most abundant and highly conserved nucleus-restricted lncRNAs whose expression is misregulated in many cancers, including breast cancer. We have used the *MMTV-PyMT* mouse model of luminal-B breast cancer to characterize the role of *Malat1* in primary breast cancer and its subsequent metastasis. *Malat1* lncRNA was knocked down via subcutaneous administration of antisense oligonucleotides (ASOs) over a period of 7 wk, after which animals were sacrificed, and primary tumors and lungs were removed for molecular and histological analyses. *Malat1* ASO treatment resulted in ~60% knockdown in the primary tumor concomitant with a significant reduction in tumor progression rate, as well as a change in the differentiation status. Detailed histopathological analysis of ASO-treated tumors showed an increase in well-differentiated ductular tumors, whereas scrambled ASO-treated tumors progressed to solid carcinomas. Most interestingly, a marked decrease was observed in the incidence of lung metastases: ~70% fewer metastatic nodules in *Malat1* ASO-treated animals than in scrambled ASO-treated animals. Furthermore, *Malat1* ASO-treated ex-vivo-generated mammary gland organoids from *MMTV-PyMT* mice resulted in an inhibition of branching morphogenesis, which recapitulates the invasive process that initiates metastases in vivo. RNA-Seq analysis of the primary tumors and tumor-derived organoids treated with *Malat1* ASO showed up-regulation of genes involved in differentiation and down-regulation of genes involved in migration and proliferation. Furthermore, *Malat1* knockdown also resulted in altered pre-mRNA splicing of many genes including critical transcription factors. Together, our data indicate that *Malat1* lncRNA regulates critical processes in breast cancer pathogenesis and represents a promising therapeutic target.

### Identification of lncRNAs Involved in Breast Cancer Progression

S. Diermeier, K.-C. Chang [in collaboration with S.M. Freier, F. Rigo, and C. Frank Bennett, Isis Pharmaceuticals]

Recent findings of the ENCODE consortium revealed that as much as ~80% of the human genome can be transcribed, in contrast to only 2% being translated into proteins. The noncoding RNA transcripts can be subdivided into several groups, with lncRNAs representing the largest and most diverse class. With breast cancer being the most frequent malignancy in women worldwide, we set out to investigate the role of lncRNAs as potential prognostic and therapeutic targets in mammary tumor formation, progression and metastasis. Here, we focus on the luminal-B and the *HER2/neu*-positive subtypes of breast cancer, as both are associated with high prevalence and bad prognosis. To identify the complement of lncRNAs with altered expression levels, we extracted carcinomas from mouse mammary tumor virus–polyomavirus middle-T antigen (*MMTV-PyMT*) and *MMTV-Neu-NDL* mice and performed a comparative RNA-Seq screen. The transcriptome of wild-type, nulliparous mammary glands served as control. We isolated RNA either directly from cryosections of the tumor tissue or from organotypic epithelial cultures that have been grown in a three-dimensional extracellular matrix. Our RNA-Seq results reveal drastic changes in both the coding and the lncRNA transcriptome. To narrow down the candidate list of lncRNAs with the highest therapeutic potential, we sorted the data computationally according to expression fold-change, intergenic genome location, and lncRNA size. Following these criteria, we identified a number of lncRNAs that are up-regulated in both the *PyMT* and the *HER2/neu* model, as well as some that are subtype-specific transcripts. These lncRNAs represent a diverse group in terms of expression levels, genomic locations, length, and biotypes. ASOs were designed that independently target several of the newly identified lncRNAs. Knockdown assays were performed in primary mammary tumor cells and organotypic epithelial cultures, resulting in a knockdown efficiency of 45%–99% for the tested transcripts. Within 6 days of ASO-mediated knockdown of several of the lncRNAs, we observed phenotypic changes of the organoids such as reduced cell migration and/or cell

growth. Future studies will focus on the molecular characterization of these lncRNAs and elucidation of their potential role(s) in mammary tumor progression and metastasis.

### Quantitative Analysis of Chromatin Interactions upon Copy-Number Variation at Mouse 4E2

C. Zepeda-Mendoza, M. Eckersley-Maslin [in collaboration with S. Mukhopadhyay and A. Mills, Cold Spring Harbor Laboratory; E.S. Wong and P. Flicek, EMBL-EBI; N. Harder, R. Eils, and K. Rohr, University of Heidelberg; E. Splinter, E. de Wit, and W. de Laat, Hubrecht Institute; T. Ried, National Cancer Institute; A.M. Sengupta, Rutgers University]

Circular chromosome conformation capture (4C) has provided important insights into three-dimensional (3D) genome organization and its critical impact on the regulation of gene expression. We developed a new quantitative framework based on polymer physics for the analysis of paired-end high-throughput sequencing 4C (PE-4C-Seq) data. Our pipeline corrects for PE-4C-Seq data biases, normalizes data, and computes the contact probability between a viewpoint fragment and all other fragments in the chromosome, allowing the quantitative testing of differences in chromatin contacts and chromatin compaction. In particular, this method is especially suited for the analysis of complex chromosome modifications such as copy-number variants (CNVs). We applied this new approach to study chromatin interactions within and around a 4.3-Mb engineered deletion in mouse region 4E2, which is orthologous to human 1p36. Allele-specific PE-4C-Seq experiments of deletion and wild-type (WT) chromosomes revealed several local and

long-range differentially interacting regions (DIRs) in the deletion-containing chromosome (*Df*). A surprising reduction of local chromatin compaction was identified for viewpoints located near the telomeric end in the *Df* chromosome. This result points to a new higher-order architectural change upon the occurrence of the 4.3-Mb deletion, affecting more than 1 Mb of the terminal part of the *Df* chromosome. Additional experiments will test the hypothesis that chromatin-tethering points exist within or neighboring the CNV region (such as LADs), which could cause the intervening chromatin to extend upon the occurrence of the 4.3-Mb deletion. Interestingly, DIRs were enriched in differentially expressed (DE) genes as detected by RNA-Seq, as well as CTCF- and Smc1-binding sites—proteins that had been previously characterized as structural determinants of chromatin organization. Selected DIRs were validated by 3D DNA fluorescence in situ hybridization (FISH), demonstrating the robustness of our analysis framework. Together, these data revealed the complex effects that CNVs can exert on chromatin structure and function.

### PUBLICATIONS

- Bell RD, Long X, Lin M, Bergmann JH, Nanda V, Cowan SL, Zhou Q, Han Y, Spector DL, Zheng D, Miano JM. 2014. Identification and initial functional characterization of a human vascular cell enriched long non-coding RNA. *Arterioscler, Thromb Vasc Biol* **34**: 1249–1259.
- Bergmann JH, Spector D. 2014. Long non-coding RNAs: Modulators of nuclear structure and function. *Curr Opin Cell Biol* **26**: 10–18.
- Eckersley-Maslin MA, Spector DL. 2014. Random monoallelic expression: Regulating gene expression one allele at a time. *Trends Genet* **30**: 237–244.
- Eckersley-Maslin MA, Thybert D, Bergmann JH, Marioni JC, Flicek P, Spector DL. 2014. Random monoallelic gene expression increases upon embryonic stem cell differentiation. *Dev Cell* **28**: 351–365.



# MOLECULAR BIOLOGY OF PAPILOMAVIRUSES

A. Stenlund S. Schuck

The papillomaviruses are a group of viruses that infect and transform the basal epithelium, inducing proliferation of the cells at the site of infection. The resulting tumors (warts) are in most cases benign and will usually regress after some time, but certain types of human papillomaviruses (HPVs) give rise to tumors that are prone to progress toward malignancy, especially frequently cervical carcinoma. Indeed, HPV infection appears to be a necessary cause of invasive cervical carcinoma and thus represents one of the few firmly established links between viral infections and the development of cancer.

An impediment to the study of papillomaviruses has been the inability to define simple *in vitro* cell culture systems for analysis of the viral life cycle. These viruses normally require specialized differentiating cells that only with difficulty can be generated in cell culture. However, for a bovine papillomavirus (BPV-1), a convenient cell culture system exists where viral gene expression, oncogenic transformation, and viral DNA replication can be studied. Thus, BPV has become a useful model for these aspects of the viral life cycle. The DNA replication properties of the papillomaviruses show some unique and interesting characteristics. As part of their normal life cycle, these viruses can exist in a state of latency, which is characterized by maintenance of the viral DNA as a multi-copy plasmid in infected cells. The copy number of the viral DNA is tightly controlled, and the viral DNA is stably inherited under these conditions. Papillomaviruses therefore provide a unique opportunity to study plasmid replication in mammalian cells. In addition, the viral DNA replication machinery represents one of the most promising targets for antiviral therapy.

In previous years, we have reported the characterization of the papillomavirus replicon and the identification of the viral components that are required for viral DNA replication. In recent years, we have directed our attention toward the biochemical events that are associated with initiation of DNA replication. We are studying the biochemical properties of the viral E1 and E2 proteins and how these two proteins interact with the viral origin of DNA replication and with the

cellular replication machinery to generate initiation complexes. Our studies demonstrate that the E1 protein has all the characteristics of an initiator protein, including *ori* recognition, DNA-dependent ATPase activity, and DNA helicase activity. The transcription factor E2, whose precise function has remained more elusive, appears to serve largely as a loading factor for E1. Through direct physical interactions with both E1 and the *ori*, E2 provides sequence specificity for the formation of the initiation complex.

We are currently attempting to elucidate how the E1 and E2 proteins orchestrate the precise biochemical events that precede initiation of DNA replication at the viral *ori*. These events include binding of the initiator to the *ori*, the initial opening of the DNA duplex (melting), and the assembly and loading of the E1 replicative helicase at the replication fork. Our studies so far indicate that these activities are generated in an ordered process that involves the sequential assembly of E1 molecules on the *ori*. This sequential assembly generates different complexes with different properties that in turn recognize *ori*, destabilize the double helix, and function as the replicative DNA helicase.

## A Conserved Regulatory Module at the Carboxyl Terminus of the E1 Helicase Domain Controls E1 Helicase Assembly

---

The papillomavirus E1 protein is a representative of a subgroup of hexameric helicases. However, E1 is not a dedicated helicase but serves several additional functions in the viral life cycle. E1 takes part in initiation of viral DNA replication in several different ways: The E1 DNA-binding domain provides sequence-specific binding to the viral origin of replication (*ori*); a particular form of E1, a double trimer, can then generate local melting of the viral *ori*. Once the DNA is melted locally, an E1 double hexamer is formed on the melted DNA and unwinds the DNA in front of the replication fork.

Obviously, the utilization of a single polypeptide to provide multiple functions presents many challenges.

In addition to accommodating multiple different activities in one polypeptide, mechanisms have to exist that allow switching between the different activities, as well as mechanisms that prevent the different activities from interfering with each other. For example, in the DT to DH transition, E1 switches from site-specific DNA binding and complex formation dictated by E1-binding sites in the *ori*, to formation of hexamers dictated by the oligomerization domain and binding to single-stranded DNA (ssDNA). How this feat is accomplished is very poorly understood.

We were interested in identifying parts of the E1 protein that control the critical transition from DT to DH. We were intrigued by a fairly well-conserved 25–30-residue-long peptide at the carboxyl terminus of E1, which according to structural studies does not constitute a part of the AAA+ helicase domain structural unit, although it is attached to the helicase domain. We performed a mutational analysis of this peptide in the context of the carboxy-terminal half of the E1 protein (E1<sub>308–605</sub>). This fragment contains the oligomerization and helicase domains that have been structurally characterized.

We made several interesting observations using mutations in the carboxy-terminal peptide. When the entire peptide was removed, the resulting protein had wild-type activity for hexamer formation on ssDNA, ATPase, and DNA helicase activities, demonstrating that the peptide is not required for these processes. However, point mutations affecting the peptide had dramatic phenotypes. On the basis of the sequence

content, we divided the carboxy-terminal peptide into two components, an acidic region and a C-tail. Point mutations or deletions of the C-tail resulted in the loss of hexamer formation, ATPase activity, and DNA helicase activity, whereas further deletion of the acidic region resulted in restoration of wild-type activity. These results demonstrated that the acidic region functions as a negative regulator of hexamer formation, whereas the C-tail can counteract this negative effect.

Because these results demonstrated that the carboxy-terminal peptide is not required for the activities of the helicase and oligomerization domain fragment, we tested the activity of the carboxy-terminal peptide in the context of full-length E1. Interestingly, mutations in the acidic region resulted in the failure to form the E1 DT complex, indicating that the formation of the DT complex requires that the acidic region blocks hexamer formation. Although the DT can form in the absence of C-tail, the transition from the DT to the DH does take place unless the C-tail is present. These results strikingly define the carboxy-terminal peptide as having two specific functions: The acidic region is specifically required for the formation of a DT complex, and the acidic region is specifically required for the formation of the DH complex.

## PUBLICATIONS

Lee S-J, Syed S, Enemark EJ, Schuck S, Stenlund A, Ha T, Joshua-Tor L. 2014. Dynamic look at DNA unwinding by a replicative helicase. *Proc Natl Acad Sci* **111**: 827–835.

# DNA REPLICATION AND CHROMATIN INHERITANCE

**B. Stillman**    W. Gao    J. Jansen    S. Ni  
                      M. Hossain    A. Mazurek    Y.-J. Sheu  
                      N. Kara        S. Nandi       P. Wendel

The process of replicating the DNA each time a cell divides is fundamental to all living organisms and remains the main focus of our studies. Using the yeast *Saccharomyces cerevisiae* system that has long been a mainstay of our research, we continue to study how pre-replicative complexes (pre-RCs) are established at origins of DNA replication before they are used for the initiation of DNA synthesis during the S phase of the cell division cycle. These studies were made possible with the ability to reconstitute pre-RC assembly in vitro with purified proteins, but increasingly we have focused on structural studies of assembly intermediates in collaboration with Dr. Huilin Li at Brookhaven National Laboratory and Stony Brook University and Dr. Christian Speck at Imperial College in London. However, the yeast system is not the perfect paradigm for understanding DNA replication in all eukaryotes because we have found that there are considerable differences between yeast and human cells. Therefore, we have continued our detailed studies of pre-RC assembly in human cells, as well as the surprisingly diverse activities of the subunits of the human origin recognition complex (ORC), one of the critical pre-RC assembly proteins.

## Studies on the Assembly of the Pre-Replicative Complex

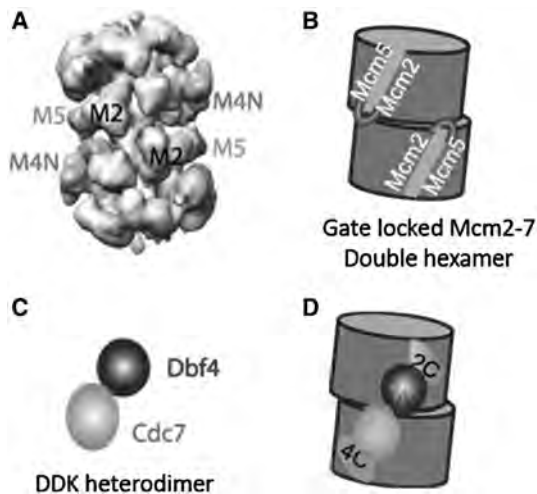
---

A number of years ago, we reconstituted the pre-RC that assembles at all origins of DNA replication prior to entry into S phase of the cell division cycle and initiation of DNA replication. In a continuing collaboration with Dr. Huilin Li and Dr. Christian Speck, we have investigated the structures of pre-RC assembly intermediates using the yeast *S. cerevisiae* system as a model for other eukaryotes. ORC, which in yeast is a stable six-subunit complex that exists throughout the entire cell division cycle, first binds to defined origins of DNA replication and recruits the Cdc6 protein

that is related in structure to the largest subunit of ORC, the Orc1 subunit. The ORC-Cdc6-DNA complex assembly involves significant re-arrangement of ORC subunits. This complex then recruits the hexameric Mcm2-7 proteins that are bound to the chaperone Cdt1, forming the ORC-Cdc6-Cdt1-Mcm2-7 (OCCM) complex on DNA. This protein complex is assembled in the presence of an analog of ATP that cannot be easily hydrolyzed so as to prevent ATPase activity and recruitment of the second Mcm2-7 hexamer, to form a double Mcm2-7 hexamer that is loaded onto the DNA and constitutes with ORC the pre-RC. Cryo-electron microscopy of the OCCM showed that it has one ORC-Cdc6 complex that sits on top of the Mcm2-7 hexamer with Cdt1 still bound to Mcm2-7. ORC-Cdc6 undergoes a substantial re-arrangement compared to the ORC-Cdc6-DNA complex. The ORC-Cdc6, when bound in the absence of Mcm2-7-Cdc1, is a flat ring of the six AAA+ proteins (Orc1, Orc2, Orc3, Orc4, Orc5, and Cdc6) formed with DNA going through the center. When the complex engages the Mcm2-7-Cdt1 complex, it distorts to form a slightly concave, spiral structure that encircles the DNA with a pitch of  $\sim 34 \text{ \AA}$ —the same as the pitch of the helix of double-stranded DNA. The structure is reminiscent of the structure of replication factor C (RFC) loading PCNA (proliferating cell nuclear antigen) that was determined by John Kuriyan, Michael O'Donnell, and colleagues, proteins that we discovered as components of the DNA replication fork in eukaryotes. The similarity of AAA+ protein loading ring-shaped protein onto DNA suggests a common mechanism whereby ATP-dependent proteins change the shape of other proteins and arrange them to surround the DNA double helix.

The product of the pre-RC assembly process is a double hexamer of Mcm2-7 proteins that surround the double helix and is the precursor of assembly of two active DNA-unwinding helicases that move with each DNA replication fork. Cryo-electron





**Figure 1.** Cryo-EM structure of the Mcm2-7 double hexamer (A) and models of how the structure locks on DNA (B) and how the DDK kinase binds (C,D). (From Sun et al 2014.)

microscopy (cryo-EM) analysis of the Mcm2-7 double hexamer shows that each hexamer assembles head to head with the double-stranded DNA passing through the middle of the barrel-shaped structure. One Mcm2-7 hexamer is twisted relative to the other, showing that the two gates between the Mcm2 and Mcm5 subunits of each hexamer are not aligned; this creates a lock that does not allow the DNA to exit the middle of the structure. The juxtaposition of the two hexamers in a head-to-head configuration brings together the amino termini of the Mcm2, Mcm4, and Mcm5 proteins that form the binding site of the DDK kinase that consists of two subunits, Cdc7 and Dbf4, that we have shown control, along with cyclin-dependent protein kinase (CDK), the initiation of DNA replication and assembly of the active helicase. The DDK kinase phosphorylates the Mcm4 subunit and relieves repression of the initiation of DNA replication by an amino-terminal fragment of the Mcm4 protein.

## The Origin Recognition Complex in Human Cells

ORC in human cells is far more complicated than the yeast paradigm suggests because it is not a single, six-subunit complex that is stable throughout the cell division cycle. We have shown previously that the large subunit of ORC, ORC1, is degraded at the G<sub>1</sub>- to

S-phase boundary by an SKP2-dependent ubiquitin ligase-mediated process and that assembly of ORC occurs during the G<sub>1</sub> phase of the cell division cycle. Recent studies using time-lapse microscopy have shown that in human cells expressing fluorescently tagged ORC1, the ORC1 binds to chromatin as cells enter into mitosis with the breakdown of the nuclear envelope. The chromatin-bound protein is inherited into the two daughter cells bound to chromatin and is present as the cell nucleus reassembles. ORC1 then displays dynamic nuclear localization patterns as cells progress through G<sub>1</sub> phase, and the ORC1 protein displays a punctate pattern late in G<sub>1</sub> phase with late-replicating regions of the genome. ORC1 binds the other ORC subunits in early- to mid-G<sub>1</sub> phase. The initial binding of Orc1 to mitotic chromosomes requires two regions of the carboxy-terminal domain of ORC1, sequences that show similarity to the mitotic chromosome binding domain within the transcriptional pioneer protein FOXA1. Two separate motifs within the ORC1 sequence that has similarity to FOXA1 are sufficient to bind mitotic chromosomes. Depletion of ORC1 prevents the assembly of the minichromosome maintenance (Mcm2-7) helicase proteins on chromatin in G<sub>1</sub> phase. These recent studies suggest that ORC1 acts as a nucleating center for ORC assembly and then pre-RC assembly by binding to mitotic chromosomes, followed by gradual removal from chromatin during G<sub>1</sub> phase. Because FOXA1 is a pioneer transcription factor that is inherited into daughter cells and is required for activation of gene transcription in the newly born cells, we suggest that, analogously, ORC1 is a pioneer factor for DNA replication. Because we have shown that ORC1 is also required for maintenance of heterochromatin and binds to the HP1 heterochromatin protein, it is possible that ORC1 binding to mitotic chromosomes also helps establish repressive chromatin in newly born daughter cells—a hypothesis we are testing.

## PUBLICATIONS

Hettmer S, Li Z, Billin AN, Barr FG, Cornelison DDW, Ehrlich AR, Guttridge DC, Hayes-Jordan A, Helman LJ, Houghton PJ, et al. 2014. Rhabdomyosarcoma: A critical review of current challenges and their implications for developing therapies. *Cold Spring Harb Perspect Med* 4: a025650. doi: 10.1101/cshperspect.a025650.

- Mazurek A, Youngkyu Park Y, Miething M, Wilkinson JE, Gillis J, Lowe SW, Vakoc CR, Stillman B. 2014. Acquired dependence of acute myeloid leukemia on the Dead-Box RNA helicase DDX5. *Cell Rep* **7**: 1–13.
- Sheu Y-J, Kinney JB, Lengronne A, Pasero P, Stillman B. 2014. Domain within the helicase subunit Mcm4 integrates multiple kinase signals to control DNA replication initiation and fork progression. *Proc Natl Acad Sci* **111**: E1899–E1908.
- Sun J, Fernandez-Cid A, Riera A, Tognetti S, Yuan Z, Stillman B, Speck C, Li H. 2014. Structural and mechanistic insights into Mcm2-7 double-hexamers assembly and function. *Genes Dev* **28**: 2291–2303.
- In Press*
- Hossain M, Stillman B. 2015. Meier Gorlin Syndrome. In *The initiation of DNA replication in eukaryotes* (ed. Kaplan DL). Springer, New York. (In Press.)
- Kara N, Hossain M, Prasanth SG, Stillman B. 2015. Orc1 binding to mitotic chromosomes precedes spatial patterning during G<sub>1</sub> phase and assembly of the origin recognition complex in human cells. *J Biol Chem* **290**: 12355–12369 (in press).
- Sun J, Yuan Z, Stillman B, Speck C, Li H. 2015. Structure and function studies of replication initiation factors. In *The initiation of DNA replication in eukaryotes* (ed. Kaplan DL). Springer, New York. (In Press.)

# CHROMATIN REGULATORS AS CANCER DEPENDENCIES

C.R. Vakoc   A. Bhagwat   F. Mercan   J. Minder   C. Shen   E. Wang  
A. Hohmann   J. Milazzo   J. Roe   J. Shi   Y. Xu

A universal attribute of cancer cells is an aberrant transcriptional program that supports a malignant phenotype. To initiate and thereafter maintain such oncogenic patterns of gene expression requires a diverse ensemble of regulatory factors, which includes sequence-specific transcription factors, chromatin-modifying enzymes, and various other cofactors. Our research directly evaluates the hypothesis that cancer cells can be eliminated by targeting certain transcriptional regulators, with an emphasis on the factors that modify chromatin. We utilize genetic screening strategies to reveal vulnerabilities in various forms of cancer and pursue detailed mechanistic analysis to reveal opportunities for therapeutic intervention.

## Mechanistic Studies of Brd4 in Acute Myeloid Leukemia Maintenance

A. Bhagwat, J. Roe, C. Shen [in collaboration with L. Joshua-Tor and D. Pappin, Cold Spring Harbor Laboratory]

Original studies from our laboratory in 2011 identified the BET bromodomain protein Brd4 (bromodomain containing protein 4) as a leukemia dependency and drug target in acute myeloid leukemia. In 2013, Brd4 inhibitors entered Phase I clinical trials in leukemia patients, and in 2014, the first complete responses were reported in leukemia patients receiving these agents. A major focus of our current work is to define the disease-relevant molecular mechanism of Brd4 function in leukemia. We have pursued this question using a host of genetic and biochemical approaches and have made substantial progress in defining a “Brd4 pathway” composed of critical *cis*- and *trans*-acting components that fuel uncontrolled cell growth in leukemia.

One area of initial focus has been to define the key *cis* elements used by Brd4 to regulate its important target genes in leukemia cells. Using chromatin immunoprecipitation followed by next-generation sequencing, we have completed a genome-wide survey of Brd4-occupied sites along the leukemia genome. One

remarkable observation from these studies has been the realization that a number of Brd4-dependent genes possess large clusters of distal enhancer elements that exhibit profound occupancy of Brd4. These large clusters of enhancers, recently termed superenhancers, appear to be the functionally relevant mode of regulation by Brd4 for a subset of its target genes. For example, *Myc* is among the most Brd4-dependent genes for its expression and *Myc* is located near one of the most profoundly Brd4-occupied enhancers in the leukemia genome. Interestingly, the *Myc* enhancers closely correspond to a region found previously as a site of recurrent focal amplification in leukemia patients, which has clarified the mechanism of transcriptional regulation performed by Brd4 to support leukemia maintenance.

An additional area of investigation has been to define the recruitment mechanism that directs Brd4 recruitment to specific *cis* elements in the leukemia genome. Through an epigenomic analysis of Brd4-occupied sites, we have uncovered a suite of hematopoietic transcription factors (PU.1, Fli1, Erg, Myb, C/EBP $\alpha$ , C/EBP $\beta$ ) that determine the locations of Brd4 occupancy. Remarkably, introducing these transcription factors into fibroblast cells can reprogram Brd4 occupancy into its leukemia-specific configuration. This effect occurs through an ability of the transcription factors to recruit the lysine acetyltransferase proteins p300/CBP to their occupied sites, which in turn acetylates various protein substrates at enhancers and promoters to promote bromodomain-dependent Brd4 recruitment. One of the key substrates of p300/CBP is likely to be histone H3 and H4; however, we have also derived evidence that Erg acetylation might also facilitate Brd4 recruitment. Collectively, this work reveals a chromatin-based signaling cascade composed of hematopoietic transcription factors, p300/CBP, and Brd4, which supports leukemia maintenance and is suppressed by BET bromodomain inhibition.

Another area of mechanistic investigation has been to identify candidate effectors of Brd4 in facilitating transcriptional activation. Through immunoprecipitation coupled with mass-spectrometry analysis, we have

identified several Brd4-associated proteins. By filtering of candidates through genetic criteria of having critical roles in leukemia maintenance, we have initiated studies of Nsd3, which is a SET domain-containing histone H3K36 methyltransferase that has not been well characterized previously. In ongoing studies, we have found that Nsd3 functions as a bona fide effector for Brd4 in supporting leukemia maintenance. We find that the genome-wide localization of Nsd3 matches that of Brd4 and that chemical inhibition of BET bromodomains leads to displacement of Nsd3 at promoter and enhancer regions. Unexpectedly, we have discovered that Nsd3 performs its essential function in leukemia cells independently of its SET domain, and we are currently investigating how Nsd3 promotes transcriptional activation independently of its catalytic domains.

We have also found that Brd4 utilizes the multisubunit Mediator complex as an effector at a subset of its occupied sites, which includes the Myc enhancers. Remarkably, a short treatment with BET inhibitors causes a large-scale displacement of several Mediator subunits from the Myc enhancers and from various other regulatory sites in the genome. In addition, we find the genetic perturbation of several Mediator subunits leads to leukemia cell differentiation, which mimics the effects of targeting Brd4. Collectively, these studies suggest that Brd4 facilitates Mediator recruitment to promote transcriptional activation of certain target genes. This represents progress in defining downstream components of Brd4 function that are functionally suppressed by BET bromodomain inhibition.

### Chemical Inhibition of SWI/SNF as a Therapeutic Approach in Acute Myeloid Leukemia

A. Hohmann, J. Minder, F. Mercan [in collaboration with L. Joshua-Tor, Cold Spring Harbor Laboratory]

Prior research from our laboratory has implicated the SWI/SNF complex as a unique vulnerability in acute myeloid leukemia. In the absence of any available drugs that target SWI/SNF, we have pursued the identification of therapeutic approaches that would be capable of modulating SWI/SNF function in leukemia cells. On the basis of proteomic and genetic evidence, we have identified one SWI/SNF subunit, called Brd9 (bromodomain containing protein 9) as an attractive target

for chemical inhibition. More specifically, our genetic evidence has implicated the bromodomain pocket of Brd9 as supporting leukemia maintenance and, based on this finding, chemistry collaborators have initiated the development of selective small-molecule inhibitors of Brd9. Notably, these inhibitors limit leukemia cell proliferation and mimic many of the cellular effects of RNA interference (RNAi)-based Brd9 inhibition. Furthermore, we have engineered a mutant Brd9 cDNA that no longer binds to the small molecule but retains functionality. Using this novel reagent, we have shown that Brd9 is the sole cellular target of this novel inhibitor that is responsible for antileukemia effects. Having established genetic and chemical evidence to support Brd9 as a drug target in leukemia, we now seek to advance Brd9 inhibition as a therapeutic strategy in leukemia. To this end, we are pursuing experiments that define the molecular function of Brd9 in leukemia cells, thus revealing how Brd9 contributes to the overall functionality of the SWI/SNF complex and to the maintenance of a leukemia-supporting gene regulatory network. In addition, we are investigating the genetic basis for primary sensitivity and resistance to Brd9 inhibition using existing human acute myeloid leukemia cell lines. These experiments will likely be complemented by CRISPR-based genetic screens that seek to identify genes that, when mutated, confer resistance to Brd9 inhibition, thus providing critical insight into the adaptive resistance mechanism. Finally, we will perform preclinical trials of Brd9 inhibitors in genetically engineered mice and in human xenograft models of disseminated acute myeloid leukemia. This research holds great promise in revealing basic mechanisms for an unstudied chromatin reader protein as well as a strong translational potential for exploring Brd9 inhibition as an epigenetic therapy in acute myeloid leukemia.

### TRIM33 Allows the Survival of B Lymphoblastic Leukemia Cells by Deactivating a Single Enhancer Element

E. Wang [in collaboration with J. Kinney, Cold Spring Harbor Laboratory]

We have recently extended our epigenetic screening platform into the high-risk blood cancer acute lymphoblastic leukemia of B-cell origin (B-ALL). This disease model is driven by the *BCR-ABL* oncogene and is associated with a dismal prognosis. We evaluated

the role of all chromatin regulators in the proliferation of these leukemia cells and have pinpointed a TRIM domain containing protein (TRIM33) in the pathogenesis of this disease. TRIM33 protein acts to support survival of B-ALL cells largely by repressing tissue-specific enhancer function. Notably, knocking down expression of TRIM33 has minimal effects on heterologous cell types. We have recently completed genome-wide studies evaluating chromatin occupancy of TRIM33 in B-ALL and are expanding these studies into other cell lineages. From this work, we hope to gain a general model for how TRIM33 governs leukemia-specific transcriptional regulation.

A remarkable attribute of TRIM33 in B-ALL is that it preferentially occupies a small number of *cis* elements across the genome. Two discrete regions are heavily occupied by TRIM33, which includes a distal upstream enhancer located 117 kb upstream of the Bim promoter. We find that TRIM33 is recruited by a dense cluster of PU.1, a hematopoietic transcription factor, which in turn deactivates the underlying enhancer. A stunning observation in this study is that CRISPR-based deletion of the Bim 117-kb element renders TRIM33 dispensable for B-ALL survival. This indicates that TRIM33 performs an essential function through a single genomic binding site. To our knowledge, this indicates a level of regulatory precision not observed previously for a mammalian regulator of transcription. Finally, our findings also nominate TRIM33 inhibition as a therapeutic strategy in B lymphoid malignancies.

### TFIID Complex as an Oncogenic Regulatory Hub in Leukemia

Y. Xu

A major theme in our research investigating chromatin regulators as cancer dependencies is that major regulatory hubs in transcriptional control can be perturbed in a selective manner to eradicate cancer cells, which reflects the essential role of such machineries to implement oncogenic transcriptional programs. On the basis of this rationale, we have begun to probe the requirement of TFIID in cancer maintenance. TFIID is a multisubunit complex composed of the TATA-binding protein (TBP) as well as various TBP-associated factors (TAFs) that function in promoter sequence recognition and as a coactivator for DNA-binding

transcription factors. We have hypothesized that specific cancers might be hypersensitive to perturbation of certain TAF subunits of TFIID to elicit selective transcriptional effects on cancer-relevant genes. Custom short hairpin RNA (shRNA) libraries have been constructed targeting each TFIID subunit and have been introduced broadly into an assortment of different cancer cell lines. Out of this comparative screening has come the identification of a submodule of TFIID that uniquely supports the biology of leukemia. In future experiments, we will define genetically and biochemically how this submodule drives leukemia pathogenesis.

### A Domain-Focused CRISPR Strategy for Cancer Drug Target Discovery

J. Shi, J. Milazzo, E. Wang [in collaboration with J. Kinney, Cold Spring Harbor Laboratory]

CRISPR-Cas9 technology holds great promise for genetic screening and for the discovery of therapeutic targets. To achieve optimal performance in negative selection screens, it is critical for CRISPR to generate homozygous loss-of-function mutations in a highly efficient manner. We have developed a CRISPR strategy that exploits this principle and simultaneously reveals protein domains that support cancer maintenance. By targeting CRISPR-mediated mutagenesis to exons encoding functional protein domains, negative selection phenotypes are achieved that are an order of magnitude stronger than those observed through mutagenesis of 5' exons. We have also developed a deep sequencing-based method for target validation that effectively excludes off-target effects. This approach has been benchmarked by mutagenizing lysine methyltransferase, lysine demethylase, bromodomain, ATPase, lysine deacetylase, and lysine acetyltransferase domains in MLL-AF9 leukemia cells, which has confirmed known dependencies and identified additional disease requirements. A broad application of these procedures may allow a comprehensive annotation of targetable protein domains that sustain cancer cell viability.

### PUBLICATIONS

Alpatov R, Lesch BJ, Nakamoto-Kinoshita M, Blanco A, Chen S, Stützer A, Armache KJ, Simon MD, Xu C, Ali M, et al. 2014.

- A chromatin-dependent role of the fragile X mental retardation protein FMRP in the DNA damage response. *Cell* **157**: 869–881.
- Bhagwat AS, Vakoc CR. 2014. A new bump in the epigenetic landscape. *Mol Cell* **53**: 857–858.
- Chen C, Liu Y, Rappaport AR, Kitzing T, Schultz N, Zhao Z, Shroff AS, Dickins RA, Vakoc CR, Bradner JE, et al. 2014. MLL3 is a haploinsufficient 7q tumor suppressor in acute myeloid leukemia. *Cancer Cell* **25**: 652–665.
- Hohmann AF, Vakoc CR. 2014. A rationale to target the SWI/SNF complex for cancer therapy. *Trends Genet* **30**: 356–363.
- Mazurek A, Park Y, Miething C, Wilkinson JE, Gillis J, Lowe SW, Vakoc CR, Stillman B. 2014. Acquired dependence of acute myeloid leukemia on the DEAD-box RNA helicase DDX5. *Cell Rep* **7**: 1887–1899.
- Roe JS, Vakoc CR. 2014. C/EBP $\alpha$ : Critical at the origin of leukemic transformation *J Exp Med* **211**: 1–4.
- Shi J, Vakoc CR. 2014. The mechanisms behind the therapeutic activity of BET bromodomain inhibition. *Mol Cell* **54**: 728–736.
- Xu Y, Vakoc CR. 2014. Brd4 is on the move during inflammation. *Trends Cell Biol* **24**: 615–616.

## CANCER: GENETICS

---

**Christopher Hammell's** lab 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 identify the components that function in controlling normal developmental timing, they use the small nematode *Caenorhabditis elegans*, applying forward and reverse genetic approaches. This past year, the team identified the gene *LIN-42* as a key regulator of developmental timing. In contrast to the extreme robustness of cell-fate lineage in *C. elegans*, in which specification of developmental programs is hard-wired, 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 re-wiring events to occur, with the idea that these components would make ideal therapeutic targets to complement existing clinical strategies.

**Gregory Hannon** is a pioneer in the study of RNA interference (RNAi), a process in which double-stranded RNA molecules induce gene silencing. Hannon and colleagues have elucidated key elements of the RNAi machinery. During the past several years, the Hannon lab has focused on roles of small RNAs in germ cells, which tend to have the most elaborate set of small RNA pathways of any cell type. They have discovered an essential role for small RNAs, called Piwi-interacting RNAs (piRNAs), that are critical for proper oocyte development and guard the genome against transposable elements. The lab has conducted screens, one in the fruit fly germ line and another in somatic cells, to search for new components of the pathway that generates piRNAs. They have identified dozens of genes that are required for piRNA production, offering insight into how germ cells ensure genomic integrity. The Hannon lab also strives to understand the biology of cancer cells, with a focus on breast and pancreatic cancer. They have led the way in using RNAi to study cancer biology and genetics, generating libraries of short-hairpin RNAs that have been widely applied in gene silencing studies. These libraries can then be used to identify new therapeutic targets for specific disease subtypes. In addition, they are exploring the roles of small RNAs as oncogenes and tumor suppressors and are using genetic approaches to understand the biology of resistance to currently used cancer therapies. Another research thrust of Hannon's team exploits the power of next-generation sequencing to understand the biology of the mammalian genome. Their efforts range from the identification of new classes of small RNAs to understanding human evolution and diversity, including an emphasis on the evolution of the epigenome and its role in driving cell-fate specification.

**Alea Mills** 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 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. The Mills lab has 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. This past year, the team identified the Chd5 protein as a key regulator of chromatin remodeling during sperm development. The lab 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 lab has generated a mouse model which 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  $\Delta$ Np63, reprograms stem cells of the skin to cause carcinoma development—the most prevalent form of human cancer. Modulation of these proteins may offer new ways to treat human malignancies in the future.

**Scott Powers'** work focuses on gene alterations that cause cancer and factors that influence responses to specific anticancer drugs. His lab uses technologies that probe the entire genome to identify candidate cancer genes and evaluate their functional role in cell transformation and tumor biology. They also use whole-genome technologies to guide development of novel cancer diagnostics and therapeutics. Using DNA copy-number analysis, the Powers group pinpoints novel amplified oncogenes and then applies functional studies to address the mechanisms by which they are implicated in oncogenesis. They have successfully applied this approach in breast, liver, colon, and lung cancers. Powers has also had an important role in the development of a distinctive CSHL approach to functional study of cancer genes. Called integrative oncogenomics, it is a rapid, large-scale screen for genes that are deleted or amplified in human cancers and suspected of being tumor suppressors, in the case of deletions, or oncogenes, in the case of amplifications.

**Michael Wigler's** work provides a new paradigm for understanding and exploring human disease. The Wigler lab 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 (with Josh Huang and Pavel Osten) to characterize neuronal subtypes, somatic mutation, and monoallelic expression. The Wigler lab'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.



# GENE REGULATION IN DISEASE AND DEVELOPMENT

C.M. Hammell C. Aguirre-Chen C. Carlston K. Doerfel

The major focus of our lab is to use the model organism *Caenorhabditis elegans* to understand how normal gene regulation contributes to proper development and how misregulation of expression leads to disease. Using forward genetics, we have made strides in identifying several novel genes and mechanisms that coordinate the specific timing of gene expression with the elaboration of normal temporal patterning. These include the identification of the *blmp-1* gene, encoding a ZnF transcription factor, which promotes the pulsatile expression of many *C. elegans* genes that control temporal cell fate. In addition, in the last year, we have initiated very fruitful collaborations with the McCombie and Wigler laboratories here at CSHL to develop an in vivo model for studying the roles of genes implicated in neuropsychiatric disorders in establishing and maintaining normal neuronal architecture.

## Oscillatory Gene Expression and Developmental Patterning

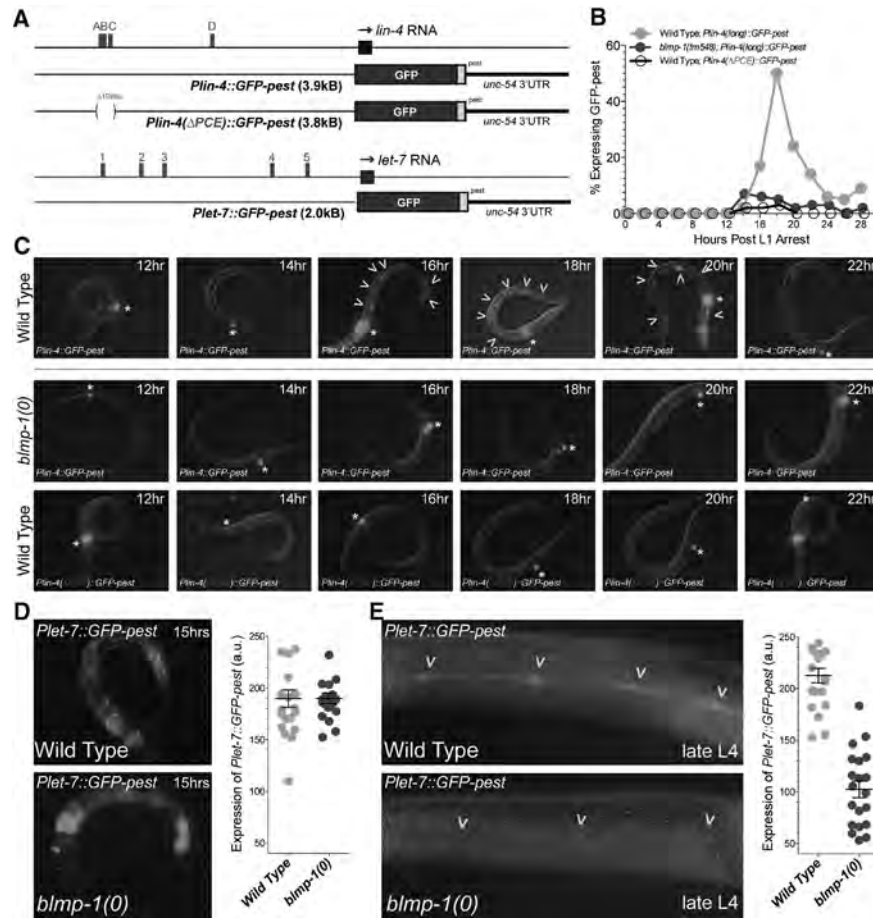
K. Doerfel, C.M. Hammell, C. Carlston

In normal development, the identity, behavior, and biological function of individual cells are coordinated with its neighbors from distinct tissues. The defined structure and specification of tissues endows them with unique functional and physiological properties that exceed the capabilities of individual cells. Setting up these architectures is an incredibly complex process and requires aspects of spatial and temporal gene regulation to be coordinated to control facets of cell proliferation and cell fate determination. Although the last few decades have illuminated a variety of conserved regulatory mechanisms that mediate spatial gene regulation, our understanding of how temporal patterning is set up lags behind. From what we do know, the genes and mechanisms that control developmental timing are very well conserved—and disruptions in this regulation result in a variety of human diseases, including cancer.

To identify genes that control temporal patterning, we have taken advantage of the model organism, *C. elegans*, which displays an essentially invariant cell lineage and is amenable to genetic studies. In the last year we have developed a variety of genetic strategies to identify regulatory components that modulate gene expression levels throughout development. During these studies, we found that many of the genes that control normal temporal patterning (including microRNAs [miRNAs]) are expressed in highly dynamic patterns. These expression patterns are repetitive and are correlated with the normal growth and behavior of developing animals; disruption of these patterns of expression leads to alterations in temporal cell fate.

The major finding from these studies has revealed that *C. elegans* encodes a developmental clock that demarcates both absolute and relative time and coordinates temporal gene expression with cell fate specification. One of the central components of this clock is *lin-42*, the *C. elegans* homolog of the human and *Drosophila Period* gene implicated in controlling circadian gene regulation. We found that a major function of *lin-42* during larval development is to maintain normal miRNA expression. Specifically, *lin-42* is required to dampen the oscillatory expression patterns of these genes at the transcriptional level, and *lin-42* mutants display precocious heterochronic phenotypes as a consequence of this defect. These findings were published in *PLoS Genetics*.

In an effort to identify the regulatory architecture required to generate oscillatory patterns of gene expression and additional components of this developmental clock, we identified the regulatory elements required for this pattern of transcription and initiated a candidate-based screen to identify *lin-42* suppressors. These efforts demonstrated that the oscillatory patterns of heterochronic miRNA expression required conserved *cis*-acting elements present in the promoters of oscillating genes, and we also identified the conserved ZnF transcription factor, BLMP-1, as a novel component of the heterochronic pathway (Fig. 1).



**Figure 1.** BLMP-1 and *cis*-regulatory DNA elements in miRNA promoters are required for oscillatory gene expression. (A) A schematic of the *lin-4* and *let-7* transcriptional reporters. (B) Transgenic animals containing the *lin-4::GFP-pest* transgene were synchronized at the L1 stage, allowed to begin postembryonic development, and monitored for GFP expression. (C) Micrographs of transgenic animals from *b* demonstrating that the wild-type *lin-4* promoter is sufficient for driving pulsatile transcription and that this activity requires promoter elements and the BLMP-1 ZnF transcription factor. (D,E) Oscillatory transcription of the *let-7::GFP-pest* reporter in intestine does not require *blmp-1* activity, but pulsatile expression patterns in hypodermal tissues do.

*blmp-1* mutant animals display mild heterochronic phenotypes and underaccumulate mature miRNAs during development. The *cis*-regulatory elements that are essential for oscillatory transcription are bound by BLMP-1 *in vitro*, and these elements are sufficient to generate oscillatory expression patterns of a heterologous gene (Fig. 2).

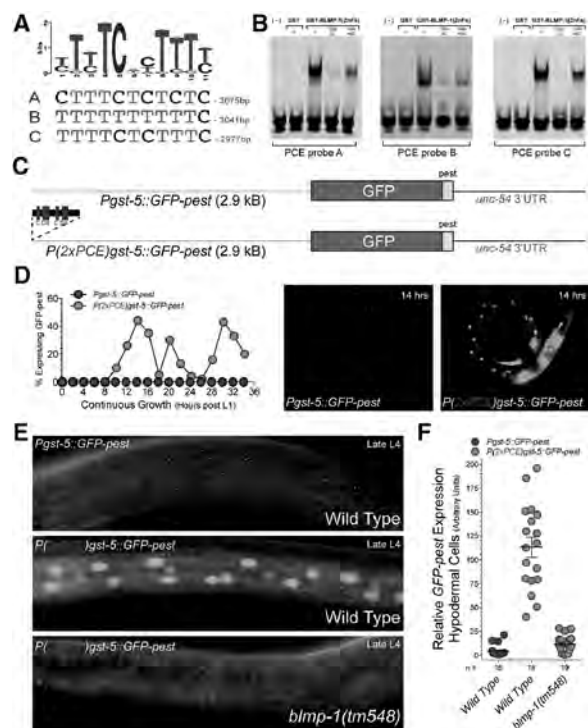
These results suggest that the transcriptional activation of heterochronic genes requires two modes of transcriptional activation: (1) a system that drives a low level of constitutive transcription and (2) a system that is dependent of BLMP-1 activation to drive highly pulsatile expression that is coupled to growth and nutritional status of developing animals. Finally, we

determined that the BLMP-1-dependent activation of heterochronic genes is essential for animals to recover normal temporal patterning after developmental diapause.

### In Vivo Screening of Genes Implicated in Human Neuropsychiatric Disorders to Identify Genes Required for the Establishment or Maintenance of Neuronal Architecture

C. Aguirre-Chen

High-throughput sequencing and genome-wide association studies of DNA derived from patients with



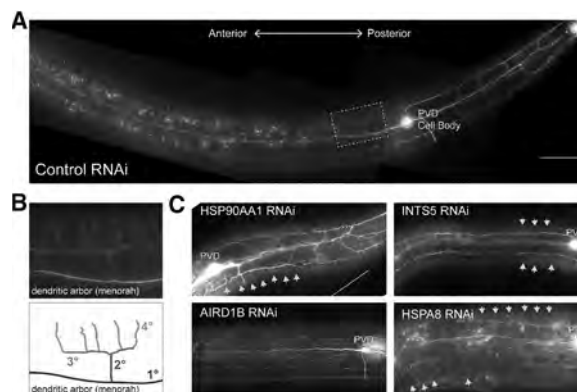
**Figure 2.** BLMP-1 binds to a conserved sequence found in the promoters of many genes that display pulsatile expression patterns, and these elements are sufficient for oscillatory transcription. (A) The experimentally derived BLMP-1 consensus sequence (PCE; pulse control element) and the three essential versions of PCE elements contained within the *lin-4* promoter. (B) The recombinant zinc finger domains of BLMP-1 directly bind the three *lin-4* elements in vitro. (C) A schematic of the *gst-5::GFP-pest* and the *(2xPCE)gst-5::GFP-pest* reporters. (D) Transgenic animals containing these reporters were monitored in a manner similar to that of the GFP-pest reporters in Figure 1B. (E) The 2xPCE element is also sufficient to drive pulsatile transcription in the hypodermis of later larval staged animals, and this activity requires BLMP-1.

various neuropsychiatric disorders have led to the identification of hundreds of separate mutations that are potentially causal. This rapid advance in the identification of inherited and de novo mutations for complex diseases such as autism and schizophrenia has created a huge bottleneck for understanding the roles of these candidate genes in normal and disease processes. This is in part due to the sheer number of genes that have already been identified, the complexity of the nervous system itself, and the expense of initiating model organism-based strategies for understanding the disease process.

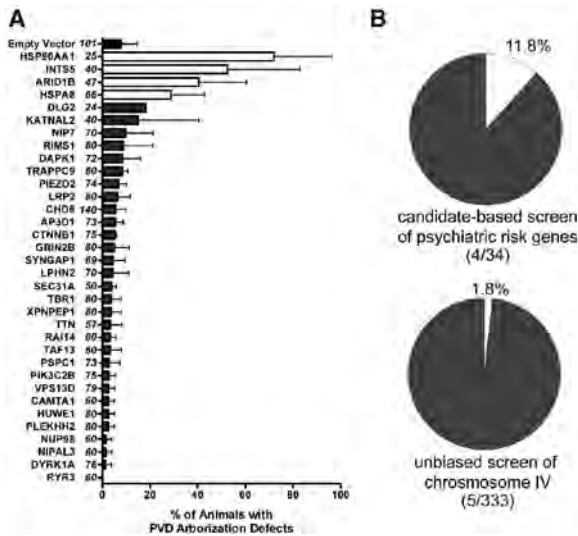
We, in collaboration with the McCombie and Wigler laboratories, hypothesized that one of many

potential mechanisms that may tie these neuropathologies together would be one in which the disease phenotype is caused by alterations in neuronal wiring. To test this hypothesis, we developed a rapid screening protocol to discover roles for the genes identified in the sequence-based family approaches in the establishment or maintenance of neuronal architecture.

This model takes advantage of the genetic and experimental tractability of *C. elegans* and allows gene function to be assayed in living animals. Our model incorporates transgenic animals that express green fluorescent protein (GFP) in two neurons: the PVD, located on the lateral sides of the animal. These neurons display highly elaborate and specific dendritic arbors that are both developmentally regulated and amenable to RNA interference (RNAi) (Fig. 3). To determine if mutation or inactivation of genes previously implicated in schizophrenia or autism results in structural defects of dendritic arbors, we depleted the *C. elegans* homologs of these genes in transgenic animals via RNAi. Working from a list of 50 candidates of which 34 had clear *C. elegans* homologs, knock down of four control genes produced clear dendritic arborization phenotypes (Figs. 3 and 4). When compared to an unbiased screen of *C. elegans* genes from chromosome IV, this candidate-based screen



**Figure 3.** An experimental assay to identify genes required for normal dendritic arborization. (A) Animals expressing the *wds52* transgene express GFP specifically in the PVD neurons. (B) The dendritic arbors of the PVD neuron display a highly repetitive and developmentally regulated “menorah” structure that is formed throughout postembryonic development. (C) Depletion of the *C. elegans* orthologs of several candidate genes results in specific disruptions in dendritic architecture. (Closed light grey arrowheads) Hyper-branching of the 4<sup>o</sup> dendrites; (grey line arrowheads) Hyper-branching of the 2<sup>o</sup> branches of the menorah structure.



**Figure 4.** A summary of the neuropsychiatric candidate gene screen using the *C. elegans* PVD reporter. (A) A bar graph depicting the average percentage of animals displaying abnormal dendritic arborization in the RNAi screen. (B) A comparison of the results from our screen and an unbiased screen of the *C. elegans* genes found on chromosome IV.

is enriched for positive hits by at least 10-fold. Most of the phenotypes produced by RNAi were recapitulated in animals that harbor loss of function alleles of the indicated genes. We are currently exploring other novel uses of this model, including using this system to establish epistasis between genes that alter dendritic arborization.

## PUBLICATIONS

- Aguirre-Chen C, Hammell CM. 2014. Cell death machinery makes life more robust. *Elife* 3.
- Goh WS, Seah JW, Harrison EJ, Chen C, Hammell CM, Hannon GJ. 2014. A genome-wide RNAi screen identifies factors required for distinct stages of *C. elegans* piRNA biogenesis. *Genes Dev* 28: 797–807.
- Perales R, King DM, Aguirre-Chen C, Hammell CM. 2014. LIN-42, the *Caenorhabditis elegans* PERIOD homolog, negatively regulates microRNA transcription. *PLoS Genet* 10: e1004486.
- Zinovyeva AY, Bouasker S, Simard MJ, Hammell CM, Ambros V. 2014. Mutations in conserved residues of the *C. elegans* microRNA Argonaute ALG-1 identify separable functions in ALG-1 miRISC loading and target repression. *PLoS Genet* 10: e1004286.

# RNA INTERFERENCE MECHANISMS AND APPLICATIONS

<b>G. Hannon</b>	G. Battistoni S. Boettcher D. Bressan K. Chang B. Czech M.J. Delas Vives	C. dos Santos O. El Demerdash N. Erard D. Fagegaltier I. Falcianti S. Fenoglio	W.S. Goh P. Guzzardo A. Haase E. Harrison E. Hodges S. Knott	M. Kudla E. Lee F. Muerdter J. Preall C. Rebbeck N. Rozhkov	E. Rozhkova L. Sabin E. Wagenblast V. Vagin K. Wasik Y. Yu	X. Zhou
------------------	---	---	---	--	---	---------

Our lab continues to work along three major lines of enquiry. First, we study small RNA biology with a particular focus on the roles of small RNAs in germ cells. Second, we work toward an understanding of unsolved mysteries in cancer, with a particular focus on breast and pancreatic cancer. Third, we continue to work to break technological barriers through the development of new approaches.

This has been a year of transition in the lab. Postdoctoral fellows John Preall and Vasily Vagin left for positions at Pfizer and CSHL, respectively. Students Marek Kudla, Paloma Guzzardo, and Felix Muerdter began postdoctoral positions, and Felix and Paloma also started life together as a married couple. With a great deal of sadness, we said goodbye to Jo Leonardo, who has worked alongside me as my PA for the past decade. Momma Jo, as she was often called, was a large part of the glue that held the lab together. She will be greatly missed, and we all wish her well in her new position at the California Institute of Technology, where at least she will find familiar faces in lab alumni.

We had a number of visiting students this year, as do we always. Eugene Jun Wen, Martin Fabry, Sara Gutierrez Angel, Annika Gable, Stephen Dickopf, Mar del Soto, and Jiaqui Gu all made important contributions to our work.

This year also marked the start of the transition of my laboratory to the University of Cambridge, after nearly 23 years at CSHL. Clare Rebbeck, Ben Czech, Iliaria Falcianti, Eugene Jun Wen, Nic Erard, Joaquina Delas Vives, Dario Bressan, Giorgia Battistoni, and Megan Keane have all come to join me in this new adventure, although Giorgia and Joaquina will remain students at CSHL.

Examples of ongoing work by members of the laboratory are presented below.

## Development of a Sequencing Protocol to Achieve Single-Base Resolution

G. Battistoni

---

DNA methylation is one of the best-studied epigenetic marks. RNA species are similarly methylated, but the functional role of this modification is largely unknown. Recent evidence suggests that *N*6-methyladenosine (m6A), one such RNA modification, has a pivotal role in gene expression regulation. The currently available sequencing techniques are not able to achieve single-base resolution. My project aims at the development of a sequencing protocol that could achieve such resolution. Analogously to bisulfite sequencing, where methylcytosine is chemically converted to uracil, our technique is based on the enzymatic conversion of m6A to inosine, which leads to an A > C transversion in the resulting cDNA and sequencing reads. During my time in the Hannon laboratory, I have purified the enzyme required for this protocol and optimized the conditions for the enzymatic conversion on single nucleosides. I have also produced some preliminary results using a methylated RNA oligonucleotide.

## Development of a Suite of Optogenetic Tools

D. Bressan

---

During the past year, we started the development of a suite of optogenetic tools to read, modify, and write the genetic and epigenetic status of a cell, which we named “laserTAG.” The project originated from the work I did during my thesis research to develop a light-dependent method for the space-specific purification of biomolecules. In this technique, a small light-sensitive molecule can be specifically bound to a target protein *in vivo* by a pulse of focused laser

light. The “tagged” protein can then be recovered and analyzed in several ways, merging the advantages of high-resolution *in vivo* analysis (i.e., microscopy) with those of a high-throughput technique such as DNA sequencing.

In the past year, we have expanded this technique to allow not only the recovery of proteins, but also their manipulation. We plan to use small chemical linkers to trigger protein interactions and drive cellular processes such as gene expression and to build a “photosensitive viral vector” able to introduce new genes into cells following light activation. This suite of tools would be transformative for several fields of science, most notably neuroscience.

Our project benefits from funding by the National Institutes of Health through both an R21 technology development grant and a major “B.R.A.I.N” initiative grant for transformative technologies in neuroscience.

## Target Discovery to Overcome Drug Resistance in Her2-Positive Breast Cancer

K. Chang

Breast cancers currently fall into three broad therapeutic categories: tumors positive for estrogen receptor (ER) expression, tumors positive for the HER2 (*erbB2/neu*) receptor tyrosine kinase oncogene, and tumors lacking ER and progesterone receptors as well as the HER2 alteration (also referred to as “triple-negative” tumors). Targeted therapies exist for both ER- and HER2-positive diseases. These are effective in many appropriately classified patients; however, some tumors fail to respond, and in many initially responsive patients, treatments eventually lose potency (acquired resistance).

HER2 is amplified and overexpressed in 20%–25% of human breast cancers. Patients whose breast cancers contain this alteration typically have an aggressive form of the disease with significantly shortened disease-free and overall survival. Trastuzumab-targeted therapy to block HER2 function is efficacious in both metastatic and early HER2-amplified breast cancers. Not all patients whose tumors contain the HER2 alteration respond to trastuzumab. Less than 35% of the patients with HER2-positive metastatic breast cancer respond to trastuzumab as a single agent, and a number of those that initially respond will acquire drug resistance within 24 to 36 months. Drug resistance has been correlated with genetic and/or

epigenetic alterations that activate pathways which reduce the dependence on established drivers observed in responsive tumors. Our objective is to use RNA interference (RNAi) as a path to breast cancer therapy.

We have carried out more than 24 genome-wide short hairpin RNA (shRNA) screens to find essential genes in tumor-derived cell lines representing all three therapeutic treatment groups of breast cancer (ER-positive, HER2-positive, and triple-negative disease) and to identify genetic determinants of therapy resistance for trastuzumab. Our goal is to mine this large data set of genome-wide breast cancer vulnerabilities to discover therapeutic targets to combat acquired trastuzumab resistance for HER2-driven breast cancer. To identify genes conferring secondary (acquired) resistance to trastuzumab from the data sets, we set a stringent statistical cutoff and filtered for shRNAs that depleted only upon trastuzumab treatment in the drug-resistant line, SKTR, but not in either of the two drug-sensitive lines, SKBR3 and EFM192A. This produced a list of 25 genes, which included those from phosphoinositol-3 kinase–mammalian target of rapamycin (PI3K-mTOR) signaling (PI4K2A, Raptor, insulin receptor, and EIF4A), RNA processing (PRPF8, U2AF1, and LSM6), mitotic checkpoint (BUB1B), and genes of unknown function. Identification of the insulin receptor and members of the PI3K-mTOR signaling pathway fulfills our expectation of finding these genes in this screen because they are known to be functionally associated with trastuzumab resistance. All 25 genes were validated to confer drug sensitivity in trastuzumab-resistant cells *in vitro*. TNFSF11/RANKL (ligand of the receptor activator of nuclear factor  $\kappa$ B), a gene of significant relevance to breast cancer, was one of two most highly depleted hits. Denosumab is a humanized monoclonal antibody designed to inhibit RANKL for treating various bone-related conditions. This drug was approved to treat giant cell tumors of the bone, breast cancer patients on adjuvant aromatase inhibitor therapy, and postmenopausal women with risk of osteoporosis, and to prevent skeletal-related events in patients with bone metastases from solid tumors. The dysregulation of the RANKL-RANK system is the major cause of osteoporosis in postmenopausal women. Appropriate RANKL signaling is also required for the formation of a lactating mammary gland, and both RANKL and RANK are expressed under the control of progesterone, prolactin, and the parathyroid hormone

protein-related peptide (PTHrP). Recent data implicate RANKL and RANK in the control of metastasis of breast cancer cells to the bone. We are currently expanding our validation studies to test the efficacy of this strategy to overcome trastuzumab resistance using a panel of drug-resistant/sensitive breast tumor cells *in vivo*. We will also use patient tumor-derived xenograft mouse models of trastuzumab resistance to functionally validate RANKL and other target genes to find the best strategy to overcome this disease.

### Molecular Characterization of piRNA Biogenesis in *Drosophila*

B. Czech

The Piwi-interacting RNA (piRNA) pathway provides an elegant defense mechanism that protects the genetic information of animal germ cells from the deleterious effects of molecular parasites known as transposable elements. This conserved small RNA silencing system consists of Piwi-clade Argonaute proteins and their associated 24- to 28-nucleotide RNA-binding partners, the piRNAs. piRNAs typically descend from acute transposon mobilization events and from piRNA clusters, which provide a genomic memory of ancestral transposon activity, through the coordinated activity of several processing factors. Although a candidate nuclease, called Zucchini, and its putative partner protein Minotaur have been identified, the precise molecular mechanism that converts transposon and cluster transcripts—but not coding mRNAs—into mature piRNAs remains largely elusive.

We have recently carried out a reverse genetic screen aimed at identifying the full repertoire of genes that participate in piRNA-mediated genome defense. Follow-up experiments have revealed a number of previously unknown factors that are crucial for piRNA biogenesis. Through genetic experiments combined with high-throughput sequencing, we identified two proteins, GASZ and PIMP, that are essential for the production of piRNAs from cluster transcripts. Interestingly, both proteins share their cellular localization on the outer surface of mitochondria with Zucchini and Minotaur. Using a combination of molecular, structural, genetic, and biochemical approaches, we are currently dissecting the specific roles of each of these factors individually, as well as their interplay, in order to understand piRNA processing in full detail.

### Mouse Mammary Glands

C.O. dos Santos

During the past year, we have focused on understanding the effects of pregnancy on the mammary gland epigenome. We performed whole-genome bisulfite sequencing of purified mouse mammary gland cell types before (nulliparous) and after pregnancy (parous). Analysis of nulliparous and parous methylomes revealed localized, but substantial, changes in DNA methylation. Many of these changes were localized near genes with known roles in milk production, cell proliferation, and apoptosis. In addition, we saw enrichment for STAT DNA motifs in regions that lost methylation during pregnancy. Further analyses demonstrated a great overlap of Stat5a ChIP-Seq (chromatin immunoprecipitation–high-throughput sequencing) peaks with parous-specific regions, suggesting that Stat activity heavily influenced the parous epigenome. Finally, we associated pregnancy-induced changes in DNA methylation with the kinetics of gene reactivation in a subsequent pregnancy, implying that the mammary gland harbors a long-term epigenetic memory of prior pregnancy cycles. We are currently investigating the correlation of these pregnancy-induced modifications with the risk of developing breast cancer, with the long-term goal of devising strategies to target these areas for cancer prevention and treatment.

### Combinatorial Drug Discovery through Two-Dimensional RNAi Screens

N. Erard

Loss-of-function RNAi screens are a powerful tool used to discover the oncogenic factors that allow cancer cells to thrive. These screens are driven by the premise that oncogenic changes alter the dependencies of cells, making them vulnerable to the loss of driving oncogenes and to addictions that their transformed state creates. Although this paradigm has proven fruitful in identifying some drug targets, there are patients who inevitably fail to respond or who acquire resistance to single-target therapies. Acquired resistance can be explained by the redundancy and plasticity of cellular pathways: Cells can rely on parallel pathways to escape a targeted therapy and continue their growth unimpeded.

To overcome the shortfalls of single-target therapy, it is necessary to turn toward combinational agents targeting critical nodes of multiple pathways. We have developed multiplexed multidimensional shRNA screens, where two genes are knocked down simultaneously in each cell. This strategy will allow us to identify gene pairs that could be targeted together to maximize therapeutic benefits.

As a proof of concept, a gene set of approximately 120 highly expressed druggable genes was targeted in four melanoma cell lines. The pairwise combinatorial library has a total complexity of ~150,000, making a pooled screen feasible. Hits in this depletion screen will be lethal pairs of genes that are consistently depleted across all cell lines. Such gene pairs will be validated in vitro, and a short list of most lethal pairs will be studied in vivo.

### Germline Genes and Functional Piwi/piRNAs in *Drosophila* Primary Cultures Relevant to Cancer

D. Fagegaltier

In a STARR Cancer Consortium collaboration, we have derived a large collection of immortalized cell lines from *Drosophila* embryos expressing an activated oncogene or mutant for a tumor suppressor gene.

High-content information contained in time series profiling, as opposed to single static measurements, potentially allows the detection of coregulated sets of genes that are functionally related. To evaluate the power of this approach in a multicellular organism, we generated transcriptome time series during the immortalization process of seven independent cell lines expressing an activated form of Ras. Stabilization of the cultures over time is associated with up-regulation of genes involved in cell cycle regulation, epigenetic state, and enrichment of neurogenic and myogenic tissue types. Using correlation analysis of the expression of the Twist muscle master transcription factor profile in the time series, we deconvoluted a transcriptional profile signature reminiscent of adult muscle progenitor cells (AMPs) and identified several candidate genes we further validated with potential roles in myogenesis/myoblast proliferation/self-maintenance. Our study demonstrates the power of RNA-Seq time series of evolving cell populations to identify functionally related genes.

In addition, we use these cell lines to uncover the role of germline PIWI proteins and piRNAs in

cancer: PIWI proteins are detected in several human and mouse tumors, yet no direct link has been established between cancer and a piRNA response to genome instability. We have identified and characterized in detail cell lines expressing a complete piRNA machinery including Aubergine (AUB) and Argonaute3 (AGO3), proteins known to amplify the piRNA response exclusively in ovarian germ cells. We demonstrated that the piRNA pathway in these tumorigenic cell lines of somatic origin is fully functional and fulfills a crucial role in regulating not only transposable elements (TEs) in the context of carcinogenesis, but numerous genes relevant to cancer as well. This work provides the first evidence of a functional role for the piRNA pathway in cancer and identifies Ras and SHW as two major signaling pathways upstream of the response. This in vitro system represents an invaluable tool to allow the identification in the long term of novel therapeutic targets.

### Role of piRNAs in Mouse Spermatogenesis

I. Falciatori

The piRNAs are a class of very diverse small noncoding RNAs present in the germline of all the organisms studied so far. They work in association with PIWI proteins and they are essential for fertility. In the mouse, dysregulation of the piRNA pathway invariably leads to male sterility, with slightly different phenotypes depending on the exact components that are perturbed. During mouse embryonic development, premeiotic piRNAs associate with PIWI proteins to form an RNA-based innate immune system able to defend the genome from the attack of genomic parasites such as transposable elements. This function involves both the direct cleavage of transposable element transcripts and the transcriptional silencing of transposons mediated by DNA methylation. Although it is well established that DNA methylation of certain young transposable elements is dependent on the Piwi protein MIWI2 and its associated piRNAs, the exact molecular mechanism by which this occurs has yet to be fully elucidated. We seek to identify protein partners or downstream targets of MIWI2 that can participate to this process.

MIWI2 is only expressed during a very short window of embryonic development, coinciding with a wave of de novo DNA methylation that occurs in



germ cells. The low number of germ cells at this stage poses quite a challenge to further understanding the molecular mechanisms behind the MIWI2-induced DNA methylation. To overcome this hurdle, we are optimizing the differentiation of mouse or human embryonic stem cells (ESCs) into germ cells. This will allow us to collect a higher number of germ cells at the right stage and to confirm or exclude the role of possible candidates in this process using shRNA technologies. As an alternative strategy, we produced transgenic mouse lines in which a MIWI2-GFP (green fluorescent protein) fusion is expressed under the control of the *Miwi* promoter. MIWI is expressed starting at meiosis, and it is associated with a class of piRNAs (meiotic piRNAs) that differs in origin and sequence from those usually associated with MIWI2. We thought that the ectopic expression of MIWI2 in meiotic and postmeiotic germ cells could have provided some insight into the targeting mechanism for DNA methylation because MIWI2 should now be directed to ectopic locations by its association with the meiotic piRNA. Unfortunately, although MIWI2 was expressed in the expected MIWI-positive cell types, the expression was rather low, and we confirmed that MIWI2 in this context was not loaded with pachytene piRNAs. We believe that this is due to the absence of proteins that are present during embryonic development and are necessary to direct the loading of piRNAs into MIWI2.

### Exploiting RNAi to Identify a Putative Target for Pancreatic Cancer Treatment

S. Fenoglio, M. Hemann, G. Hannon

Although pancreatic cancer occurs at low incidence in the United States, it is nevertheless currently the fourth leading cause of cancer-related death in the country. The high mortality rate is the result of poor diagnostic tools and ineffective treatment: In fact, most of the patients present at diagnosis with an advanced and metastatic disease. For this reason, resection of the tumor mass is impractical in the majority of cases, leaving chemotherapy and radiotherapy as the only options, which are applied with very low success rate as the improvement in survival achieved by any treatment is minimal.

Few therapeutic options are currently available for the treatment of pancreatic cancer: Gemcitabine is

administered as standard care; alternatively, a combination of four different drugs (folfinirox) can be used as first-line therapy for those patients who can tolerate the high toxicity of the treatment. Finally, the FDA recently approved the combination therapy of albumin-conjugated paclitaxel with gemcitabine. This last approach to pancreatic cancer treatment aims to improve drug delivery to the tumor and represents only one of the multiple formulations under clinical evaluation.

Because of the lethality of pancreatic cancer, it is crucial to achieve a better understanding of the disease in order to increase therapeutic efficacy. To address this urgent clinical need, we performed an *in vivo* RNAi screening in a model system of pancreatic cancer to identify genetic dependencies that could be exploited for the treatment of the disease. The technology of RNAi is a powerful tool to discover the factors necessary for tumor growth by specifically knocking down one protein per cell, revealing the phenotypic consequences of the loss of gene function.

The model system we adopted is a cancer cell line derived from a genetically engineered mouse model of pancreatic cancer, and it carries molecular alterations that are relevant to the disease. The oncogenic driver is mutant *kras*, which is present in >95% of patients, and the cells lack *tp53*, a gene that is altered in ~60% of cases. This cell line can be transplanted orthotopically in the pancreas of syngeneic immunocompetent mice, which allows for the evaluation of the immune system contribution to gene dependencies. We believe that it is necessary to include the tumor microenvironment as a variable in the experiments, because the interactions with the environment shape the behavior of cancer cells in response to treatment, here replicated by RNAi-mediated knockdown of genes. Accordingly, in our model, the tumors are resistant to gemcitabine *in vivo*, although we observe sensitivity when the cells are cultured *in vitro*.

We identified one gene, *cyclinE2*, that is required for tumor growth exclusively *in vivo*, and we are now investigating the mechanism responsible for this effect. Upon loss of *cyclinE2*, we observed changes in proliferation and induction of cellular senescence, a program that results in permanent cell cycle arrest. The phenotypes are a consequence of the role of *cyclinE2* in cell cycle regulation; what is surprising is the exclusive dependency on its function for a cellular process, the cell cycle, which has remarkable redundancy.

## piRNA-Directed Cleavage of Meiotic Transcripts Regulates Spermatogenesis

W.S. Goh

MIWI catalytic activity is required for spermatogenesis, indicating that piRNA-guided cleavage is critical for germ cell development. To identify meiotic piRNA targets, we augmented the mouse piRNA repertoire by introducing a human meiotic piRNA cluster. This triggered a spermatogenesis defect by inappropriately targeting the piRNA machinery to mouse mRNAs essential for germ cell development. Through analysis of such de novo targets, we derived a signature for pachytene piRNA target recognition. This enabled identification of both transposable elements and meiotically expressed protein-coding genes as targets of native piRNAs. Cleavage of genic targets began at the pachytene stage and resulted in progressive repression through meiosis. We also find evidence that pachytene piRNAs undergo the ping-pong cycle to drive the silencing of mRNA targets. Our data support the idea that meiotic piRNA populations must be strongly selected to enable successful spermatogenesis, both driving the response away from essential genes and directing the pathway toward mRNA targets that are regulated by small RNAs in meiotic cells.

## Optimization of shRNA Efficacy and the Creation of Fifth-Generation shRNA Reagents Targeting Human and Mouse Genomes

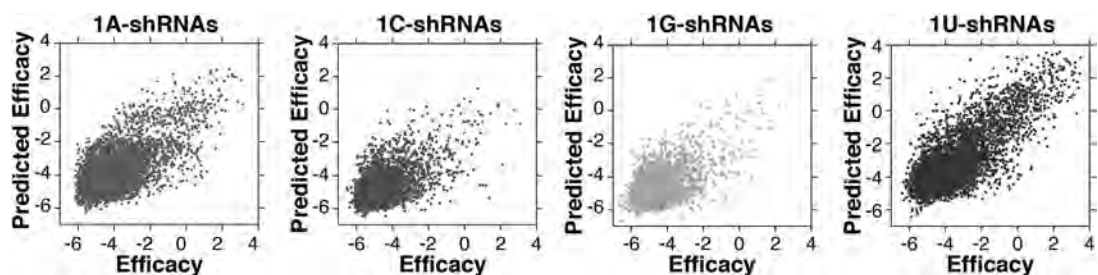
S. Knott

We have published our version-4 and version-5 builds of genome-wide shRNA libraries. As mentioned in previous reports, we performed the shRNA “sensor assay”

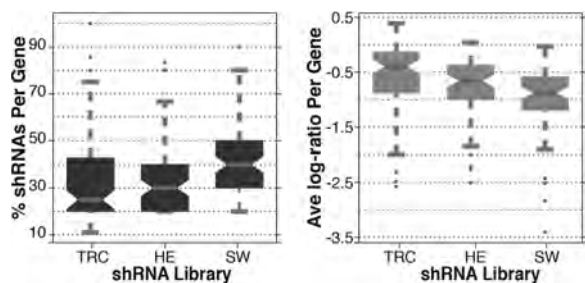
(Fellmann et al., *Mol. Cell* 41: 733 [2011]) to develop a data set of 250,000 shRNA efficacy data points. Using standard linear regression tools, we then determined what pairs and triplets of nucleotides within the shRNA guide were most predictive of shRNA strength. These variables were then included during machine learning to develop a computational algorithm that could predict well on a set of withheld data points (predicted scores showed 0.72 correlation with true scores; Fig. 1).

To benchmark the algorithm (which we call shERWOOD), we designed shRNAs (10 per gene) targeting a set of genes that, based on their gene ontology annotations, were likely to be essential for cell proliferation in vitro. We also included in this set the olfactory genes that are not expressed in cancer cells and could thus serve as negative controls. We also developed libraries targeting these genes using the current Broad TRC collection and our version-3 collection. These three libraries were then used to perform multiplexed shRNA screens in the pancreatic cancer cell line A385. To develop a set of essential genes, we identified all genes for which at least two hairpins per library “hit” in the screen. We then assessed library quality based on the percentage of shRNAs targeting these genes hit for each library and also the rate of depletion that was achieved on average for each of these genes. We found that the shERWOOD shRNAs targeting these genes had a higher hit rate and a greater level of depletion than the other two libraries (Fig. 2).

For many genes, it is difficult to predict potent shRNAs due to their length or GC content. Shorter genes provide a lesser search space and potent shRNAs are typically A/T-rich. On the basis of our current knowledge, the 5' bases of RISC (RNA-induced silencing complex)-loaded shRNA guides are inaccessible and do not contact the target sequence. Furthermore,



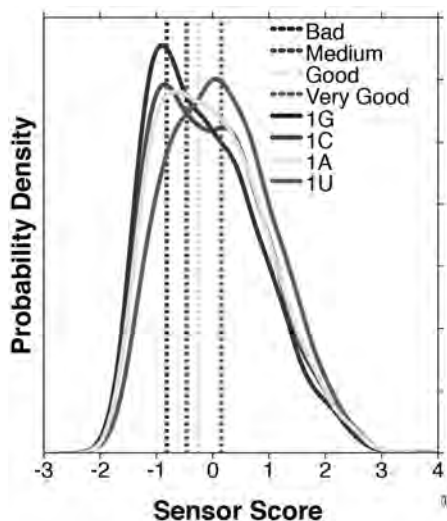
**Figure 1.** Predicted scores as compared to true scores for shRNAs included in the Fellmann et al. (2011) data set. Nine separate algorithms were trained, each time eliminating from the training set one of the nine genes in the data set. The plots indicate the predicted scores for each gene when it was withheld from training. Data points are separated based on the first base of the shRNA guide.



**Figure 2.** Percent shRNAs per essential gene depleted (*left*) and average level of depletion (*right*) for essential genes for each of the TRC, V3, and shERWOOD shRNA libraries.

it is well known that shRNAs with a uracil at this position tend to be the most potent. Thus, we hypothesized that for any target sequence, we could alter the endogenous guide to have a 5' U and that this would allow for potent shRNAs to be identified for short or GC-rich genes. To test this hypothesis, we performed a sensor assay using constructs that contained a 5' U endogenous and artificial guides and only endogenous targets. To our surprise, when the scores of these shRNAs were stratified based on their first base, the endogenous 5' U guides were most potent (Fig. 3). However, some of the nonendogenous 5' U shRNAs did provide potent knockdown.

On the basis of the above findings, we trained a machine learning algorithm to discern which of the

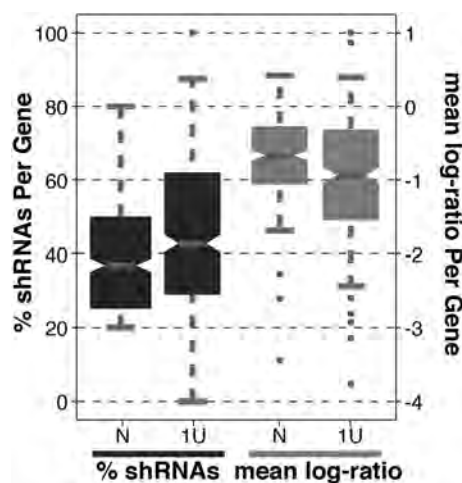


**Figure 3.** Sensor scores for endogenous 5' U shRNAs (1U) and for nonendogenous 5' U shRNAs (1A, 1C, and 1G). Also included in the plot are the average scores of benchmark shRNAs that were found to be bad, medium, good, and very good.

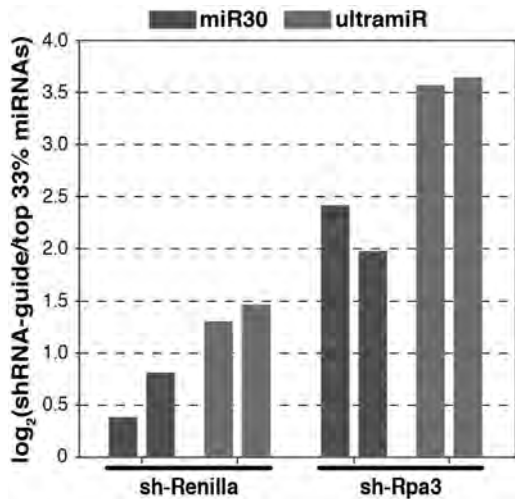
5' U shRNAs were most efficacious. We then applied this algorithm to an shRNA screen performed with only 5' U shRNAs (containing both endogenous and nonendogenous shRNAs) and to the screen described above. From this analysis, it was discovered that by applying this additional filter step, one can achieve a higher rate of shRNAs that “hit” and a greater degree of depletion (Fig. 4).

Finally, recent studies into microRNA (miRNA) sequence conservation have indicated that the placement of the restriction sites in the altered miR30 scaffold that we use to clone shRNAs with might be deleterious to their processing. To determine if this was the case, we developed a method by which shRNAs could be placed into the endogenous scaffold and named this setup UltramiR. To test if small RNA processing was improved with this altered scaffold, we performed small RNA sequencing on cells that had been infected with two different shRNAs, where each was harbored within the original miR30 scaffold or the new UltramiR scaffold. Processing was assessed as the log ratio of the shRNA guide count as compared to the 66th quantile of the endogenous miRNA counts. This analysis determined that the placement of shRNAs into the UltramiR scaffold results in a roughly twofold increase in processing (Fig. 5).

To better ascertain what advantage one has by using the UltramiR scaffold, we placed the shERWOOD

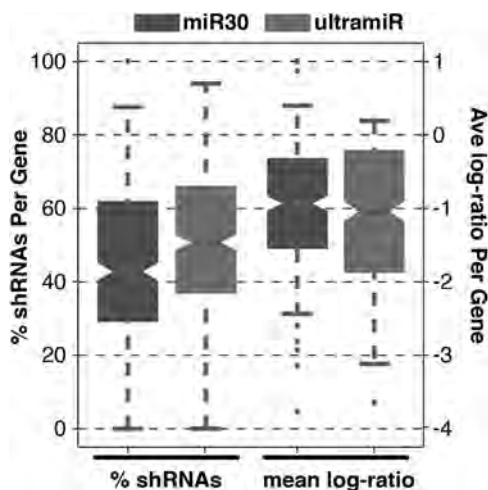


**Figure 4.** Percent shRNAs that were significantly depleted in the initial shERWOOD shRNA screen and in a screen combining this initial screen, a full 1U strategy, and the additional algorithmic filter. Average depletion of shRNAs targeting essential genes is also plotted on the left.



**Figure 5.** Measurement of the small RNA levels of processed shRNAs targeting Renilla Luciferase and mouse Rpa3 when they are placed in the altered miR30 scaffold or when they are placed into UltramiR. Counts were normalized by the average counts of endogenous miRNAs in the 66th quantile (based on counts).

shRNA libraries mentioned above into the altered scaffold and performed the shRNA screens again. Following screening, the 1U filter was applied and the results of the UltramiR were compared with the corresponding miR30 results. This analysis showed that when the shRNAs were placed into the UltramiR scaffold, a greater percentage of shRNAs per essential

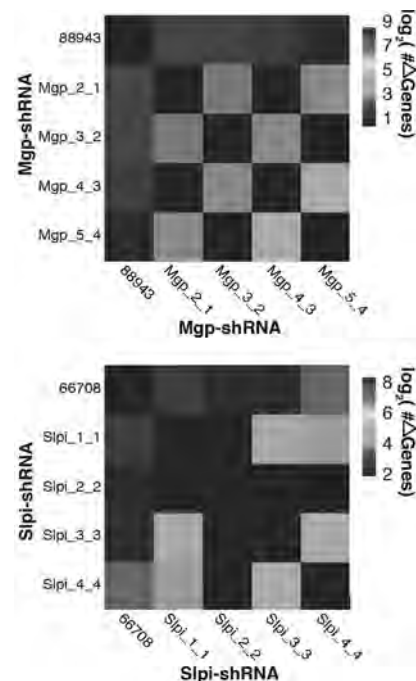


**Figure 6.** Percent shRNAs that were significantly depleted in the initial screen 1U + filter shRNA screen and corresponding UltramiR shRNA screen. Average depletion of shRNAs targeting essential genes is also plotted on the left.

gene significantly depleted and the average rate of depletion was greater (Fig. 6).

Finally, to assess how the increased processing achieved with the UltramiR scaffold affects off-targeting, we performed RNA-Seq on cells that had been infected with potent shERWOOD-UltramiR shRNAs targeting mouse Serpine2 and Slpi and with U6-driven TRC (The RNAi Consortium) shRNAs targeting the same genes. By comparing libraries from cells with the same genes targeted, we were able to distinguish between on-target and off-target effects by looking at the median profiles. shRNAs that had the most off-target effects would have corresponding libraries that were most distant from the median. Inter-library comparison determined that the cells infected with shERWOOD-UltramiR shRNAs differed by about 20 differentially expressed genes. In comparison, the U6-driven TRC constructs differed greatly from these profiles, with more than 100 genes differentially expressed (Fig. 7).

In summary, we have developed an shRNA selection strategy that produces potent shRNAs with low off-targeting effects. We have produced individually



**Figure 7.** The number of differentially expressed genes (more than twofold change and FDR < 0.05) identified through pairwise comparisons of the cell lines corresponding to Mgp and Slpi knockdown by the shERWOOD-1U selected shRNAs and TRC shRNAs 88943 and 66708.

arrayed sequence-verified versions of these libraries. The current human library contains ~50K molecules and the current mouse library contains ~35K molecules. They are housed within retroviral vectors where the shRNA is expressed from the LTR. The vectors also constitutively express neomycin and zsGreen. This work has been published by Knott et al. (2014).

## The Role of miRNAs in Mouse Embryonic Stem Cell Differentiation

M. Kudla

---

miRNAs are part of the posttranscriptional mechanism of gene expression that modulates protein translation. Despite multiple investigations of miRNAs taking part in development, the global image is still elusive. Using a genome-wide, next-generation sequencing HITS-CLIP (high-throughput sequencing-cross-linking immunoprecipitation) technique, we have generated snapshots of miRNA inhibition in the differentiation of embryonic stem (ES) cells to neuronal progenitor cells. Our data confirm that during this differentiation, miRNA increasingly targets brain-specific genes as the cells assume neuronal progenitor identity. This may be an indication of the involvement of miRNAs in the complex feedback mechanism required for establishing and maintaining neuronal identity. During differentiation, several miRNA families become up-regulated, a fact that is emphasized particularly by poor diversity of ES cell miRNAs that are dominated primarily by the miR-290 family. In particular, we observe up-regulation of miR-19 and miR-367 and observe the corresponding up-regulation of binding of those miRNAs to the transcripts, as indicated by the CLIP data. The subset of miR-19- and miR-367-targeted transcripts is also rich in genes involved in neuronal functions. This study is unique in its capacity for investigating both cause and effect in miRNA regulation, because we can track both miRNA level increase and the corresponding increase of binding to the matching transcripts.

## DCIS Project

C. Rebbeck

---

This past year, we have been working on building the foundations toward a long-term project on early-stage

breast cancer (ductal carcinoma in situ [DCIS]). In addition to securing a large grant for this project, we have almost completed a pilot study looking at DCIS lesions in 30 patients (as part of a collaboration with Duke University Medical Center). We are comparing their expression profiles and copy-number variations with other areas and cell types from the same patient and also across patients. Among these patients, some have progressed to invasive breast cancer and we have taken data from these invasive areas also. During this pilot study, we have developed and improved tissue-staining methods, clinical annotation collaborations, laser-capture microscopy methods, and sequencing techniques.

In addition, we have been working on a project involving dormant cancer; we have spent the past year working on low-input sequencing of fixed cells (1–10) and methods for identifying and isolating dormant cells. Part of this project has involved initiating and developing nanobodies, small antibodies made in alpacas, for the purpose of improved antibody staining.

I have also been working on tumor-associated macrophages, comparing gene expression profiles across these cells taken from *in vivo* tumors. I have identified the gene targets and have then proceeded to do *in vitro* knockdown experiments with shRNAs.

## A Molecular Characterization of Breast Cancer Progression and Metastasis

E. Wagenblast, S. Knott

---

Breast cancer is one of the most common cancers among women. About one in eight women will develop invasive breast cancer during their lifetime. Despite recent advances in treatment of advanced breast cancer, the majority of patients with metastatic disease will succumb to their disease. During metastasis, primary tumor cells evolve the capacity to intravasate into the lymphatic or vascular systems, extravasate into a target organ such as the lung, and colonize secondary sites. Previous studies have demonstrated that individual cells within complex populations show heterogeneity in their capacity to form metastatic lesions. However, no model system has yet emerged that allows genetically modified cells to be followed in a polyclonal context throughout each stage of metastatic disease progression. We developed a

mouse model of breast tumor heterogeneity and found that distinct clones within a mixed population display specialization—for example, dominating the primary tumor, contributing to metastatic populations, or showing tropism for entering the lymphatic or vascular systems. We correlated these stable properties to distinct gene expression profiles. Those clones that efficiently entered the vasculature expressed two secreted proteins, *Serpine2* and *Slpi*, which were necessary and sufficient to program these cells for vascular mimicry. Vascular mimicry is a phenomenon wherein highly aggressive tumor cells form channels to distribute blood to hypoxic (lacking oxygen) regions of the tumor. Our data indicate that *Serpine2* and *Slpi* drive the formation of extravascular networks, which ultimately lead to the formation of distant metastases. Thus, these two secreted proteins, and the phenotype they promote, may be broadly relevant as drivers of metastatic progression in human cancer.

### RNF17 Blocks Promiscuous Activity of PIWI Proteins in Mouse Testes

K. Wasik

PIWI proteins and their associated piRNAs protect germ cells from the activity of mobile genetic elements. Two classes of piRNAs—primary and secondary—are defined by their mechanisms of biogenesis. Primary piRNAs are processed directly from transcripts of piRNA cluster loci, whereas secondary piRNAs are generated in an adaptive amplification loop termed the ping-pong cycle. In mammals, piRNA populations are dynamic, shifting as male germ cells develop. Embryonic piRNAs consist of both primary and secondary species and are mainly directed toward transposons. In meiotic cells, the piRNA population is transposon-poor and largely restricted to primary piRNAs derived from pachytene piRNA clusters. The transition from the embryonic to the adult piRNA pathway is not well understood. We have shown that RNF17 shapes adult meiotic piRNA content by suppressing the production of secondary piRNAs. In the absence of RNF17, ping-pong occurs inappropriately in meiotic cells. Ping-pong initiates piRNA responses not only against transposons, but also against protein-coding genes and long noncoding RNAs (lncRNAs), including genes essential for germ cell development. Thus, the sterility of *Rnf17*

mutants may be a manifestation of a small-RNA-based autoimmune reaction.

### *Macrostomum lignano*: Building a Toolbox for a New Model Species

K. Wasik

Flatworms, such as *Macrostomum lignano*, have attracted scientific attention for centuries due to their astonishing, almost unheard of in the animal kingdom, regenerative abilities. Most flatworms can regenerate their entire body, or at least a large majority of damaged or amputated organs, thanks to an abundance of somatic stem cells, termed neoblasts. This regenerative potential is a rare occurrence among bilaterally symmetrical animals. To contribute to the understanding of the impressive regenerative abilities of flatworms, we, in collaboration with the Schatz lab here at CSHL, have generated and analyzed a reference genome sequence for *M. lignano*. Our assembly incorporates a 100× coverage of PacBio data, making it the most complete genome assembly in the Lophotrochozoan clade. We have also generated and annotated a *Macrostomum* transcriptome from intact and regenerating worms. Both will serve as a reference for future studies of this fascinating clade and as a resource to help advance the field of regenerative research, where the choice of model organisms is widely restricted by the fact that only a few organisms are able to regenerate multiple tissues in vivo. Stem cells possess great therapeutic potential and neoblasts provide a very accessible model to study somatic pluripotency.

### Piwi-Mediated Transcriptional Gene Silencing

Y. Yu

I have focused on understanding the mechanism of Piwi-mediated transcriptional gene silencing. In brief, we took advantage of a  $\lambda$ N/BoxB tethering system to mimic piRNA targeting. Surprisingly, artificial tethering of  $\lambda$ N-Piwi to a luciferase reporter containing BoxB sites failed to silence the expression of the reporter. Moreover, the reporter remained actively transcribed even when Piwi was tethered simultaneously with another piRNA silencing effector, Asterix/DmGTSF1.

This is in direct contrast with the observation that guiding Piwi complexes via artificial piRNAs can lead to repression of their complementary targets. Taken together, our data suggest that additional factor(s) and/or conformation changes are required to license Piwi complexes to repress transcription.

Through mining the data from several independent genome-wide RNAi screens for factors required for transposon silencing, we identified a novel protein-coding gene that can influence global transposon expression in a fashion similar to Piwi when knocked down by RNAi. The effect is not due to the defects of piRNA biogenesis because levels of piRNAs remained unchanged and Piwi maintained its nuclear localization. Strikingly, enforced tethering of this protein to nascent mRNA transcripts causes cotranscriptional silencing of the source locus and the deposition of repressive chromatin marks. Because of its function in the piRNA pathway and relationship to Asterix, we have named this protein Panoramix, the mentor who empowers Asterix to perform his feats of strength. Importantly, we found that both Eggless, an H3K9 methyltransferase, and HP1a, an H3K9me2/3 binder, are required for Panoramix-mediated transcriptional silencing. Finally, we demonstrated that Panoramix binds to genomic loci that harbor transposon sequences in a manner that depends upon Asterix. We propose that Panoramix forms one of the missing links between the piRNA pathway and the general silencing machinery that it recruits to enforce transcriptional repression to protect germline from deleterious transposons.

## Functional Characterization of PIWI Proteins and piRNAs in the Flatworm *Macrostomum lignano*

X. Zhou

The Argonaute/PIWI family of proteins is the essential component of RNAi pathways in eukaryotes. The PIWI subfamily is predominantly expressed in the germline and interacts with a class of small noncoding RNA known as piRNA in multiple species. piRNA pathways have crucial roles in germline development and maintenance of genomic integrity. However, in flatworms such as *M. lignano*, the germline is not the only cell type that expresses PIWI proteins. A PIWI protein, Macpiwi1, has been shown to be predominantly expressed in both germline and somatic stem

cells in *M. lignano*. The stem cells in flatworms, known as neoblasts, are the only proliferating cells in the organisms. They are crucial for postembryonic development, adult homeostasis, and the remarkable regeneration ability of flatworms. The previous study of Macpiwi1 has shed light on PIWI's importance in stem cell and germline maintenance. Knockdown of Macpiwi1 has been shown to lead to elimination of neoblasts, collapse of the germline, and death of the worms. However, the molecular mechanism of PIWI and piRNA pathways in flatworms remains unclear. First, we identified another PIWI protein, Macpiwi2. Like Macpiwi1, Macpiwi2 is predominantly expressed in the germline and somatic neoblasts. Knockdown of Macpiwi2 caused the death of worms, whereas the germline remained unaffected, indicating different mechanisms between Macpiwi1 and Macpiwi2. Macpiwi1 knockdown induced a global decrease of piRNA production, and interestingly, Macpiwi2 knockdown led to a global elevation of piRNAs. Repression of piRNA production from transposable element (TE)-related transcripts resulted in up-regulation of TEs, whereas the opposite trend was observed in Macpiwi2 knockdown. These results imply the importance of piRNA pathways and TE regulation in germline and stem cell functions in *M. lignano*.

## PUBLICATIONS

- Chang K, Marran K, Valentine A, Hannon GJ. 2014. Generation of transgenic *Drosophila* expressing shRNAs in the miR-1 backbone. *Cold Spring Harb Protoc*. doi:10.1101/pdb.prot080767.
- Deng L, Shang L, Bai S, Chen J, He X, Trevino RM, Chen S, Li X, Meng X, Yu B, et al. 2014. microRNA100 inhibits self-renewal of breast cancer stem-like cells and breast tumor development. *Cancer Res* **74**: 6648–6660.
- Fagegaltier D, König A, Gordon A, Lai EC, Gingeras RT, Hannon GJ, Shcherbata HR. 2014. A genome-wide survey of sexually dimorphic expression of *Drosophila* miRNAs identifies the steroid hormone-induced miRNA let-7 as a regulator of sexual identity. *Genetics* **198**: 647–668.
- Goh WS, Seah JW, Harrison EJ, Chen C, Hammell CM, Hannon GJ. 2014. A genome-wide RNAi screen identifies factors required for distinct stages of *C. elegans* piRNA biogenesis. *Genes Dev* **28**: 797–807.
- Knott SR, Maceli AR, Erard N, Chang K, Marran K, Zhou X, Gordon A, El Demerdash O, Wagenblast E, Kim S, et al. 2014. A computational algorithm to predict shRNA potency. *Mol Cell* **56**: 796–807.
- Lu D, Shang L, Bai S, Chen J, He X, Martin-Trevino R, Chen S, Li X, Meng X, Yu B, et al. 2014. microRNA 100 inhibits self-renewal of breast cancer stem-like cells and breast tumor development. *Cancer Res* **74**: 6648–6660.

- Malaspinas AS, Lao O, Schroeder H, Rasmussen M, Raghavan M, Moltke I, Campos PF, Sagredo FS, Rasmussen S, Gonclaves VF, et al. 2014. Two ancient human genomes reveal Polynesian ancestry among the indigenous Botocudos of Brazil. *Curr Biol* **24**: R1035–R1037.
- Meikar O, Vagin VV, Chalmel F, Sostar K, Lardenois A, Hammell M, Jin Y, Da Ros M, Wasik KA, Toppari J, Hannon GJ, Kotaja N. 2014. An atlas of chromatoid body components. *RNA* **20**: 483–495.
- Molaro A, Falciatori I, Hodges E, Aravin AA, Marran K, Rafii S, McCombie WR, Smith AD, Hannon GJ. 2014. Two waves of de novo methylation during mouse germ cell development. *Genes Dev* **28**: 1544–1549.
- Rechavi O, Hourri-Ze'evi L, Anava S, Goh WS, Kerk SY, Hannon GJ. 2014. Starvation-induced transgenerational inheritance of small RNAs in *C. elegans*. *Cell* **158**: 277–287.
- Stein P, Rozhkov NV, Li F, Cardenas FL, Davydenk O, Vandivier LE, Gregory BD, Hannon GJ, Schultz RM. 2014. Essential role for endogenous siRNAs during meiosis in mouse oocytes. *PLoS Genet* **11**: e1005013.
- Yang M, Haase AD, Huang FK, Coulis G, Rivera KD, Dickinson BC, Chang CJ, Pappin DJ, Neubert TA, Hannon GJ, Boivin B, Tonks NK. 2014. Dephosphorylation of tyrosine 393 in Argonaute 2 by protein tyrosine phosphatase 1B regulates gene silencing in oncogenic RAS-induced senescence. *Mol Cell* **55**: 782–790.



# TRACKING DOWN CANCER CELL BY CELL

J. Hicks   N. Anaparthi   H. Qi  
              T. Baslan        M. Riggs  
              H. Cox            A. Stepansky  
              S. Li

Our laboratory is focused on further developing high-throughput applications for the single-cell genomic technologies developed at Cold Spring Harbor during the last several years in the Wigler and Hicks research groups. This technology permits the rapid molecular profiling of clinical cancer samples, with the primary goal of understanding the cellular complexity of cancer initiation and progression, and, through these studies, to discover and implement molecular biomarkers into clinical practice.

A core motivation for our work on single-cell genomic analysis is the now accepted understanding that tumors and their resulting metastases are nearly always heterogeneous in both phenotype and genotype. The distinct genotypes define populations of cells, each potentially exhibiting a different response to anticancer therapy as well as the potential to follow novel evolutionary paths as the cancer progresses. Although “bulk” DNA sequencing of tumor tissue can provide important information about the gross makeup of a tumor, novel subpopulations will often be obscured. Single-cell genomics provides a means to resolve the complex mixtures of cells within a tumor and enable us to identify otherwise cryptic cell types that may emerge as treatment-resistant lineages as the patient undergoes therapy. The goal is to characterize the specific genetic changes as “biomarkers” that can provide prognostic information for the patient and establish guideposts that can help direct treatment decisions for the physician.

The ultimate goal of this work is the achievement of “precision medicine”: matching the right therapies with the right patients at the right time. To this end, our group has established collaborations with oncologists and pathologists at clinical cancer centers in order to apply molecular analysis—in particular, next-generation DNA sequencing (NGS) and copy-number variation (CNV) analysis—to profile cancer at its most fundamental level, the single cell.

## Single-Cell Profiling of Circulating Tumor Cells

H. Cox, M. Riggs, N. Anaparthi [in collaboration with P. Kuhn, University of Southern California]

---

Most patients being treated for advanced cancer have a small number of epithelial cells in their bloodstream that are apparently shed from the primary or metastatic sites. These so-called “circulating tumor cells” can be identified among tens of millions of white blood cells by immunohistochemical staining for specific protein markers and provide a potential window into the deep recesses of metastatic cancer. During the last year, we have optimized a method for genomic analysis of both cells and free DNA captured from blood samples using a novel approach developed by our collaborator, Dr. Peter Kuhn of the University of Southern California, that permits the observation of every single cell, including both epithelial and mesenchymal “stem-like” cells, in an unselected manner from an entire blood sample, along with so-called “cell-free” DNA. The combination of Dr. Kuhn’s blood sampling technology with our single genome technology constitutes a breakthrough in the analysis of cancer from blood samples.

In collaboration with the group at the University of Southern California, we have analyzed 800 blood samples from more than 500 patients in various stages of breast, prostate, melanoma, and ovarian cancer. In many cases, we have been able to document the progress of disease and the response to treatment over time during the course of treatment and are building up a database of treatment responses associated with biomarkers extracted from the genomic profiles. Two of these data sets for prostate (Dago et al. 2014) and melanoma (Ruiz et al. 2015) have been published recently and several more are papers are in preparation.

## Single-Cell Technology: Analysis of Archived FFPE Pathology Samples

T. Baslan, H. Cox [in collaboration with L. Martelotto and J. Reis-Filho, Memorial Sloan-Kettering Cancer Center]

Up to now, for reasons associated with preparation methods, single-cell profiling has been limited to cultured cells and fresh or fresh-frozen specimens. However, in clinical practice, surgical specimens and biopsy “cores” taken during diagnostic procedures are not preserved by freezing but by formaldehyde fixation and paraffin embedding (FFPE). This standard method preserves tissue for routine histopathology, but it causes enough chemical damage to DNA and RNA to make archived tissue inaccessible to our single-cell approach. This limitation made huge collections of archived cancer samples inaccessible to our methods for identifying tumor heterogeneity. However, as part of our biomarker discovery work in collaboration with the Reis-Filho laboratory in the Dept. of Pathology at Memorial Sloan-Kettering Cancer Center in New York, Dr. Luciano Martelotto used DNA-modifying enzymes to repair the DNA of 3–5-year-old FFPE samples to a level where isolated nuclei could be individually sorted into 96-well plates and the DNA amplified for sequencing. The method now yields results comparable to our previous data from frozen cells and opens up the entire archive of clinical trial samples to our method of analysis. The research was accepted for presentation at the 2015 Annual Meeting of the American Association of Cancer Research (AACR).

## Methods for Optimizing Single-Cell Genomic Profiling

T. Baslan, H. Cox

Single-cell genomic profiling has many valuable applications in cancer research, from analyzing the results of mouse model experiments to tracking patient response to drugs in clinical trials. However, one limit to this utility is in the cost of preparing cells and sequencing the DNA and RNA. With support from the Breast Cancer Research Foundation, Timour Baslan has pursued a project designed to significantly reduce the cost of single-cell copy-number profiling. In the course of this project, he has developed a new method that combines DNA amplification with Illumina

sequencing library preparation into a two-step process that cuts the time for preparing 96 single cells in half. He has applied this method to cancer cell lines and to breast tumor biopsy tissue, showing that it is superior to our previously used methods. Furthermore, he has performed analytical experiments to determine the exact minimum levels of DNA-sequencing depth that are required to permit the elucidation of useful copy-number profiles. Reducing DNA sequencing depth to the minimum necessary allows us to profile 96 cells on one sequencing lane, thereby tripling our data output for less cost overall.

An article describing these methods has been accepted for publication in *Genome Research* and will appear early in 2015.

## Biomarker Discovery: Retrospective CNV Analysis of Matched DCIS and Invasive Breast Cancer

With support from a 3-year grant from the Susan G. Komen Foundation, we have embarked upon a new breast cancer project in collaboration with Drs. Jorge Reis-Filho, Britta Weigel, and Tari King in the Department of Pathology at Memorial Sloan-Kettering Cancer Center. Many breast lesions discovered during mammographic screening fall into the precancerous category of ductal carcinoma in situ, or DCIS, most of which do not progress to invasive disease. This diagnosis leaves many patients with the agonizing choices that fall between no treatment and drastic treatment. The goal of this work is to identify biomarkers that would be able to distinguish indolent from malignant cancers at a very early stage, thereby saving both human and financial costs involved in treating growth that would not turn into life threatening disease.

We hypothesized that if DCIS is composed of mosaics of genetically diverse tumor cell clones and that the process of invasion is an evolutionary bottleneck, then examination of the subclones in invasive breast cancer (IBC) could reveal genetic events that may result in the acquisition of invasive behavior. To test this hypothesis and search for genetic biomarkers, we began a study of 50 DCIS and matched IBC collected by Dr. King and macrodissected in the Reis-Filho laboratory. These are being analyzed by a combination of comparative exome sequencing and single-cell profiling to determine whether there are mutations or

**Table 1.** Mutations identified in the DCIS and matched IBC.

ID	DCIS		IBC		Shared DCIS-IBC	
	Nonsynonymous	Indel	Nonsynonymous	Indel	Nonsynonymous	Indel
SK01	40	0	73	3	31	0
SK02	59	1	59	1	43	1
SK03	25	1	35	1	17	1
SK04	114	1	105	1	25	1
SK05	42	3	28	4	17	3
SK06	37	0	19	0	8	0

copy-number variants that are associated with invasive behavior.

We next focused on the mutational landscape of the six DCIS–IBC pairs. In the DCIS samples, a median of 41 nonsynonymous mutations (range 25–114), and in the IBC samples, a median of 47 nonsynonymous mutations (range 19–105) were found (Table 1).

In all six DCIS–IBC pairs analyzed, we identified somatic mutations that were present in both the DCIS and the adjacent IBC components. These encompassed somatic mutations affecting known “driver” breast cancer genes (Mansson et al. 1990), including an *AKT1* E17K mutation (case SK01), *PIK3CA* H1047R hot-spot mutations (cases SK02 and SK05), a *GATA3* S407fs mutation (SK03), a *MAP2K4* splice-site mutation (SK06), and a *TP53* G245S mutation (SK04). Interestingly, however, we have also identified mutations that were restricted either to the DCIS or to the IBC. For example, in case SK03, we found an *ATRX* S1154\* truncating mutation in the IBC, which could not be identified in the matched DCIS of the same patient. In case SK05, an *ALKS928fs* frameshift mutation and a *PKD2* R872\* mutation were restricted to the IBC, whereas an *ESR1* F461V mutation was only found in the DCIS.

These initial results indicate that the invasive and noninvasive forms of the disease separate early in the evolution of the lesion and support the notion that the change from indolent to invasive may be due to a genetic progression and that certain mutations may be genetic markers for progression. We are therefore encouraged to continue this study through the entire set of available tissue samples.

### Bioinformatics of Single-Cell Genomics

H. Qi, S. Li [in collaboration with J. Kendall (Wigler Lab), Cold Spring Harbor Laboratory]

A natural fallout from our efforts to reduce the cost of single-cell sequencing is the general problem of trying

to maximize the biological information that can be extracted from a limited amount of sequence data. Our single-cell method for copy-number profiling uses informatically created “bins” across the genome to aggregate sparse sequencing reads into a statistically useful number for each locus. Our resolution for copy-number breakpoints is thus limited to a range of 3–5 bin positions, depending on read count. Huan Qi, working with Jude Kendall, has attacked the problem of using a Bayesian statistics approach that can effectively assign a probability measure for the location of a breakpoint relative to any gene in the reference genome, independent of previously established bin boundaries. This work is being written up for publication.

In a separate project, Shanshan Li is working on ways to call single-nucleotide variant (SNV) mutations in DNA isolated from patient blood samples. Blood from cancer patients contains small amounts of tumor DNA mixed with variable amounts of normal DNA routinely shed from cells. The goal is to determine the amount of tumor DNA in any sample from sparse sequencing and use that value to determine the depth of sequencing that will be necessary to identify a potential novel mutation arising from the tumor at any desired confidence level. This project is just getting under way and is supported by a grant from the Breast Cancer Research Foundation.

### PUBLICATIONS

- Dago AE, Stepansky A, Carlsson A, Luttgen M, Kendall J, Baslan Kolatkar A, Wigler M, Bethel K, Gross ME, Hicks J, Kuhn P. 2014. Rapid phenotypic and genomic change in response to therapeutic pressure in prostate cancer. *PLoS One* **9**: e101777.
- Miething C, Scuoppo C, Bosbach B, Appelmann I, Nakitandwe J, Ma J, Wu G, Lintault L, Auer M, Premsrirut P, et al. 2014. PTEN action in leukemia dictated by the tissue microenvironment. *Nature* **520**: 402–406.
- Pratt ED, Stepansky A, Hicks J, Kirby BJ. 2014. Single cell copy number analysis of prostate cancer cells captured with geometrically enhanced differential immunocapture microdevices. *Anal Chem* **86**: 11013–11017.

# MAMMALIAN FUNCTIONAL GENOMICS

A. Mills   Y. Chang   A. Jaganathan  
D.-W. Hwang   Y.E. Son

Our research is focused on elucidating the genetic/epigenetic basis of neurodevelopmental syndromes and cancer. We have discovered new genes that impact these ailments, have revealed mechanisms for how the encoded proteins normally work, and determined what goes wrong during the disease process. These findings have had a major impact in the scientific community and have affected how clinicians analyze and treat patients with these syndromes.

Major discoveries:

- Identifying *p63* as a gene affecting development, aging, and cancer.
- Defining the genetic basis of autism.
- Discovering *CHD5* as a gene that prevents cancer.
- Determining that *Chd5* loss causes male infertility.

## Identifying *p63* as a Gene Affecting Development, Cancer, and Aging

---

We discovered *p63*, a gene that looked very similar to *p53*—a well-studied gene known to be defective in over half of all human cancers. Besides looking similar to *p53*, it was not clear how *p63* worked. We found that a lack of *p63* leads to aging. We discovered that *p63* is needed for stem cell renewal and that when *p63* is depleted, rapid aging features take place, including curvature of the spine, hair loss, and severe skin lesions. Yet there is a fine balance, as an excess of one version of *p63* ( $\Delta Np63\alpha$ ) causes carcinoma—the most prevalent type of human cancer. In contrast, we found that a different version of *p63* (TAp63) prevents cancer. We showed that TAp63 inhibits tumor growth and that it can do so even when *p53* is absent. This work was surprising, as it had always been believed that *p53* was absolutely essential for inhibiting cancer. Instead, we found that TAp63 can do the job alone. We are currently working on strategies to turn *p63* on, which might be useful in the clinic.

When we first identified *p63*, we found that it was needed for development, as its loss in mice causes

malformations of the limbs, skin, and palate. This work provided a clue to clinicians searching for the cause of human syndromes in which children have birth defects, including anomalies of the hands and feet, abnormal skin, and severe cleft palate. Our discovery that *p63* is needed for development led to the discovery that *p63* mutations cause seven different human syndromes involving birth defects affecting the limbs, skin, and palate. We generated the first mouse models for one of these syndromes, called ectrodactyly, ectodermal dysplasia, clefting (EEC) syndrome. These models yielded new insight into why some children with EEC syndrome 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. We are currently working to understand how *p63* regulates stem cell biology and how its perturbation leads to cancer.

## Discovering *CHD5* as a New Cancer-Preventing Gene

---

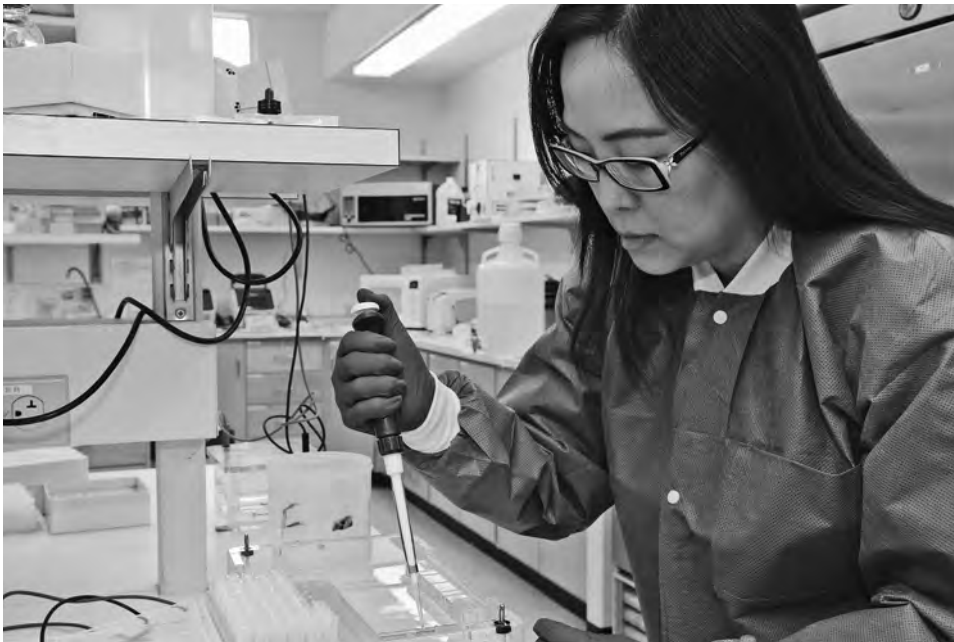
We discovered *CHD5* as a tumor suppressor mapping to human 1p36—a region of our genomes that is often deleted in cancer cells. 1p36 deletions occur in many different types of human cancers, including those of the epithelia, brain, and blood. Although this suggested that a cancer-suppressing gene resided in this region, its identity remained a mystery. To tackle this problem, we generated mice with deletions and duplications of the genomic region corresponding to 1p36. Using chromosome engineering technology—a strategy with which we can generate precise chromosome rearrangements in the mouse—we pinpointed a region of the genome with potent tumor suppressive activity. Using genetic and molecular approaches, we identified *CHD5* as the tumor suppressor gene in the region and found that its product worked as a “circuit breaker” for a cancer-preventing network. In addition, we discovered that *CHD5* was frequently deleted in human

glioma. We found that CHD5 uses its plant homeodomains to bind histone 3 and that this interaction is essential for Chd5 to function as a tumor suppressor. Our work paved the way for further discoveries, and it is now known that *CHD5* is mutated in human cancers of the breast, ovary, and prostate, as well as in melanoma, glioma, and neuroblastoma. Furthermore, recent reports indicate that *CHD5* status predicts whether anticancer therapy will be effective; indeed, patients with high levels of *CHD5* have much better overall survival than those with low levels. Within the past year, we have shown that Chd5 is essential for packaging DNA and that loss of Chd5 causes improperly packaged DNA that is prone to DNA damage (Li et al. 2014). Intriguingly, 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 male infertility (Li et al. 2014; Li and Mills 2014). We also

discovered that Chd5 is expressed highly in neurons and that Chd5 plays a pivotal role in the brain, suggesting that inappropriate DNA packaging contributes to neurodevelopmental syndromes such as autism (Li and Mills 2014b). We are currently delving deeper into the mechanisms whereby Chd5-mediated regulation of chromatin affects gene expression cascades that regulate neuronal stem cells and how deregulation of this process sets the stage for neurodevelopmental syndromes and cancer.

#### PUBLICATIONS

- Li W, Mills AA. 2014a. Packing for the journey: CHD5 remodels the genome. *Cell Cycle* **13**: 1833–1834.
- Li W, Mills AA. 2014b. Architects of the genome: CHD proteins in cancer, developmental disorders, and neurological syndromes. *Epigenomics* **6**: 381–395.
- Li W, Wu J, Kim S-Y, Zhao M, Hearn SA, Zhang MQ, Meistrich ML, Mills AA. 2014. Chd5 orchestrates chromatin remodeling during sperm development. *Nat Commun* **5**: 3812.



Y.E. Son

# CANCER GENES

S. Powers J. Li A. Mofunanya M. Rao

Our main focus has been to use genome-wide DNA copy-number profiling to identify recurrent amplified driver genes and to perform in vitro and in vivo analysis to validate their functional role in oncogenesis. Our discoveries in this area include gene amplification of *PPM1D* and *ACK1* in breast and prostate adenocarcinomas and more recently activation of *FGF19* and *POFUT1* by genomic amplification and overexpression in hepatocellular carcinoma. During the last few years, we have become increasingly interested in the emerging possibilities for developing new genome-based methods for cancer detection and diagnostics. To enable initiatives along these lines, our base of operations is now at Stony Brook Medicine, where we have developed collaborations with pathologists and oncologists to test genomic methods for early detection of cancer. Currently, there are three ongoing research projects in our laboratory, each being spearheaded by an individual lab member.

## Oncogenomic Screening for Linked Driver Dependencies in Breast Cancer

A. Mofunanya [in collaboration with K. Scott, Baylor University; B. Hahn, Broad Institute; and M. Hemann, Massachusetts Institute of Technology]

---

We are part of a network of laboratories (CTD<sup>2</sup>; <https://ocg.cancer.gov/programs/ctd2>) that are funded by the National Cancer Institute to work individually and collaboratively to use cancer genome data for the discovery and development of human cancer therapeutics. Many of the best targets for cancer treatment are encoded by the genes that when altered drive cancer progression and that when inhibited selectively block the survival of cancer cells with that particular alteration (linked driver-dependency, e.g., amplification of *HER2* drives breast cancer development; amplified *HER2* tumors are selectively dependent upon *HER2* function). Following this paradigm, I have performed functional genomic screening of aberrantly expressed candidate breast cancer genes along with controls (102 in total) and am now validating hits to determine which gene(s) to focus on for in-depth characterization.

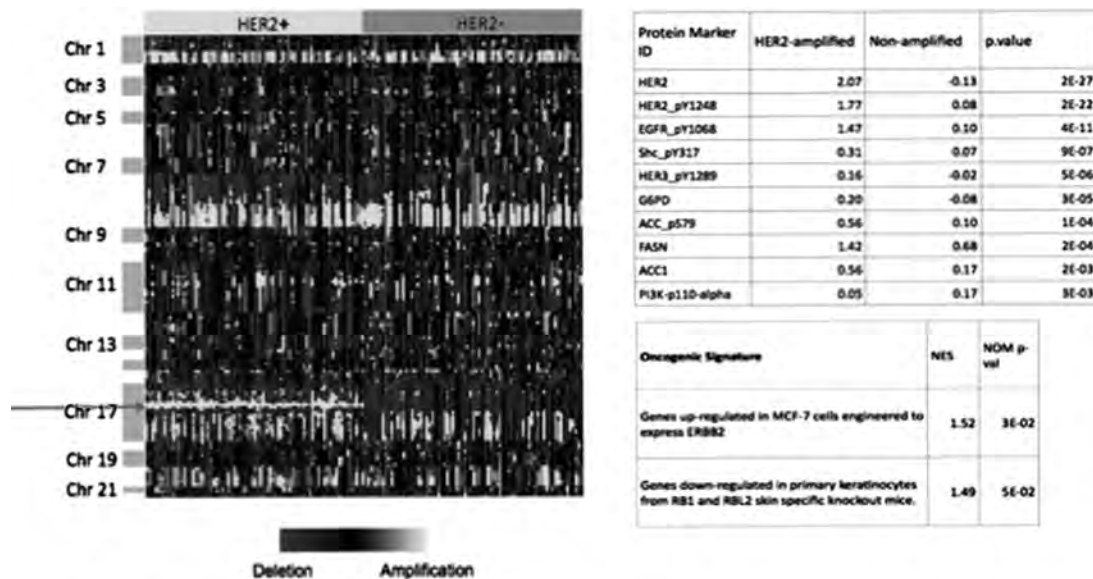
Two driver screens were performed: one using human breast epithelial cells (MCF10A; screened in three-dimensional culture) and the other using murine breast epithelial cells (NMuMG; screened in vivo). Approximately one-half of the genes significantly affected growth of MCF10A, most of them negatively and a smaller set positively, with *IL19* and *WWTR* being the highest-ranked. Surprisingly, although one-half of the genes also significantly affected the tumorigenicity of NMuMG cells, almost all did so negatively; only three increased tumorigenicity, including *ALDHB2*, *NKX2.1* (which negatively affected MCF10A), and *WWTR*. The small overlap of hits between the two systems as well as the opposite effects of *NKX2.1* may reflect the ductal, ER+ status of MCF10A and the basal, ER- status of NMuMG. Dependency screens were performed with custom-synthesized short hairpin RNA (shRNA) libraries with eight independent shRNAs designed to target each of the 102 genes. The libraries were transfected as pools into four different breast cancer cell lines and assayed for effects on both growth under normal cell culture conditions and anchorage-independent growth. Next-generation sequencing of the bar-coded shRNAs were used to estimate biological effects. To convert individual shRNA results into a gene-score, Stouffer's weighted Z-score was computed. *WWTR1* showed the highest dependency overall and is the strongest dependency in the triple-negative cell line CAL-120.

## Comparative Analysis Using FLLat Reveals Oncogenomic Properties of Individual DNA Copy-Number Features

J. Li [in collaboration with J. Pollack and R. Tibshirani, Stanford University; and W. Zhao and G. Mills, MD Anderson Cancer Center]

---

The ability to translate complex cancer genome data sets to improved diagnosis and treatment strategies remains challenging, particularly for acquired copy-number alterations (CNAs). We have been collaborating with R. Tibshirani on extending his FLLat modeling of



**Figure 1.** Comparison of closely matched tumors that differ only with regard to the *HER2* amplicon. Using linked proteomic and gene expression data, comparative analysis indicated that the types of changes associated with *HER2* overexpression in model systems are also found in human cancers using the comparative approach we developed.

DNA copy-number alterations to identify oncogenic properties of specific DNA copy-number alterations (a fused lasso latent feature model for analyzing multi-sample aCGH data; Nowak et al. *Biostatistics* 12: 776 [2011]). The purpose of FLLat is to increase the interpretability of massive DNA copy-number data sets by modeling them as a much smaller set of copy-number features. This modeling accurately preserves the important details of DNA copy-number profiles. One of our goals in using FLLat has been to develop auxiliary computational tools that would allow us to do comparative biology of closely matched tumors that differ only with regard to a single DNA copy-number feature. To test the method we developed, we applied it to the *HER2* amplicon which is probably the best understood copy-number feature in breast cancer (Fig. 1). These results suggest that comparative biology can lead to a deeper understanding of poorly understood copy number alterations.

### Genomic Test for Early Detection of High-Grade Serous Ovarian Cancer

M. Rao [in collaboration with M. Pearl and K. Shroyer, Stony Brook University]

We aim to develop a genomic test that can find high-grade serous ovarian cancer (HGS-OVCA) cells that

are shedding from the suspected precursor serous tubal intraepithelial carcinoma located at fallopian tube fimbriae into the uterine cavity. These precursors have been shown to have already undergone the two key features of high-grade serous ovarian cancer, dominant *TP53* mutations and DNA copy-number alterations (Hunter et al., *Am J Pathol* 184: 1871 [2014]). We are basing our test on the most up-to-date characterization of the genomic features of HGS-OVCA that distinguish it from normal, healthy tissue. HGS-OVCA is unusual among human cancer types in that virtually every tumor has undergone extensive copy-number alterations. Despite the extensive heterogeneity of copy-number breakpoints, there is uniformity in terms of peak regions amplified or deleted. By modeling and simulation studies, we have chosen 10 DNA probes that detect different chromosomal abnormalities, each of which distinguishes HGS-OVCA from normal tissue. Together, this panel of 10 probes is predicted to detect >99% of HGS-OVCA cases, using digital polymerase chain reaction (PCR). Our plan is to develop a test that would involve gentle rinsing of the uterine cavity with saline at the time of collection of a routine cervical Pap test specimen. Cells collected by this approach would subsequently be analyzed for genomic aberrations using highly sensitive single-molecule DNA tests with our panel of 10 probes. Because

TIC cells are shed from the fallopian tube into the uterine cavity, developing very sensitive methods to detect and subsequently treat these lesions should prevent most cases of HGS-OVCA.

## PUBLICATIONS

- Catanzaro JM, Sheshadri N, Pan JA, Sun Y, Shi C, Li J, Powers RS, Crawford HC, Zong WX. 2014. Oncogenic Ras induces inflammatory cytokine production by upregulating the squamous cell carcinoma antigens SerpinB3/B4. *Nat Commun* **5**: 3729.
- Rudalska R, Dauch D, Longerich T, McJunkin K, Wuestefeld T, Kang TW, Hohmeyer A, Pesic M, Leibold J, von Thun A, et al. 2014. In vivo RNAi screening identifies a mechanism of sorafenib resistance in liver cancer. *Nat Med* **20**: 1138–1146.
- In Press*
- Li J, Chanrion M, Sawey E, Wang T, Chow E, Tward A, Su Y, Xue W, Lucito R, Zender L, Lowe SW, Bishop JM, Powers S. 2015. Reciprocal interaction of Wnt and RXR- $\alpha$  pathways in hepatocyte development and hepatocellular carcinoma. *PLoS ONE* (in press).
- Powers S. 2015. Cooperation between MYC and companion 8q genes in hepatocarcinogenesis. *Hepatology* (in press).



# CANCER AND HUMAN GENETICS

M. Wigler	R. Aboukhalil	D. Esposito	J. Kendall	J. McIndoo	J. Troge
	J. Alexander	T. Forcier	Y.-H. Lee	S. Mukhopadhyay	Z. Wang
	J. Allen	E. Grabowska	A. Leotta	L. Rodgers	B. Yamrom
	P. Andrews	I. Hakker	B. Ma	M. Ronemus	C. Yoon
	M. Bekritsky	R. Kandasamy	S. Marks	J. Rosenbaum	L. Zhang

Our lab studies human cancer and genetic disorders from a genomics perspective, mainly in collaborative efforts with Dan Levy and Ivan Iossifov here at CSHL. The cancer effort focuses on breast and prostate cancer and involves collaborative clinical studies to discover mutational patterns predicting treatment response and outcome. The lab also develops methodologies for single-cell DNA and RNA analysis to detect cancer cells in tissues and bodily fluids. This has major potential applications to the early detection of cancer, monitoring recurrence, and response to therapy during and after treatment. Single-cell analysis has led to insights about the clonal evolution and heterogeneity of cancers and may lead to a better understanding of initiation, progression, and metastasis. These single-cell methods are also being applied to problems in neurobiology.

The research on genetic disease is aimed at determining the role of de novo mutation in pediatric disorders. The work has mainly involved participation in a large study of autism organized by the Simons Foundation Autism Research Initiative. Recently published work has confirmed and extended the team's earlier observations on the role of de novo copy-number variation in autism, with similar results in other disorders such as congenital heart disease. The large-scale whole-exome sequencing project was done in collaboration with W. Richard McCombie at CSHL, as well as Evan Eichler and Jay Shendure (University of Washington, Seattle) and Matthew State (University of California, San Francisco).

## The Molecular Genetics of Autism

---

We are examining the genetic basis of autism spectrum disorders (ASDs) by studying de novo (DN) mutations occurring in affected individuals. We have combined

the efforts of three teams that analyzed 2517 families from the Simons Simplex Collection (SSC). This collection consists of families with only one affected child (simplex) and is therefore not depleted for causal DN mutations. These data were published in Iossifov et al. (2014). By comparing affected to unaffected siblings, we estimate that 13% of DN missense mutations (that change protein sequences) and 42% of DN likely gene-disrupting (LGD) mutations contributed to 12% and 9% of diagnoses, respectively. Including large-scale copy-number changes, coding DN mutations contributed to ~30% of diagnosis simplex and 45% of female diagnoses. Virtually all LGD mutations occurred opposite normal alleles. LGD targets in affected females significantly overlapped the targets in males of lower IQ, but neither overlapped significantly with targets in males of higher IQ. We estimate that LGD mutation in about 400 genes can contribute to the "joint" class of affected females and males of lower IQ, with an overlapping and similar number of genes vulnerable to causative missense mutation. LGD targets in the joint class overlap with published targets for intellectual disability and schizophrenia and are enriched for chromatin modifiers, genes bound by the Fragile X mental retardation protein (FMRP), and genes expressed during embryonic development.

## Genetic Models for Autism

---

A limitation in studying the genetics of autism has been the lack of accurate models. Many genetic mechanisms have been proposed for autism, including transmitted and DN mutation, incorporating both small and large effects. On the basis of data from our genomics efforts, we have built a theoretical population genetic model and computational framework that can incorporate

various high-level epidemiological and genetic observations, such as incidence, gender bias, recurrence rate in families with one or multiply affected children, twin concordance, DN mutation rate, and other factors. We have developed the simplest model that successfully captures much of the epidemiology of autism, involving a single class of genes and three parameters: penetrance for males (i.e., the likelihood that an individual carrying a mutation will have the disorder), penetrance for females, and rate of new mutation. By constraining male incidence to an exact value, this model has just two free parameters. We are in the process of refining this model to take into account all observations from our data as well as our published data sets, allowing variants of small effect.

We observed in our molecular genetic study of the SSC that strong purifying selection—evidenced by a surprising lack of deleterious variants—was acting on many genes implicated in autism. We term such genes “vulnerable” to mutation. Taking advantage of our large data set, as well as those from other studies, we have expanded our observations. From the known rates of DN mutation, it is evident that only half of the targets of DN LGD mutation in autistic children are likely to be causal. On the basis of our analysis, we can use vulnerability to mutation to prioritize the list of autism candidate genes, because those that are depleted for rare LGDs in unaffected individuals are more likely to be related to autism.

## De Novo Mutation in Sporadic Human Disease

---

Congenital heart disease (CHD) is one of the most common malformations in humans—found in nearly 1% of live births—and is largely sporadic, with evidence of a significant genetic component. Utilizing the same approach as in our autism study, we studied 213 families in collaboration with Drs. Dorothy Warburton and Wendy Chung at Columbia University Medical Center. As in the other studies, we found strong evidence that DN CNVs may cause CHD, with greater than 10% of affected children having a germline variant—a rate five times higher than was seen in a control population. This study is now published (Warburton et al. 2014), and it suggests that rare DN mutation of high risk may contribute to many types of sporadic disorders in humans.

## Algorithms for Genomic Analysis

---

In collaboration with Dan Levy, Ivan Iossifov, and Mike Schatz, we developed tools for genome analysis by refining existing methods and creating novel algorithms when “off-the-shelf” software is inadequate. Insertions and deletions (indels), as well as other genome rearrangements, contribute significantly to genetic variation and disease. Many small indels can be detected by existing software, but in some genomic contexts—such as within microsatellites—indels can be difficult to discern. We have developed “Scalpel” based on local reference-free assembly, and a “split-read” method called “Mumdel” to find almost-unique matches to align sequence reads to the genome. Further development has allowed us to create a database of all discontinuities observed in samples. Within this population-based database, we can readily distinguish rare events of significance. We are now refining this algorithm to make it more generally applicable to different types of sequencing data.

## Molecular Methods

---

We have continuously worked on developing methods to perform measurements with greater resolution, or more efficiently, or to make new types of measurements. In the area of genomics, this includes algorithm development as described above as well as bench methods. In this regard, we have two ongoing projects. First, we are developing a method to maximize the genomic “real estate” available on the current generation of sequencing instruments. Termed SMASH sequencing, this consists of making genomic copy-number measurements based on sequencing and lowering experimental costs using sample barcodes, pooled samples, and DNA fragmentation. In combination, we use a custom algorithm to map and count sequence tags. Our results suggest that SMASH performs equally as well as whole-genome sequencing and is a dramatically cheaper option than any other method for accurately determining copy-number variation. In addition, it allows genome-wide coverage on a scale not available using hybridization-based technologies, such as microarrays.

Second, we have developed a bench protocol for mutational tagging using random deamination of cytosine at a tunable conversion rate—MuSeq.

DNA templates are imbued with a unique mutational identifier that allows individual molecules to be discerned and larger contigs to be constructed by matching unique patterns. This addresses some of the prominent shortcomings of existing short-read technology, namely, errors in sequencing and polymerase chain reaction (PCR), which limit utility in many applications. We have been able to create a stable and reproducible protocol, as well as algorithms that can successfully perform assembly. Potential applications range from single-cell transcriptome isoform profiling, precise measurements of copy number, error-free DNA sequencing, sensitive detection of rare populations, haplotype phasing, to long-range genome assembly. The latter is of particular interest, as it will allow us to uncover structure in genomic regions that are currently difficult to resolve, such as the 16p11.2 locus, the HLA region, and the Y chromosome.

### Single-Cell Analysis of Human Neurons

---

Neuronal cells of the brain are the most functionally diverse of any organ in the human body, but the totality of that diversity is not fully understood. Moreover, the diversity of neurons in the brain may be increased by somatic mutational mechanisms occurring during the last few cycles of neuronal differentiation—before neuronal fate is fixed. With recent advances in single-cell sequencing technology, it is possible to sequence DNA and RNA from single neurons in the brain. In collaboration with Dan Geschwind at the University of California, Los Angeles, we are exploring the roles of somatic mutation in neurons derived from embryonic human brains, categorizing subtypes and specific transcription patterns, and exploring the contribution of monoallelic expression to neuronal diversity and cognitive-behavioral variation. With Josh Huang here at CSHL, we are conducting a similar study with special focus on GABAergic interneurons.

### Genomic Landscape of Prostate Cancer

---

Prostate cancer (PCa) is the most common genitourinary malignancy among men. It often presents as a multifocal disease characterized by tremendous biologic heterogeneity with a variable clinical course

ranging from indolent to lethal disease. Current guidelines for prognostication and treatment strategies rely on clinic pathological parameters such as serum PSA levels and histological appearance. Developing genomic parameters such as copy-number variation, genome instability, and clonality will enhance current methods for prognostication and disease management. We are using a single-cell DNA (scDNA) genomics approach to obtain high-resolution profiles of the genetic alterations that occur in single prostatic cancer cells. Our methods are very general and are applicable to many kinds of cancer and many types of biological sampling.

In collaboration with Drs. Ashutosh Tewari of Weill Cornell Medical College and Herbert Lepor at New York University Medical School, we have shown that analysis of the genomes of just a few hundred cells can provide a landscape of the evolution of prostatic neoplasia. A comparison to Gleason scores suggests that early stages of the disease comprise cells with unstable genotypes, with little if any clustering into emergent clones with consensus genotypic markers. In contrast, advanced-stage disease displays not only cells with divergent and unstable genotypes, but also emergent clones of cells with common chromosomal deletions and amplifications that spread throughout the organ. Efforts are under way to determine if genome scores derived from single-cell genomic analysis may improve risk stratification in PCa. This study has continued from last year with a doubling of samples and confirmation of earlier results.

We are working out a new scDNA protocol that will be less expensive and also give us nucleotide-level information from each cell, with the aim of using rates of mutation to aid in prognostication and staging as well as phylogenetic mapping. In collaboration with Alex Krasnitz here at CSHL, we are developing a single-cell “viewer” that integrates scalable single-cell profile data with phylogeny, sector of origin, and histopathology. A prototype is already written and is being used routinely to observe and analyze data.

### PUBLICATIONS

- Dago AE, Stepansky A, Carlsson A, Luttgen M, Kendall J, Baslan Kolatkar A, Wigler M, Bethel K, Gross ME, Hicks J, Kuhn P. 2014. Rapid phenotypic and genomic change in response to therapeutic pressure in prostate cancer inferred by high content analysis of single circulating tumor cells. *PLoS One* **9**: e101777.
- Iossifov I, O’Roak BJ, Sanders S, Ronemus M, Krumm N, Levy D, Stessman H, Witherspoon K, Vives L, Patterson K. 2014. The

- contribution of *de novo* coding mutations to autism spectrum disorder. *Nature* **515**: 216–221.
- Levy D, Wigler M. 2014. Facilitated sequence counting and assembly by template mutagenesis. *Proc Natl Acad Sci* **111**: E4632–E4637.
- Narzisi G, O’Rawe JA, Iossifov I, Fang H, Lee Y-H, Wang Z, Wu Y, Lyon GJ, Wigler M, Schatz MC. 2014. Accurate *de novo* and transmitted indel detection in exome-capture data using microassembly. *Nat Methods* **11**: 1033–1036.
- Ronemus M, Iossifov I, Levy D, Wigler M. 2014. The role of *de novo* mutation in the genetics of autism spectrum disorders. *Nat Rev Genet* **15**: 133–141.
- Warburton D, Ronemus M, Kline J, Jobanputra V, Williams I, Anyane-Yeboah K, Chung W, Yu L, Wong N, Awad D, et al. 2014. The contribution of *de novo* and rare inherited copy number changes to congenital heart disease in an unselected sample of children with conotruncal defects or hypoplastic left heart disease. *Hum Genet* **133**: 11–27.



J. Troge

## CANCER: SIGNAL TRANSDUCTION

---

**Mikala Egeblad** and colleagues study tumors and, in particular, the contributions of 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 stromal cell types and the 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 lab'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 lab 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 lab recently 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). The Egeblad lab has collaborated with Scott Powers' group to understand how normal cells surrounding a tumor promote cancer growth. They have found that normal cells signal to tumors through multiple pathways, and blocking these signals together has the greatest effect on inhibiting tumor growth—offering a new strategy to fight cancer.

**Darryl Pappin's** lab develops chemical and computational methods for analysis of proteins and peptides. These are fundamental tools for proteomics, and they are vital in many fields of biological investigation. Proteins and peptides are typically analyzed via mass spectrometry, a method that involves fragmenting samples by colliding them with gas atoms in a vacuum. Masses of the resulting fragments are measured, and computer algorithms match the results with known or predicted molecules whose amino acid sequences are either known or inferred. Pappin has developed search engines for mass spectrometry data that enable investigators to sift hundreds of thousands of experimental spectra at a time for database matches. He also seeks to reduce sample complexity via an approach he calls chemical sorting. This includes the use of chelation to enrich phosphopeptides from the total peptide pool and the use of specific affinity-tagged small-molecule inhibitors to segregate classes of kinases or phosphatases for more specific mass spectroscopic analysis.

Despite their large variety of genetic abnormalities, cancer cells have been found to be extremely sensitive to the reversal of certain mutations. **Raffaella Sordella** and colleagues study why cells in certain cancers are responsive to the inhibition of one particular gene or gene product. Why, for instance, do non-small-cell lung cancer (NSCLC) cells that have a particular mutation in the EGFR (epidermal growth factor receptor) respond dramatically to its inhibition by the drug Tarceva? This occurs in 15%–20% of patients, the great majority of whom, within 1–3 years, develop resistance. Various mutations have been implicated in about half of resistant patients. Sordella and colleagues have discovered a new resistance mechanism in a subpopulation of NSCLC cells that are intrinsically resistant to Tarceva. These tumor cells were observed to secrete elevated amounts of a growth factor called transforming growth factor- $\beta$  (TGF- $\beta$ ), which in turn increases secretion of interleukin-6 (IL-6), an immune signaling molecule. Significantly, these effects were independent of the EGFR pathway. The team therefore hypothesizes that inflammation is one of the factors that can render a tumor cell resistant to treatment with Tarceva. In other work, Sordella collaborates with the Krainer lab to study whether alternative splicing has a role in the failure of

p53-mediated senescence to halt oncogenesis in certain lung cancers. In a discovery cited as one of 2014's "signaling breakthroughs of the year," Sordella's team discovered a new variant of p53, which they named p53 $\Psi$ . In contrast to normal p53, this splice variant has pro-growth and even pro-metastatic effects, which it produces by interacting with a protein at the membrane of the mitochondria.

**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<sub>2</sub>S) under conditions of ER 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.

**Lloyd Trotman's** recent research path begins at his discovery some years ago that the loss of a single copy of a master tumor suppressing gene called PTEN is sufficient to permit tumors to develop in animal models of prostate cancer. His team later found that complete loss of PTEN paradoxically triggers senescence, an arrested state that delays or blocks cancer development in affected cells. These findings explained why many patients only display partial loss of this tumor suppressor when diagnosed with prostate cancer. Now the team is researching ways to restore the PTEN protein levels in these patients. This therapeutic approach could slow disease progression and thus greatly reduce the need for surgical removal of the prostate or of similar drastic interventions that carry the risks of incontinence and impotence. Their second approach to combat prostate cancer is to model the lethal metastatic disease in genetically engineered mice. They are developing a novel approach, RapidCaP, that allows for quick generation and visualization of metastatic disease. The efficacy of existing and novel late-stage therapies, such as antihormonal therapy, can then be tested and optimized in these animals. At the same time, the Trotman lab is exploring the genome alterations associated with metastatic disease and with resistance to therapy. To this end, they use single- and multicell genome sequencing techniques developed at CSHL by Drs. Wigler and Hicks.

**David Tuveson's** lab uses mouse and human tissue models of neoplasia to explore the fundamental biology of these diseases and thereby identify new diagnostic and treatment strategies. His team's main focus is pancreatic cancer, a lethal malignancy that has eluded clinical solutions despite intensive study. The lab's approaches at CSHL run the gamut from designing new model systems of disease to inventing new therapeutic and diagnostic platforms for rapid evaluation in preclinical and clinical settings. For example, they have adopted a new method of culturing tissue fragments indefinitely in cell culture, enabling deep analysis with genetic and pharmacological probes. In addition, therapeutic experiments in mouse models have revealed an important role of

redox metabolism and stromal interactions on influencing therapeutic response. They have used the mouse model system to identify the mechanism of a promising drug treatment for pancreatic cancer. The lab found that using the drug in combination with more standard chemotherapeutic drugs stopped the tumor growth and lengthened life span for the mice, suggesting that the combinatorial therapy may help overcome the drug resistance that is so commonly found in cancers. Tuveson's lab also has a strong link to clinical trials locally and internationally, with confirmation in early phase trials the ultimate goal. Collectively, their strategy in the preclinical and clinical arena is codified as the "Cancer Therapeutics Initiative," and this initiative will provide these same approaches to the entire CSHL cancer community.

**Linda Van Aelst's** lab 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, lab members have obtained evidence that *OPHN1* has a critical role in mediating mGluR-LTD, 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. This year, the team demonstrated that *OPHN1* is involved in AMPAR trafficking; when it is mutated or missing, long-term potentiation is impaired. 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 results in an increase in the number of proliferating neuronal progenitors and defects in the genesis of neurons. The team has also shown that DOCK7 has a central regulatory role in the process that determines how and when a radial glial cell progenitor "decides" to either proliferate or give rise to cells that will differentiate into pyramidal neurons. They recently discovered that DOCK7 activation of ErbB4 is essential for normal chandelier cell development. These lines of research provide novel insight into mechanisms that coordinate the maintenance of the neural progenitor pool and neurogenesis.

**Hongwu Zheng's** lab aims to define the complex biology of malignant glioma pathogenesis, with the ultimate goal of translating the developed knowledge into patient benefits. Although eerily similar in terms of their self-renewal capacity and distinct phenotypic plasticity, malignant glioma cells conspicuously lack the terminal differentiation traits possessed by their normal counterparts—neural progenitors. With the use of multiple approaches combining human cancer genomics, animal modeling, and stem cell biology, Zheng has unraveled the causal relationship between aberrant differentiation and ensuing gliomagenesis. Perhaps more importantly, his team has demonstrated that forced restoration of differentiation capacity within glioma cells can drastically attenuate their tumorigenic potential. This finding fits well with the team's overall strategy, which is to target differentiation control pathways as a novel avenue for malignant glioma treatment. To this end, they have sought (1) to develop various animal models to recapitulate the human glioma pathogenesis and utilize them to trace and investigate in vivo tumor initiation/progression and (2) to identify key pathways/players controlling normal and neoplastic neural progenitor cell renewal and fate determination.

# THE INFLUENCE OF THE TUMOR MICROENVIRONMENT ON DRUG RESISTANCE AND METASTASIS

M. Egeblad   A. Almeida   M. Fein   V. Kuettnner   J. Park   R. Wysocki  
J. Cappellani   Y. Jia   L. Maiorino   M. Shields

Solid tumors are aberrant tissues, and like organs, they are composed of cancer cells and stroma. The stroma is the supportive framework of the organs and includes the extracellular matrix (ECM) and fibroblasts, adipocytes, cells of the vascular system, and immune cells. Interactions between epithelium and stroma are essential for normal organ development as well as for tumor progression. In solid tumors, the stroma is also known as the tumor microenvironment.

We study how the tumor microenvironment influences cancer cells in the context of drug resistance and metastasis. We use mouse models of breast and pancreatic cancer together with real-time spinning-disk microscopy in living mice (intravital imaging). This allows us to study how cancer cell proliferation, migration, and survival are influenced by stromal components in real time.

## The Influence of the Tumor Microenvironment on Response to Therapy

J. Jia, R. Wysocki, M. Fein, J. Cappellani

Breast cancer will recur in 20% of patients within 10 years after apparently successful treatment with chemotherapy. When resistance to chemotherapy develops, no other effective treatment options are available. How resistance develops is still an open question with three main theories proposed for the origin of the resistant cells: (1) from subpopulations of cells already within the tumor, (2) from cells with stem-cell-like properties, or (3) through mutations or epigenetic changes. Understanding which of these mechanisms drive resistance is critical to reducing recurrence.

Surprisingly, little is known about how cancer cells in intact tumors respond to classical chemotherapy, although these drugs have been used for decades. Most knowledge of the responses has been obtained from cell culture or xenograft animal experiments, but such experiments are often not predictive of drug responses in patients. We previously used intravital imaging to

show that monocytes are recruited to tumors ~30 h after chemotherapeutic treatment with doxorubicin, after cancer cell death was observed. We determined that monocytes were recruited to tumors through activation of the chemokine receptor CCR2. We further determined that the infiltration of these monocytes contributed to chemoresistance, as the effect of doxorubicin and cisplatin on tumors transplanted to mice lacking CCR2 was significantly better than that of tumors in normal mice.

We noted that when tumors recurred in mice lacking CCR2, they were of lower histological grade than tumors relapsing in wild-type mice. We have now found that the tumors relapsing in mice where host cells lack CCR2 are more differentiated toward a luminal epithelial cell type than the rapidly relapsing tumors in wild-type mice. We are currently investigating how activation and expression of transcription factors involved in breast epithelial cell differentiation are altered during response to chemotherapy. To identify the cellular and molecular mechanisms, we are also using lineage tracing combined with intravital imaging over weeks, using mammary imaging windows, together with microgenomics approaches.

## Effects of Myeloid Cells on Breast Cancer Metastasis

J. Park, R. Wysocki, M. Fein

The prognosis of metastatic breast cancer is poor. More than a century ago, Dr. Stephen Paget observed that metastases develop preferentially in certain organs, suggesting that factors external to cancer cells influence metastasis and that targeting such factors might reduce the ability of cancer cells to take seed. It is now recognized that communications between cancer cells through growth factors and cytokines have an important role in the process of metastasis.

Traditionally, studies of metastasis have relied primarily on measurements made at the endpoint of the



process, the establishment of micro- or macrometastases. However, the metastatic process is dynamic and characterized by the ability of cancer cells to move from one part of the body to another: Cells exit the primary tumor, invade the local tissue, enter blood or lymphatic vessels, and are transported to a distant site where they exit the vessels and move into the tissue. Thus, a different level of understanding of metastasis might be achieved using technologies that can follow these dynamic processes *in vivo*.

We use intravital imaging with mouse models of breast cancer and classical genetic manipulation to determine how interactions between cancer cells and stromal cells influence metastasis. We have compared the microenvironment of tumors formed from the metastatic 4T1 and the nonmetastatic 4T07 cell lines, isolated from the same breast tumor. We have identified significant differences in the types and amount of chemokines that are secreted by the cancer cells and in the nature of the myeloid cell infiltrate between the metastatic and nonmetastatic tumors. Strikingly, tumors grow slower and metastasis is greatly reduced in mice that lack the receptor for one of the chemokines that specifically is secreted by metastatic cancer cells and acts on neutrophils, one of the least studied inflammatory cell types in cancer metastasis. Ongoing studies are addressing how the chemokine–chemokine receptor signaling axis regulates invasion and promotes metastasis. Preliminary data strongly suggest that cancer cells activate an unusual cell-death mechanism in the neutrophils, leading to the release of strong proteolytic enzymes. We have determined that pharmacological inhibitors blocking the neutrophil cell death pathway are effective at reducing neutrophil-promoted invasion *in vitro* and metastasis *in vivo*.

## Collagen Architecture in Pancreatic Cancer Progression

M. Shields, L. Maiorino, J. Cappellani

---

Tumors are often discovered as nodules that are harder than the surrounding tissue. This reflects the changes in ECM stiffening and architecture. The interstitial ECM consists of macromolecules such as fibrillar collagens, fibronectin, and proteoglycans. Type I collagen is the major fibrillar collagen in many tissues and forms a scaffold that provides stability. Type I collagen

also has signaling functions mediated by, for example, integrins and discoidin domain receptors. The synthesis and proteolytic remodeling of the fibrillar type I collagen increases in many tumors, most notably in pancreatic tumors.

Collagen cross-linking in mouse models of mammary carcinoma delays tumor onset and slows tumor progression. We also found that the collagen architectural structure became abnormal with progression of breast cancer: Whereas collagen fibers are curly and oriented in parallel to normal or hyperplastic epithelium, there is a progressive change in the fibers so that they are straighter and mostly perpendicular to the tumor border in the late stages. This changed architecture promotes cell invasion by enabling cells to migrate along the collagen fibers or by enhancing integrin signaling. Enzymes of the lysyl oxidase (LOX) family are important for the collagen cross-linking that results in straighter fibers, and these enzymes have been shown to promote cancer progression and metastasis of breast cancer.

Pancreatic cancer has higher levels of type I collagen and of the LOX-like 2 enzyme than breast cancer, and we therefore hypothesized that collagen cross-linking would have a significant effect on reducing tumor initiation and metastatic spread. However, using both pharmacological inhibitors and genetic approaches, we instead found an increase in the growth of pancreatic tumors and in metastatic spread. Intravital imaging in mice showed that pancreatic cancer cells also migrate along collagen fibers and that collagen architecture is altered by cross-linking inhibitors in pancreatic cancer. This suggests that collagen architecture or the cancer cells' response to collagen is fundamentally different between breast and pancreatic cancer. In addition, we have determined that intracellular signaling pathways, downstream from integrins, are up-regulated after inhibition of LOX in pancreatic cancer cells.

## Coactivating Innate and Adaptive Immune Cells

A. Almeida [in collaboration with S. Adams, New York University]

---

Innate and adaptive immune cells (e.g., macrophages and T cells, respectively) infiltrate breast tumors. Activated, cytotoxic T cells can eliminate cancer cells, but

cancer cells counteract T-cell activation and killing by engaging inhibitory receptors. Checkpoint blockade immunotherapeutics target these inhibitory receptors, leading to durable clinical responses in several metastatic malignancies, especially those with tumor infiltrating lymphocytes (TILs). In triple-negative breast cancer, high levels of TILs are an independent predictor of survival. Consistently, durable responses were recently reported with checkpoint inhibitors for this breast cancer subtype. However, in hormone receptor (HR)-positive breast cancer, TILs do not predict positive outcomes, and checkpoint inhibitors have no effect. Instead, macrophage infiltration correlates with *poor* prognosis and such macrophages suppress T-cell activation. Together, these data suggest that approaches that target macrophages are needed for HR-positive breast cancers.

Our preliminary data show that *in vitro*, macrophages activated by lipopolysaccharide (LPS, a Toll-like receptor [TLR] agonist) and interferon- $\gamma$  (IFN- $\gamma$ ) efficiently kill ~90% of cancer cells within 48 h. Importantly, they also potently activate T cells. LPS cannot be used in humans due to severe side effects, but other synthetic TLR agonists can be used clinically.

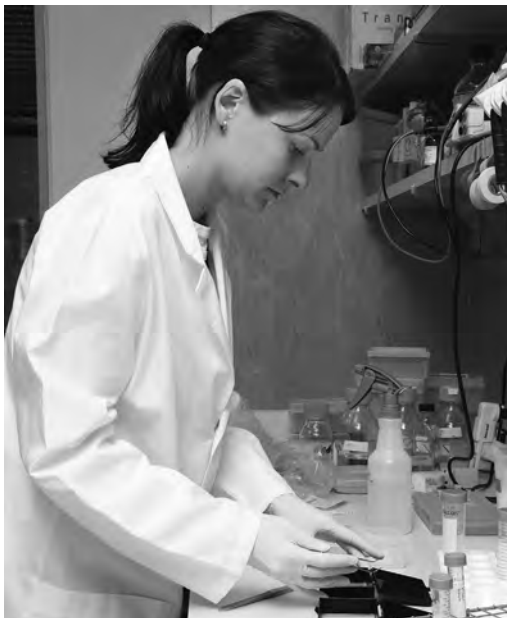
To dissect the pathways regulating macrophage/T cell/cancer cell interaction and to define the optimal

conditions for coactivation of macrophages and T cells, we have established an immune coculture assay (ICCA). To interrogate immune cell/cancer interactions in tumors, we have developed a sophisticated intravital imaging approach: confocal, time-lapse microscopy of cancer cells, macrophages, and T cells *in vivo*, using a metastatic mouse model of luminal B breast cancer, mouse mammary tumor virus–polyomavirus medium T (MMTV-PyMT)-mCherry-OVA. Ovalbumin (OVA) antigen serves as an artificial tumor antigen, and OVA-recognizing T cells infiltrate the tumors, but are suppressed by macrophages.

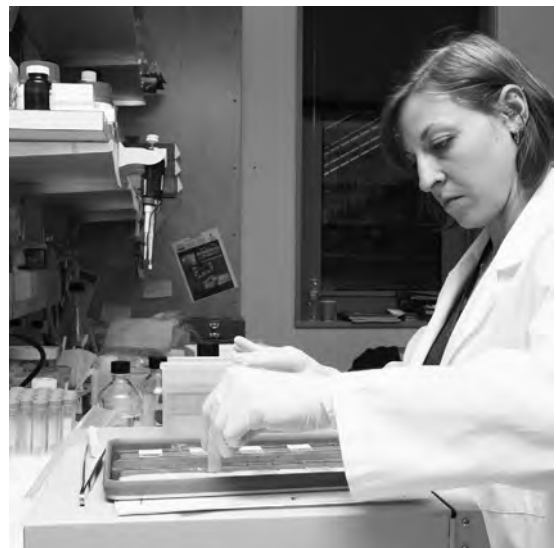
The immunotherapeutic strategy of stimulating macrophages' and T cells' tumoricidal activities in parallel has the potential to (1) overcome the immune suppressive environment of HR-positive breast cancer and (2) eradicate metastasized breast cancer cells while sparing normal cells.

## PUBLICATIONS

- Ewald AJ, Egeblad M. 2014. Cancer: Sugar-coating cell signaling. *Nature* **51**: 298–299.
- Fein MR, Wysocki RW, Egeblad M. 2014. Cancer in the spotlight: Using intravital imaging in cancer research. In *Advances in intravital microscopy* (ed. Weigert R). Springer, New York.



V. Kuettnner



M. Fein

# PROTEOMICS AND MASS SPECTROMETRY LABORATORY

D. Pappin   S. Bencze   K. Rivera  
              C.M. Gauss   J. Snider  
              T. Hassan    N.S. Turna  
              S. Peacock   J. Wilson

## A Software Tool to Evaluate Correspondence between Extracted Ion Chromatographic Peaks and Peptide-Spectrum Matches in Shotgun Proteomics

P. Leopold, D. Pappin, S. Peacock, K. Rivera, C. Ruse

---

Chromatographed peptide signals form the basis of additional data processing that produces functional information derived from data-dependent bottom-up proteomics experiments. We first ranked LC/MS (liquid chromatography/mass spectrometry) parent ions by the quality of their extracted ion chromatograms. Ranked extracted ion chromatograms were then used as a filter to improve the quality of MS/MS spectra submitted for database searches. In trial experiments, we initially identified approximately 5000 proteins when considering detector shifts of <7 ppm in the first pass of database searching. High-quality parent ions for which the database search yielded no hits were then used as candidates for subsequent unrestricted analysis for searching for posttranslational modifications (PTMs). Following this rational approach, we were able to identify more than 5000 *additional* spectrum matches from modified peptides and even confirmed the presence of acetylaldehyde-modified His/Lys residues. We have developed a logical workflow that scores data-dependent selected ion chromatograms and leverages information about semianalytical LC/LC dimensions prior to MS. The method can be successfully used to identify unexpected modifications in peptides with excellent chromatography characteristics, independent of fragmentation pattern and activation methods.

## Biochemical Characterization of a Lys/Arg-N Protease for MS Applications

J. Wilson, D. Pappin

---

Directed in vacuo peptide fragmentation using MS can generate specific ion fragment series that simplify

MS/MS collision spectra. Put simply, this is most readily achieved by controlling where charge is located on the peptide, either enzymatically or by chemical modification. To produce predominantly amino-terminal ions, the charge must be placed at or near the amino termini of peptides. This can be elegantly achieved using a proteolytic enzyme with amino-terminal cleavage specificity at basic amino acids, and we have discovered a novel thermostable protease with amino-terminal arginine and lysine specificity. The enzyme can be used as a substitute for trypsin, generates predominantly amino-terminal ions, and completes digestions in 1–2 h at 60°C.

The enzyme exhibits a temperature optimum of ~60°C and broad pH specificity centered at pH 8.5, with at least 50% activity remaining at pH 5 and pH 11. The protease is most active in low-ionic-strength MS-compatible acetate buffers, and it requires both calcium and zinc. Specificity for amino-terminal cleavage at arginine and lysine was ~>95%, with the remaining 5% amino-terminal to larger aliphatic residues. In whole-cell lysates, the numbers of identified peptides and proteins were similar to those of trypsin, yet generated in approximately 20 times shorter digestion times; 80% of the enzyme activity was retained in 0.5% SDS, and specificity was equal for both lysine and arginine residues. As expected, the b-ion series in MS/MS spectra were significantly more intense than y-ions, and relative b/y ion intensities were almost exactly reversed compared to trypsin. In complex mixtures, a shift to lower peptide charge states was also observed. This tendency was partly reversed by modification with iTRAQ (isobaric tags for relative and absolute quantitation), probably due to increased distance between charges, although iTRAQ labeling efficiency was unchanged.

This active metalloprotease is suitable for extremely rapid (1–2-h) digestion of samples for high-throughput MS analysis. The peptides produced by this enzyme always have major basic centers at the peptide

amino termini, producing more easily interpretable MS/MS fragmentation spectra by generating a dominant b-ion series. In contrast to trypsin, ion current is not spread between b- and y-ions, also increasing detection sensitivity.

### p53 Mutations Change Phosphatidylinositol Acyl Chain Composition

S. Bencze [in collaboration with A. Naguib, G. Bencze, D. Engle, C. Chio, D. Tuveson, and L. Trotman, Cold Spring Harbor Laboratory]

Phosphatidylinositol phosphate (PIP) second messengers relay extracellular growth signals through the phosphorylation status of the inositol sugar, a signal transduction system that is known to be deregulated in cancer. In stark contrast to PIP inositol head-group phosphorylation, changes in phosphatidylinositol (PI) lipid acyl chains in cancer have remained ill-defined. In collaboration with the Trotman and Tuveson laboratories, we applied an MS-based method capable of unbiased high-throughput identification and quantification of 20 different cellular PIs that vary in acyl chain composition. Using this multiple reaction monitoring (MRM) approach, we found that PI lipid chains represent a cell-specific fingerprint and are unperturbed by serum-mediated signaling. This is in complete contrast to the inositol head group. We also found that mutation of Trp53 results in PIs containing reduced-length fatty acid moieties, from 38- to 36- and 34-carbon acyl chains. Our results suggest that the anchoring tails of lipid second messengers form an additional layer of PIP signaling in cancer that operates independently of PTEN/PI3-kinase activity but is instead linked to p53.

### Function of N-Acetyltransferases in Neurological Disorders

This work was done in collaboration with M. Doerfel and G. Lyon (Cold Spring Harbor Laboratory).

The large majority (>85%) of human proteins are acetylated at their amino-terminal amino acids, making N( $\alpha$ )-terminal acetylation the most abundant protein posttranslational modification (PTM). Despite this, very little is known about the cellular effects or functions of this modification. In humans, six distinct amino-terminal amino-acetyltransferases

(NATs) catalyze the transfer of an acetyl group from acetyl-CoA to the N( $\alpha$ )-terminal amino group of their specific target proteins.

The major human acetyltransferase, NatA, consists of an auxiliary subunit, Naa15, and a catalytically active subunit, Naa10. Recently, the Lyon laboratory described two families with a lethal X-linked disorder of infancy called Ogden syndrome. This disorder comprises a distinct combination of an aged appearance, craniofacial anomalies, hypotonia, global developmental delays, cryptorchidism, and cardiac arrhythmias. Using X-chromosome exon sequencing and a recently developed variant annotation, analysis, and selection tool (VAAST), they identified a c.109T > C (p.Ser37Pro) variant in Naa10 as the probable disease-causing variant.

To characterize the effects of this mutation on complex formation of NatA and to identify new interaction partners, Naa10 wild-type or S37P mutants were immunoprecipitated from HEK293 cells. The isolated complexes were labeled with distinct iTRAQ reagents and analyzed by LC/MS. The data revealed a strong interaction of Naa10 with ribosomal components, supporting the idea that NTA occurs cotranslationally on newly synthesized proteins as they emerge from the ribosome. Furthermore, the interaction of the S37P mutant toward specific interaction partners was found to be strongly reduced, which remains to be confirmed in future *in vitro* experiments.

### Dephosphorylation of Tyrosine 393 in Ago2 by PTP1B Regulates Gene Silencing in RAS-Induced Senescence

This work was done in collaboration with M. Yang and N. Tonks (Cold Spring Harbor Laboratory).

Oncogenic RAS (H-RAS(V12)) induces premature senescence in primary cells by triggering production of reactive oxygen species (ROS), but the molecular mechanism was unclear. The Tonks laboratory investigated whether inhibition of protein tyrosine phosphatases by ROS contributed to H-RAS(V12)-induced senescence. They identified protein tyrosine phosphatase 1B (PTP1B) as a major target of H-RAS(V12)-induced ROS. Inactivation of PTP1B was both necessary and sufficient to induce premature senescence in H-RAS(V12)-expressing IMR90 fibroblasts. We identified phospho-Tyr 393 of argonaute 2

(Ago2) as a direct substrate of PTP1B using a trapping mutant and MS. Phosphorylation of Ago2 at Tyr 393 inhibited loading of microRNAs (miRNAs) and thus inhibited miRNA-mediated gene silencing, which counteracted the function of H-RAS(V12)-induced oncogenic miRNAs. The data illustrated that premature senescence in H-RAS(V12)-transformed primary cells was a consequence of oxidative inactivation of PTP1B and inhibition of miRNA-mediated gene silencing.

### Role of Nrf2 in Pancreatic Cancer

This work was done in collaboration with C. Chio and D. Tuveson (Cold Spring Harbor Laboratory).

Drug resistance is a major cause of lethality in pancreatic cancer. This has been ascribed to augmented cell survival pathways and poor drug delivery. Nrf2 is a transcription factor that promotes both pancreatic ductal adenocarcinoma (PDAC) progression and drug resistance, so antagonizing the Nrf2 pathway may be clinically advantageous. Because transcription factors are difficult to target therapeutically, the Tuveson laboratory has sought to comprehensively characterize the mechanisms used by Nrf2 to promote PDAC so that more feasible approaches to counter the effects of Nrf2 in PDA may be developed.

Activation of the Nrf2 program is known to elicit a more reducing intracellular environment. Because protein oxidation can influence protein activity and stability, one hypothesis is that certain mediators of Nrf2 function may be “redox sensor” proteins that contain labile cysteine thiol groups. Accordingly, we used pancreatic ductal organoids to develop a highly sensitive proteomic method that identifies proteins containing cysteines that undergo selective reduction because of the Nrf2 antioxidant program.

We previously developed a novel proteomics workflow that uses selective capture of cysteine peptides in conjunction with iTRAQ isobaric labeling to determine quantitative changes in free (reduced) cysteine residues on a proteome-wide scale that also normalizes these changes to expression levels of the respective

proteins. A proof-of-concept experiment was performed using mouse embryonic fibroblasts treated with BSO or DEM (two oxidizing agents known to induce Nrf2 target genes); 98% of isotope-enriched affinity tag (ICAT)-enriched peptides contained cysteines, demonstrating the specificity of the ICAT-labeling reagent. We found 349 cysteine-containing peptides modified by both of these two treatments from a total of 3662 proteins and 4572 cysteine-containing peptides, showing that the approach was effective in quantifying free cysteine changes upon redox alterations. Proteins oxidized by both agents included those previously reported to contain redox sensitive sites, such as Pkm1 and Pkm2, thioredoxin, and galectin-1.

To identify key Nrf2-dependent, redox-sensitive effector proteins/pathways, three pairs of N (normal), P (preneoplastic) and T (tumor) organoids bearing the wild-type or knockout Nrf2 allele were then analyzed to evaluate differences in the total and reactive cysteine proteomes. The analysis yielded 3416 cysteine labeled peptides. To identify key Nrf2-dependent, redox-sensitive pathways, the total proteome (iTRAQ) and cysteine proteome (ICAT) in N, P, and T organoids with or without Nrf2 were analyzed using gene set enrichment analysis (GSEA). Eight of 10 of the most enriched pathways encompassed highly oxidized proteins that were involved in translational control. These included core translational machinery components, such as ribosomal subunits, as well as numerous regulatory factors that control translation efficiency (GSEA  $p$  values range from  $p < 0.04$  to  $p < 0.1$  for the top 10 enriched pathways). Individual targets are currently being assessed by model-remodel (MRM) approaches to evaluate the oxidative states of the redox-sensitive cysteines.

### PUBLICATIONS

Yang M, Haase AD, Huang FK, Coulis G, Rivera KD, Dickinson BC, Chang CJ, Pappin DJ, Neubert TA, Hannon GJ, Boivin B, Tonks NK. 2014. Dephosphorylation of tyrosine 393 in argonaute 2 by protein tyrosine phosphatase 1B regulates gene silencing in oncogenic RAS-induced senescence. *Mol Cell* **55**: 782–790.

# ROLE OF CROSS-TALK BETWEEN INTERLEUKIN-6 AND TRANSFORMING GROWTH FACTOR- $\beta$ 1 IN CHEMORESISTANCE

R. Sordella   D. Pal   N. Shirole  
S. Senturk   Z. Yao

Cancers are largely a consequence of accumulating genomic damage during life. Because some of the same mutations that are driving the onset and the progression of tumors could be druggable, we can design therapies that will spare normal cells and specifically target only the cancer cells. This approach is usually referred to as “targeted therapy” or “precision medicine.” Many targeted therapies have been and currently are being developed. Yet their success in the clinic has been hampered by the development of resistance. Many studies in fact have shown that despite initial responses, targeted therapies overall are failing.

To better understand the mechanism of resistance to targeted therapy, we used as a paradigm lung cancer tumors harboring oncogenic mutations in the enhanced epidermal growth factor (EGFR) genes. Previously, we showed that although tumors with these mutations are initially highly sensitive to the inactivation of EGFR by small-molecule inhibitors such as erlotinib and gefitinib, these same tumors soon become refractory to the effect of the drug treatment. In recent years, we and other investigators have identified multiple mechanisms of resistance to EGFR small-molecule inhibitors. These comprise secondary EGFR mutations (T790M) that block the binding of EGFR inhibitors to their targets, amplification of *c-Met*, and the acquisition of mesenchymal-like properties. In the latter case, we showed that resistance to inhibition of EGFR is due to the activation of a transforming growth factor- $\beta$ /interleukin-6 (TGF- $\beta$ /IL-6) axis. In the past year, we have focused our effort in trying to understand how TGF- $\beta$  regulates IL-6 expression and how TGF- $\beta$  protects cells from apoptosis.

## ZC3H12A Down-Regulation by TGF- $\beta$ Is Critical for Increasing IL-6 Expression and Chemoresistance

TGF- $\beta$  family members are secreted proteins that control many cellular functions, including proliferation

and cellular differentiation. They have a role in cancer, immunity, fibrosis, heart disease, diabetes, and degenerative diseases such as Parkinson’s disease and multiple sclerosis. TGF- $\beta$ 1, TGF- $\beta$ 2, and TGF- $\beta$ 3 ligands function as the primary mediators of TGF- $\beta$  signaling. These ligands bind to the type-2 TGF- $\beta$  receptor (TGFBR2), which causes recruitment and phosphorylation of TGFBR1 (also known as ALK5), resulting in downstream signaling activation through the phosphorylation of R-SMAD proteins (SMAD2 or SMAD3). This ultimately leads to transcriptional activation or transcriptional repression of several genes via oligomerization of R-SMAD proteins with a variety of transcription cofactors including SMAD4. In addition to this signaling cascade, often referred to as “canonical” TGF- $\beta$  signaling, stimulation of cells with TGF- $\beta$  can induce the activation of alternative signaling modules such as the NF- $\kappa$ B, PI3K/AKT, RHOA, and MAPK pathways in a SMAD-independent manner.

The outcome of the activation of these signaling pathways in cancer cells is highly contextual and can result in either suppression of cell proliferation or induction of cellular migration and invasion. In premalignant stages, for example, the TGF- $\beta$  pathway has been shown to have tumor-preventive roles by induction of apoptosis as well as senescence. On the other hand, during malignant transformation, many of these tumor suppressive functions are instead lost, and the TGF- $\beta$  pathway eventually promotes tumor growth, invasion, and metastatic dissemination of cancer cells and importantly drives resistance to standard and targeted chemotherapy.

One key component of TGF- $\beta$ -driven resistance is the expression of the cytokine IL-6. Given the importance of this axis in our studies, we investigated the molecular insights underlying the regulation of IL-6 expression by TGF- $\beta$ . A better understanding of this mechanism could aid in the development of novel and more effective cancer therapies.

Although initial studies indicated that TGF- $\beta$  could induce the expression of IL-6 via JunD homodimers

in lung fibroblasts, we found that this mechanism has a minimal role in cancer cells. In particular, we demonstrated that transcriptional repression of ZC3H12A (also known as MCPIP1 or Regnase-1) by TGF- $\beta$  was critical. In fact, we found that TGF- $\beta$  inhibition of ZC3H12A transcription increased IL-6 expression via a bimodal mechanism encompassing increased mRNA stability and/or NF- $\kappa$ B-mediated mRNA transcription. Interestingly, we observed that the activation of this signal transduction mechanism could also have a role in mediating primary resistance to EGFR inhibition in tumors harboring oncogenic EGFR mutations. Our findings imply that co-inhibition of the NF- $\kappa$ B and IL-6 axes could improve current cancer therapies in lung cancer patients harboring EGFR mutations. Excitingly, a clinical trial has been recently opened to evaluate this hypothesis.

NF- $\kappa$ B is a sequence-specific transcription factor that is best known as a major regulator of inflammatory and innate immune responses. Hence, the repression of ZC3H12A by TGF- $\beta$  not only could have a key role in regulating intrinsic/cell autonomous mechanisms of resistance to targeted therapy, but could also be important in fine-tuning the dialog between malignant cells and their microenvironment.

### TGF- $\beta$ Down-Regulation of $\beta$ -TRCP Is Required for Chemoresistance and Metastasis

We have previously shown that the induced expression of IL-6 by TGF- $\beta$  is required and sufficient for antagonizing anti-apoptotic pathways activated as a consequence of EGFR inhibition in cells addicted to oncogenic EGFR mutations.

To better understand the molecular mechanisms underpinning the protection of cells from apoptosis by TGF- $\beta$ /IL-6, we examined the expression of Bcl-2 family members in cells in response to TGF- $\beta$  treatment. So far, 15 mammalian Bcl-2 family members have been identified, which can be divided into three subfamilies: the Bcl-2 pro-survival subfamily (Bcl-2, Bcl-XL, Bcl-w, Mcl-1, and A1); the pro-apoptotic Bax subfamily (Bax, Bak, and Bok); and the BH3 pro-apoptotic subfamily (Bad, Bid, Bik, Blk, Hrk, BNIP3, PUMA, NOXA, and BimL).

We initially examined expression of this family of proteins in A549 cells, an NSCLC-derived cell line in

which TGF- $\beta$  has been previously shown to induce epithelial-to-mesenchymal transition. We found increased expression of the pro-apoptotic proteins BAX, BAD, and BIM-EL and of the anti-apoptotic protein MCL-1.

To test whether *mcl-1* was required for the survival of cells in TGF- $\beta$  treatment, we silenced *mcl-1* expression via small interfering (siRNA)-mediated knockdown and analyzed the effect on viability in A549 cells upon treatment with TGF- $\beta$ . Although knockdown of *mcl-1* had minimal effect in control cells (scramble siRNA) or in TGF- $\beta$ -naïve cells, it resulted in a notable decrease of viable cells in populations exposed to TGF- $\beta$ . Consistent with a role of *mcl-1* in antagonizing pro-apoptotic stimuli, we observed a large sub-G<sub>1</sub> fraction in cells treated with TGF- $\beta$  and increased expression of cleaved caspase 3 upon *mcl-1* knockdown.

Given the central role of *mcl-1* in opposing TGF- $\beta$ -induced apoptosis, we next undertook studies aimed at understanding the regulation of *mcl-1* expression by TGF- $\beta$ . Differently from the pro-apoptotic Bcl-2 family members Bax, Bad, and BIMEL, we did not observe increased expression of *mcl-1* mRNA in cells treated with TGF- $\beta$ . In principle, this could indicate an alternative transcriptional-independent regulation of *mcl-1* by TGF- $\beta$ . Hence, we investigated whether TGF- $\beta$ , rather than regulating the transcription of *mcl-1*, could instead influence the stability of the MCL-1 protein. Levels of MCL-1 and of EGFR were analyzed following inhibition of the translational machinery with cycloheximide. Western blot analysis of MCL-1 levels from lysates of cells treated for 15, 30, and 60 min with cycloheximide demonstrated striking differences in the abundance of MCL-1 in cells in the presence of TGF- $\beta$  when compared to the untreated cells. Because no difference in the expression of EGFR was observed, this implies that TGF- $\beta$  did not affect the translational machinery but rather that MCL-1 protein is inherently more stable in TGF- $\beta$ -treated cells.

The ubiquitin proteasome system (UPS) is a major regulator of protein abundance in cells. In general, the transfer of ubiquitin molecules onto target proteins constitutes a signal that results in the targeting of the modified proteins for proteolysis by the proteasome system. Thus, we asked whether the increased stability of *mcl-1* that we observed in cells was mediated by impaired ubiquitination/degradation. We found that TGF- $\beta$  regulation of Mcl-1 occurs mainly at the level of protein stability by halting *mcl-1* degradation via the UPS.

More specifically, we observed that TGF- $\beta$  control of MCL-1 stability was mediated by the down-regulation of  $\beta$ -TRCP expression.  $\beta$ -TRCP has important roles in regulating cell cycle checkpoints, in protein translation, in cell growth, and in multiple transcriptional programs by regulating, for example, the

NF- $\kappa$ B and WNT pathways. In light of the current literature, our findings were particularly interesting because they suggest a potential role for the inhibition of  $\beta$ -TRCP in mediating some of the biological effects of TGF- $\beta$  such as EMT cell cycle control beyond the expression of MCL-1.



# PROTEIN TYROSINE PHOSPHATASES AND THE CONTROL OF SIGNAL TRANSDUCTION

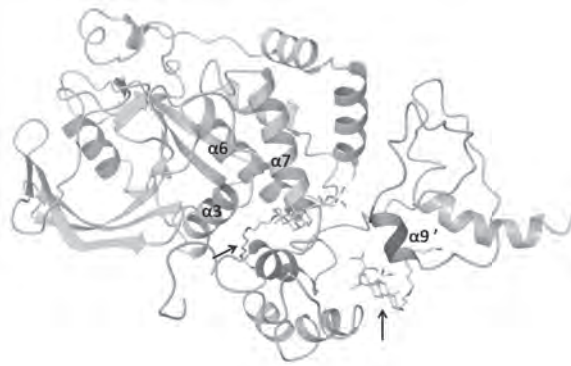
N.K. Tonks   S. Akshinthala   F. Chaudhary   N. Krishnan   M. Ramesh   D. Song   S. Zhang  
G. Bencze   G. Fang   L. Li   I. Rus   M. Yang   W. Zheng

An important aspect of the work of the Tonks lab is to devise creative new approaches to exploiting the protein tyrosine phosphatase (PTP) family of enzymes as targets for therapeutic intervention in major human diseases, including diabetes, obesity, and cancer. At the core of this effort is the study of PTP1B, the prototypic protein tyrosine phosphatase, which Nick Tonks first discovered 25 years ago. It is now established that PTP1B not only plays a role in attenuating insulin signaling, but also plays a *positive* role in promoting signaling events associated with breast tumorigenesis. This suggests that inhibition of PTP1B function may represent a novel therapeutic strategy to address not only diabetes and obesity, but also mammary tumorigenesis and malignancy.

The reversible addition and removal of phosphate to proteins, which is termed protein phosphorylation, is the central feature of the mechanism of signal transduction—the process by which cells respond to stimuli in their environment. The activities of the enzymes that mediate the addition (kinases) and removal (phosphatases) of phosphate groups are coordinated in signal transduction pathways to mediate the cellular response to environmental stimuli, and the function of these enzymes is frequently disrupted in human diseases, including diabetes and cancer. The ability to modulate such signal transduction pathways selectively with drugs holds enormous therapeutic potential. Drugs that target the protein kinases represent breakthroughs in cancer therapy. For example, HER2 is a member of the family of kinases and is amplified and/or overexpressed in several cancers, in particular in ~25% of breast cancer, where it is associated with poor prognosis. The humanized HER2-directed antibody Herceptin (trastuzumab) is an example of a “rational cancer therapy” for treatment of HER2-positive metastatic breast tumors. It targets HER2 as a unique marker of the cancer cell. Although Herceptin is used frequently and is being presented as a treatment of choice, the overall success

rate is low and patients develop resistance to the therapy. The problem of acquired resistance has become an obstacle to the successful application of kinase-directed therapies in general. Therefore, despite the obvious potential, it is anticipated that new alternative therapies, administered alone or in combination with kinase-directed drugs, will represent the way forward. The challenge is to identify such alternative therapies. Considering the intimate cooperation between kinases and phosphatases in the regulation of signal transduction under normal and pathophysiological conditions, the Tonks lab focuses on the PTPs, which have been garnering attention as potential therapeutic targets, but remain a largely untapped resource for drug development.

In following established approaches to developing small-molecule drugs that bind to the active site of an enzyme, industry has found PTPs to be challenging targets for therapeutic development. Although it was possible to generate potent, specific, and reversible inhibitors of PTP1B, such molecules were highly charged, due to the chemistry underlying PTP-mediated catalysis, and thus of limited drug development potential. Consequently, new approaches are required to exploit this important target effectively and reinvigorate drug discovery efforts. One strategy that would avoid targeting the active site of PTP1B would be to look for allosteric inhibitors, which bind at a site remote from the catalytic center, but induce conformational changes in the enzyme that result in inhibition. The Tonks lab has characterized a small-molecule natural product, TRODUSQUEMINE, as such an allosteric inhibitor of PTP1B. This is a unique mechanism of inhibition that will have been missed in the efforts to target PTP1B that have been conducted to date in industry. They have demonstrated that this inhibitor antagonizes the function of HER2 in cell and animal models of breast cancer. In particular, it essentially abrogates metastasis of HER2-positive tumor cells in the NDL2 mouse model of breast cancer.



**Figure 1.** Structure model for PTP1B (residues 1–405) with two molecules of the allosteric inhibitor MSI-1436 (marked with arrows) docked to the two putative binding sites. Important structural features within PTP1B are highlighted.

It is important also to note that TRODUSQUEMINE (also known as MSI-1436) has already been tested in a Phase 1 study that involved 88 obese patients and was found to be extremely well tolerated. Of particular significance is the fact that the effects observed in mice by the Tonks lab were achieved at ~20% of the maximum dose that has been administered to patients. During the last year, discussions have been initiated with the FDA, and a collaboration between the Tonks lab and the Montefiore Cancer Center at North Shore LIJ has been set up to take TRODUSQUEMINE/MSI-1436 into clinical trials for HER2-positive cancer. In an initial meeting with the FDA, it was indicated that a Phase 1 study of safety and tolerability of TRODUSQUEMINE/MSI-1436 as a single agent in metastatic breast cancer patients would have to be performed. The preparations for this clinical trial are now under way and will represent a focus of effort in 2014.

Although TRODUSQUEMINE/MSI-1436 demonstrates efficacy in an injectable format, it also has limited oral bioavailability. Industry has set oral bioavailability as a hurdle for the next generation of therapies for diabetes. During the last year, the Tonks lab has identified a derivative of TRODUSQUEMINE/MSI-1436 for which they have demonstrated oral bioavailability in a weight loss study in a high-fat diet-induced mouse model of obesity. At this time, they are defining the mechanism of action of this inhibitor. Furthermore, they have identified novel chemical entities that display properties similar to those of TRODUSQUEMINE/MSI-1436, which are currently being validated as inhibitors of PTP1B to identify further new drug candidates.

The Tonks lab is also trying to harness a physiological mechanism for regulation of PTP1B function as a distinct approach to development of therapeutics, particularly in the context of diabetes and obesity. PTP1B is a major regulator of the signaling pathways initiated by insulin and leptin, which controls appetite. Gene-targeting studies demonstrated that PTP1B-null mice are healthy, display enhanced insulin sensitivity, do not develop type-2 diabetes, and are resistant to obesity when fed with a high-fat diet. Furthermore, depletion of PTP1B expression with antisense oligonucleotides elicits anti-diabetic and anti-obesity effects in rodents, as well as human subjects. The Tonks lab discovered that the activity of PTP1B is attenuated by reversible oxidation of an essential cysteinyl residue at the active site of the enzyme. The architecture of the PTP active site is such that this essential cysteinyl residue displays unique properties that favor its role in catalysis but also render it prone to oxidation. Insulin stimulation of mammalian cells leads to enhanced production of intracellular  $H_2O_2$ , which causes reversible oxidation of PTP1B and inhibition of its enzymatic activity, which in turn concomitantly promotes the signaling response to insulin. They have shown that mild oxidation of PTP1B, such as occurs in response to insulin, results in profound conformational changes in the active site of the enzyme that transiently inhibit substrate binding and catalysis. These structural changes, however, are reversible and the enzyme can be reduced back to its active state. Therefore, reversible oxidation of PTP1B in response to insulin provides a mechanism for fine-tuning the signaling response to the hormone. They hypothesized that a conformation-sensor antibody that recognizes the reversibly oxidized form of PTP1B selectively may stabilize the inactive state, inhibit its reactivation by reducing agent, and thereby inhibit phosphatase activity. Using antibody phage display, they generated such conformation-sensor antibodies and demonstrated that expression of these antibodies in cells enhanced insulin-induced signal transduction. Their data provide proof-of-concept that stabilization of the oxidized, inactive form of PTP1B with appropriate therapeutic molecules may offer a novel paradigm for phosphatase drug development. This has the additional advantage that, if one assumes that in responding to insulin the cell targets for oxidation the pool of PTP1B that is important for physiological regulation of the signaling response, then this strategy

will target that pool specifically, possibly also reducing complications of side effects that may accompany inhibition of the native enzyme as a whole. The antibodies themselves are unlikely to be of use as therapeutics; however, the Tonks lab has now developed an assay that will permit screening of a small-molecule library on an industrial scale to identify drug-like molecules that mimic the effects of the antibodies. A pilot-scale screen has already been completed, demonstrating the feasibility of this strategy and identifying candidate small molecules that stabilize specifically the oxidized form of PTP1B. During the last year, this approach has led to the initiation of a collaboration with a large pharmaceutical company. Studies are under way to test the potential for conformation-sensor antibodies that recognize selectively the oxidized form of PTP1B to promote the signaling response to leptin. If these antibodies are shown also to potentiate leptin signaling, it is anticipated that a full-scale screen of the company's small-molecule collection will be initiated to exploit fully this unique approach to development of PTP inhibitors, thereby opening up a new strategy for therapeutic intervention in a major disease.

Efforts in the Tonks lab to define PTP function, and wherever possible establish links to human disease, are not restricted to PTP1B. Examples of progress in additional areas include the following.

### **Characterization of PTP Function in the Control of Migration and Invasion of Mammary Epithelial Cells**

---

The Tonks lab has characterized missing in metastasis (MIM), which is a scaffold protein that is down-regulated in multiple metastatic cancer cell lines compared to nonmetastatic counterparts. MIM regulates cytoskeletal dynamics and actin polymerization, and it has been implicated in the control of cell motility and invasion. MIM has also been shown to bind to a receptor PTP, PTP $\delta$ , an interaction that may provide a link between tyrosine-phosphorylation-dependent signaling and metastasis. They used small hairpin RNA (shRNA)-mediated gene silencing to investigate the consequences of loss of MIM on the migration and invasion of the MCF10A mammary epithelial cell model of breast cancer. They observed that suppression of MIM by RNA interference (RNAi) enhanced migration and invasion of MCF10A cells, effects that

were mediated by enhancing the stability and quantity of PTP $\delta$ . Furthermore, analysis of human clinical data indicated that PTP $\delta$  was elevated in breast cancer samples when compared to normal tissue. They demonstrated that SRC is a direct substrate of PTP $\delta$  and, upon suppression of MIM, they observed changes in the phosphorylation status of the SRC protein tyrosine kinase—the inhibitory site (Y527) was hypophosphorylated, whereas the activating autophosphorylation site (Y416) was hyperphosphorylated. Thus, the absence of MIM led to PTP $\delta$ -mediated activation of SRC. Finally, the SRC inhibitor SU6656 counteracted the effects of MIM suppression on cell motility and invasion. This demonstration of PTP $\delta$ -dependent activation of SRC in cells depleted of MIM suggests a new therapeutic strategy for targeting metastasis.

### **Investigation of the Role of Thioredoxin in Reversible PTP Oxidation**

---

Despite great progress in defining the mechanism of reactive oxygen species (ROS) production, the controls over the levels of ROS in cells and the scope of potential targets of this modification, the mechanisms underlying reduction and reactivation of the oxidized PTPs have remained relatively underexplored. It is essential that PTPs recover their active form when ROS concentrations decrease so as to limit the response to growth factors/hormones and prevent uncontrolled stimulation of signaling pathways. For the PTPs, redox-active enzymes and small molecules, such as thioredoxin (TRX), glutaredoxin, or glutathione, have been implicated in reduction and reactivation. In a collaborative study with Arne Östman's lab, they demonstrated that TRX was important for reduction and reactivation of PTP1B in the context of PDGF signaling. In a separate study, they used a combination of biochemical analysis, mechanism-based trapping, and RNAi-induced suppression to demonstrate the importance of TRX1 for the reduction and reactivation of PTP1B and PTEN. Although TRX is best known as a disulfide reductase, they demonstrated that it plays an important role in the reduction and reactivation of the oxidized form of PTP1B, which features a cyclic sulphenamide modification of the active-site cysteine residue. Furthermore, they have shown that these effects of TRX have the potential to fine-tune the signaling response to insulin. Finally,

in addition to highlighting the important role of TRX in reduction and reactivation of PTPs, the TRX trapping mutants illustrate another approach to defining the importance of reversible oxidation in the regulation of PTP function in general and tyrosine-phosphorylation-dependent signaling in a broad array of signaling contexts.

### Identification of the Anti-Inflammatory Compound BAY 11-7082 as a Potent Inhibitor of PTPs

BAY 11-7082 is an anti-inflammatory compound that has been reported to inhibit I $\kappa$ B kinase activity. The compound has an  $\alpha,\beta$ -unsaturated electrophilic center, which confers the property of being a Michael acceptor; this suggests that it may react with nucleophilic cysteine-containing proteins, such as PTPs. In this study, the Tonks lab demonstrated that BAY 11-7082 was a potent, irreversible inhibitor of PTPs. Using mass spectrometry, they demonstrated that BAY 11-7082 inactivated PTPs by forming a covalent adduct with the active-site cysteine. Administration of the compound caused an increase in protein tyrosine phosphorylation in RAW 264 macrophages, similar to the effects of the generic PTP inhibitor sodium orthovanadate. These data illustrate that BAY 11-7082 is an effective pan-PTP inhibitor with cell permeability, revealing its potential as a new probe for chemical biology approaches to the study of PTP function. Furthermore, the data suggest that inhibition of PTP function may contribute to the many biological effects of BAY 11-7082 that have been reported to date.

### Investigation of the Regulation and Function of Receptor PTP $\alpha$ in Breast Cancer Models

HER2/ERBB2 is a member of the epidermal growth factor (EGF) receptor family of protein tyrosine kinases (PTKs), which is amplified and/or overexpressed in ~25% of breast cancer and associated with poor prognosis. Using MCF10A mammary epithelial cells that ectopically express an “activatable” ErbB2 chimera as a model, they demonstrated that NOX-dependent production of hydrogen peroxide occurs rapidly

following ERBB2 activation, leading to reversible oxidation of the transmembrane receptor phosphatase PTP $\alpha$ . This inhibited the activity of PTP $\alpha$ , contributing to increased ERBB2 signaling. Furthermore, the suppression of PTP $\alpha$  by shRNA led to ERBB2-dependent increased cell migration, which was characterized by prolonged interaction of GRB7 with ERBB2, increased association of ERBB2 with a  $\beta$ 1 integrin-rich complex, and was dependent on GRB7-SH2 domain interactions. Interestingly, the human *GRB7* gene is commonly co-amplified with *ERBB2* in breast cancer, and GRB7 has been implicated in cell migration. Supporting this concept, the ERBB2-dependent migration in PTP $\alpha$ -knockdown cells was suppressed by a GRB7-SH2 domain inhibitor. They demonstrated that PTP $\alpha$  dephosphorylated FAK specifically on Tyr<sup>407</sup>, and FAK Tyr<sup>407</sup> phosphorylation was enhanced in cells following ERBB2 activation when PTP $\alpha$  was suppressed, which contributed to the recruitment of vinculin to FAK. Collectively, these data support a role for PTP $\alpha$  in regulating motility of mammary epithelial cells in response to ERBB2 signaling, consistent with a role for the phosphatase in breast cancer.

### PUBLICATIONS

- Krishnan N, Koveal D, Miller DH, Xue B, Akshinthala SD, Kragelj J, Jensen MR, Gauss CM, Page R, Blackledge M, Muthuswamy SK, Peti W, Tonks NK. 2014. Targeting the disordered C terminus of PTP1B with an allosteric inhibitor. *Nat Chem Biol* **10**: 558–566.
- Pulido R, Baker SJ, Barata JT, Carracedo A, Cid VJ, Chin-Sang ID, Davé V, den Hertog J, Devreotes P, Eickholt BJ, et al. 2014. A unified nomenclature and amino acid numbering for human PTEN. *Sci Signal* **7**: pe15.
- Schwertassek U, Haque A, Krishnan N, Greiner R, Weingarten L, Dick TP, Tonks NK. 2014. Reactivation of oxidized PTP1B and PTEN by thioredoxin 1. *FEBS J* **281**: 3545–3558.
- Yang M, Haase AD, Huang FK, Coulis G, Rivera KD, Dickinson BC, Chang CJ, Pappin DJ, Neubert TA, Hannon GJ, Boivin B, Tonks NK. 2014. Dephosphorylation of tyrosine 393 in argonaute 2 by protein tyrosine phosphatase 1B regulates gene silencing in oncogenic RAS-induced senescence. *Mol Cell* **55**: 782–790.
- In Press*
- Chaudhary M, Lucito R, Tonks NK. 2015. Missing-in-Metastasis regulates cell motility and invasion via PTP $\delta$ -mediated changes in SRC activity. *Biochem J* **465**: 89–101.
- Fan G, Wrzeszczynski KO, Fu C, Su G, Pappin DJ, Lucito R, Tonks NK. 2015. A quantitative proteomics-based signature of platinum sensitivity in ovarian cancer cell lines. *Biochem J* **465**: 433–442.

# UNDERSTANDING PROSTATE CANCER METASTASIS

L. Trotman   A. Ambrico   D. Nowak  
H. Cho   C. Stahlhut  
T. Herzka   K. Watrud  
A. Naguib

Resistance to therapy of metastatic prostate cancer (CRPC) is responsible for the annual deaths of about 250,000 U.S. men. Although there is considerable progress in the development of improved anti-hormone therapy for treatment of metastatic disease, this standard of care approach will invariably fail at some point.

Our focus is to understand the mechanisms driving human prostate cancer in its most lethal form: metastatic disease. We have studied the human genetics behind the transition from indolent to lethal metastatic prostate cancer and combined it with viral transgene delivery into prostate. With this approach, we have succeeded in generating a unique, fast, and faithful mouse model for metastatic prostate cancer. We have termed this system RapidCaP because it allows us to generate any genetically mutant mouse prostate cancer with a much accelerated time frame compared to breeding-based approaches. We now use RapidCaP for analysis and therapy of metastatic disease. At the same time, we aim to better define PTEN/PI3 kinase signaling and derive new types of markers that indicate high metastatic potential of disease.

## The RapidCaP GEM Model for Metastatic Prostate Cancer

H. Cho, D. Nowak, T. Herzka, K. Watrud, A. Ambrico, C. Stahlhut [in collaboration with J. Wilkinson, University of Michigan; J. Bradner, Dana Farber Cancer Center; B. Robinson, Weill Cornell Medical College; M. Castillo-Martin and C. Cordon-Cardo, Mt. Sinai School of Medicine]

We previously identified the loss of the p53 senescence program as a critical barrier to metastatic progression of PTEN/PHLPP mutant prostate cancer. On the basis of these findings, we were able to transform GEM modeling of prostate cancer by developing the RapidCaP system for analysis and therapy of endogenous metastatic prostate cancer.

Although considerable progress has been made in the development of improved drugs for treatment of metastatic disease, it remains a huge challenge to predict which patients will best respond to a given approach. The RapidCaP system is highly versatile as genetic alterations are introduced by infection of the mouse prostate through a transgene-carrying virus. In a first approach, we used RapidCaP to model loss of the *Pten* and *p53* genes, a signature that has been found in 50% of human castration-resistant metastatic prostate cancer. In contrast to the classic approaches of studying (human) tumors *in* a mouse, our approach allows us to study prostate cancer *of* a mouse. RapidCaP makes use of existing genetic engineering tools with conditional tumor suppressor alleles. We trigger disease by viral Cre delivery and monitor progression of disease by integration of the luciferase transgene and noninvasive visualization using bioluminescence imaging. It exceeds current GEM prostate models by producing focal disease that retains intact histopathology of surrounding stroma and normal tissue. In contrast, traditional conditional knockout of the entire prostatic epithelium, and targeting the same genes, results in lethal primary disease burden within months. This artificial scenario is without a human counterpart and has so far hampered analysis of progression phenotypes.

The innovative RapidCaP system is important in four areas: (1) it matches the high speed at which human candidate cancer genes are discovered because mutant animals are generated by viral infection, not by breeding; (2) it allows us to trace disease progression and therapy in live animals without the need for costly high-tech imaging modalities such as magnetic resonance imaging (MRI) or positron emission tomography (PET); (3) primary and secondary lesions arise natively and *in situ* in fully immune-competent mice and entirely native tissue microenvironments; and (4) most pertinent to our future work, it allows us to test the causality of genetic mutations in the endogenous metastatic process.

RapidCaP recapitulates the therapy failure seen in patients, as shown by emergence of refractory metastatic disease after castration. We have used the RapidCaP system to define new mechanisms behind lethal CRPC2, specifically identifying *Myc* as a driver of *Pten* mutant metastasis. This finding has opened up novel rational approaches for CRPC therapy using MYC-suppressing bromodomain inhibitors. Further demonstrating the power of the new model system, we showed using in vivo RNA interference (RNAi) technology that the RNA-regulating protein SHQ1, a human genomics-derived candidate cancer gene residing in a common deleted chromosomal region, is a driver of *Pten* mutant prostate metastasis. To gain further insights, we studied the mechanisms behind activation of the *Myc* oncogene after *Pten* loss.

The system now affords us new insights into lethal cancer evolution. We also use it to test whether and how novel therapies can enhance or even outperform the 80-year-old standard of care—hormone deprivation therapy—and turn lethal prostate cancer into a curable disease.

### The PTEN/PI3 Kinase Pathway in Action

A. Naguib, T. Herzka, W. Zheng [in collaboration with D. Pappin and L. Joshua-Tor, Cold Spring Harbor Laboratory]

PTEN is the major negative regulator of phosphoinositol-3 kinase (PI3K) signaling with cell-specific functions that go beyond tumor suppression. We aim to better define PTEN/PI3K signaling at its most fundamental level. This has led us to probe PTEN function in the cytoplasm, where the vast majority of PTEN resides. We use live imaging, super resolution microscopy, and fluorescence photobleaching to map and capture PTEN/PI3K signaling in motion. To better understand modification of the lipid second

messengers that communicate PI3K signaling, we use mass spectrometry. Our approaches are geared toward rapid analysis in tissues, with the goal in mind of detecting aberrant growth signals on a time scale that is much faster than standard genomics-based approaches.

Finally, our analysis of phosphatidylinositol-based signaling has shown that the lipid tails of these molecules reveal information on the *p53* mutation status of a cell. This finding is of particular interest as we developed a mass-spectrometry-based detection protocol that yields results within minutes. Thus, our successful analysis is based on a technology that is much faster than sequencing-based mutation analysis and thus could be used to guide the definition of margins of a patient's cancer during surgery.

### PUBLICATIONS

- Chen M, Nowak DG, Trotman LC. 2014. Molecular pathways: PI3K pathway phosphatases as biomarkers for cancer prognosis and therapy. *Clin Cancer Res* **20**: 3057–3063.
- Cho H, Herzka T, Zheng W, Qi J, Wilkinson JE, Bradner JE, Robinson BD, Castillo-Martin M, Cordon-Cardo C, Trotman LC. 2014. RapidCaP, a novel GEM model for metastatic prostate cancer analysis and therapy, reveals *myc* as a driver of *Pten*-mutant metastasis. *Cancer Discov* **4**: 318–333.
- Labbé DP, Nowak DG, Deblois G, Lessard L, Giguère V, Trotman LC, Tremblay ML. 2014. Prostate cancer genetic-susceptibility locus on chromosome 20q13 is amplified and coupled to androgen receptor-regulation in metastatic tumors. *Mol Cancer Res* **12**: 184–189.
- Newton AC, Trotman LC. 2014. Turning off AKT: PHLPP as a drug target. *Annu Rev Pharmacol Toxicol* **54**: 537–558.
- Pulido R, Baker SJ, Barata JT, Carracedo A, Cid VJ, Chin-Sang ID, Davé V, den Hertog J, Devreotes P, Eickholt BJ, et al. 2014. A unified nomenclature and amino acid numbering for human PTEN. *Sci Signal* **7**: pe15.
- Reyes G, Niederst M, Cohen-Katsenelson K, Stender JD, Kunkel MT, Chen M, Brognard J, Siercecki E, Gao T, et al. 2014. Pleckstrin homology domain leucine-rich repeat protein phosphatases set the amplitude of receptor tyrosine kinase output. *Proc Natl Acad Sci* **111**: E3957–E3965.

## CANCER MEDICINE LABORATORY/CANCER THERAPEUTICS INITIATIVE

D. Tuveson	B. Alagesan	D. Engle	E. Lee	H. Tiriach
L. Baker	E. Elyada	M. Ludwig	K. Wright	
G. Biffi	M. Feigin	E. Olund	G. Yordanov	
D. Cheng	D. Filippini	T. Oni	K. Yu	
I.I. Chio	A. Handly-Santana	Y. Park		
V. Corbo	C.-I. Hwang	M. Ponz-Sarvisse		

The lethality of pancreatic ductal adenocarcinoma (PDAC) is due to the advanced stage of disease at the time of diagnosis and the poor efficacy of current therapies. During the past 12 years, our lab has pioneered preclinical model systems to investigate both of these problems, and we have translated this information to clinical investigation. Most recently, we codeveloped primary tissue models of both human and murine pancreatic cancer, termed “organoids.” We expect that over the next several years, organoids will reveal fundamental insights into pancreatic cancer progression and the tumor microenvironment while also serving as a useful platform and renewable biorepository for the discovery of novel diagnostic and therapeutic targets. Accordingly, organoids are currently being evaluated as a transformative technology for our new Preclinical Experimental Therapeutics-X Facility (PET<sup>x</sup>), with an immediate task of determining whether organoids can facilitate early-stage human trials.

### Organoid Models of Pancreatic Cancer

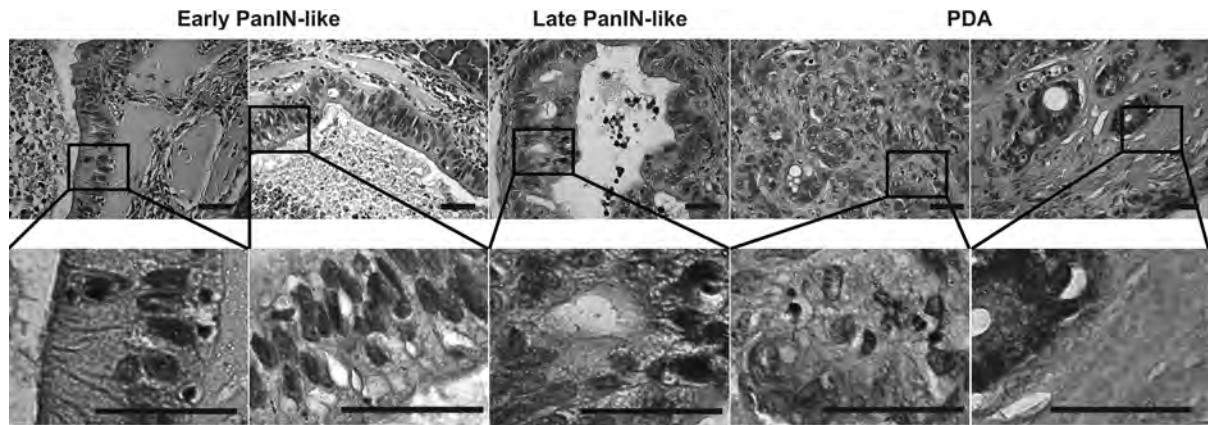
This work was done in collaboration with H. Clevers (Utrecht, The Netherlands); D. Pappin, M. Hammell, and J. Hicks (Cold Spring Harbor Laboratory)

Organoids are three-dimensional primary epithelial cell cultures formed from benign and malignant murine and human pancreatic tissues (Boj SF, et al., in press). Organoids can be generated from small human tissue samples such as endoscopic needle biopsies. Biopsies are more easily and commonly obtained in the clinic than resected tumors, so the ability to study cells taken from biopsies provides a feasible route to influence patient management. With our panel of organoids, we can perform diverse global molecular analyses. Our own transcriptomic, genomic, metabolomic, and proteomic analyses have identified new

pathways that are altered in pancreatic cancer, including pathways involved in lipid metabolism. Additionally, analysis of our initial set of human pancreatic cancer organoids has revealed marked interpatient genetic diversity, which we are now able to study *ex vivo*. Strikingly, single-cell assessment of our human organoids has also uncovered inpatient subclonal diversity. To determine the extent of interpatient and inpatient heterogeneity, we have initiated a project with multiple international sites to collect and characterize 300 human organoids. Single-cell analyses will be performed on some of these specimens to determine the pattern of co-occurring genetic alterations. Using this information, we can define distinct genetic subtypes of pancreatic cancer that may represent unique therapeutic vulnerabilities. By deriving new organoid cultures from individual cells isolated from those multiclonal, heterogeneous organoids, we will be able to further study individual subclones within a tumor.

Using our mouse organoid system, we discovered that primary pancreatic ductal adenocarcinoma (PDAC) in the classic *Kras<sup>+/LSL-G12D</sup>; Trp53<sup>+/LSL-R172H</sup>; Pdx1-Cre* (KPC) mouse model of pancreatic cancer did not harbor loss of heterozygosity (LOH) of the wild-type *Trp53* allele and maintained normal karyotypes. In contrast, organoids derived from metastases and two-dimensional, monolayer cell lines generated from primary-tumor-derived organoids did undergo *Trp53* LOH and displayed abnormal karyotypes. Importantly, human organoids derived from primary PDAC all had abnormal karyotypes. This demonstrates that prior work from our laboratory and others using the KPC mouse system may best model the biology and therapeutic response of early cancers, a finding which prompts us to develop more advanced mouse models of PDAC for current and future work.

Intriguingly, following orthotopic transplantation, organoids recapitulated the stages of pancreatic tumor



**Figure 1.** Following orthotopic transplantation of human tumor organoids into *Nu/Nu* mice, human tumor organoids recapitulated the full progression of pancreatic cancer. Early-stage preinvasive lesions are initially formed and later progress to invasive carcinomas. Tumor sections are stained with hematoxylin and eosin. Higher magnification images of *insets* are shown below. Bar, 50  $\mu$ m.

progression, forming structures resembling early preinvasive carcinomas, which eventually progressed to invasive carcinomas. Thus, organoids represent the first robust, transplantable model of human pancreatic cancer progression (Fig. 1). Remarkably, within tumors derived from transplanted organoids, both the neoplastic and tumor microenvironment composition evolved simultaneously. We are currently determining whether these orthotopically engrafted organoid (OGO) models accurately reflect therapeutic responses. We expect that the OGO models will be a rich resource for studying the fundamental biology of PDAC, discovering new biomarkers, and identifying new treatment strategies.

To determine whether organoids are more predictive of therapeutic response than standard models of human pancreatic cancer, we are comparing organoids to monolayer cultures, patient-derived-xenograft models (PDXs), and patient-derived circulating tumor cells (CTCs). As part of our work to investigate treatment response, we have translated one of our findings into a clinical trial at Memorial Sloan

Kettering Cancer Center (PI, Kenneth Yu). This trial is evaluating the ability of pegylated hyaluronidase (PEGPH20) to enhance the delivery of cetuximab. In addition to evaluating patient response and drug delivery, we will prepare human organoids from resected tumors obtained from patients enrolled in this trial for follow-up studies. In the future, we plan to generate organoid cultures from pretreatment biopsies, as well as serial biopsies from patients under therapy, to explore mechanisms of primary and acquired drug resistance that will inform new treatment strategies.

## PUBLICATIONS

Sherman MH, Yu RT, Engle DD, Ding N, Atkins AR, Tiriach H, Collisson EA, Connor F, Van Dyke T, Kozlov S, et al. 2014. Vitamin D receptor-mediated stromal reprogramming suppresses pancreatitis and enhances pancreatic cancer therapy. *Cell* **159**: 80–93.

*In Press*

Boj SF, Hwang C-I, Baker LA, Chio IC, Engle DD, Corbo V, Jager M, Ponz-Sarvis M, Tiriach H, Spector MS, et al. 2015. Organoid models of human and mouse ductal pancreatic cancer. *Cell* **160**: 324–338.



# RAS AND RHO GTPASES AND THE CONTROL OF SIGNAL TRANSDUCTION

L. Van Aelst   B. Elmer   M. Wang  
Y. Tai   Y.-T. Yang  
C.-L. Wang   J.-R. Yu

Research in my laboratory is focused on the role of Ras and Rho GTPase family members in signal transduction. Ras and Rho family members have key roles in cellular activities controlling cell growth control, differentiation, and morphogenesis. Alterations in Ras and Rho functions have been causally linked to both cancer and brain disorders, including mental retardation, schizophrenia, and epilepsy. Our interests lie in understanding how defects in Ras- and Rho-linked proteins contribute to the development of these disease processes. Toward this end, my lab has continued to define the functions of selected GTPases and their regulators and effectors in models of cancer and neurodevelopmental disorders. Below are highlighted selected key projects that have been carried out during the past year.

## Multifunctional Role of the X-linked Intellectual Disability Protein OPHN1 at the Hippocampal CA1 Synapse

*Oligophrenin-1 (OPHN1)*, which encodes a Rho-GTPase activating protein, was the first identified Rho-linked intellectual disability (ID) gene. It was initially identified by the analysis of a balanced translocation t(X;12) observed in a female patient with mild ID. Subsequent studies revealed the presence of *OPHN1* mutations in families with ID associated with cerebellar hypoplasia, lateral ventricle enlargement, and/or epilepsy. Most of the *OPHN1* mutations identified to date have been shown, or predicted, to result in OPHN1 loss of function; however, the pathophysiological role of *OPHN1* has remained poorly understood.

We have begun to unveil the function of OPHN1 at both the presynaptic site and postsynaptic site of the hippocampal CA3-CA1 synapse. We previously described how during early development, presynaptic OPHN1 is important for efficient retrieval of synaptic vesicles, whereas postsynaptic OPHN1 has a key

role in activity-dependent maturation and plasticity of excitatory synapses, suggesting the involvement of OPHN1 in normal activity-driven glutamatergic synapse development. Interestingly, we more recently found that postsynaptic OPHN1 also has a critical role in mediating a form of plasticity (mGluR-LTD [long-term depression]) that relies on the activation of group I metabotropic glutamate receptors, which consist of mGluR1 and mGluR5 in CA1 hippocampal neurons. Alterations in this form of plasticity have been linked to drug addiction and cognitive disorders. A key characteristic of mGluR-LTD is its dependence on rapid protein synthesis; however, the identities of the proteins mediating LTD have remained largely elusive. We obtained evidence that OPHN1 expression is translationally induced in dendrites of CA1 neurons within 10 min of mGluR activation and that this response is essential for mGluR-dependent LTD. Specifically, acute blockade of new OPHN1 synthesis impedes mGluR-LTD and the associated long-term decreases in surface AMPARs ( $\alpha$ -amino-3-hydroxy-5-methyl-4-isoxazole propionic acid receptors). Interestingly, the rapid induction of OPHN1 expression is primarily dependent on mGluR1 activation and is independent of fragile-X intellectual disability protein (FMRP).

We further demonstrated that OPHN1's role in mediating mGluR-LTD is dissociable from its role in basal synaptic transmission. mGluR-LTD and the associated long-term decreases in surface AMPARs are dependent on OPHN1's interaction with Endophilin A2/3, whereas regulation of basal synaptic strength requires OPHN1's Rho-GAP activity and association with Homer 1b/c proteins. As to how OPHN1 could mediate the strengthening of synapses via interactions with Homer 1b/c and RhoA, we previously showed that OPHN1 becomes enriched in spines in response to synaptic activity and NMDAR (*N*-methyl-D-aspartate receptor) activation, where by locally modulating RhoA/Rho kinase activities (i.e., in the proximity of

AMPARs), it contributes to the stabilization of AMPARs. Homer 1b/c proteins, on the other hand, were shown to participate in the positioning of the endocytic zone (EZ) near the postsynaptic density (PSD). The close juxtaposition of the EZ and the PSD enables localized endocytosis and recycling of AMPARs that serves to maintain a mobile pool of surface AMPARs required for synaptic potentiation. Interestingly, we recently found that a physical interaction between OPHN1 and Homer 1b/c is crucial for the positioning of the EZ adjacent to the PSD, and we obtained evidence that this interaction is important for OPHN1's role in controlling activity-dependent strengthening of excitatory synapses. Disruption of the OPHN1–Homer 1b/c interaction caused a displacement of EZs from the PSD, along with impaired AMPAR recycling and reduced AMPAR accumulation at synapses, under both basal conditions and conditions that can induce synaptic potentiation. Thus, these data indicate that OPHN1 exerts its effects on synapse strengthening by serving at least two distinct roles. Via its interaction with Homer 1b/c, OPHN1 facilitates the recycling and thereby maintenance of a mobile pool of surface AMPARs, whereas via its RhoGAP activity, OPHN1 contributes to the stabilization of synaptic AMPARs.

Collectively, our findings point to a multifunctional role for OPHN1 at CA1 synapses. Independent of its role in activity-driven glutamatergic synapse development, regulated OPHN1 synthesis has a critical role in mGluR-dependent LTD. Thus, it is conceivable that, on one hand, OPHN1 might have an important role in synapse maturation and circuit wiring during early development. On the other hand, the regulated OPHN1 synthesis could operate during adulthood to weaken synapses in response to behaviorally relevant stimuli.

### Regulation of Chandelier Cell Cartridge and Bouton Development via DOCK7-Mediated ErbB4 Activation

DOCK7 is a member of the evolutionarily conserved DOCK180-related protein superfamily. The DOCK180 family, of which in total 11 mammalian members have been identified (termed DOCK1/DOCK180 to DOCK11), emerged as a distinct class of Rac and/or Cdc42 GTPase guanine nucleotide exchange factors (GEFs). This class of proteins has been

implicated in diverse cell-type-specific processes. We initially identified DOCK7 as a novel activator of Rac, which is highly expressed in the developing brain. Significantly, mutations in DOCK7 have been reported in individuals with epileptic encephalopathy and cortical blindness; however, the role(s) of DOCK7 in neuronal development and/or function has remained largely elusive.

We previously described how DOCK7 has a critical role in the polarization and genesis of newborn pyramidal neurons and how it does so by promoting Rac activity and antagonizing TACC3 (transforming acidic coiled-coil-containing protein 3) function, respectively. Interestingly, in more recent studies examining the expression patterns of members of the DOCK180 family in GABAergic interneurons, we intriguingly observed the presence of DOCK7, among other parvalbumin (PV)-expressing interneurons, in chandelier cells (ChCs) of adolescent/adult mouse brains. ChCs, typified by their unique axonal morphology, are the most distinct interneurons present in cortical circuits. Through their distinctive axonal terminals, called cartridges, these cells selectively target the axon initial segment of pyramidal cells and control action potential initiation; yet, the mechanisms that govern the characteristic ChC axonal structure have remained elusive. The main obstacles have been a lack of unique biochemical ChC markers and versatile methods to target and manipulate gene expression in these cells. This prompted us to develop a method to manipulate gene expression in ChCs so that DOCK7's role in their development could be assessed. On the basis of recent evidence indicating that progenitors in the ventral medial ganglionic eminence (vMGE) provide a source of ChCs, we reasoned it should be possible to target gene expression in nascent ChCs by means of in utero electroporation directed toward the vMGE, and we found that this is indeed the case. Using this technique, we demonstrated a critical role for DOCK7 in ChC cartridge/bouton development. In particular, knockdown of DOCK7 caused a disorganization of ChC cartridges and a decrease in the density and size of ChC boutons. DOCK7 overexpression elicited essentially the opposite phenotypes. An organized network of ChC cartridges was formed, and both the density and size of boutons were increased. We further found that DOCK7 functions as a cytoplasmic activator of the schizophrenia-associated ErbB4 receptor tyrosine kinase and importantly that DOCK7 modulates ErbB4 activity to control

ChC cartridge/bouton development. Indeed, an ErbB4 mutant with enhanced kinase activity was able to overcome the phenotypes associated with DOCK7 knock-down, whereas silencing of ErbB4 prevented the phenotypes elicited by DOCK7 overexpression. Thus, our findings define DOCK7 and ErbB4 as key components of a pathway that controls the morphological differentiation of ChCs, with implications for the pathogenesis of schizophrenia.

### TGF- $\beta$ /Smad Signaling through DOCK4 Facilitates Lung Adenocarcinoma Metastasis

The cytokine transforming growth factor- $\beta$  (TGF- $\beta$ ) has an important, albeit complex, role in epithelial tumorigenesis. During early stages of tumorigenesis, TGF- $\beta$  typically functions as a tumor suppressor; however, at later stages, it can act as a potent promoter of multiple steps of the metastatic process, which comprise local motility/invasion, entry of cancer cells into the blood stream (intravasation), exit from the blood vessels (extravasation), and colonization of distant organs. The relevance of TGF- $\beta$  signaling for disease progression has been particularly recognized in tumors where cancer cells retain the core TGF- $\beta$ -signaling components, as is frequently the case in breast and lung cancers. However, a major remaining challenge is the identification of TGF- $\beta$  target genes that drive the different steps of metastasis, especially since TGF- $\beta$  modulates gene expression in a highly cell- and context-specific manner. Although some progress to this end has been made in the context of breast cancer metastasis, the genes and mechanisms that mediate the prometastatic effects of TGF- $\beta$  in lung adenocarcinoma (ADC) remain largely unknown.

To explore the molecular mechanisms that could mediate the prometastatic effects of TGF- $\beta$  in lung ADC, we took a candidate gene approach and started off by scrutinizing members of the DOCK180-related protein superfamily. We found that in lung ADC cells, expression of DOCK4, but not other DOCK180-family members, is rapidly and robustly induced by TGF- $\beta$  in a Smad-dependent manner. Moreover, we found that DOCK4 is a direct TGF- $\beta$ /Smad target

gene and, importantly, that high DOCK4 expression correlates with activated TGF- $\beta$  signaling and poor prognosis in human lung ADC. These findings prompted us to assess the role of DOCK4 in mediating the prometastatic effects of TGF- $\beta$  in lung ADC in vivo in a mouse model. We found that DOCK4 induction is essential for TGF- $\beta$ -driven lung ADC metastasis. Specifically, we observed that blockade of TGF- $\beta$ -induced DOCK4 attenuates the ability of lung ADC cells to extravasate into distant organ sites, resulting in a marked reduction in metastatic burden in mice. At a cellular level, we obtained evidence that TGF- $\beta$ -induced DOCK4 signaling facilitates extravasation by stimulating lung ADC cell protrusive activity, motility, and invasion, without affecting epithelial-to-mesenchymal transition (EMT), and, intriguingly, that it does so by driving Rac1 activation. So far, Rac1 has only been linked to TGF- $\beta$  via a noncanonical pathway. Thus, our findings identified the atypical Rac1 activator DOCK4 as a novel, key component of the TGF- $\beta$ /Smad pathway that promotes lung ADC cell extravasation and metastasis.

### PUBLICATIONS

- Murray DW, Didier S, Chan A, Paulino V, Van Aelst L, Ruggieri R, Tran NL, Byrne AT, Symons M. 2014. Guanine nucleotide exchange factor Dock7 mediates HGF-induced glioblastoma cell invasion via Rac activation. *Br J Cancer* **110**: 1307–1315.
- Nakano-Kobayashi A, Tai Y, Kasri N, Van Aelst L. 2014. The X-linked mental retardation protein OPHN1 interacts with Homer1b/c to control spine endocytic zone positioning and expression of synaptic potentiation. *J Neurosci* **34**: 8665–8671.
- Soderling SH, Van Aelst L. 2014. Principles driving the spatial organization of Rho GTPase signaling at synapses. In *Ras superfamily of small G proteins: Biology and mechanisms 1* (ed. Wittinghofer A), pp. 395–419. Springer, New York.
- Tai Y, Janas JA, Wang C-L, Van Aelst L. 2014. Regulation of chandelier cell cartridge and bouton development via DOCK7 mediated ErbB4 activation. *Cell Rep* **6**: 254–263.
- In Press*
- Penzo M, Robert V, Tucciarone J, Darvas M, Huang ZJ, Palmiter R, De Bundel D, Wang M, Van Aelst L, Parada L, He M, Li B. 2015. The paraventricular thalamus controls a central amygdala fear circuit. *Nature* (in press).
- Yu J-R, Tai Y, Jin Y, Hammell MC, Wilkinson JE, Jae-Seok Roe J-R, Vakoc CR, Van Aelst L. 2015. TGF- $\beta$ /Smad signaling through DOCK4 facilitates lung adenocarcinoma metastasis. *Genes Dev* **29**: 250–261.

# UNDERSTANDING MALIGNANT GLIOMA PATHOGENESIS

H. Zheng   S. Klingler   C. Wu  
F. Li                      L. Zhang

Our laboratory studies the molecular and developmental programs that control initiation and progression of malignant glioma, the most common and lethal type of brain tumor. Our long-term goal is to define the complex biology of malignant glioma pathogenesis with the aim of identifying novel therapeutic and pharmacological opportunities that will improve treatment options for glioma patients. During the past year, we have continued to focus our research around two major areas.

1. We are developing genetically engineered animal models to recapitulate the process of human glioma pathogenesis and utilizing these animal models as tools to investigate *in vivo* tumor initiation and progression, as well as their response to various therapeutic treatments.
2. We are also identifying the genetic and epigenetic pathways governing normal and neoplastic neural stem and glioma cell-fate determination.

## **EGFR-Dependent and -Independent Mechanisms Underlie EGFR-Targeted Therapeutic Resistance in Malignant Glioma**

The epidermal growth factor receptor (EGFR) belongs to a family of receptor tyrosine kinases, and it is highly amplified, mutated, and overexpressed in human malignant gliomas, making it a compelling candidate for targeted therapies. Despite the prevalence and growth-promoting functions of EGFR, therapeutic strategies including EGFR tyrosine kinase inhibitors (TKIs) to target its kinase activity have not been translated into profound beneficial effects in glioma clinical trials. The glioma patient's poor response to EGFR TKIs is in contrast to the remarkable therapeutic benefits of these inhibitors in lung cancer patients carrying EGFR mutations. The reason behind these differential responses has remained poorly understood.

We have begun to uncover the mechanisms underlying the EGFR-targeted therapeutic resistance in

malignant gliomas. Notably, although glioma and lung cancer EGFR mutants both exhibit ligand-independent autophosphorylation and activation, one major difference between them is the distribution of mutations within the EGFR-coding sequence; lung cancer mutations generally reside in the EGFR intracellular kinase domain, whereas glioma mutations mainly cluster in the extracellular domain. These tumor-type-specific EGFR mutational patterns likely reflect different functional requirements during tumor initiation and progression. In our studies to define their response to EGFR TKIs, we demonstrated that the kinase activities of glioma-specific EGFR mutants can be efficiently inhibited just like those of the lung cancer. These observations prompted us to reason that the poor response of glioma patients to EGFR TKIs might be due to inefficient drug penetration through the blood-brain barrier or that EGFR might not be required for glioma cell survival and propagation. To address these questions, we collaborated with scientists from Columbia University and Weill Cornell Medical College to develop a novel malignant glioma mouse model driven by the inducible expression of a glioma-specific mutant EGFR. Using an approach combining genetic and pharmacological interventions, we further demonstrated that oncogenic EGFR indeed serves a tumor-maintenance role in EGFR-driven malignant gliomas. Genetic suppression of mutant EGFR expression led to significant tumor regression. This is in contrast to the meager effect of EGFR TKI treatment despite the fact that the inhibitors can efficiently inhibit mutant EGFR kinase activity within the gliomas. Our findings suggest that the poor response to EGFR TKIs in glioma clinical trials was not due to the inhibitors' inability to suppress EGFR phosphorylation but rather that glioma was less dependent on EGFR kinase activity. To achieve optimal efficacy against EGFR-driven glioma, the next generation of targeted therapeutics will need to consider other EGFR functional domains as well as its kinase activity.

In addition to the EGFR-dependent resistance mechanism, our studies further uncovered an EGFR-independent mechanism responsible for glioma relapse

following genetic ablation of mutant EGFR expression. In collaboration with scientists from the University of Texas M.D. Anderson Cancer Center, we demonstrated that a subpopulation of glioma cells that survived oncogenic EGFR ablation and were responsible for relapse existed prior to treatment. Interestingly, we found that the EGFR-independent relapsed tumors were strongly enriched for mesenchymal features. Addition of a phosphoinositol-3 kinase/mammalian target of rapamycin (PI3K/mTOR) inhibitor could significantly delay the relapse and prolong survival. Together, our research reveals important mechanistic insights into glioma therapeutic resistance development and provides a platform for testing therapies targeting EGFR signaling-driven malignant gliomas.

### Characterization of the Function of ATRX during Glioma Pathogenesis

---

*α-thalassemia/mental retardation syndrome X-linked (ATRX)*, encoding a Rad54-like ATP-driven DNA translocase, belongs to the SWI/SNF family of chromatin remodelers. *ATRX* was initially identified to be the gene involved in the X-linked *α-thalassemia/mental retardation syndrome*. Its germline mutation causes a form of syndromal mental retardation with multiple developmental abnormalities. The first link of *ATRX* with cancer came from the observation that a cohort of patients presenting with myelodysplastic syndrome (MDS) concurrently harbor *ATRX* mutation. More recently, frequent somatic loss-of-function *ATRX* mutations have been identified in human pediatric and adult gliomas, pancreatic neuroendocrine tumors (PanNETs), osteosarcomas, and others, indicating the tumor suppressor roles of *ATRX*. Functionally, it has been shown that *ATRX*, in combination with its interaction partner Death domain-associated protein 6 (DAXX), can facilitate replication-independent histone variant H3.3 incorporation into pericentric, telomeric, and other heterochromatin regions with high repeat sequences; however, its link to tumorigenesis has remained poorly understood.

In collaboration with Scott Lowe and Chris Vakoc's groups here at CSHL, we previously initiated a project to probe the epigenetic networks controlling neural stem and/or glioma cell self-renewal and differentiation. By conducting a comprehensive screen of a customized small hairpin RNA (shRNA) library targeting epigenetic regulators, we identified multiple chromatin modulators including *ATRX* whose suppression attenuates the differentiation capacity of neural stem and glioma cells. Suppression of *Atrx* protein expression strongly promotes glioma formation from a mouse neural stem cell inactivated of *p53* and *Pten* tumor suppressor genes. More recently, we demonstrated that deletion of *Atrx* in neural stem cells could block the neuronal lineage terminal differentiation but had minimal impact on glial lineages. This is consistent with our studies examining *Atrx* expression pattern in adult mouse brains, in which we found that *Atrx* protein levels increase upon neuronal differentiation with strongest expression confined to postmitotic neuron cells. Our findings therefore indicate that *ATRX* has a critical role controlling neuronal maturation, and its inactivation may facilitate lineage reversion and glioma formation.

It is well known that cancer cells must overcome the progressive telomere attrition that occurs with each cell division in order to grow indefinitely. Unlike the majority of cancer cells that circumvent the replicative mortality through reactivation of telomerase, *ATRX* mutant tumors seemingly rely on a telomerase-independent mechanism termed the alternative lengthening of telomeres (ALT) pathway for telomere maintenance. In collaboration with scientists from the Wistar Institute in Philadelphia, we obtained evidence that *ATRX* is involved in telomere maintenance by regulating telomerase activity. Deletion of *ATRX* expression in human glioma cells leads to progressive cellular senescence or crisis that can be rescued by introduction of exogenous telomerase expression. Additionally, we found that *ATRX* functions to repress aberrant telomeric recombination and has a key role in telomere homeostasis. Collectively, our findings point to a multifunctional role for *ATRX* during glioma pathogenesis.

# NEUROSCIENCE

---

How does the brain encode stimuli from the outside world, within and across sensory modalities, to generate specific perceptions that trigger complex behaviors? How is the brain shaped by sensory experience and contextual information, and what modifications occur in neuronal circuits that allow us to adapt to unexpected changes in our environment? These are the questions guiding the work of **Florin Albeanu**, who is searching for neuronal mechanisms underlying behavior by investigating circuit dynamics in the olfactory bulb, the olfactory cortex, and the corresponding premotor areas of mice. Advances in optical imaging, patterned illumination, and optogenetics, combined with electrophysiological recordings, enable Albeanu and colleagues to monitor and/or alter patterns of activity at unprecedented synaptic and millisecond resolution, in real time, as animals are engaged in various behaviors. For survival, rodents need to identify the smells of objects of interest such as food, mates, and predators, across their recurring appearances in the surroundings, despite apparent variations in their features. Furthermore, animals aptly extract relevant information about the environment across different sensory modalities, combining olfactory, visual, or auditory cues. By recording neuronal activity in the input and output layers of the olfactory bulb and cortex, as well as feedback from cortical areas and neuromodulatory signals, Albeanu and his team aim to understand computations these circuits perform and how this information is decoded deeper in the brain and leads to action.

The study of decision making provides a window into the family of brain functions that constitute cognition. It intervenes between perception and action and can link one to the other. Although much is known about sensory processing and motor control, much less is known about the circuitry connecting them. Some of the most interesting circuits are those that make it possible to deliberate among different interpretations of sensory information before making a choice about what to do. **Anne Churchland**'s lab investigates the neural machinery underlying decision making. Lab members use carefully designed paradigms that encourage experimental subjects to deliberate over incoming sensory evidence before making a decision. Recent results show that rats and humans have a statistically similar decision-making ability. To connect this behavior to its underlying neural circuitry, they measure electrophysiological responses of cortical neurons in rodents as they perform designated tasks. The lab's current focus is on the parietal cortex, which appears to be at the midpoint between sensory processing and motor planning. Churchland and colleagues also use theoretical models of varying complexity to further constrain how observed neural responses might drive behavior. This approach generates insights into sensory processing, motor planning, and complex cognitive function.

Research in **Josh Dubnau**'s lab is concentrated on two different questions. First, Dubnau and his team are investigating mechanisms of memory using *Drosophila* as a model system. A second area of research is focused on uncovering mechanisms of neurodegeneration that underlie amyotrophic lateral sclerosis (ALS) and frontotemporal lobar degeneration (FTLD). Work in the Dubnau lab has suggested a novel hypothesis to explain neurodegeneration in these disorders. They discovered that awakening retrotransposons in the genome of some brain cells might be responsible for causing cell death. Retrotransposons are virus-like repetitive elements that are encoded in the genome and are capable of replicating and inserting into new chromosomal positions. This can lead to DNA damage and cell death by a process known as apoptosis. The lab is investigating this hypothesis for ALS/FTLD using a multidisciplinary approach that includes experimental work with fly, mouse, cell culture, and human postmortem tissue. Computational analyses of genomic data are performed in collaboration with Molly Hammell's group. If the retrotransposon hypothesis is

correct, it will change the trajectory of neurodegeneration research and have obvious clinical impact. Retrotransposon RNAs and proteins are promising new diagnostic markers and potentially important therapeutic targets.

**Grigori Enikolopov** and colleagues study stem cells in the adult brain. They have generated several models to account for how stem cells give rise to progenitors and, ultimately, to neurons, and they are using these models to determine the targets of antidepressant therapies, to identify signaling pathways that control generation of new neurons, and to search for neuronal and neuroendocrine circuits involved in mood regulation. Recent experiments suggest a new model of how stem cells are regulated in the adult brain, with a focus on stem cells' decision on whether to divide—and embark on a path of differentiation—or remain quiescent. This model explains why the number of new neurons decreases with advancing age—potentially leading to impairments in memory and depressed mood. It also explains why multiple brain trauma and prolonged neurodegenerative disease may lead to accelerated decrease of cognitive abilities. In other research, the team has identified stem cell targets of various therapies used for treating depression and developed a general platform to determine the effect of drugs and therapies and predict their action. The team is now focusing on the signaling landscape of neural stem cells and on their interaction with the surrounding niche. Enikolopov's group is also focusing on other types of stem cells in the organism. Their latest discovery, with a team at Cornell University, relates to a new type of stem cell in the ovary that normally heals the ovarian tissue after an oocyte is released, but easily transforms to become malignant and generate tumors. The team is now using these discoveries to reveal how stem cells relate to neural and oncological disorders.

**Hiro Furukawa's** lab 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 lab 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 has unraveled the overall architecture and discovered and mapped several regulatory sites in specific classes of NMDA receptors—progress that now opens the way to the development of a potentially new class of drugs to modulate the receptor's activity.

**Josh Huang** and colleagues study the assembly and function of neural circuits in the neocortex of the mouse. The neocortex comprises a constellation of functional areas that form a representational map of the external (sensory, social) and internal (visceral, emotional) world. These areas are strategically interconnected into elaborate networks which support dynamic operations that process information and guide intelligent behavior. The group's overarching hypothesis is that, at the cellular level, cortical processing streams and output channels are mediated by a large set of distinct glutamatergic pyramidal neurons (PyNs), and functional PyN ensembles are regulated by a diverse yet distinct set of GABAergic interneurons (INs). The Huang lab systematically builds cell-type genetic tools that integrate a full set of modern techniques for exploring neural circuits. Building on their success in genetic targeting of GABA INs, they have recently extended this

effort to PyNs. Their current research program is beginning to integrate studies of GABA INs and PyNs toward understanding the development and function of specific cortical circuits underlying behavior. Among GABA interneurons, the chandelier cell is one of the most distinctive cell types that controls PyN firing at the axon initial segment. Huang and colleagues are studying the developmental specification of chandelier cells, their activity-dependent circuit integration, and their functional connectivity. Regarding pyramidal neurons, they are systematically characterizing the developmental origin, axon projection, and input connectivity of multiple classes of genetically defined PyN types, focusing on the forelimb motor cortex. They combine a range of approaches that include genetic and viral engineering, cell-type gene expression, genetic fate mapping, imaging, electrophysiology, and behavior analysis. With this progress, they are beginning to integrate their studies in the context of the motor cortex control of volitional forelimb movements and motor learning.

**Adam Kepecs** and colleagues are interested in identifying the neurobiological principles underlying cognition and decision making. They use a reductionist approach, distilling behavioral questions to quantitative behavioral tasks for rats and mice that enable the monitoring and manipulation of neural circuits supporting behavior. Using state-of-the-art electrophysiological techniques, they first seek to establish the neural correlates of behavior and then use molecular and optogenetic manipulations to systematically dissect the underlying neural circuits. Given the complexity of animal behavior and the dynamics of neural networks that produce it, their studies require quantitative analysis and make regular use of computational models. The team also has begun to incorporate human psychophysics to validate its behavioral observations in rodents by linking them with analogous behaviors in human subjects. Currently, the team's research encompasses study of (1) the roles of uncertainty in decision making, (2) the division of labor among cell types in prefrontal cortex, (3) how the cholinergic system supports learning and attention, and (4) social decisions that rely on stereotyped circuits. A unifying theme is the use of precisely timed cell-type- and pathway-specific perturbations to effect gain and loss of function for specific behavioral abilities. This past year, Kepecs and colleagues discovered in experiments with rats that confidence is a measurable quantity; they identified a brain region, the orbitofrontal cortex, whose function is required for the animals to express confidence in their decisions. In other notable work, the lab has identified a class of inhibitory neurons that specializes in inhibiting other inhibitory neurons in cerebral cortex and is recruited during behavior when rewards and punishment are delivered. Through manipulations of genetically and anatomically defined neuronal elements, the team hopes to identify fundamental principles of neural circuit function that will be useful for developing therapies for diseases such as schizophrenia, Alzheimer's disease, and autism spectrum disorder.

**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.

Dysfunction of synapses in the brain is believed to have an important role in the pathogenesis of major psychiatric disorders, including depression, anxiety, and schizophrenia. But what are the causes? Where in the brain does the dysfunction occur? How does it result in the behavioral



symptoms of illness? To address these issues, **Bo Li** and colleagues are studying, in animals, normal synaptic plasticity underlying adaptive behaviors and synaptic aberrations responsible for maladaptive behaviors that are related to depression, anxiety, and schizophrenia. Their long-term goal is to develop methods allowing the manipulation of activity in specific brain circuits in order to change disease-related behaviors. Li's group uses a variety of methodologies, including patch-clamp recording and calcium imaging of labeled neurons, two-photon imaging of spine morphology and tagged receptors, in vivo stereotaxic virus injection, RNA interference (RNAi)-based gene silencing, activation of specific axon terminals using light-gated cation channels, activation or silencing of specific brain regions using transgenes, and assessment of the behavioral consequences of certain manipulations. A project focusing initially on a gene called *ErbB4* seeks to determine the genetic causes of attention deficit, a cognitive impairment that is consistently observed in schizophrenia. During the last two years, Li and his lab members demonstrated that neurons in a tiny area of the mammalian brain called the central amygdala encode fear memory and control fear expression. These and related findings lay the foundation for future work aimed at understanding the circuit mechanisms of anxiety disorders, particularly post-traumatic stress disorder (PTSD). The team has also identified a group of neurons in the medial prefrontal cortex (mPFC) that determines how mice respond to stress. These neurons, hyperactive in the depressed brain, are weakened in mice with natural stress resilience.

**Partha Mitra's** group studies complex neurobiological systems using a combination of experimental and computational approaches. The primary experimental focus continues to be the Mouse Brain Architecture (MBA) project, with a goal to generate a mesoscale connectivity map. The approach is to microinject anterograde and retrograde neuronal tracer substances at target locations systematically selected on a grid in the brain and then histologically process, image, and digitally reconstruct whole brains and visualize all labeled connections. Currently, more than 800 tracer-injected mouse brains are viewable via the on-line digital microscope on the MBA project web portal. Mitra's neuroinformatics research involves the development of analytical tools and informatics infrastructures to process, analyze, and integrate large volumes (totaling about a petabyte) of neurobiological data in various brain connectivity projects. In collaborative studies, the team applies methods of whole-brain digital neuroanatomy developed for the MBA project to normal and disease model brains in various species. A joint study with Josh Huang to obtain a brain-wide census of genetically targeted GABAergic interneurons in the mouse brain during development and in genetic models of autism is drawing to a conclusion. New work with Pavel Osten seeks to obtain an accurate whole-brain census of neuron classes in cell-type-specific reporter mice and to develop neuroinformatics infrastructure to analyze and integrate such data. With Sacha Nelson (Brandeis University), the team is engaged in a cross-species study of neuronal cell types in transgenic strains of rats and mice. The lab also has collaborations with Marcello Rosa (Monash University, Australia) and Hideyuki Okan (RIKEN, Japan) on a project to develop the first atlas of the brain of a New World primate, the marmoset. In collaboration with Balaraman Ravindran, Mitra has initiated a Center for Computational Brain Research at the Indian Institute of Technology (IIT), Madras. In parallel, Mitra is undertaking theoretical work at the interface among physics, engineering, and biology. In this line of investigation, Mitra and his collaborators bring methods from statistical physics and control theory to bear on problems in complex interconnected biological systems, including whole-brain connectivity.

**Pavel Osten's** lab works on identification and analysis of brain regions, neural circuits, and connectivity pathways that are disrupted in genetic mouse models of autism and schizophrenia. Osten hypothesizes that (1) systematic comparison of multiple genetic mouse models will allow determination of overlaps in pathology—neural circuit endophenotypes—responsible for the manifestation of neuropsychiatric disorders and (2) neural circuit-based classification of autism and

schizophrenia will provide key circuit targets for detailed mechanistic studies and therapeutic development. Osten and colleagues have developed the first systematic approach to the study of neural circuits in mouse models of psychiatric diseases, based on a pipeline of anatomical and functional methods for analysis of mouse brain circuits. An important part of this pipeline is high-throughput microscopy for whole-mouse brain imaging, called serial two-photon (STP) tomography. They have used this method to describe the first whole-brain activation map representing social behavior in normal mice. They are currently focusing on using this approach to study brain activation changes in two mouse models of autism, the 16p11.2 *df/+* mouse model, which shows an increased propensity to seizures and hyperactivity, and the *CNTNAP2* knockout mouse model, which shows abnormal social behavior.

**Stephen Shea's** lab 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 lab 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 or GCs) that has long been hypothesized to be crucial for memories but has resisted direct study. They have developed methods for recording that are 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 lab. The two labs 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 lab is using imaging techniques to determine how memories are stored among broad neuronal ensembles, at a different level of the system. Recently, the lab 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 lab 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 lab to extend this work by directly linking these deficits to the action of the gene *MeCP2* in the auditory cortex.

What is a memory? When we learn an association, information from two different sensory streams somehow becomes linked together. What is this link in terms of neural activity? For example, after a few bad experiences, we learn that the "green" smell of an unripe banana predicts its starchy taste. How has the neural response to that green smell changed so it becomes linked to that taste? What are the underlying mechanisms—what synapses change strength, what ion channel properties change? These are the questions that drive research in **Glenn Turner's** laboratory. His team addresses these questions by tracking neural activity using a combination of different techniques. Using electrophysiological methods, they can examine individual neurons with very high resolution, monitoring synaptic strength and spiking output. They have also developed functional imaging techniques to monitor the activity of the entire set of cells in the learning and memory center of the fly brain. This comprehensive view of neural activity patterns enables them to actually predict the accuracy of memory formation in separate behavioral experiments. This year, the Turner lab was able to map the activity of a particular region of the brain that is associated with

learning and memory. They found that a remarkably small number of neurons are required for flies to distinguish between odors. The Turner lab also studied the role of a specific type of cells known as Kenyon cells, which receive input via several large claw-like protrusions. These neurons use their claws to recognize multiple individual chemicals in combination in order to remember a single scent. By examining the effects of learning-related genes on these processes, they can in the future connect their network-level view of memory formation to the underlying molecular mechanisms that govern the basic cellular and synaptic changes that drive learning.

**Anthony Zador** and colleagues study how brain circuitry gives rise to complex behavior. Work in the lab 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 lab 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 lab, 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.

**Yi Zhong’s** lab studies the neural basis of learning and memory. The team works with fruit fly models to study genes involved in human cognitive disorders, including neurofibromatosis, Noonan syndrome (NS), and Alzheimer’s disease. Mutations leading to a lack of function of the neurofibromatosis 1 (*NF1*) gene cause noncancerous tumors of the peripheral nervous system as well as learning defects. The lab’s analyses of *Drosophila NF1* mutants have revealed how expression of the mutant gene affects a pathway crucial for learning and memory formation. The *NF1* gene and a gene called *corkscrew*, implicated in NS, share a biochemical pathway. Recently, the lab succeeded in linking changes in this pathway due to specific genetic defects in NS with long-term memory deficiencies. In fly models, they discovered the molecular underpinnings of the “spacing effect”—the fact that memory is improved when learning sessions are spaced out between rest intervals. Zhong’s team also has succeeded in reversing memory deficits in mutant flies, work suggesting that longer resting intervals for Noonan’s patients might reverse their memory deficits. They also identified a means of reversing memory loss in fruit flies while suppressing brain plaques similar to those implicated in Alzheimer’s disease by blocking epidermal growth factor receptor (EGFR) signaling, a pathway commonly targeted in cancer. Separately, having discovered that memory decay is an active process regulated by the Rac protein, the team has proposed that Rac’s role in erasing memory is related to its influence on downstream cytoskeleton remodeling agents. This year, the Zhong lab explored how neurons control our response to different scents, offering insight into how the brain distinguishes between food odors that are attractive and repulsive. They determined that neurons expressing a particular peptide were only activated by food odors and that the amount of activation predicted how strongly a fly was attracted to a particular odor.

# UNDERSTANDING NEURONAL CIRCUITS IN THE MAMMALIAN OLFACTORY BULB

D. Albeanu   F. Anselmi   H.G. Chae   P. Alexandre Garcia Da Silva   M. Koh   G. Otazu  
A. Banerjee   M. Davis   P. Gupta   F. Marbach   B. Rebouillat

The focus of our research group is to understand how neuronal circuits encode and interpret inputs from the environment across different contexts, senses, and brain states to generate meaningful behaviors. Toward this end, we use optogenetic methods (such as fast multiphoton laser-scanning imaging of genetically encoded neuronal activity reporters or patterned illumination of light-gated neuronal activity switches) coupled with electrophysiological measurements (extracellular and intracellular recordings). We want to understand (1) how inputs get processed at different synapses of the underlying neuronal circuits, (2) how these representations change with the state of the system and its circuits (awake vs. anesthetized, naïve vs. behaving, learning), and (3) what changes in the activity patterns of well-defined neuronal circuits contribute to specific behaviors. The broad scope of this effort is observing how perceptions arise.

We use the rodent olfactory system as a model and monitor neuronal inputs, outputs, and feedback loops across different layers of the circuit. We modulate activity of select neuronal populations in a reversible manner in awake animals to understand how fundamental sensory encoding problems are solved by the olfactory bulb (OB) and olfactory (piriform) cortex networks.

## Characterizing the Input–Output Transform of the Olfactory Bulb

---

In the OB, sensory neurons expressing the same type of olfactory receptor converge in tight focus, forming ~2000 clusters of synapses called glomeruli. The layout of glomeruli on the bulb is highly reproducible across individuals with a precision of 1 part in 1000. However, nearby glomeruli are as diverse in their responses to odors as distant ones, lacking an apparent chemotopic arrangement (Soucy et al., *Nat Neurosci* 12: 210 [2009]). From each glomerulus, a

few dozen mitral cells (principal output neurons of the OB) carry the output further to the olfactory cortex and several other brain areas. Such “sister” mitral cells typically have only one primary dendrite that projects to a single glomerulus, but they can sample inputs on their primary and secondary dendrites from functionally diverse glomeruli via several types of interneurons. Using optogenetic manipulations and digital micro-mirror device (DMD)-based patterned photostimulation, we found that sister cells carry different information to the cortex: average activity representing shared glomerular input, and phase-specific information that refines odor representations and is substantially independent across sister cells (Dhawale et al., *Nat Neurosci* 13: 1404 [2010]).

Activity in the bulb is a rich mix of excitation and inhibition, via both direct inputs and feedback connections. Many different classes of interneurons interact with the sensory afferents and/or the output neurons via both short- and long-range connections. Their connectivity patterns and roles in olfactory processing, however, remain largely unknown. We use a Cre/loxP approach to express reporters (i.e., synaptopHluorin, GCaMP6) and light-gated switches of neuronal activity (ChR2, Halorhodopsin, Arch) in different bulbar neuronal types. We monitor the inputs (glomerular activity patterns) and the outputs (mitral cell firing) in response to numerous odorants across a wide range of concentrations (three orders of magnitude) as we systematically perturb the activity of the input nodes (glomeruli) and of different interneuron types. Patterned photostimulation allows us to bypass odor stimulation and gain precise spatiotemporal control over the inputs by directly activating glomerular patterns of choice. We record bulbar outputs via multitetrode recordings and patch clamp, or we use optical imaging readouts via multiphoton microscopy in vivo to understand what computations the olfactory bulb performs.

We are currently investigating the roles of two classes of interneurons: dopaminergic/GABAergic cells

(DAT+) that broadcast long-range signals in the glomerular layer and granule cells (GCs) that establish reciprocal synapses with mitral/tufted cells and receive rich cortical feedback input.

### Long-Range Interactions between DAT+ Cells and External Tufted Cells Implement Gain Control and Gate the Output of the Glomerular Layer of the OB

---

Odors elicit distributed activation of input nodes (glomeruli) on the OB. This necessitates long-range interactions between coactive glomeruli. Long-range projecting DAT+ cells provide the earliest opportunity for such cross-talk. DAT+ cells, in the glomerular layer, receive inputs from olfactory sensory neurons and/or external tufted (ET) cells and release both GABA and dopamine, synapsing onto ET cells as far as tens of glomeruli away (Kiyokage et al., *J Neurosci* 30: 1185 [2010]). Computational models (Cleland et al., *Trends Neurosci* 33: 130 [2007]) have suggested that these cells may be involved in long-range normalization of bulb outputs, but to date, their function in the intact brain has not been investigated. We find that DAT+ cells convey concentration-dependent odor signals by broadcasting focal glomerular inputs throughout the OB. Using optogenetic and pharmacogenetic approaches, we find that DAT+ cells influence distant mitral/tufted (M/T) cells by modulating local excitatory interneurons (ET cells). Synaptic action of DAT+ cells enables long-range suppression of ET cells. In addition, we find electrical coupling between DAT+ and ET cells, which may facilitate local synchronization of ET cell responses. Our results indicate that ET cells are essential gatekeepers of glomerular output and prime determinants of M/T responsiveness. Furthermore, using a cell-specific ablation approach, we show that long-range DAT + -to-ET action implements gain control and contrast enhancement via M/T cell odor response decorrelation, by allowing multiple coactive inputs to alter the output of any given glomerulus (Fig. 1).

### Characterization of Granule Cell Odor Responses in Awake Head-Fixed Mice

---

GCs mediate both lateral and recurrent inhibition by forming reciprocal synapses with M/T, the principal output neurons of the OB. GCs receive glutamatergic

inputs both from M/T cells and from feedback axons originating in the olfactory cortex. GCs are the most numerous cells in the OB (~90%), outnumbering the M/T cells by two orders of magnitude, and further represent the main target of direct feedback from the olfactory cortex.

Computational models and behavioral studies have suggested critical roles for these cells in olfactory processing and learning. However, very little is known about their response properties in vivo, because of technical difficulties in electrophysiological recordings from these small-sized neurons.

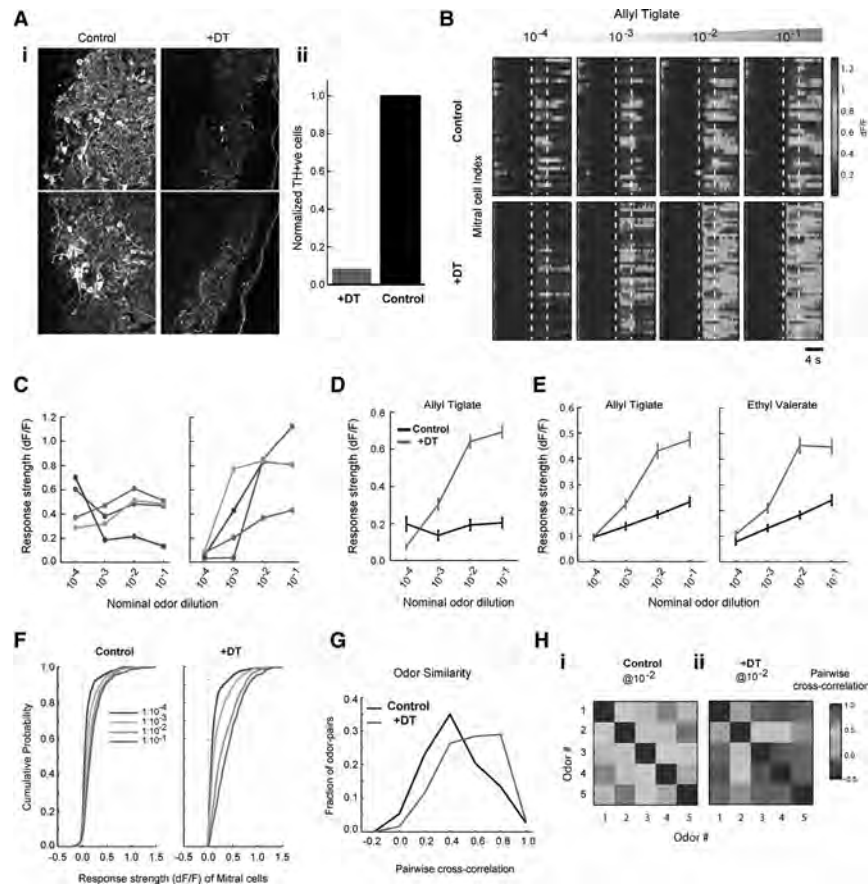
To begin characterizing the odor response properties of GCs, we used a genetically encoded calcium indicator (GCaMP6f) in the GCs (CST-Cre) and multiphoton imaging to monitor the odor-evoked responses of GCs (250–350  $\mu\text{m}$  deep) in awake head-fixed mice (Fig. 2). GCs showed robust spontaneous activity and were sparsely activated upon odor presentation, displaying a diverse range of enhanced and suppressed, ON, OFF, and ON–OFF responses. Enhanced responses were more common than suppressed responses (65% vs. 35%). Furthermore, a significant fraction (~25%) of GCs exhibited characteristic enhanced OFF responses, independent of stimulus duration. Pairwise analysis of GCs monitored simultaneously indicates that neighboring GCs are as diverse in their odor responses as pairs of distant cells. No spatial clustering of similarly odor-responding GCs was apparent within a 350- $\mu\text{m}$  range. Enhanced responsive GCs increased monotonically in number across concentrations. Suppressed and OFF responsive GCs varied in a complex fashion with concentration.

To dissect the contribution of the corticobulbar feedback to GC activity, we are using pharmacological and optogenetic manipulations of the cortical input in tandem with multiphoton imaging of GCs and MTs. Furthermore, we are currently monitoring how GC responses evolve during learning and performing of olfactory detection and discrimination tasks.

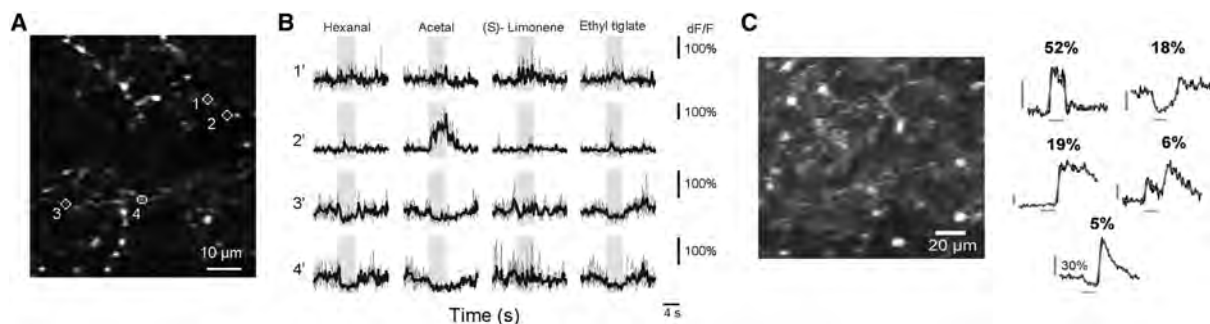
### Dissecting the Spatial and Temporal Features of the Glomerular Code in Awake Behaving Animals via Optogenetic Strategies

---

Odorants activate precise spatiotemporal glomerular activity patterns on the OB surface. These patterns



**Figure 1.** DAT+ cells implement gain control and decorrelate mitral cell odor responses. (A) TH+ immunohistochemistry in the olfactory bulb (OB) of a DAT-Cre  $\times$  Thyl-GCaMP3.0 mouse injected with AAV2.9-FLEX-DTR-GFP virus on the dorsal aspect of the OB followed by diphtheria toxin (DT) intraperitoneal injection. Confocal images of the TH+ signal in two representative fields of views (FOVs) on the ventral (*left*) and dorsal (*right*) aspect of a given slice. Average counts of TH+ cells from all dorsal FOVs and from all ventral FOVs in DTR-injected mice, normalized by the number of visible glomeruli (100% for control and 8.3% for +DT,  $n = 1259$  cells, eight FOVs, two mice). Residual signals mostly consisted of neuropil. (B) Baseline-subtracted, normalized GCaMP3.0 ensemble mitral cell responses to increasing concentrations of Allyl tiglate. (*Top*) Day 0 before injection of DT (control,  $n = 33$  cells). (*Bottom*) Different FOVs in the same OB on Day 7 after DT injection (+DT,  $n = 44$  cells). Each row represents an individual mitral cell (ROI) in the same FOV. Gray color indicates odor presentation (4 s). (C) (*i*) Distribution of onset latencies of mitral cell odor responses for control ( $1.55s \pm 1.51s$ ,  $n = 3660$  cell-odor pairs) and +DT condition ( $1.2s \pm 1.39s$ ,  $n = 3000$  cell-odor pairs). Numbers denote mean and standard deviation. Both the mean (two-sample  $t$ -test,  $p < 10^{-18}$ ) and the variance ( $F$ -test,  $p < 10^{-5}$ ) are significantly smaller under the +DT condition. (*ii*) Odor-evoked response (dF/F) of four example mitral cells from B, as a function of odor concentration. (*iii*) Mean odor-evoked response (dF/F) of all mitral cells from B as a function of odor concentration for Allyl tiglate on Day 0 (black,  $n = 33$  cells) and Day 7 after +DT administration (gray,  $n = 44$  cells). (D) Mean odor-evoked response (dF/F) of all mitral cells pooled across experiments as a function of odor concentration for Allyl tiglate and Ethyl valerate. Concentration response curves are shown for control (black,  $n = 168$  cells, seven FOVs, seven mice) and +DT conditions (light gray,  $n = 150$  cells, seven FOVs, three mice). (E) Cumulative distribution of odor response strength of all mitral cells pooled across experiments, to five odors as a function of concentration under the control condition ( $n = 168 \times 5$  cell-odor pairs for each concentration) and after +DT injection ( $n = 150 \times 5$  cell-odor pairs for each concentration). (F) Cumulative distribution of slopes of the ensemble mitral cell concentration response curves separately fitted for each FOV and odor, under the control and +DT conditions ( $n = 35$ , seven FOVs  $\times$  five odors). (G) Populations sparseness of mitral cell odor representations for each concentration under the control (black,  $n = 168$  cells, seven FOVs) and +DT conditions (gray  $n = 150$  cells, seven FOVs). Each bar denotes the mean population sparseness for five odors at each concentration. (H) (*i*) Matrix of correlation coefficients between the neural response spectra (length of vector = number of recorded neurons) for each odor pair at 1:100 dilution under the control and +DT conditions. (*ii*) Distribution of pairwise correlation coefficients between all stimulus pairs ( $n = 190$  pairs, five odors at four concentrations) under the control and +DT conditions.



**Figure 2.** Cortical feedback and granule cell dynamics (A) Multiphoton imaging of cortical feedback boutons labeled with GCaMP5, 250  $\mu\text{m}$  deep from OB surface. (B) Example odor responses from individual boutons selected in A in an awake animal. (C) GCs in an awake head-fixed CST-Cre mouse expressing GCaMP6f in the OB. (Right) Odor response types and their frequency in GC population. (From top left) Types are enhanced ON, suppressed ON, enhanced OFF, enhanced ON–OFF, suppressed ON – enhanced OFF.

constitute the first representation of odor input into the olfactory system and hence must contain all meaningful attributes of the incoming odor stimulus, such as odor identity, intensity, and temporal dynamics. Different odors activate distinct overlapping and nonoverlapping glomerular maps. In principle, these odor maps can vary in several possible features, such as number (how many glomeruli are activated), space (which glomeruli), intensity (absolute and relative levels of activity of individual glomeruli), and timing (onset with respect to each other and the order of their activation). Previous experiments suggest that presenting different odorants (A vs. B), mixtures (A and B), or different concentrations of the same odorant (A) may all result in modulation of odor maps along many (if not all) features. It is unknown how the downstream circuitry (M/T cells, olfactory cortex) interprets these variations and assigns meaning to them, in order to segregate odor identity from changes in concentration or components within/from a mixture.

One way to understand how the circuit makes sense of changes in the glomerular activity patterns to extract relevant information is to systematically alter features of odor maps and study the concurrent changes in neuronal outputs and olfactory behavior. Hitherto, however, this has not been possible, primarily because of the inability to activate and modulate individual glomeruli in a controlled manner using odorants. We are using optogenetic tools to bypass odorant stimulation and simulate odor-like glomerular activity patterns, or alter them, by directly activating/inhibiting glomeruli using light in transgenic mice that express ChR2 or Arch in all olfactory sensory neurons (OSNs) or in a single type of OSN (Olf151). Furthermore,

using patterned illumination, we are activating/inhibiting select subsets of glomeruli with single-glomerulus precision in terms of intensity, onset time, or duration. This approach confers unprecedented ability to effect feature-specific perturbations in a glomerular pattern of choice. We aim to understand which features of an odor map are used by a behaving mouse under naturalistic conditions to identify odors and their attributes such as concentration, temporal fluctuations, or spatial location. Furthermore, we are testing the resolution at which mice can in principle detect variations within specific features of odor maps in a strictly controlled artificial regime. While doing so, we will also monitor neuronal responses at multiple layers in the olfactory system to compare the behavioral and neuronal detection thresholds and to understand the neuronal correlates of olfactory behavior.

To this end, we are training head-fixed and water-restricted mice to identify a target stimulus (“A”) as being unique from other nontarget (“not A”) stimuli by licking a “Left” versus “Right” water reward port placed on either side of the mouth. Animals are reinforced with water reward for correct choices and a mild air puff as punishment upon licking of the wrong port. We are using wide-field imaging (intrinsic, fluorescence) and two-photon microscopy to monitor the glomerular input maps evoked by the target (“A”) and the non-target stimuli (“not A”) as mice learn and perform the task. These activity patterns are further used as templates to design photostimulation masks for perturbation of stimulus-induced activity patterns during behavior. Once the task has been learned to >85% accuracy, we reversibly perturb the stimulus “A” activity pattern and determine the behavioral

threshold at which an altered activity pattern is classified as “not A”.

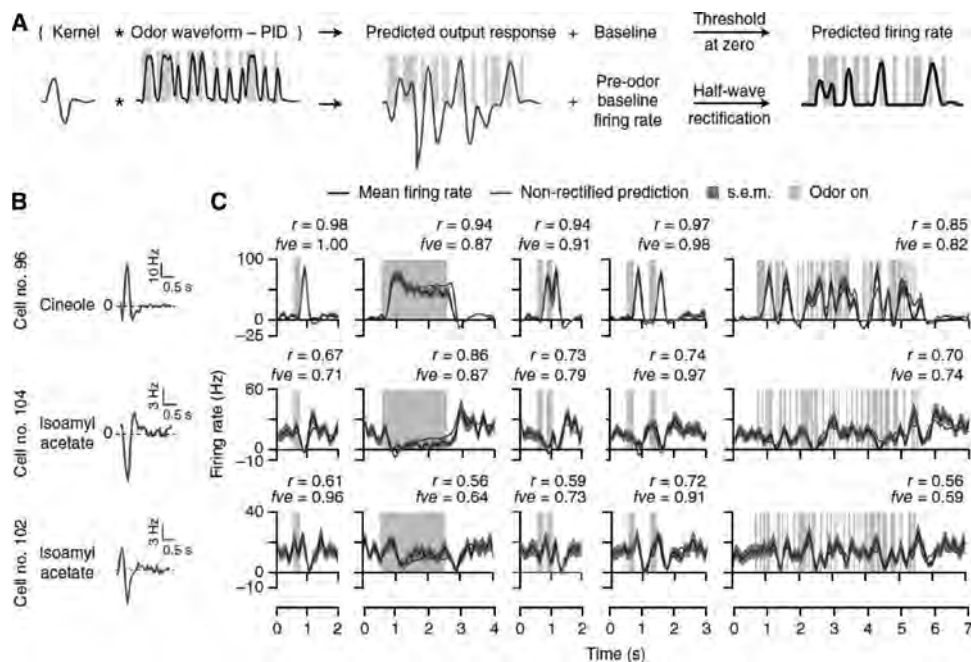
## OB Coding of Puffs, Mixtures, and Sniffs

The olfactory system receives intermittent and fluctuating inputs arising from dispersion of odor plumes and active exploration by the animal. The highly nonlinear olfactory transduction machinery, and complex excitatory–inhibitory circuitry, is believed to generate a nonlinear population trajectory of neuronal activity that distinguishes odorants. In contrast, we recently showed (Gupta et al., 2015) that individual M/T cells linearly sum inputs across odors and time. By decoupling odor sampling from respiration, we showed that M/T cell responses to arbitrary odor waveforms are well described by an odor-specific impulse response convolved with the odorant’s temporal profile (Fig. 3). This linear convolution applied to respiratory airflow predicts the classical inhalation-locked firing of OB neurons and several other reported response

properties of M/T cells. This raises interest in further understanding how this linear transform is achieved given the complexity of local microcircuits in the OB and feedback from other brain areas. These results suggest that the OB linearly processes fluctuating odor inputs, thus simplifying downstream decoding of stimulus identity and temporal dynamics.

## Investigating the Roles of Cortical Feedback in Invariant Odor Perception

Invariant perception refers to the identification of a sensory stimulus or an object of interest in a generalized fashion across its variable and recurring presentations. This is a common feature across all sensory modalities and becomes particularly preeminent in olfaction. Natural odor scenes are composed of odor plumes originating from multiple sources and traveling at fluctuating intensities that span several orders of magnitude. Despite this turbulent nature of the stimulus, rodents readily identify odors essential for their survival against varying



**Figure 3.** Schematic of the linear model. (A) The M/T cell response is calculated by convolving the odor kernel with the odor waveform (measured via PID). Negative firing rates are rectified after addition of the pre-odor baseline. Vertical bars mark odor ON periods. (B) Estimated odor kernels of three example M/T cells. Odors used are indicated at the left of each kernel. Odors were delivered at 1% saturation. (C) Comparison of model predictions and observed responses of the M/T cells in B to various time-varying odor stimuli. Vertical gray bars mark odor ON periods. Black lines show the observed mean firing rate across 12 trials. Gray bands indicate s.e.m. Red lines show the nonrectified prediction from the model using the estimated kernels shown in B.



odor-rich backgrounds and faithfully track a fluctuating odor stream to its source. Several models have attempted to explain invariant perception in different sensory modalities. Some of these models are based on the conventional view of sensory processing that relies solely on feed-forward information flow, whereas others argue that feed-forward mechanisms are insufficient and suggest the involvement of feedback between different neuronal layers. Anatomical studies show massive feedback projections from higher brain areas to the sensory periphery that often outnumber direct sensory inputs. The interplay of feed-forward and feedback signals has been proposed to be fundamental for learning and memory recall.

The OB receives rich glutamatergic projections from the piriform cortex. However, the dynamics and importance of these feedback signals remain unknown. We use multiphoton calcium imaging to monitor cortical feedback in the OB of awake mice and further probe its impact on bulb output. We found that the responses of feedback boutons are sparse, odor-specific, and often outlasted stimuli by several seconds. Odor presentation both enhanced and suppressed the activity of boutons. However, any given bouton responded with stereotypic polarity across multiple odors, preferring either enhancement or suppression. Feedback representations were locally diverse and differed in dynamics across bulb layers. Pharmacological inactivation of the piriform cortex increased odor responsiveness and pairwise similarity of mitral cells, but had little impact on tufted cells (Fig. 4). We propose that cortical feedback differentially impacts the two output channels of the bulb by specifically decorrelating mitral cell responses to enable odor separation.

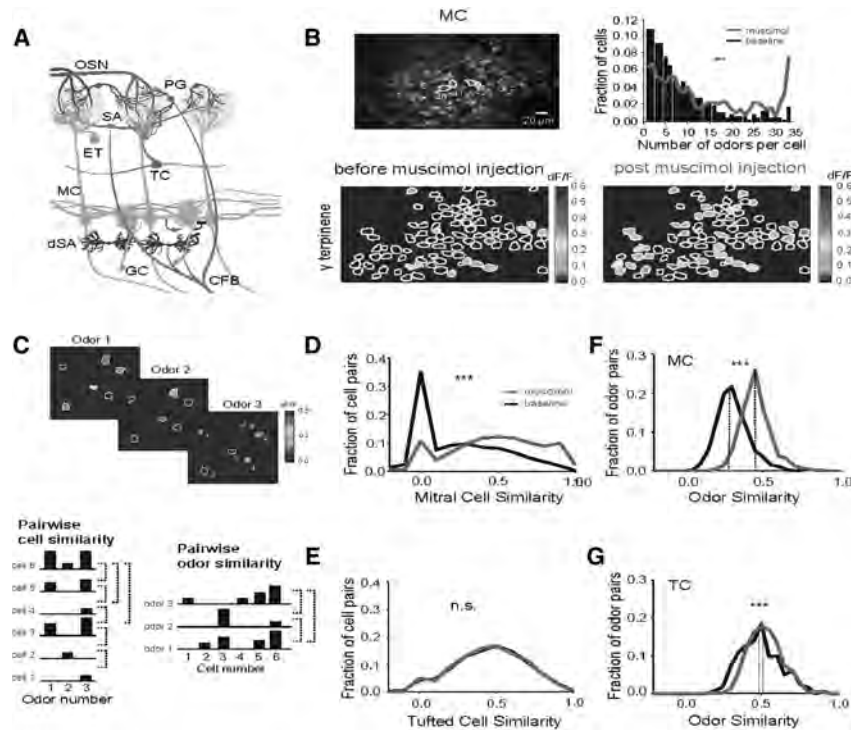
We further aim to understand the role of corticobulbar feedback (Fig. 4) in the identification of an odor, invariant of fluctuations in its concentration, and/or timing and presence of background odors. We are taking advantage of optogenetic tools and patterned illumination techniques recently developed by our group to reversibly manipulate the activity of cortical feedback fibers and their targets in awake head-fixed mice, as they learn and perform invariant odor perception tasks. To mimic naturalistic odor conditions, we are making use of custom-designed odor delivery systems that reproducibly deliver arbitrary time-varying patterns of individual odors and mixtures. To understand the underlying neuronal mechanisms, we monitor the activity of feedback inputs and their targets via high-speed scanning multiphoton microscopy and multielectrode recordings.

## Monitoring Cholinergic Input in the OB in Mice Engaged in Attention Tasks

---

Cholinergic inputs from the basal forebrain have been shown to enhance encoding and discrimination of visual, auditory, and somatosensory cues by modulating attention and learning processes. Although rich cholinergic projections innervate the mouse olfactory bulb (MOB), to date, little is known about their contribution to olfactory behaviors. Pharmacological blockade of ACh action has been shown to impair olfactory behavior, whereas augmenting acetylcholine (ACh) levels improved odor discrimination. A clear understanding of the underlying mechanisms of ACh action and their spatiotemporal statistics, however, has been limited by the low yield of recordings from the basal forebrain, as well as difficulties in ascertaining the downstream targets of the recorded neurons, given the wide-spread and long-range nature of cholinergic projections.

To circumvent these issues, we are taking advantage of genetically engineered mice (Chat-Cre) and viral strategies to target expression of calcium indicators (GCaMP6) in cholinergic neurons and to optically monitor the activity patterns of the projection fibers to the MOB in behaving animals (Fig. 5). We are pursuing two complementary strategies to gain access to the activity of ensembles of cholinergic neurons: (1) multiphoton imaging in awake head-fixed mice (Fig. 5C,D), which allows observation of large number of cholinergic projections in the MOB, with axonal resolution, enabling precise spatiotemporal characterization of their activity patterns within, as well as across, different circuit layers (glomerular vs. mitral cell layer), and (2) fiber optic imaging in freely moving mice (Fig. 5E–G). Although limited to assessment of average population activity of cholinergic inputs, this approach can be easily coupled with electrophysiological means to simultaneously monitor the OB output (mitral/tufted cells). Combining insights obtained via these two approaches will reveal how ACh action in the OB is linked to the timing and nature of olfactory stimuli across different behavioral states. We are investigating how cholinergic surges in the OB during odor detection, discrimination, and multimodal reversal learning tasks (switching reward contingencies between olfactory and auditory stimuli) modulate the output of the OB and relate to behavioral performance (Fig. 5).

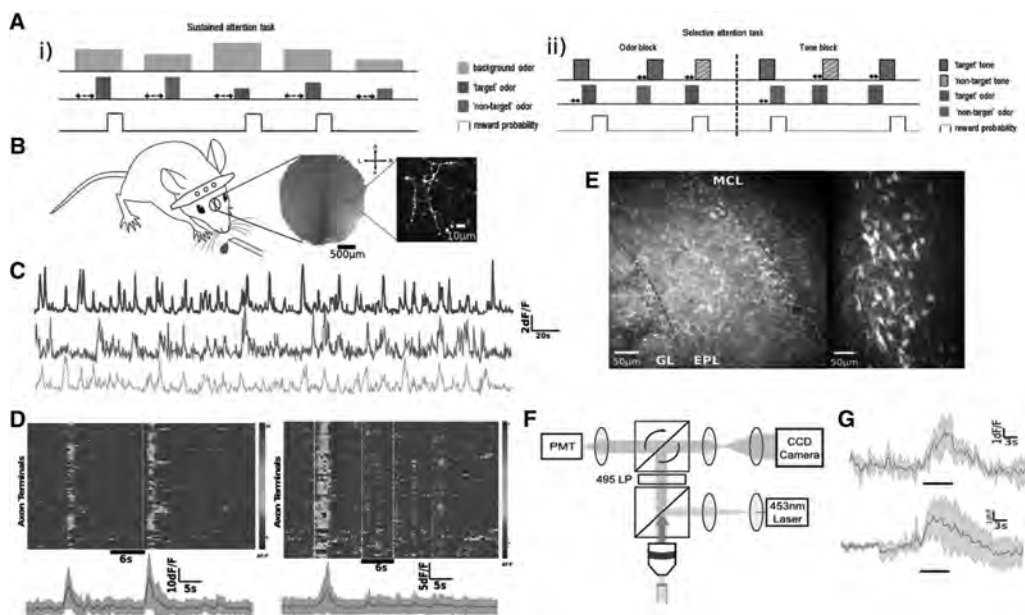


**Figure 4.** Cortical-bulbar feedback decorrelates the OB output. (A) Olfactory bulb (OB) circuit and neuronal types; OSN, olfactory sensory neurons; PG, periglomerular cells; SA, superficial short axon cells; ET, external tufted cells; TC, tufted cells; MC, mitral cells; dSA, deep short axon cells; GC, granule cells; CFB, cortical feedback fibers synapsing onto inhibitory interneurons across the OB. (B, *Top left*) Average resting fluorescence of an example field of view in the MC layer ( $\sim 220 \mu\text{m}$  from surface). (*Bottom*) Gray map showing average fluorescence change in response to  $\gamma$  terpinene in the field of view before (*left*) and after (*right*) muscimol injection (injected in anterior piriform cortex to suppress cortical activity and thus disable cortical feedback). (*Top right*) Histogram of the number of odors individual mitral cells responded to before (black bars) and after (gray trace) muscimol injection; Asterisks indicate significance level ( $p < 0.001$ , Wilcoxon-signed rank test). (C) Schematic exemplifying pairwise cell similarity and pairwise odor similarity calculations for a given field of view. (*Left*) Cartoon showing responses of six identified regions of interest (ROIs) (gray outlines) within a given field of view across three odors; degrees of grayness indicates the average response amplitude ( $dF/F$ ) for each ROI. (*Center*) An odor response spectrum (ORS) is calculated for each ROI (cell) as the vector containing the average  $dF/F$  for each odor; pairwise cell similarity is calculated as the uncentered correlation between the ORS vectors for each pair of cells (indicated by dotted lines). (*Right*) A cell response spectrum (CRS) is calculated for each odor as the vector containing the average  $dF/F$  for each cell upon presentation of the given odor; pairwise odor similarity is calculated as the uncentered correlation between the CRS vectors for each pair of odors. (D–G) Histogram of pairwise cell (*left*) and odor (*right*) similarity of mitral cells (MC) and tufted cells (TC) before (black, baseline) and after (gray) muscimol injection; asterisks indicate significance level ( $p < 0.001$ , Wilcoxon signed rank test); dotted lines indicate the median.

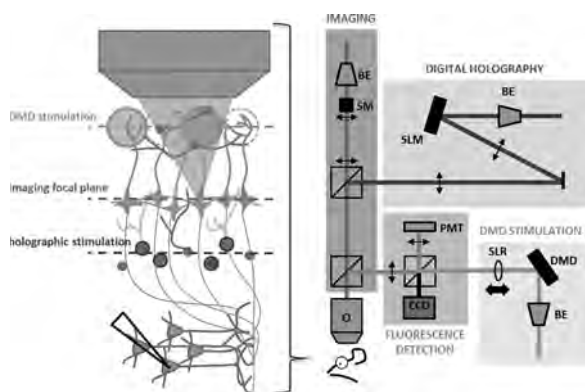
## Implementing DMD-Based and Digital Holography Methods to Investigate in a Closed-Loop Fashion the Spatiotemporal Integration Rules in the OB and Cortex

We are implementing strategies that will enable non-invasive, functional dissection of neuronal networks, with cellular resolution, in behaving animals. This will be brought about via a closed-loop strategy involving real-time control of activity of select neurons

with simultaneous monitoring of the concomitant effects of these manipulations on neuronal outputs within the circuit and elsewhere in the brain. Briefly, we are using digital holography methods via spatial light modulators to optogenetically control neurons of interest at single-cell level and DMD-based methods to control cell-type-specific populations across large brain regions (Fig. 6). This allows us to both replicate and systematically manipulate stimulus-evoked activity patterns in a circuit. We are



**Figure 5.** (A) Sustained attention (i) and selective attention (ii) tasks. (B) 1-mo-old cranial window over MOB; cholinergic axons expressing GCaMP5 imaged at 100- $\mu$ m depth (150  $\times$  150  $\mu$ m field of view). (C) Recorded traces of spontaneous activity of cholinergic axons from three different ROIs (4 min, mean  $dF/F_0 = 200\%$ ). (D) Two examples of cue-evoked (white lines) transients of cholinergic axon activity in awake animals. Rows represent different axon boutons. (Bottom) Mean (black line) and standard deviation (gray line) of all responsive axon boutons (mean  $dF/F$  for transients = 200%). (E) Spread of cholinergic axons expressing GCaMP5 in different layers of the MOB; injection site showing cholinergic cell bodies. (F) Diagram of the fiber optic imaging setup. (G) Two proofs of principle odor-evoked responses acquired with optical fibers from cortical feedback axons expressing GCaMP3 (left, 6-sec odor; (black bar) mean  $dF/F_0 = 3\%$ , four repeats each) recorded in the main OB.

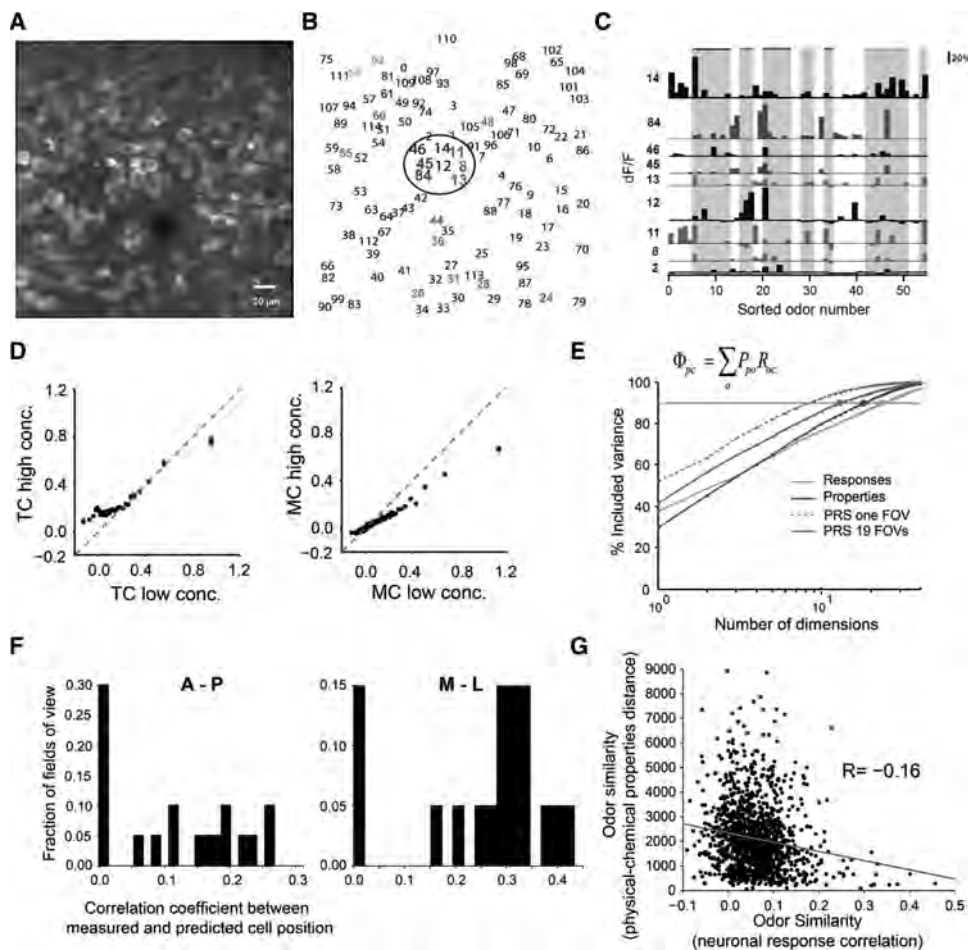


**Figure 6.** Combined imaging and photostimulation setup. (Left) Experimental configuration, combining scanning two-photon imaging, DMD photostimulation, and holographic photostimulation. (Right) Schematic of the experimental microscope. DMD stimulation is used to create spatial-temporal light patterns on the surface (<100  $\mu$ m). Digital holography is used to photostimulate deeper (<500  $\mu$ m) into the brain with cellular resolution. Calcium activity is monitored in an independent optical plane by two-photon imaging and by electrodes in downstream brain regions. BE, Beam expander; SM, scan mirrors; O, objective; PMT, photomultiplier; SLR, camera lens; DMD, digital micromirror device; SLM, spatial light modulator; CCD, charge coupled device.

simultaneously using two-photon calcium imaging and electrophysiology within the same and different brain regions (OB vs. olfactory cortex) to dissect how the alteration of select circuit elements, or their specific properties, affects the output of the network. This closed-loop approach will make it possible to determine the spatiotemporal integration rules within the OB and the olfactory cortex, investigate the relevance of spike-time codes, and reveal underlying decoding schemes.

## Mapping Odor Space onto Neuronal Representations in the Olfactory System

Unlike many other sensory systems, low-dimensional metrics for characterizing stimuli have remained elusive for olfaction, and it is unclear what features of chemical stimuli are represented by neurons. We would like to relate neural activity in the early olfactory system of mice to the physicochemical properties of odorants. We imaged odor-evoked responses



**Figure 7.** Mapping odor space on mitral cell activity patterns (A) Average resting fluorescence multiphoton image in the mitral cell layer. (B) Numbers indicate the relative positions of 112 mitral cell bodies in the imaged field of view. (C) Example odor response spectra of mitral cell bodies within an arbitrarily picked 75- $\mu$ m diameter region in A, sorted according to functional chemical groups (tiglates, thiazoles, ethyl esters, aldehydes, ketones, alcohols, and acids). (D) Scatter plots showing odor-induced change in tufted (*left*) and mitral (*right*) cells with increasing odor concentration. (E) Dimensionality of odorant properties; M/T neuronal responses and M/T receptive fields (properties response strength, PRS,  $\Phi$ ). (F) A LASSO/jackknife predictor yields a set of AP and ML cell positions that are weakly correlated with observed values. (G) Odor pairwise similarity across a set of 1660 physicochemical properties versus neuronal responses.

in identified tufted and mitral cells in awake mice using multiphoton microscopy. Similarity of odors computed through a comprehensive set of 1664 physicochemical features was a poor predictor of similarity of the corresponding neuronal representation by mitral or tufted cells. Dimension reduction revealed that fewer dimensions (~12) in the combined neurochemical space than in neural space alone (~22) were required to explain 90% of variance in neural responses across the population, suggesting that factors other than the physicochemical

properties we considered are required to fully explain the neural responses (Fig. 7). Responsive mitral and tufted cells were spatially dispersed, and cells within a local region were functionally heterogeneous, with only limited and variable dependence of M/T cell position on odorant characteristics. Our data indicate that novel descriptors are needed to link chemical space to neuronal representations and that odor information leaves the OB in a mosaic pattern, with substantial local diversity (in collaboration with the Koulakov lab here at CSHL).

## Other Collaborative Projects with CSHL Groups

---

We are collaborating with other CSHL scientists on the following projects: DLP-based patterned stimulation to functionally map the motor cortex (Z.J. Huang); sequencing the OB—bridging the gap between glomerular odors responses and odor receptor sequences by identifying the molecular identity of glomeruli (A. Koulakov, J. Lee, A. Zador); fiber-optic-based approach to monitor neuronal activity in punishment and reward neuronal circuits during behavior (B. Lee) (we are focusing on monitoring and manipulating activity in the lateral division of the central amygdala and the insula cortex); developing light-sheet-based approaches for fast optical reconstruction of neuronal

circuits (P. Osten); optical monitoring and manipulation of neuronal activity in genetically and anatomically defined cortical circuits in animal models of cognition (A. Zador).

### *In Press*

- Anselmi F, Banerjee A, Gupta P, Albeanu DF. 2015. *Patterned stimulation in the brain*. Springer International Publishing. (In press.)
- Banerjee A, Marbach F, Anselmi F, Koh M, Garcia de Silva P, Davis MB, Delevich K, Oyibo H, Gupta P, Li B, Albeanu DF. 2015. An interglomerular circuit gates glomerular output and implements gain control in the mouse olfactory bulb. *Neuron* (in press).
- Gupta P, Albeanu DF, Bhalla US. 2015. Olfactory bulb coding of odors, mixtures and sniffs is a linear sum of odor time-profiles. *Nat Neurosci* **18**: 272–281.
- Otazu GH, Chae HG, Davis MB, Albeanu DF. 2015. Cortical feedback decorrelates the olfactory bulb output in awake mice, *Neuron* **86**: 1461–1477.

# INVESTIGATING NEURAL CIRCUITS FOR SENSORY INTEGRATION AND DECISION MAKING

A. Churchland   A. Brown   A. Licata   D. Nunes Raposo  
L. Chartarifsky   F. Najafi   O. Odoemene  
M. Kaufman   H. Nguyen   M. Ryan

Making use of sensory information requires more than simply relaying incoming signals from the sensory organs. It requires interpreting information classifying it, drawing inferences, and ultimately using the context of behavioral goals to make a decision about its meaning. A decision is a commitment to a particular choice or proposition at the expense of competing alternatives. In some situations, decisions involve integration of evidence—that is, they make use of multiple pieces of information from the environment or from memory. These decisions can provide a framework in which to investigate complex cognitive processes and open a window into higher brain function in general.

Although previous experiments have begun to reveal how neural systems combine evidence to make decisions, they have left a critical gap in our understanding. Specifically, very little is currently known about the neural mechanisms that make it possible to combine information from multiple sensory modalities for decisions. The gap is apparent, although it is clear from behavioral observations that neural systems can combine multisensory information: When parsing speech in a crowded room, for example, the listener makes use of both auditory information (the speaker's vocal sounds) and visual information (the speaker's lip movements). Understanding the neural mechanisms of multisensory integration is critical for two reasons. First, it is essential for a complete understanding of sensory perception because real-world stimuli rarely affect a single sense in isolation. Therefore, understanding how the brain interprets incoming information requires understanding how the brain merges information from different senses. Second, it is likely of clinical importance that several developmental abnormalities appear to be related to difficulties in integrating sensory information. For example, abnormalities in multisensory

processing are a hallmark of subjects with autism spectrum disorder. Impairments in multisensory processing are also observed in subjects with a collection of sensory abnormalities known together as sensory processing disorder and may also be evident in patients with Rett syndrome and dyslexia. Understanding the neural mechanisms of multisensory integration could inform treatment of those conditions. Our long-term goal is to understand how the brain can make decisions that integrate inputs from our multiple senses, stored memories, and innate impulses. Our projects in previous years constituted the very first steps toward achieving this goal: We developed a multisensory behavior that could be carefully measured in both humans and rodents (Raposo et al., *J Neurosci* 32: 3726 [2012]; Shepard et al., *J Vision* 13: 1 [2013]). In addition, we built on the foundation by measuring the responses of neurons in rodents engaged in the behavior, affording insight into the neural circuits underlying multisensory decisions.

## The Role of Parietal Cortex in Multisensory Decision Making

M. Ryan, A. Licata [in collaboration with D. Raposo, Champalimaud Neuroscience Program, Lisbon, Portugal; M. Kaufman, Stanford University]

The goal of this project is to gain a deeper understanding of the neural circuits that enable integration of visual and auditory inputs for decision making. David Raposo has been a leader in the lab in two techniques this year: He measures electrophysiological responses from neurons in the posterior parietal cortex (PPC) of rats engaged in decision-making behavior. He presented this work at two conferences (Raposo et al. 2013). In addition, David inactivated

those same neurons and measured the effect on behavior (Raposo and Churchland, *Soc Neurosci Abs* 2013). Matt Kaufman has worked alongside David to develop sophisticated analyses of the data. John Shepard has also contributed to the project by recording additional neurons that are included in our growing population of cells. John, David, and Matt are working as a group; their main finding is that PPC neurons reflect random combinations of stimulus features, but they can be decoded at the population level to provide the animal with ongoing estimates of incoming sensory stimuli. The three are working collectively to write up the exciting results from these experiments for publication.

### Population Dynamics across Cortex

This work was done in collaboration with M. Kaufman (Stanford University)

---

The goal of this project is to understand how population activity changes from one neural structure to the next to support behavior. To tackle this question, Matt has brought a new technique to the lab: two-photon imaging. Using this technique, we can measure the responses of 80–100 neurons simultaneously. This approach will revolutionize the kinds of questions we can address about decision making, especially when used in conjunctions with emerging mathematical techniques for analysis. We have benefited greatly in this project from technical support from the Albeanu lab, our neighbors in the Marks building, and valued collaborators. This year, Matt was supported by a grant from the Swartz Foundation. He presented data from our *Nature Neuroscience* paper (on which he was the co-first author) at the Annual Meeting of the Swartz Foundation in Seattle, Washington.

### Multisensory Integration Mechanisms

This work was done in collaboration with L. Chartarifsky (Watson School of Biological Sciences)

---

The goal of this project is to understand how the brain estimates the reliability of sensory stimuli and uses that estimate to guide decision making. To reveal mechanisms that support this behavior, Lital Chartarifsky will use multisensory stimuli (called “conflict stimuli”) in which information from the

two modalities differ. In Lital’s experiments, these will be auditory and visual stimuli that differ in their number. She will leverage a large body of human literature that has explored these behaviors and will bring them to a powerful rodent model to elucidate neural mechanisms.

### Decision Making in Mice

H. Nguyen [in collaboration with F. Najafi, University of Pennsylvania; O. Odoemene, Watson School of Biological Sciences]

---

We have developed a decision-making paradigm using mice that will allow us to take advantage of genetic tools. These tools will also allow us to probe the neural circuitry that underlies decision making. In the past year, Onyekachi Odoemene has sought to establish which neural structures in mice are critical for decision making. He is using cutting-edge techniques to inactivate different structures and examine the effect on behavior. By using different strains of transgenic mice, Onyekachi is able to target particular cell types within each area and examine how they contribute to behavior. This project had benefited greatly from our collaborator Dr. Z. Josh Huang, who has provided us with materials for experiments and also with invaluable advice on understanding inhibitory circuitry. Support has been provided by a new research technician, Hien Nguyen, who has developed new training paradigms for mice and designed new devices for measuring neural responses.

A new postdoctoral fellow in the lab, Farzaneh Najafi, is the newest member of the team studying decision making in mice. Her expertise is in imaging neurons in animals during behavior and connecting the two using advanced analyses. She will bring these skills to the decision-making paradigms we use here, with an emphasis on understanding animal’s strategies and how they influence behavior and neural responses.

### Bayesian Number Estimation

A. Brown

---

This project aims to understand whether humans use a probabilistic approach to number estimation, and it is part of an international collaboration with

Dr. Alexandre Pouget (University of Geneva). This work challenges a long-standing assumption about human number estimation—that numbers are estimated as scalar quantities. Our experiments suggest that the assumption is not correct. Instead, humans appear to represent numbers probabilistically, and a signature of this representation is evident when they combine number estimates from two modalities.

We completed the project this year and have submitted a manuscript for publication.

#### PUBLICATIONS

Raposo D, Kaufman MT, Churchland AK. 2014. A category-free neural population supports evolving demands during decision-making. *Nat Neurosci* 17: 1784–1792.



# MECHANISMS OF NEURODEGENERATION AND MEMORY FORMATION

J. Dubnau   R. Borges Monroy   W. Li   M.F. Shih  
L. Krug   L. Prazak

There are two main areas of research in my lab. First, we are using neurogenetic and genomics approaches to investigate a novel hypothesis to explain neurodegenerative disorders such as frontotemporal lobar degeneration and amyotrophic lateral sclerosis. Second, we are investigating genetic and neuronal mechanisms of memory formation using *Drosophila* as a model system.

## Impact of Retrotransposons on Aging and Neurodegeneration Underlying Frontotemporal Lobar Degeneration and Amyotrophic Lateral Sclerosis

R. Borges, L. Krug, L. Prazak, J. Dubnau [In collaboration with M. Hammell and Y. Jin, Cold Spring Harbor Laboratory]

Transposable elements (TEs) are mobile genetic elements that provide a massive reservoir of potential genetic instability and toxicity. We have advanced the novel hypothesis that deregulated TE/ERV (endogenous retrovirus) expression may contribute to TDP-43-mediated neurodegenerative disorders (Li et al. 2013). We have found that TE transcripts are derepressed during the normal aging process and that this leads to active mobilization of transposons, resulting in new insertions in the genome of neurons. We used mutations in *Drosophila* Argonaute 2 (dAgo2) to create a situation in which TEs are derepressed in the brain. Such mutants exhibit precocious TE elevation in young animals and exacerbated age-dependent derepression. This is accompanied by rapid age-dependent memory impairment and shortened life span. These findings support the hypothesis that protracted loss of TE silencing contributes to age-dependent decline in neuronal physiological function.

To investigate the hypothesis that TE derepression contributes to neurodegenerative disorders, we focused (in collaboration with Molly Hammell and Ying Jin) on TDP-43, an RNA-binding protein involved

in a broad spectrum of neurodegenerative disorders including amyotrophic lateral sclerosis (ALS) and frontotemporal lobar degeneration (FTLD). As a first foray into the role of TDP-43 in transposon control, we mined a series of deep sequencing data sets of protein–RNA interactions and of gene expression profiles and uncovered broad and extensive binding of TE transcripts to TDP-43 (Li et al., *PLoS One* 7: e44099 [2012]). We also found that association between TDP-43 and many of its TE targets is reduced in FTLD patients. Finally, we discovered that a large fraction of the TEs to which TDP-43 binds become derepressed in mouse TDP-43 disease models. On the basis of these findings, we propose the hypothesis that TE misregulation contributes to TDP-43-related neurodegenerative disease. This hypothesis is now being tested using a *Drosophila* model and examining post-mortem human tissue from ALS patients.

We have used an established *Drosophila* model of TDP-43 neurodegeneration to test the hypothesis that transposons contribute to the underlying pathology. On the basis of the bioinformatics analyses of sequencing studies in mouse, rat, and human (see above), we predict that aggregation pathology of the TDP-43 protein will disrupt the normal mechanisms of transposon silencing, leading to a storm of awakened transposons. Indeed, we find that overexpression of human TDP-43 in *Drosophila* leads to activation of gypsy transposon expression (L. Krug et al., in prep.). Interestingly, the TDP-43 pathology in glial cells shows more dramatic gypsy expression than manipulation solely in neurons. This is accompanied by apoptotic cell death, progressive defects in locomotion, and shortened life span. Moreover, using reporter systems, we have established that TDP-43 expression leads to a disruption of the normal cellular systems to silence transposons. Together, these findings support the idea that a storm of unregulated transposons contributes to TDP-43-mediated neurodegeneration. This idea is now being tested using post-mortem tissue from ALS and FTLD patients.

## Age-Dependent Cognitive Decline and Retrotransposon Activation

L. Prazak

On the basis of our observation that certain retrotransposons become highly active during normal brain aging, we are investigating the hypothesis that retrotransposon-generated DNA damage contributes to age-related neuronal decline. Using the toolbox for genetic manipulations that are available in *Drosophila*, we are (1) investigating the normal mechanisms of retrotransposon silencing in the brain, (2) testing the impacts of age and stress on these cellular mechanisms, and (3) testing whether retrotransposon activation has a causal role in age-related cognitive impairment.

## Neurogenetic Mechanisms of Olfactory Memory

M.F. Shih

An understanding of memory, indeed of all behavioral phenotypes, will require a multidisciplinary approach to forge conceptual links between the relevant genetic/cell signaling pathways and neural circuits. Work in genetic model systems such as *Drosophila* can

contribute to our understanding in several ways. First, by enabling discovery of genes and genetic pathways underlying normal memory as well as pathological or degenerative cognitive disorders, model systems provide entry points for dissection of cellular mechanisms that are often conserved. Second, systematic manipulation of gene function within relevant anatomical circuits allows a conceptual integration of findings from cellular, neuroanatomical, and behavioral levels.

Our current work on this topic is focused on forging links between gene function and neural circuitry. We are using cell-type-specific genomic approaches to identify memory-induced transcriptional changes within each of the neuronal cell types that form part of the circuit for olfactory memory processing in *Drosophila*.

## PUBLICATIONS

- Dubnau J, ed. 2014. *Behavior genetics of the fly*. Cambridge University Press, Cambridge.
- Dubnau J, Li W, Prazak L. 2014. Transposon activation during aging and neuronal decline. US Patent 20,140,377,758.
- Shih MFM, Dubnau J. 2014. Converting *Drosophobia* into *Drosophila*. *Nat Neurosci* **17**: 1430.
- Singh PK, Bourque G, Craig NL, Dubnau J, Feschotte C, Gunderson KL, Malik HS, Moran JV, Peters JE, Slotkin RK. 2014. Mobile genetic elements and genome evolution. *Mobile DNA* **5**: 26.



L. King



M.-F. Shih

# STEM CELLS, SIGNAL TRANSDUCTION, AND BRAIN FUNCTION

G. Enikolopov T. Michurina J.H. Park N. Peunova

Our main focus is on stem cells in the adult organism and on signals controlling their maintenance, division, and fate. Most of our effort is directed at stem cells of the adult brain and on the signaling landscape of the neural stem cell niche. We are also interested in stem cells in nonneural tissues, with a particular focus on those tissues and organs that are involved in complex physiological circuits and major physiological and behavioral responses of the organism. We also study how diffusible signaling molecules affect stem cell maintenance, differentiation, and interactions with their local environment. Furthermore, we work to generate new genetic tools that would allow tracking multiple signaling events in a cell or a tissue. Besides the main goal of elucidating the molecular logic of stem cell control, we consider these studies as an entry point for designing therapeutic interventions to ameliorate the effects of aging or disease.

## Neural Stem Cells

---

The key focus of our group is on adult neurogenesis. Production of new neurons is important for behavior, pathophysiology, aging, and neural tissue repair in adult humans and animals. New neurons are born from neural stem cells that are maintained at specific locations in the adult brain. Continuous drive toward a differentiated state starts as soon as an activated stem cell produces progeny destined to become a neuron. Neural stem cells are the only source of new neurons in the adult brain (in contrast to the situation in a number of tissues, e.g., liver and pancreas, where fully differentiated cells can be activated to produce similar differentiated cells). Therefore, our understanding of the features and the role of new neurons depends on the ability to identify adult stem cells, trace their lineage, and reveal basic mechanisms governing their maintenance, division, differentiation, and death.

Our main approach is generating animal models that would enable visualization of stem cells and their environment and monitoring of their signaling landscape. The use of such model reporter mouse lines

allowed us to develop a new model for the quiescence, maintenance, and division of the adult hippocampal stem cells (*Cell Stem Cell*, 2011). Our results indicate that adult neural stem cells may remain quiescent for their entire postnatal life; however, when activated, they rapidly divide several times in quick succession to bud off daughter cells that eventually yield neurons, while the remaining stem cell differentiates into a mature astrocyte, thus leaving the stem cell pool. Hence, most hippocampal stem cells can be described as a “single-use” or a “disposable” unit, used in adulthood only once and then abandoned in its stem cell capacity. We also found that astrocytic differentiation of hippocampal stem cells is tightly coupled to their division, that all dividing stem cells of the hippocampus convert into astrocytes, and that new astrocytes of the dentate gyrus (DG) derive from these stem cells. We found that the decrease in the number of new neurons that accompanies aging is driven by the disappearance of stem cells via their division-coupled astrocytic differentiation. This continuous loss of stem cells underlies age-dependent diminished production of new neurons and may contribute to age-related cognitive impairment.

During the last several years, we have applied our model and our reporter lines to determine the classes of stem and progenitor cells that are affected by various pro- and antineurogenic factors. We found that in the adult brain, stem cell output can be increased in different ways and that various pro- and antineurogenic stimuli affect defined targets in the neurogenesis cascade. Importantly, we found that each mode of augmented production of new neurons may have different effects on the pool of stem cells, with important implications. For instance, some antidepressant treatments (e.g., fluoxetine or deep brain stimulation) do not affect stem cells, but instead target rapidly amplifying progenitor population. Another antidepressant therapy (electroconvulsive shock) increases the number of asymmetric divisions of stem cells without recruiting additional stem cells (and therefore leads to an increased number of new neurons without additional loss of stem cells). Some clinically used compounds increase recruitment of normally quiescent

stem cells in division, potentially leading to an increase in new neurons at the expense of premature exhaustion of the stem cell pool. These findings highlight the potential clinical relevance of our studies and may have direct implications for human therapy. Our results warn that any drug or therapeutic treatment that involves changes in neurogenesis should be investigated for the precise mechanisms of those changes, because seemingly identical outcomes may be induced by very different mechanisms and with very different long-term consequences.

Complementing our investigation of the basic mechanisms of stem cell regulation, much of our current effort focuses on developing new methods to study adult neurogenesis. These approaches include the following:

1. *Triple S phase labeling of dividing stem cells.* We have developed a new method for analyzing the division of stem and progenitor cells. This method is based on sequential labeling with three different nucleotide analogs on the background of green fluorescent protein (GFP)-labeled cells (i.e., quadruple labeling). Our method brings a qualitative improvement to the analysis of kinetics of cell division and to the tracing of multiple cell populations. We are currently applying this new approach to determine the changes in neural stem cell division and differentiation upon aging, physical activity, neuronal damage, and response to drugs.
2. *Serial two-photon tomography: Three-dimensional representation of neural stem cells.* We are applying a new method of serial two-photon tomography to our reporter mouse lines to produce quantitative three-dimensional representations of changes in stem and progenitor cell populations in the adult brain (in collaboration with Dr. Pavel Osten, CSHL). We can visualize and quantify the entire complement of stem and progenitor cells, as well as cells with a stem cell potential, in the major neurogenic zones and in less-studied regions of the adult brain (e.g., the circumventricular organs). We are now applying serial two-photon tomography to reveal changes due to age or treatment with neurogenesis-related drugs.
3. *Metabolic profiling of dividing cells by MRS (PLoS One, 2014).* In collaboration with Dr. Helene Benveniste (Stony Brook University), we have applied proton magnetic resonance spectroscopy (<sup>1</sup>H MRS) and a widely used spectral fitting algorithm (LCModel) to characterize the metabolic changes in the live rodent brain after electroconvulsive shock or in glioblastomas with actively dividing cells, paralleling these changes to postmortem histology. LCModel processing reveals
4. *Using MIMS to study stem cell division and turnover.* We have collaborated with Dr. Claude Lechene (Harvard Medical School), who has developed the multi-isotope imaging mass spectrometry (MIMS) approach for biological studies, to investigate division and maintenance of stem cells in the adult brain. Our results with long-term labeling of stem cells and their progeny with stable isotope <sup>15</sup>N-labeled thymidine are fully compatible with our proposed model of “disposable” adult stem cells (but not with the conventional model of recurrent stem cell quiescence). A similar approach can be used to detect the turnover of other macromolecules, and we are currently using combinations of stable isotopes to label both DNA and protein and follow metabolic changes in adult stem cells.
5. *Computational modeling and cell-counting algorithms.* In collaboration with Dr. Alex Koulakov here at CSHL, we continue to develop a general framework for computational modeling of stem and progenitor division and differentiation in the adult brain. This model incorporates several key kinetics parameters that can be independently manipulated to cover most of the complex interconversion schemes for dividing and differentiating cell populations. We use this model to determine age-related changes in specific subpopulations of quiescent and amplifying progenitor, astrocytes, neuroblasts, and neurons. Furthermore, we developed a new algorithm for counting labeled nuclei in sections and three-dimensional images of the developing and adult brain. This protocol is faster and more precise than most of the available commercial suites for cell counting and will accelerate analysis of adult neurogenesis.

## Stem Cells in Nonneural Tissues

---

We began engineering reporter mouse lines to visualize, track, and isolate neural stem cells more than 15 years ago and some of our reporter lines (e.g., Nestin-GFP mice, *J Comp Neurol*, 2004) became widely used genetic tools. These lines were originally designed to identify stem cells in the nervous system; however, we soon found the same reporter lines (Nestin-GFP, Nestin-CFPnuc, or Nestin-mCherry) help to discover stem and progenitor cells in a range of other tissues as diverse as anterior pituitary, skeletal muscle, testis, hair follicles, liver, pancreas, retina, or bone marrow.

We identified some of those nonneuronal stem cell species in our group, while other types were studied in collaboration with other research groups. Of the findings with nonneural stem cells, we will mention most recent results with subtypes of pericytes in the skeletal muscle, liver oval cells (*Cell*, 2014), and stem cells in the adult adrenal medulla (*Stem Cells*, 2015).

## NO, Development, and Differentiation

---

During the last two decades, we have investigated the diversity of biological functions mediated by nitric oxide (NO) and discovered the essential role that NO has in development, differentiation, and stem cell regulation. NO potentiates weak calcium signals (*Nature*, 1993), mediates neuronal differentiation (*Nature*, 1995), regulates *Drosophila* imaginal disk development (*Cell*, 1996), controls brain development (*J Neurosci*, 2001) and adult neurogenesis (*Proc Natl Acad Sci*, 2003), and coordinates cell proliferation and cell movements during early development (*Cell Cycle*, 2007). We have uncovered a novel function of NO and thus our current focus is on the role it plays in the development and function of cilia in multiciliated cells of the mucociliary epithelium. We found that in such diverse settings as mouse trachea and embryonic frog skin, NO synthase is crucial for the main function of the motile cilia: to beat in concert and generate directed flow of fluid across tissue. In the trachea and bronchi, this flow helps to clear pathogens, foreign particles, and toxic chemicals from the airways. We identified evolutionary conserved signaling pathways controlling cilia distribution, polarity, and function and found that genetic defects in cilia function can be rescued by NO-releasing compounds. Our recent results suggest that NO may similarly contribute to cilia function in other types of ciliary epithelium; furthermore, they suggest that certain inborn and acquired human ciliopathies and related disorders may also be associated with decreased availability of NO and may benefit from NO-based therapies.

## Genetically Encoded Sensors

---

We have recently launched a new direction, aimed at visualizing multiple signaling fluxes in real time using genetically encoded sensors and elucidating the signaling landscape of the neural stem cell niche. Our particular

focus is on reactive oxygen species (ROS), because redox signaling emerges as a regulatory network that is critical for supporting metabolism and homeostasis of stem cells and contributing to aging and disease. To that end, we are developing a new platform for genetically encoded redox sensors characterized by greatly improved quantum yield, contrast, sensitivity, and expanded palette. Our system reports changes in redox status in various cell compartments using four separate channels, as well as changes in redox signals in parallel to other types of signals (e.g., calcium, pH, or NAD(H)), thus approaching our goal of multiparametric monitoring of signal fluxes and cell activity. Currently accomplished projects (performed in collaboration with Dr. Vsevolod Belousov, a Visiting Scientist in the group) related to genetically encoded sensors include the following:

1. *Genetically encoded fluorescent indicator for imaging NAD<sup>+</sup>/NADH ratio* (*BBA*, 2014). The ratio of NAD<sup>+</sup> to NADH is a key indicator that reflects the overall redox state of the cells. We developed and validated a new genetically encoded probe for NAD(H) using redox-sensitive protein domains from *Thermus aquaticus*. Our RexYFP sensor has several advantages over existing sensors such as small size and optimal affinity for different compartments. We demonstrated that the affinity of the probe enables comparison of NAD<sup>+</sup>/NADH in compartments with low (cytoplasm) and high (mitochondria) NADH concentration. Thus, RexYFP is an efficient sensor, suitable for detecting the NAD(H) redox states in different cellular compartments.
2. *Sensor-generator pairs to evaluate redox production and signaling* (*ARS*, 2014). Yeast D-amino acid oxidase (DAO), activated by externally added D amino acids (e.g., D-alanine), can serve as a genetically encoded producer of ROS. We found that DAO, fused to or cotransfected with a genetically encoded H<sub>2</sub>O<sub>2</sub> indicator (HyPer), can be used for controlled production and detection of ROS in living eukaryotic cells. Our results show the utility of the HyPer-DAO genetically encoded system for redox signaling studies, indicate profound heterogeneity in ROS production dynamics between individual cells, and suggest that H<sub>2</sub>O<sub>2</sub> produced by DAO in the cytoplasm acts locally in close proximity to the enzyme.
3. *Red fluorescent genetically encoded indicator for intracellular hydrogen peroxide* (*Nat Comm*, 2014). Currently available recombinant redox sensors have green emission, which overlaps with the spectra of many other probes. Expanding the spectral range of recombinant ROS

probes would enable multiparametric in vivo ROS detection. We generated a genetically encoded red fluorescent sensor for H<sub>2</sub>O<sub>2</sub> detection, HyPerRed, and demonstrated the utility of the sensor by tracing low concentrations of H<sub>2</sub>O<sub>2</sub> produced in the cytoplasm of cultured cells upon growth factor stimulation. Moreover, using HyPerRed, we were able to detect local and transient H<sub>2</sub>O<sub>2</sub> production in the mitochondrial matrix upon inhibition of endoplasmic reticulum Ca<sup>2+</sup> uptake.

We are now applying our multicolor redox sensor platform for multiparametric imaging of signaling events in stem cells and transgenic animals.

## PUBLICATIONS

- Bilan DS, Matlashov ME, Gorokhovatsky AY, Schultz C, Enikolopov G, Belousov VV. 2014. Genetically encoded fluorescent indicator for imaging NAD<sup>+</sup>/NADH ratio changes in different cellular compartments. *Biochem Biophys Acta* **1840**: 951–957.
- Enikolopov G, Guillemier C, Wang M, Trakimas L, Steinhäuser M, Lechene C. 2014. Brain stem cell division and maintenance studied using multi-isotope imaging mass spectrometry (MIMS). *Surf Interface Anal*
- Ermakova YG, Bilan DS, Matlashov ME, Mishina NM, Markvicheva KA, Subach OM, Subach FV, Bogeski I, Hoth M, Enikolopov G, Belousov V. 2014. Red fluorescent genetically encoded indicator for intracellular hydrogen peroxide. *Nat Comm* **5**: 5222.
- Matlashov ME, Belousov VV, Enikolopov G. 2014. How much H<sub>2</sub>O<sub>2</sub> is produced by recombinant D-amino acid oxidase in mammalian cells? *Antiox Redox Sign* **20**: 1039–1044.
- Park J-H, Lee H, Makaryus R, Yu M, Smith DS, Sayed K, Feng T, Van der Linden A, Bolwig TG, Holland EC, Enikolopov G, Benveniste H. 2014. Metabolic profiling of dividing cells in live rodent brain by proton magnetic resonance spectroscopy and LC-Model analysis. *PLoS One* **9**: e94755.
- Tschaharganeh DF, Xue W, Calvisi DF, Evert M, Michurina TV, Dow LE, Banito A, Katz SF, Kastenhuber ER, Weissmueller S., et al. 2014. p53-dependent Nestin regulation links tumor suppression to cellular plasticity in liver cancer. *Cell* **158**: 579–592.
- In Press*
- Enikolopov G, Overstreet-Wadiche L, Ge S. 2015. Transgenic and viral strategies for identification of neural stem and their progeny. *Cold Spring Harb Persp Biol* **7**: in press. doi: 10.1101/cshperspect.a018814.
- Makaryus R, Lee H, Feng T, Park J-H, Nedergaard M, Jacob Z, Enikolopov G, Benveniste H. 2015. Brain maturation in neonatal rodents is impeded by sevoflurane anesthesia. *Anesthesiology* (in press).
- Mignone J, Peunova N, Enikolopov G. 2015. Nestin-based reporter transgenic mouse lines. *Methods Mol Biol* (in press).
- Park MH, Jin H-K, Lee JK, Lee H, Lee WW, Akiyama H, Herzog H, Lee SJ, Min W-K, Enikolopov G, Schuchman EH, Bae J-S. 2015. Neuropeptide Y regulates the hematopoietic stem cell microenvironment and prevents nerve injury in the bone marrow. *EMBO J* (in press).
- Rubin de Celis MF, Garcia-Martin R, Wittig D, Valencia GD, Enikolopov G, Funk RH, Chavakis T, Bornstein SR, Androutsellis-Theotokis A, Ehrhart-Bornstein M. 2015. Multipotent glia-like stem cells mediate stress adaptation *Stem Cells* (in press).

# MOLECULAR ANALYSIS OF NEURONAL RECEPTORS AND ION CHANNELS

H. Furukawa    E. Karakas    A. Romero Hernandez  
                  T. Malinauskas    N. Simorowski  
                  M. Regan            N. Tajima

The research program in our lab attempts to understand the molecular basis for the functions of receptors and ion channels at the cellular membrane that initiate cellular signal transductions involved in neurotransmission in the mammalian brain and in the process of neurological diseases and disorders including schizophrenia, depression, stroke, and Alzheimer's disease. Toward achieving our goals, we have been focusing on understanding various functional aspects of ion channels called *N*-methyl-D-aspartate receptors (NMDARs) that control intracellular calcium signaling upon stimulation by voltage and/or neurotransmitters. NMDARs are ligand-gated ion channels that regulate neurotransmission, the fundamental process whereby multiple neurons communicate with one another. Dysfunction of the NMDAR ion channels has a strong link to the neurotoxicity that results in various neurological disorders and diseases described above. We use X-ray crystallography and cryo-electron microscopy to determine three-dimensional atomic structures and test structure-based mechanistic hypotheses by site-directed mutagenesis in combination with biochemical and biophysical techniques. We also use structural information to develop compounds with therapeutic potential for treating the neurological disorders described above. Our main accomplishment in 2014 was to establish an effective heterologous expression method suitable for assembling heteromeric membrane protein complexes such as NMDARs and the completion of the first crystal structure of the intact heterotetrameric NMDAR that contains the entire extracellular domain and the transmembrane ion channel pore domain.

## Production and Biochemical Characterization of Heteromultimeric NMDAR Complexes

---

We have established a novel expression system that facilitates the assembly of heteromeric membrane pro-

teins such as NMDARs. In achieving this, we conducted various constructs of NMDARs that contain the GluN2B subunit, which is carboxy-terminally truncated, and the GluN1 subunit, which is carboxy-terminally fused to enhanced green fluorescent protein (EGFP), to allow (1) screening of host cells including Sf9 and High Five insect cells, HEK293, CHO, and Cos7 mammalian cells; (2) screening of expression systems including transient transfection, production of stable cell line, and baculovirus-mediated protein production; and (3) screening of promoters including CMV, IE1 and IE2, WSSV, and Hsp70 promoters. Expression, subunit assembly, and size homogeneity were assessed by fluorescence-coupled size-exclusion chromatography that detects the proteins of interest by fluorescence (in this case, the EGFP signal derived from the GluN1-EGFP fusion protein). Of all combinations tested, the following produced the fully assembled NMDARs: (1) HEK293 in combination with the CMV promoter and with either transient transfection or baculovirus-mediated DNA incorporation and (2) Sf9 or High Five cells in combination with the CMV or Hsp70 promoter and with baculovirus-mediated DNA incorporation. Of all host cell/promoter/expression system combinations, expression in Sf9 cells in combination with the baculovirus harboring the Hsp70 promoter gave the highest level of expression (~0.1 mg of purified proteins per liter Sf9 culture). This expression system has been instrumental to successful structural studies outlined in Aim 2. In addition, this expression system is now distributed to approximately 20 laboratories interested in expressing membrane proteins.

## Structural and Functional Studies on Intact Heteromultimeric NMDARs

---

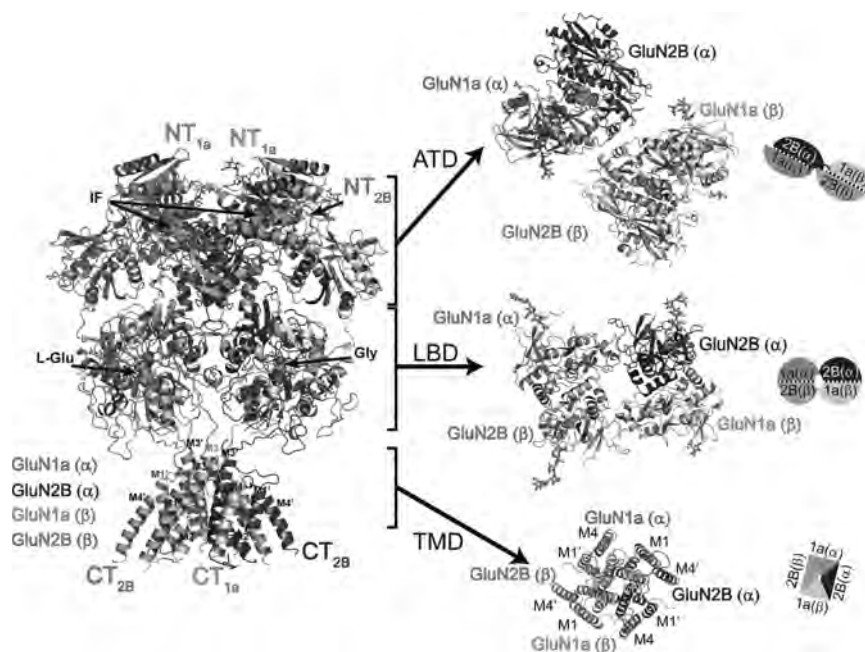
By taking full advantage of the protein production system we established above, we have completed the

work showing the first crystal structure of heterotetrameric NMDARs, which contains the entire extracellular domain including an amino-terminal domain (ATD), ligand-binding domain (LBD), and the transmembrane domain (TMD). In general, this study marked the first crystal structure of the heteromeric ion channel or of the eukaryotic heteromeric membrane proteins reconstituted in a recombinant expression system. Toward solving the crystal structure of NMDAR, we have optimized the construct to improve protein stability and crystallizability. This led to the construct GluN1a/2B<sub>cryst</sub>, which contains a number of point mutations that reduce the number of free cysteine residues and N-linked glycosylations, and truncations that eliminate the carboxy-terminal domain (CTD) and the linker between the ATD and the LBD in GluN2B.

The GluN1a/2B<sub>cryst</sub> construct was crystallized in the presence of the GluN1 agonist, glycine; the GluN2 agonist, glutamate; and an ATD-binding allosteric inhibitor, ifenprodil. The structure was initially solved at 5.7 Å by molecular replacement using the structures of the ATD and LBD as search probes. To improve the diffraction power of these crystals, we stabilized the heterotetrameric subunit interactions by

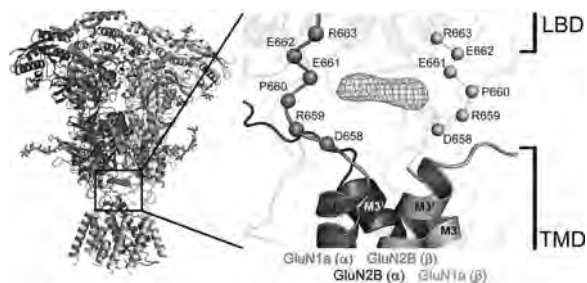
forming intersubunit disulfide cross-links at the extracellular and juxtamembrane regions based on the 5.7-Å structure above. These construct manipulations improved the diffraction limit to beyond 4 Å, which resulted in electron density sufficient to build most of the GluN1a/2B NMDAR including all the extracellular domains, TMD, linkers between ATD and LBD, and linkers between LBD and TMD.

The GluN1a/GluN2B NMDAR bound to glycine, L-glutamate, and ifenprodil is shaped like a “hot-air balloon” where a balloon and a basket correspond to the entire extracellular domains and the TMD, respectively. The structure has a clear boundary between the layers of LBD and TMD, whereas ATD and LBD appear as a single unit. Both ATD and LBD form two GluN1a-GluN2B heterodimers that are assembled as a GluN1-GluN2-GluN1-GluN2 (1-2-1-2) heterotetramer where the twofold symmetry axis runs across the center of the molecule through the interface between the two GluN2B ATDs to the middle of the ion channel pore (Fig. 1). The overall shape of GluN1a/GluN2B NMDAR is highly distinct from that of the homotetrameric GluA2 AMPAR ( $\alpha$ -amino-3-hydroxy-5-methyl-4-isoxazole propionic acid receptor), another member



**Figure 1.** First crystal structure of heterotetrameric NMDARs. We crystallized the GluN1a/GluN2B NMDARs without CTD. The heterotetramers are formed as a dimer of heterodimers in the GluN1-GluN2B-GluN1-GluN2B arrangement in each domain layer. Interestingly, dimer pairs are swapped between ATD and LBD.





**Figure 2.** Plausible calcium binding site at the juxtamembrane site. Shown here is an anomalous difference Fourier map obtained from diffraction data of crystals soaked against holmium. Holmium has been previously shown to bind calcium-binding site.

of the iGluR family, which has an overall “Y” shape. This difference in the overall architecture is attributed to the fact that ATD and LBD pack tightly through numbers of well-defined interdomain interfaces in GluN1a/GluN2B NMDARs, whereas ATD and LBD interact minimally in GluA2 AMPARs. This surprisingly “compact” architecture of GluN1a/GluN2B NMDAR stems from the intersubunit and interdomain interactions over large surface areas that are unique to NMDARs.

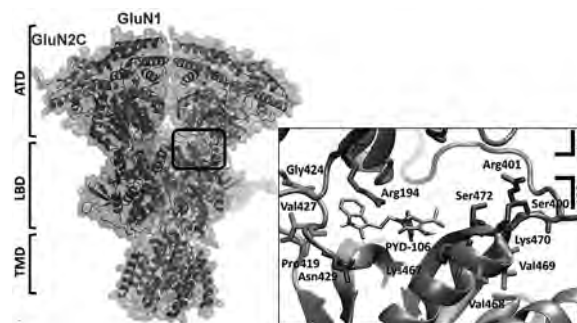
The TMD of GluN1a/GluN2B NMDAR forms the heterotetrameric ion channel with pseudo four-fold symmetry. The ion channel of GluN1a/GluN2B NMDAR is similar to the closed state of GluA2 AMPARs and KcsA bacterial potassium channels. In this plausible allosterically inhibited state, the ion channel is closed to an extent similar to the closed channels of AMPAR and KcsA. One of the key functions of NMDARs is their high permeation of calcium ions, which has a major role in neuronal plasticity as well as excitotoxicity. The crystal structure complexed with holmium, a lanthanide known to recognize calcium-binding sites in many biological macromolecules, shows the binding in between the LBD-TMD linkers of GluN1 right above the center of the ion channel (Fig. 2). A set of acidic residues in GluN1 (DRPEER motif) located in this region has been shown previously to be a critical part of high calcium flux characteristic of NMDARs. Thus, the lanthanide-binding site along with the previous electrophysiological study further confirms the physiological relevance of the current crystal structure. However, despite extensive efforts, the region of TMD that determines voltage-dependent  $Mg^{2+}$  block and  $Ca^{2+}$  permeation was not clearly resolved

in this crystal structure. A structure-based understanding of ion selectivity and flux regulation thus remains to be achieved.

The current work on GluN1a-GluN2B NMDARs marks the first crystal structure of a heteromeric ion channel. The structure will serve as a template for designing experiments that further address the complex function of NMDARs. Furthermore, the defined subunit and domain interfaces should serve as an invaluable blueprint for binding sites for potential compounds leading to design of therapeutic compounds, an example of which is described in the next section.

### Locating Novel Compound Binding Site in NMDAR

The crystal structure of the intact heterotetrameric NMDAR now provides a solid foundation for predicting compound binding sites—especially the those located at the subunit or domain interface. In collaboration with James Snyder and Stephen Traynelis at Emory University, we have used the crystal structure described above to predict the binding site of a novel compound called PYD-106 that binds GluN2C-containing NMDAR. A homology model for GluN2C was built using the above crystal structure with the program Modeller. On the basis of knowledge gained from electrophysiological data that binding of the compound may involve residues from both ATD and LBD of GluN2C, the molecular model of PYD-106 was docked to the ATD-LBD interface of the GluN1/GluN2C homology model.



**Figure 3.** Prediction of novel compound binding site. Homology model of GluN1/GluN2C NMDAR built based on the crystal structure of GluN1/GluN2B NMDAR. Docking study shows that PYD-106 may be binding at the domain interface between ATD and LBD.

The binding site residues were energy-minimized to obtain the most likely mode of compound binding (Fig. 3), which is consistent with electrophysiology data.

Overall, our work has provided an important foundation for understanding the structural and functional mechanism of NMDARs in unprecedented detail. We are continuing our effort to provide structural understanding of subtype specificity in NMDARs to further support development of therapeutic agents.

## PUBLICATIONS

- Jespersen A, Tajima N, Fernandez-Cuervo G, Garnier-Amblard EC, Furukawa H. 2014. Structural insights into competitive antagonism in NMDA receptors. *Neuron* **81**: 366–378.
- Karakas E, Furukawa H. 2014. Crystal structure of a heterotetrameric NMDA receptor ion channel. *Science* **344**: 992–997.
- Khatri A, Burger PB, Swanger SA, Hansen KB, Zimmerman S, Karakas E, Liotta DC, Furukawa H, Snyder JP, Traynelis SF. 2014. Structural determinants and mechanism of action of a GluN2C-selective NMDA receptor positive allosteric modulator. *Mol Pharmacol* **86**: 548–560.

# ASSEMBLY AND FUNCTION OF NEURAL CIRCUITS IN THE NEOCORTEX

Z.J. Huang   X. An   S. Kelly   J. Levine   M. Moissidis   J. Tucciarone  
D. Huilgol   K. Krishnan   J. Lu   A. Paul   P. Wu

We aim to understand the developmental assembly and functional organization of neural circuits in the cerebral cortex. The neocortex consists of a constellation of functional areas that form a representational map of the external (sensory, social) and internal (visceral, emotional) world. These areas are strategically interconnected into elaborate networks that support dynamic operations that process information and guide intelligent behavior. Our overarching hypothesis is that, at the cellular level, cortical processing streams and output channels are mediated by a large set of distinct glutamatergic pyramidal neurons (PyNs), and functional PyN ensembles are regulated by a diverse yet distinct set of GABAergic interneurons (INs). We systematically build cell-type genetic tools that integrate a full set of modern techniques for exploring neural circuits. Building on our success in genetic targeting of GABA INs, we have recently extended this effort to PyNs. We combine a range of approaches that include genetic and viral engineering, cell-type gene expression, genetic fate mapping, imaging, electrophysiology, and behavior analysis. This progress allows us to integrate our studies in the context of the motor cortex control of volitional forelimb movements and motor learning.

## Genetic Targeting of Pyramidal Neuron Subtypes in the Mouse Neocortex

---

A key obstacle to studying the development, organization, and function of neural circuits in the cerebral cortex is the stunning diversity of neuron types and a lack of comprehensive knowledge about their basic biology. The problem of neuronal diversity and identity in the cortex is fundamental, transcending developmental and systems neuroscience, and lies at the heart of defining the biology of cognition and psychiatric disorders. PyNs constitute ~80% of cortical neurons, are endowed with a large capacity for information coding, storage, and plasticity, and carry the output of cortical computation.

PyNs consist of diverse subtypes based on their specific laminar locations, axonal projection patterns, and gene expression profiles. Subsets of PyNs form multiple and hierarchical subnetworks of information processing, with distinct output channels to cortical and subcortical targets that subserve sensory, motor, cognitive, and emotional functions. Importantly, PyN subtypes are differentially affected in various neuropsychiatric and neurodegenerative disorders. However, the severe lack of specific and effective genetic tools for studying PyNs has hampered progress in understanding cortical circuits. We have begun to build a comprehensive genetic tool set for major PyN subtypes in the mouse in collaboration with Dr. Paola Arlotta at Harvard University. We have discovered a set of specific and combinatorial markers that distinguish major PyN subtypes. We are using intersection, subtraction, and inducible strategies to target PyN subtypes. We have already generated 10 knockin driver lines that target restricted populations of corticothalamic, corticofugal, corticostriatal, and corticocortical PyNs.

## The Progenitor Origin of Chandelier Cells

---

Diverse GABAergic INs regulate the functional organization of cortical circuits and derive from multiple embryonic sources. It remains unclear to what extent embryonic origin influences interneuron specification and cortical integration due to difficulties in tracking defined cell types. Using genetic fate mapping, we followed the developmental trajectory of chandelier cells (ChCs), the most distinct interneurons that innervate the axon initial segment of pyramidal neurons and control action potential initiation. We recently discovered that the specification of ChC identity is mediated through lineage and birth timing mechanisms in the embryonic subpallium, and young ChCs navigate over long distances with stereotyped routes and schedules to achieve characteristic laminar and areal distribution in the cortex by the end of first postnatal

week. We have been using combinatorial genetic fate mapping to explore whether there are designated ChC progenitor pools that give rise to these distinct cell types. Results in the past year suggest that there are fate-restricted progenitors that give rise to L5/6 or L2 ChCs, respectively.

### Activity-Regulated Pruning and Circuit Integration

---

We have discovered a massive and neural activity-regulated pruning of ChCs during their postnatal integration into the cortical network. Upon fate specification at cell birth, young ChCs migrate over long distances following stereotyped routes and schedules and settle mainly in cortical layers 2, 5, and 6 by the end of the first postnatal week. In the following week (from ~P7 to P14), ~60% of L2 ChCs undergo cell death throughout the cortex through the Bax-Bak apoptosis pathway. At the primary (V1) and secondary (V2) visual cortex border region, ChC apoptosis and density are regulated by contralateral callosal axons in an activity- and retina-dependent manner. Blocking activity of contralateral callosal PyNs leads to reduction of ChC pruning at V1-V2 border, and removing retinal input by neonatal monocular enucleation leads to ectopic callosal projections and reduction of ChC pruning in V1 regions that correlated with ectopic callosal axons. Our working hypothesis is that patterned retinal activities before vision onset coordinate the wiring of bilateral callosal PyNs, in part through eliminating inappropriately connected young ChCs. This might allow retinotopically matched callosal PyN axons to establish fast bihemispheric signaling and contribute to the fusion of left-right visual field. Such activity-regulated ChC pruning may represent a general mechanism that shapes interhemisphere and interarea processing pathways.

### Connectivity and Function of Chandelier Cells in a Medial Prefrontal Cortex Fear Circuit

---

Although the amygdala is central in mediating innate and learned fear, the prelimbic area (PL) of the rodent medial prefrontal cortex (mPFC) has been implicated in the regulation of fear responses in part through modulating amygdala activities. PL may integrate multiple inputs that convey contextual, internal state, and

mnemonic information in regulating the activity of basolateral amygdala (BLA), thereby achieving flexibility in fear response to changing cues, but the underlying circuitry mechanism is unclear. The supragranular layers of PL contain multiple populations of pyramidal neurons (PyNs) that project to several target areas such as striatum (STRp), contralateral PL (CCp), and BLA (BLAp). Among these, BLAp firing directly drives BLA fear neurons, but the inhibitory control of BLAp is not understood. Among diverse cortical GABA inhibitory neurons, the ChCs innervate the axon initial segment (AIS) of PyNs and likely exert a decisive control over their firing. Using cell-lineage-dependent genetic targeting, we examined the connectivity and function of ChCs in PL. Layer 2 (L2) ChCs preferentially control BLAp over CCp in PL, even when these PyNs are intermingled, suggesting a role of ChCs in regulating BLA activity. Paired recordings and retrograde rabies tracings indicate that L2 ChCs do not receive significant glutamatergic inputs from local PyNs, but they do receive multiple long-range inputs from the mediodorsal thalamus, ventral hippocampus, BLA, diagonal band (acetylcholine), and ventral tegmental area (dopamine), which may carry contextual, emotional, and brain state information. Importantly, chemical genetic suppression of PL L2 ChCs resulted in enhanced fear expression. Together, these results suggest that prelimbic ChCs may integrate multiple contextual, emotional, and motivational information to modify BLAp firing and achieve flexible regulation of fear expression.

### Input-Specific Maturation of Synaptic Recruitment of Parvalbumin Interneurons in the Primary Visual Cortex

---

Cortical networks consist of local recurrent circuits and long-range pathways from other brain areas. Parvalbumin inhibitory interneurons (PVNs) regulate the dynamic operation of local ensembles as well as the temporal precision of afferent signals. The synaptic recruitment of PVNs that support these circuit operations is not well understood. Here, we demonstrate that the synaptic dynamics of PVN recruitment in the mouse visual cortex are customized according to input source with distinct maturation profiles. Although the long-range inputs to PVNs show strong short-term depression throughout postnatal maturation, local inputs from nearby pyramidal neurons progressively

lose such depression. This enhanced local recruitment depends on PVN-mediated reciprocal inhibition and results from both pre- and postsynaptic mechanisms, including calcium-permeable AMPA ( $\alpha$ -amino-3-hydroxy-5-methyl-4-isoxazole propionic acid) receptors at PVN postsynaptic sites. Although short-term depression of long-range inputs is well-suited for afferent signal detection, the robust dynamics of local inputs may facilitate rapid and proportional PVN recruitment in regulating local circuit operations.

## PUBLICATIONS

- Fenno LE, Mattis J, Ramakrishnan C, Hyun M, Yeun Lee S, He M, Tucciarone J, Selimbeyoglu A, Berndt A, Grosenick L, et al. 2014. INTRSECT: Single-component targeting of cells using multiple-feature Boolean logic. *Nat Method* **7**: 763–772.
- Fink AP, Azim E, Croce KR, Huang ZJ, Abbott LF, Jessell TM. 2014. Presynaptic inhibition of spinal sensory feedback ensures smooth forelimb movement. *Nature* **509**: 43–48.
- Fu Y, Tucciarone JM, Espinosa JS, Sheng N, Darcy DP, Nicoll RA, Huang ZJ, Stryker MP. 2014. A cortical circuit for gain control by behavioral state. *Cell* **156**: 1139–1152.
- Huang ZJ. 2014. Toward a genetic dissection of cortical circuits in the mouse. (Review) *Neuron* **83**: 1284–1302.
- Lu JT, Tucciarone J, Lin Y, Huang ZJ. 2014. Circuit-specific maturation of synaptic dynamics in the recruitment of parvalbumin interneurons in visual cortex. *Proc Natl Acad Sci* **111**: 16895–16900.
- In Press*
- Penzo M, Robert V, Tucciarone J, He M, Huang ZJ, Li B. 2015. The paraventricular thalamus controls a central amygdala fear circuit. *Nature* (in press).

# NEUROBIOLOGY OF COGNITION AND DECISION MAKING

A. Kepecs   N. Bobrowski-Khoury   M. Lorenc   J. Sanders  
E. Demir   P. Masset   T. Sikkens  
B. Hangya   H.J. Pi   A. Vaughan  
J. Hirokawa   J.L. Pie  
D. Kvitsiani   S. Ranade

Understanding the mysteries of cognition has been an age-old aspiration of humankind. Our laboratory engages this question by studying the neurobiological basis of cognitive tasks in rodents. We use a reductionist approach to translate psychological questions into the language of neuroscience by developing quantifiable, well-controlled behavioral tasks for rodents. We then couple these tasks with targeted, high-resolution, and high-throughput monitoring and manipulation of neural circuits mediating cognitive behaviors. Given the complexity of animal behavior and the dynamics of neural networks producing it, our studies rely on computational models to guide and sharpen the neurobiological questions. Finally, we use human psychophysics to validate our behavioral observations in rodents by linking them to analogous behaviors in humans. Behavioral links to humans can also serve to identify behavioral dimensions that are predictive of disordered mental states, and our goal is to bridge our studies in animals with psychiatric disorders. Using this integrated approach, our long-term goal is to “reverse engineer” the computational and neurobiological principles underlying cognition and decision making.

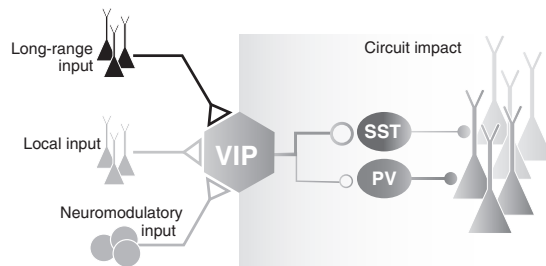
In terms of topics, our approach is multifaceted: We study (1) the roles of uncertainty and confidence in decision making, (2) foraging decisions about whether to stay or to switch, (3) the division of labor between different cell types in the prefrontal cortex, (4) how the cholinergic system supports sustained attention, and (5) social decisions that rely on shallow, stereotyped circuits. A unifying theme is the use of cell-type- and pathway-specific perturbations to effect gain and loss of function for specific behavioral abilities. Through such manipulations of genetically and anatomically defined neuronal elements, we hope to identify fundamental principles of neural circuit function that will be ultimately useful for developing therapies for diseases such as addiction, major depression, schizophrenia, Alzheimer’s disease, and autism spectrum disorder.

## Neural Basis of Decision Confidence

J. Hirokawa, P. Masset [in collaboration with A. Lak, University of Cambridge; G. Costa, and Z.F. Mainen, Champalimaud Neuroscience Program, Portugal; A. Koulakov, Cold Spring Harbor Laboratory]

Natural events in the world, and the likely consequences of our actions, are both fraught with ambiguity. In face of pervasive uncertainty, one’s degree of confidence in a belief is a critical component of cognition. This can confer benefits for a broad range of activities from the sophisticated to the mundane to the essential: managing a stock portfolio, deciding whether to carry an umbrella, or deciding between fight and flight. An essential use of confidence estimates is to guide information-seeking behaviors, learning, and attention so as to reduce the level of uncertainty. Conversely, the pathological miscalculation of confidence can contribute to a wide range of neuropsychiatric conditions, including anxiety, obsessive-compulsive disorder, and addiction.

What is the neural basis for such judgments? Previously, we discovered neurons in the orbitofrontal cortex (OFC) that signal decision confidence, and we are pursuing these observations by trying to establish the causal neural circuit basis for estimating and acting on decision confidence. We designed a new behavioral task for rats, in which we could measure confidence behaviorally on a trial-by-trial basis. Briefly, rats are trained on a simple olfactory decision task, and by delaying reward, we found that the time they are willing to wait for an uncertain reward is proportional to decision confidence, as predicted by theory. To further refine the role OFC in confidence, we are trying to “read out” confidence reports (i.e., predict the timing of leaving decisions on a trial-by-trial basis based on neural activity [establish correlation]). To demonstrate that OFC is necessary for confidence reports, we used both excitotoxic lesions



**Figure 1.** A canonical cortical circuit for disinhibition. Vasointestinal peptide (VIP) interneurons receive inputs from local neurons, long-range afferents from other cortical areas, as well as serotonergic and cholinergic neuromodulatory input. VIP interneurons generate disinhibition by inhibiting two other major inhibitory neuron types, somatostatin (SST) and to a smaller degree parvalbumin (PV) interneurons.

and pharmacological inactivation techniques and found that confidence reports were disrupted without changing decision accuracy. To establish the sufficiency of OFC for confidence reporting, we are preparing gain-of-function experiments. Because there is no obvious map in prefrontal cortex, we cannot use electrical microstimulation. Instead, we will use optogenetic activation of neurons defined by projection target using retrograde viruses and attempt to inject extra “uncertainty” into the brain by activating the appropriate population of neurons. Currently, we are mapping the projections into and out of OFC to define a precise neural circuit for computing and using decision confidence.

### From Metacognition to Statistics: Confidence Judgments in Humans

J. Sanders, B. Hangya, P. Masset

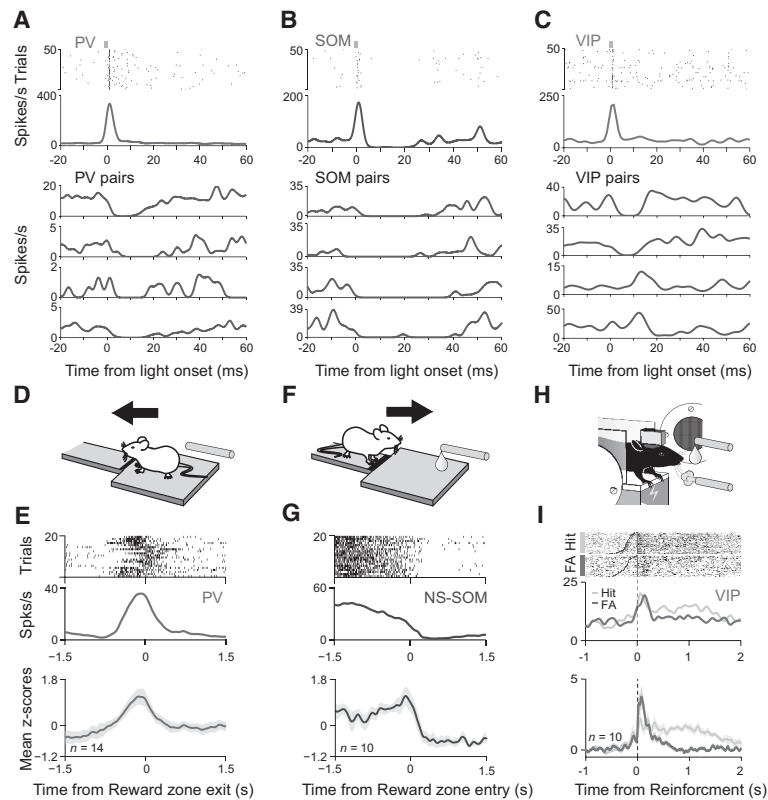
In our rodent studies of confidence, we use a computational framework to interpret behavioral and neural data. We are also interested in understanding the degree to which the statistical concept of *decision confidence* corresponds to the human notion of subjective confidence. Human confidence judgments are thought to originate from metacognitive processes that provide a subjective assessment about one’s beliefs. Confidence can be alternatively framed as an objective statistical quantity, an estimate of the probability that a chosen hypothesis was correct. Despite similar usage of this term, it remains unclear whether the objective, statistical notion of confidence is related to the subjective human feeling of confidence.

To address this issue, we developed a normative statistical framework to mathematically derive predictions

relating a statistical measure of decision confidence to observable measures of decisions. We found that human confidence reports in both sensory and general knowledge decision tasks followed the normative statistical predictions, defying several naïve intuitions about confidence. In addition, by fitting a model to choice behavior within this framework, we could quantitatively predict confidence reports. On the basis of these results, we suggest that the computation underlying the human sense of confidence is consistent with a mental implementation of statistical decision confidence.

Next, we developed a new confidence-reporting task that provides both implicit and explicit measures of decision confidence and is suitable for both rodents and humans in order to make direct comparisons. We designed a perceptual discrimination task where subjects choose which of two Poisson click streams is clicking faster and rate their confidence in their choices. On each trial, we store the precise time of each click the subject heard in each ear and acquire three measures: a reaction time (how long the subject collected evidence before choosing), a choice (which side they thought was faster), and a confidence report (how much they believe in the hypothesis that their choice was correct). Using several previously described models of choice and confidence, we tried to design a computer program that could “listen” to the same click trains and respond with the same reaction time, choice, and confidence. To our surprise, we discovered that standard models of perceptual decision making failed to reproduce the basic patterns from our normative model, that confidence reflects the strength of perceived evidence. Currently, we are examining a broader range of models to fit our psychophysical data.

We find that all of these confidence-reporting tasks share the basic patterns of confidence that is predicted by statistical confidence. On the basis of these results,



**Figure 2.** Network effects and behavioral correlates of identified interneuron classes. (A,B,C) (Top) Spike raster and peri stimulus time histogram (PSTH) of PV (A), SOM (B), and VIP (C) interneurons aligned to light onset. (Bottom) PSTH of three simultaneously recorded unidentified neurons (PV pairs, SOM pairs, and VIP pairs). (D,F) Cartoon of mouse performing the reward foraging task. (D) Mouse is leaving the rewarded zone. (E, Top two panels) Spike raster and peri-event time histogram (PETH) for an identified PV interneuron aligned to time of reward zone exit (bottom), mean z-scored response of 14 PV neurons (shaded area indicates S.E.M.). (F) Mouse is approaching a rewarded zone. (G, Top panels) Spike raster and PETH for an NS-SOM interneuron aligned to the time of reward zone entry. (Bottom) Mean responses for 10 NS-SOM neurons (shaded area represents S.E.M.). (H) Schematic of behavioral setup and auditory discrimination task. (I) (Top two panels) Raster plots and PETH of example VIP neuron. Neural activity was aligned to reinforcement (reward/punishment; dark gray). (Bottom panel) PETHs of VIP neurons ( $n = 10$ ). All VIP neurons showed a strong increase in firing rate after punishment.

we are now in a position to use these quantitative measures of decision confidence in humans, through collaborations and with neuroimaging and genetic approaches. Moreover, by showing that a single confidence measure is applicable to humans and rodents, our results strengthen the case for using the rat as a model system for studying decision confidence.

## The Representational Content of the OFC during Outcome Anticipation

J. Hirokawa, A. Vaughan, J.L. Pie

OFC is strongly implicated in decision making under uncertainty, especially when decisions require the

evaluation of predicted outcomes. Neurophysiological studies conceptualize OFC as a hub for the integration of different variables into unified value signals. Using model-based approaches, previous studies have identified a number of critical outcome-related variables in OFC, such as reward value, risk, and confidence. However, such model-based approaches are limited in their ability to characterize neuronal representations because it is difficult to address all potentially encoded variables. Here, we address the representational content of OFC using both model-based and model-free (i.e., unsupervised) approaches. First, we assessed whether and how different decision variables are integrated in OFC using a model-based approach. Second, we sought to uncover the content of



OFC representations using unsupervised techniques to assess whether decision variables arise naturally, and whether they are represented separately or as a continuum.

To address this, we trained rats in a task that combines two classic decision paradigms, one based on sensory uncertainty (perceptual decision) and one based on reward uncertainty (value-based decision) to engage OFC. We recorded neural activity and focused our analysis on the period after choice/before feedback (i.e., while the animal is anticipating the outcome of its choice). First, we constructed a normative model to account for the behavioral data. Our model combines reward value and perceptual decision confidence in a Bayesian way to yield integrated value, which is then used to make choices. The model correctly predicted the behavioral choices and reaction times observed in each animal. From neuronal data, analysis of single units revealed that individual OFC neurons can represent identifiable decision variables separately, including reward size and confidence, as well as their combination into integrated reward value. Second, we used an unsupervised clustering method to uncover a more complete representational structure of task-related activity in a model-free way. This analysis revealed non-random, structured representations of these decision variables in OFC. Taken together, we demonstrated that OFC representations of task variables are fundamentally linked to model-based decision variables such as reward value and decision confidence.

### **Computational Psychiatry: Quantitative Behavioral Phenotyping in Rodents and Humans**

P. Masset, J. Hirokawa [in collaboration with D. Rubino, University of North Carolina, Chapel Hill]

At present, the classification of mental disorders is based on clinical criteria and self-reports of symptoms. Objective and reliable diagnoses are a prerequisite for meaningful progress in establishing genetic and neural etiologies and eventually new treatments. Although the promise of a dimensional approach to characterize behavior and its dysfunctions is widely appreciated, we do not know how to identify appropriate “behavioral dimensions” that are predictive of disordered mental states. Our goal is to develop behavioral tasks that isolate specific mental computations, such as decision confidence, and use computational models to quantitatively characterize the behavioral data. We have begun

this work in rodents and healthy humans with the goal of applying them to genetic models of psychiatric diseases in rodents and patient populations.

Usually, animal models are validated mainly on the basis of genetic insights and simple behavioral measures. However, in the case of cognitive diseases, this has become challenging because it is unclear how to map human behavioral deficits to animal models. To overcome these issues, we focus on microtraits that can be easily quantified using psychometric tasks and be directly related to humans (e.g., reward sensitivity, learning rate, and decision confidence). We attempted to use such a dimensional approach to identify microtraits that may be quantitatively different between our model rats and their controls. We studied major depressive disorder (MDD), a heterogeneous condition whose pathophysiology remains unclear, with a commonly used rodent model, the congenital learned-helplessness (cLH) rat. Using different behavioral paradigms, we tested two major hypotheses about depression: anhedonia, a blunted sensitivity to reinforcers, and depressive realism, the improved calibration of the sense of confidence. We first manipulated reward contingencies by introducing a bias in either reward size or reward probability. Our results showed no significant difference between cLH and cNLH rats, which suggested that cLH rats did not display any imbalance in reward evaluation. We then assessed confidence based on the rat’s willingness to invest waiting time for uncertain outcomes. When they were most confident, the optimal strategy was to wait longer, and indeed this is what we observed. Moreover, we observed greater differences in cLHs’ waiting times with confidence, suggesting that they had more insight into their choices than controls during the task. These preliminary results highlight how a computational phenotyping approach can be used to relate animal models of psychopathologies to human behavior.

### **Impact of Drugs of Abuse on Decision Making and OFC**

J.L. Pie, J. Hirokawa

Drug abuse and addiction result in and are compounded by compromised decision-making processes. OFC has emerged as one important region for structural and functional alterations resulting from chronic drug exposure. Although opioids are widely used for treatment of chronic pain conditions, their long-term effects on

cognitive functions have not been well characterized. Morphine is a potent opiate drug that has been shown to directly affect OFC, along with other prefrontal regions with strong dopaminergic projections. Addiction to morphine renders subjects unable to make good judgments and flexible choices, which mimics the effects of OFC lesions. Therefore, we sought to find out whether the previously unexplained behavioral aspects of drug addiction could be due to defects in integration of decision confidence and reward value in the decision-making process. Specifically, we seek to identify the impact of morphine self-administration protocol on optimal decision making in the reward-biased 2AFC task.

We trained rats in the 2AFC task and exposed them to ~2 wk morphine self-administration followed by 2 wk withdrawal. These rats reliably evoked morphine self-administration, and following withdrawal, we tested for disruption of the decision-making processes as deviations in two types of choice biases. Reward size bias that changes choices modulated the propensity to choose toward the larger reward and previous outcome bias that generated changes in choice behavior based on the immediately preceding outcome. We observed distinct changes in choice bias: Reward size bias was significantly reduced, whereas previous outcome bias was not by morphine exposure. This result suggests that chronic morphine administration alters decision making selectively, disrupting either a circuit- or cell-type-specific representation of choice bias. On the basis of this evidence, we are pursuing the hypothesis that morphine self-administration leads to suboptimal decision making by disrupting integration and distribution of decision variables in OFC.

### Foraging Decisions, Anterior Cingulate Cortex and Inhibitory Neuron Types

D. Kvitsiani, S. Ranade, B. Hangya, M. Lorenc  
[in collaboration with Z.J. Huang, Cold Spring Harbor Laboratory]

Making decisions about when to engage with options as they are encountered and when to disengage and switch behavior is a fundamental problem for all decision makers. Such decisions about whether to stay and exploit an opportunity or leave and explore alternatives, presented in the language of behavioral ecology, are called foraging decisions. It is critical to appreciate that foraging decisions are not binary ones between currently available options (such as two-alternative

forced choice), but are instead decisions about when and how to engage with the available options.

We have been studying the anterior cingulate cortex (ACC) during a simple foraging task. In humans and rodents, ACC has been implicated in a variety of goal-directed behaviors including reward processing, inhibitory response control, and conflict monitoring. Electrophysiological recordings from ACC show great diversity of neuronal responses to a range of behavioral variables. To address this issue, we used optogenetics as a means to identify extracellularly recorded neurons in freely moving mice, focusing on inhibitory interneurons, which exhibit the largest diversity of cell types in cortex. We demonstrated that parvalbumin (Pv)-expressing neurons and a subtype of somatostatin (Som)-expressing neurons form functionally homogeneous populations, showing a double dissociation between both their inhibitory impact and behavioral correlates. Of a number of events pertaining to behavior, a subtype of Som neurons selectively responded at reward approach, whereas Pv neurons responded at reward leaving, encoding preceding stay duration. These behavioral correlates of Pv and Som neurons defined a behavioral epoch and a decision variable important for foraging (whether to stay or to leave), a crucial function attributed to ACC. Our results also point to a new view of inhibition, according to which interneurons encode behaviorally relevant variables and serve to control the flow of information and behavioral timescales.

### Cortical VIP Interneurons and Disinhibitory Control

H.J. Pi, B. Hangya, A. Vaughan, T. Sikkens [in collaboration with Z.J. Huang, Cold Spring Harbor Laboratory; B. Rozsa, KOKI, Hungary]

We are interested in understanding how the diversity of cortical inhibitory interneurons underlies distinct neural circuit dynamics that direct perception and behavior. We have recently identified a disinhibitory cortical circuit motif that appears to be a conduit for fast neuromodulatory action in cortex. This circuit is controlled by a class of inhibitory interneurons that express vasoactive intestinal polypeptide (VIP), which inhibits other interneurons and thereby disinhibits a subpopulation of principal neurons. VIP neurons express ionotropic nicotinic acetylcholine and serotonergic 5-HT<sub>3A</sub> receptors. Functionally, we showed that VIP interneurons in the auditory cortex

are recruited in response to specific reinforcement signals such as reward and punishment. Although these observations raise the possibility that VIP neurons transform neuromodulatory signals to cortical output, the exact neuromodulatory signals that are involved remain unknown.

What are the mechanisms by which VIP neurons are recruited at specific moments in behavior? Previous studies suggest that VIP neurons may act as a conduit for fast neuromodulatory control and thereby acquire behaviorally relevant responses. To identify which neuromodulatory systems recruit a VIP-controlled disinhibitory circuit, we took advantage of a rabies-virus-assisted retrograde tracing system that labels only monosynaptic input neurons fluorescently. We identified direct projections from cholinergic neurons in the basal forebrain and serotonergic neurons in the raphe nuclei. On the basis of these results, we are pursuing the hypothesis that VIP neurons transform fast neuromodulatory signals into cortical disinhibitory output.

## Neural Representation of Social Decisions and Rewards

E. Demir, N. Bobrowski-Khoury [in collaboration with R. Axel, Columbia University]

Social behavior is integral to animals' survival and reproduction. Social deficits are at the heart of cognitive disorders such as autism spectrum disorder that have proven profoundly difficult to study in model organisms. Mice, like humans, are social animals. To interact, cooperate, and compete with others, mice must collect information about each other's identity, fertility, and likely intent. We would like to understand how social information is represented, computed, and used by mice. In rodents, a main source of information for social decision making and reward valuation is the chemosensory system. The neural circuits supporting these tend to be shallow, from sensory input to motor actions, and highly stereotyped, enabling the systematic dissection of this system.

We started our studies by recording from part of the amygdala that processes information from the vomeronasal system, detecting behaviorally significant chemical cues such as pheromones. We have designed a behavioral task that presents different urinary proteins to our mice while we are performing electrophysiology in medial amygdala. Moreover,

we characterized molecular markers that enable us to further dissect social amygdalar circuit. Currently, we are carrying out optogenetic experiments as well as electrophysiological recordings to map behavioral relevance onto specific medial amygdala neurons.

We are also interested in understanding the basic rules that mice use to choose their partners. For this purpose, we have developed a psychophysical social behavior task, the "social carousel," inspired by perceptual psychophysics and game theoretic traditions that have been instrumental in studying other facets of cognition. Our task enables the reliable, quantitative, and high-throughput analysis of social interactions in mice. This task contrasts appetitive rewards with opportunities for social interaction. Water-deprived animals are trained to alternate between two platforms: One platform provides a small water reward, and the other allows for social interaction with caged mice drawn at random from a revolving magazine. Subjects can engage in extended social interactions with these mice, but at the expense of delaying the water reward. This task allows us to infer the "social value" of a mouse based on the trade-off between social interactions and appetitive rewards. In addition, this task is compatible with our electrophysiological studies because it is devised for precise stimulus delivery and reproducible behavioral contingencies. Currently, we are characterizing social preferences of individual mice of different strains and subspecies.

## Optogenetic Dissection of Nucleus Basalis during Sustained Attention and Reinforcement

B. Hangya, S. Ranade, M. Lorenc

The nucleus basalis (NB) is a vitally important yet poorly understood neuromodulatory system that is thought to have significant roles in cognitive functions, including learning, memory, and attention. Cognitive deficits in Alzheimer's disease, Parkinson's dementia, age-related dementias, and normal aging are correlated with the extent of deterioration of NB cholinergic neurons. Yet, despite the association of NB with higher cognitive functions and a host of disease states in humans, it is surprising how little we understand about its function during behavior. Previous research, mostly using lesions, pharmacology,

and microdialysis, revealed that NB can have strong and confusingly diverse effects on downstream targets and behavior. However, there are no functional studies that reveal what NB neurons actually do during behavior.

To overcome these challenges, we combine optogenetic and electrophysiological approaches to record from identified cholinergic projection neurons in two distinct nuclei of NB during behavior: The caudal region projecting to the auditory cortex and the horizontal diagonal band that projects to the prefrontal cortex. We have successfully developed visual and auditory versions of sustained attention tasks for mice. In these tasks, mice report target stimuli occurring at uncertain times. Performance and reaction time are used as indicators of attention. We manipulate temporal expectancy by using specific distributions for stimulus timing, thus modulating attention in a temporally precise, graded manner, also reflected in the reaction times of the animals. We recorded single-neuron activity from the NB while mice were performing the auditory sustained attention task and identified cholinergic neurons by using an optogenetic approach. We found that central cholinergic neurons responded to primary reward and punishment with remarkable speed and precision ( $18 \pm 3$  msec), unexpected for a neuromodulatory system. Responses to reward were scaled by reinforcement surprise, indicating that the cholinergic system also conveys cognitive information. These results reveal that cholinergic neurons provide a rapid, reliable, and temporally precise signaling route

for reinforcement feedback that can mediate fast cortical activation and plasticity. Such fast responses of cholinergic neurons have not been reported and cast into doubt the prevailing notion that neuromodulators are slow-acting. These data also provide support to the idea that acetylcholine is a “broadcast signal” involved in identifying salient events, possibly mediating cortical plasticity and learning.

## PUBLICATIONS

- Hangya B, Pi HJ, Kvitsiani D, Ranade SP, Kepecs A. 2014. From circuit motifs to computations: Mapping the behavioral repertoire of cortical interneurons. *Curr Opin Neurobiol* **4**: 117–124.
- Kepecs A, Fishell G. 2014. Interneuron cell types are fit to function. *Nature* **505**: 318–326.
- Kepecs A, Masset P. 2014. Categorical decisions. In *Encyclopedia of computational neuroscience* (ed. Jaeger D, Jung R), pp. 310–311. Springer, New York.
- Lak A, Costa GM, Romberg E, Koulakov AA, Mainen ZF, Kepecs A. 2014. Orbitofrontal cortex is required for behavioral report of perceptual decision confidence. *Neuron* **84**: 190–201.
- Ranade S, Pi HJ, Kepecs A. 2014. Waiting for serotonin. *Curr Biol* **24**: R803–R805.
- Sanders JI, Kepecs A. 2014. A low cost programmable pulse generator for neurophysiology and behavior. *Front Neuroeng* **7**: 43.
- Takada N, Pi HJ, Sousa VH, Waters J, Fishell G, Kepecs A, Osten P. 2014. A developmental cell-type switch in cortical interneurons leads to a selective defect in  $\gamma$  oscillations. *Nat Commun* **5**: 533.

## In Press

- Hangya B, Sanders J, Kepecs A. 2015. A mathematical framework for statistical decision confidence. *BioRxiv* (in press).
- Hangya B, Ranade S, Lorenc M, Kepecs A. 2015. Central cholinergic neurons are rapidly recruited by reinforcement feedback. *Cell* (in press).

# THEORETICAL AND COMPUTATIONAL NEUROSCIENCE

A. Koulakov D. Ferrante Y. Wei

Our laboratory develops theories of neural computation. We work in parallel on three topics. First, we formulate mathematical models for combining genetic information and experience (nature and nurture) during formation of connections between neurons. Our models describe how genes can help build neural networks and how neural activity adds a layer of plasticity to the network topology that reflects learning and experience. These models have been tested on circuits that are formed in the visual system and can be rewired through genetic, surgical, and pharmacological manipulations. Second, we have been developing the neural network theory for olfactory processing. We previously argued that information about smells can be represented in the olfactory system in the form of temporal sequences. We develop network models that can interpret these sequences. With these models, we hope to understand the function of olfactory cortex. Finally, we are working on computational models of decision making and related processes such as confidence estimation within cortical networks forming perceptual decision.

## **A Network That Performs Brute-Force Conversion of a Temporal Sequence to a Spatial Pattern: Relevance to Odor Recognition**

H. Sanders, B. Kolterman, J. Lisman, D. Rinberg, A. Koulakov

---

A classic problem in neuroscience is how temporal sequences can be recognized. This problem is exemplified in the olfactory system, where an odor is defined by the temporal sequence of olfactory bulb output that occurs during a sniff. This sequence is discrete because the output is subdivided by gamma frequency oscillations. We proposed a new class of “brute-force” solutions to recognition of discrete sequences. We demonstrated a network architecture in which there are a small number of modules, each of

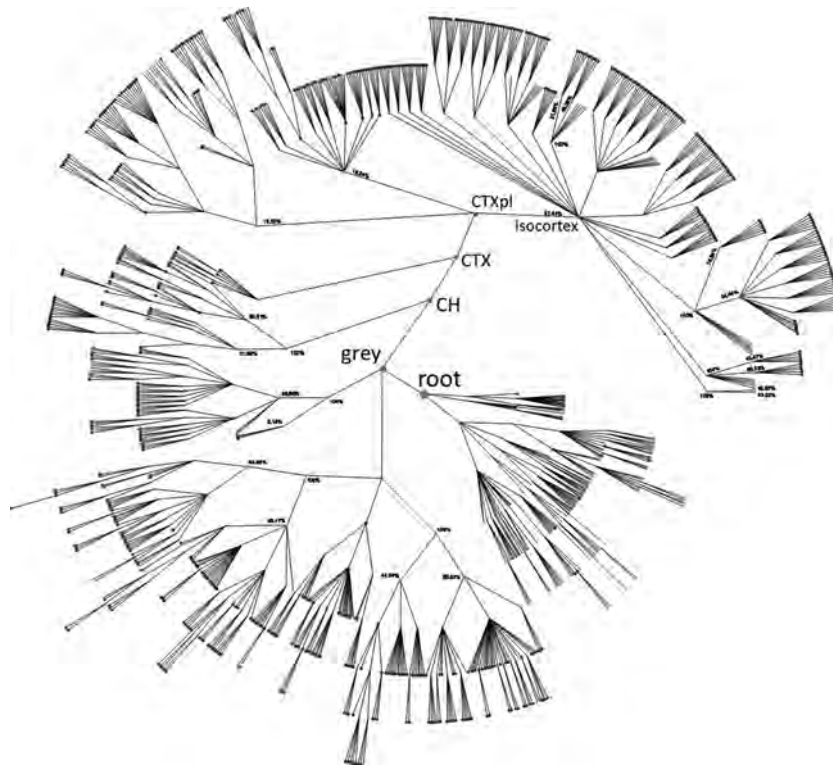
which provides a persistent snapshot of what occurs in a different gamma cycle. The collection of these snapshots forms a spatial pattern that can be recognized by standard attractor-based network mechanisms. This strategy has implications for recognizing odor-specific sequences generated by the olfactory bulb.

## **Orbitofrontal Cortex Is Required for Optimal Waiting Based on Decision Confidence**

A. Lak, G. Costa, E. Romberg, A. Koulakov,  
[in collaboration with Z. Mainen; A. Kepecs,  
Cold Spring Harbor Laboratory]

---

Confidence judgments are a central example of metacognition—knowledge about one’s own cognitive processes. According to this metacognitive view, confidence reports are generated by a second-order monitoring process based on the quality of internal representations about beliefs. Although neural correlates of decision confidence have been recently identified in humans and other animals, it is not well understood whether there are brain areas specifically important for confidence monitoring. To address this issue, we designed a postdecision temporal wagering task in which rats expressed choice confidence by the amount of time they were willing to wait for reward. We found that orbitofrontal cortex inactivation disrupts waiting-based confidence reports without affecting decision accuracy. To address this data, we have developed a computational model that both accounts for the results of the temporal wagering task and provides a statistical basis for confidence measurement. We show that the normative model can quantitatively account for waiting times based on the computation of decision confidence. These results establish an anatomical locus for a metacognitive report, confidence judgment, distinct from the processes required for perceptual decisions.



**Figure 1.** Mouse brain parcellation tree. Each brain region (PU) is represented by a node. PUs farther from the root are parts of upstream PUs.

## Statistical Model of Evolution of Brain Regions

D. Ferrante, Y. Wei, A. Koulakov

We study the distribution of brain and cortical area sizes—parcellation units (PUs; Fig. 1)—obtained for three species: mouse, macaque, and human. We find that the distribution of PU sizes is close to lognormal. We analyze the mathematical model of evolution of brain parcellation based on iterative fragmentation and specialization. In this model, each existing PU has a probability of splitting that depends on PU size only. This model shows that the same evolutionary process may have led to brain parcellation in these three species. Our model suggests that region-to-region (macro) connectivity is given by the outer product form. We show that most experimental data on non-zero macaque cortex macroconnectivity (62% for area V1) can be explained by the outer product power-law form suggested by our model. We propose a multiplicative Hebbian learning rule for the macroconnectome that could yield the correct scaling of connection strengths between areas.

## Long-Term Memory Stabilized by Noise-Induced Rehearsal

Y. Wei, A. Koulakov

Cortical networks can maintain memories for decades despite the short lifetime of synaptic strengths. Can a neural network store long-lasting memories in unstable synapses? We have studied the effects of ongoing spike-timing-dependent plasticity (STDP) on the stability of memory patterns stored in synapses of an attractor neural network. We show that certain classes of STDP rules can stabilize all stored memory patterns despite a short lifetime of synapses. In our model, unstructured neural noise, after passing through the recurrent network connections, carries the imprint of all memory patterns in temporal correlations. STDP, combined with these correlations, leads to reinforcement of all stored patterns, even those that are never explicitly visited. Our findings may provide the functional reason for irregular spiking displayed by cortical neurons and justify models of system memory consolidation. Therefore, we propose that irregular

neural activity is the feature that helps cortical networks maintain stable connections.

#### PUBLICATIONS

Lak A, Costa GM, Romberg E, Koulakov AA, Mainen ZF, Kepecs A. 2014. Orbitofrontal cortex is required for optimal waiting based on decision confidence. *Neuron* **84**: 190–201.

Sanders H, Kolterman BE, Shusterman R, Rinberg D, Koulakov A, Lisman J. 2014. A network that performs brute-force conversion of a temporal sequence to a spatial pattern: Relevance to odor recognition. *Front Comput Neurosci* **8**: 108.

Wei Y, Koulakov AA. 2014. Long-term memory stabilized by noise-induced rehearsal. *J Neurosci* **34**: 15804–15815.

# THE FUNCTION AND PLASTICITY OF CENTRAL SYNAPSES IN ADAPTIVE AND MALADAPTIVE BEHAVIORS RELATED TO PSYCHIATRIC DISORDERS

B. Li   S. Ahrens   L. Mejia   M. Stephenson-Jones  
K. Delevich   M. Penzo   K. Yu  
G. Hwang   H. Schiff   X. Zhang

Understanding the relationship among synapse, circuit, and behavior has been the focus of research in our lab. We are particularly interested in understanding the synaptic and circuit mechanisms underlying cognitive functions, as well as synaptic and circuit dysfunction that may underlie mental disorders, including anxiety, depression, schizophrenia and autism. To address these questions, we use *in vitro* and *in vivo* electrophysiology, two-photon imaging, molecular, genetic, optogenetic, and chemical-genetic methodologies to probe and manipulate the function of specific neural circuits in the rodent brain and to determine their role in behaviors, such as learning and expression of fear, learned helplessness, and attention. We are currently undertaking the following major lines of research, which are summarized below.

## The Paraventricular Thalamus Controls a Central Amygdala Fear Circuit

---

Appropriate responses to an imminent threat brace us for adversities. The ability to sense and predict threatening or stressful events is essential for such adaptive behavior. In the mammalian brain, one putative stress sensor is the paraventricular nucleus of the thalamus (PVT), an area that is readily activated by both physical and psychological stressors. However, the role of the PVT in the establishment of adaptive behavioral responses remains unclear. Here, we demonstrate in mice that the PVT regulates fear processing in the lateral division of the central amygdala (CeL), a structure that orchestrates fear learning and expression. Selective inactivation of CeL-projecting PVT neurons prevented fear conditioning, an effect that can be accounted for by an impairment in fear-conditioning-induced synaptic potentiation onto somatostatin-expressing (SOM<sup>+</sup>) CeL neurons, which has previously been shown to store fear memory. Consistently,

we found that PVT neurons preferentially innervate SOM<sup>+</sup> neurons in the CeL, and stimulation of PVT afferents facilitated SOM<sup>+</sup> neuron activity and promoted intra-CeL inhibition, two processes that are critical for fear learning and expression. Notably, PVT modulation of SOM<sup>+</sup> CeL neurons was mediated by activation of the brain-derived neurotrophic factor (BDNF) receptor tropomyosin-related kinase B (TrkB). As a result, selective deletion of either *Bdnf* in the PVT or *Trkb* in SOM<sup>+</sup> CeL neurons impaired fear conditioning, whereas infusion of BDNF into the CeL enhanced fear learning and elicited unconditioned fear responses. Our results demonstrate that the PVT–CeL pathway constitutes a novel circuit essential for both the establishment of fear memory and the expression of fear responses. Additionally, we have uncovered mechanisms linking stress detection in PVT with the emergence of adaptive behavior.

## Synaptic Modifications in the Medial Prefrontal Cortex in Susceptibility and Resilience to Stress

---

When facing stress, most individuals are resilient, whereas others are prone to developing mood disorders. The brain mechanisms underlying such divergent behavioral responses remain unclear. We used the learned helplessness procedure in mice to examine the role of the medial prefrontal cortex (mPFC), a brain region highly implicated in both clinical and animal models of depression, in adaptive and maladaptive behavioral responses to stress. We found that uncontrollable and inescapable stress induced behavioral state-dependent changes in the excitatory synapses onto a subset of mPFC neurons: those that were activated during behavioral responses as indicated by their expression of the activity reporter *c-Fos*. Whereas synaptic potentiation was linked to learned helplessness,



a depression-like behavior, synaptic weakening, was associated with resilience to stress. Notably, enhancing the activity of mPFC neurons using a chemical-genetic method was sufficient to convert the resilient behavior into helplessness. Our results provide direct evidence that mPFC dysfunction is linked to maladaptive behavioral responses to stress and suggest that enhanced excitatory synaptic drive onto mPFC neurons may underlie the previously reported hyperactivity of this brain region in depression.

### Depression of Excitatory Synapses onto Parvalbumin Interneurons in the Medial Prefrontal Cortex in Susceptibility to Stress

In response to extreme stress, individuals either show resilience or succumb to despair. The PFC is required for coping with stress, and PFC dysfunction has been implicated in stress-related mental disorders, including depression. Nevertheless, the mechanisms by which the PFC participates in stress responses remain unclear. Here, we investigate the role of parvalbumin (PV) interneurons in the mPFC in shaping behavioral responses to stress induced by the learned helplessness procedure, in which animals are subjected to an unpredictable and inescapable stressor. PV interneurons in the mPFC were probed and manipulated in knock-in mice expressing the Cre recombinase under the endogenous PV promoter. Notably, we found that excitatory synaptic transmission onto these neurons was decreased in mice showing helplessness, a behavioral state that is thought to resemble features of human depression. Furthermore, selective suppression of PV interneurons in the mPFC using hM4Di, a DREADD (designer receptor exclusively activated by designer drug), promoted helplessness, indicating that activation of these neurons during stress promotes the establishment of resilient behavior. Our results reveal a cellular mechanism of mPFC dysfunction that may contribute to the emergence of maladaptive behavioral responses in the face of adverse life events.

### The Schizophrenia-Linked Gene *ErbB4* Regulates a Thalamic Reticular Nucleus Circuit for Sensory Selection

Selective processing of behaviorally relevant sensory inputs against irrelevant inputs is a fundamental cognitive function whose impairment has been implicated in major psychiatric disorders. It is known that the thalamic reticular nucleus (TRN) gates sensory information en route to the cortex, but the underlying mechanisms remain unclear. We show in mice that deficiency of *ErbB4*, a gene linked to schizophrenia, in somatostatin-expressing TRN neurons markedly alters behaviors that are dependent on sensory selection. Whereas the performance of the *ErbB4*-deficient mice in identifying targets from distractors was improved, their ability to switch attention between conflicting sensory cues was impaired. These behavioral changes were mediated by an enhanced cortical drive onto the TRN that promotes the TRN-mediated cortical feedback inhibition of thalamic neurons. Our results uncover a previously unknown role of ErbB4 in regulating cortico-TRN-thalamic circuit function. We propose that ErbB4 sets the sensitivity of the TRN to cortical inputs at levels that can support sensory selection while allowing behavioral flexibility.

#### PUBLICATIONS

- Wang M, Perova Z, Arenkiel BR, Li B. 2014 Synaptic modifications in the prefrontal cortex in susceptibility and resilience to stress. *J Neurosci* **34**: 7485–7492.
- In Press*
- Ahrens S, Jaramillo S, Yu K, Ghosh S, Lai C, Huang ZJ, Li B. 2015. ErbB4 regulation of a thalamic reticular nucleus circuit for sensory selection. *Nat Neurosci* **18**: 104–111.
- Penzo MA, Robert V, Tucciarone J, De Bundel D, Wang M, Van Aelst L, Darvas M, Parada LF, Palmiter R, He M, et al. 2015. The paraventricular thalamus controls a central amygdala fear circuit. *Nature* **519**: 455–459.
- Perova Z, Delevich K, Li B. 2015. Depression of excitatory synapses onto parvalbumin interneurons in the medial prefrontal cortex in susceptibility to stress. *J Neurosci* **35**: 3201–3206.

# INTEGRATIVE SYSTEMS NEUROBIOLOGY

P. Mitra   N. Baltera   A. Field-Pollatou   G. Havkin   S. Michelsen   V. Pinskiy   A. Tolpygo  
D. Ferrante   N. Franciotti   F. Mechler   A. Mukherjee   S. Savoia   K. Weber

We study complex neurobiological systems using a combination of experimental and computational approaches. The primary area of experimental work in our laboratory continues to be the Mouse Brain Architecture (MBA) project, with a goal of generating a mesoscale connectivity map. Our neuroinformatics research involves the development of analytical tools and informatics infrastructures to process, analyze, and integrate large volumes of neurobiological data in various brain connectivity projects.

In collaborative studies, we apply the methods developed for the MBA project to the zebra finch, marmoset, macaque monkey, and human; here we name the major ones. A joint study with Josh Huang at CSHL (supported by the Simons Foundation) to obtain a brain-wide census of GABAergic interneurons in the mouse brain during development and in genetic models of autism is drawing to a conclusion this year. With funding from the National Institutes of Health awarded through the BRAIN Initiative program starting in 2014, we joined forces with Pavel Osten at CSHL to obtain an accurate whole-brain census of neuron classes in cell-type-specific reporter mice and develop neuroinformatics infrastructure to analyze and integrate such data. On a separate BRAIN-I grant, we join a team effort led by Sacha Nelson (Brandeis University) on a cross-species study of neuronal cell types in transgenic strains of rats and mice. We also have ongoing collaborations with Marcello Rosa (Monash University, Australia) and Hideyuki Okan (RIKEN, Japan) on a project to develop the first atlas of the marmoset brain; with Harvey Karten (University of California, San Diego) to map the zebra finch brain; and with Bijan Pesaran (New York University) to study anatomical and functional connectivity of visual areas in the macaque brain.

I have recently been appointed distinguished (visiting) professor at the Indian Institute of Technology (IIT), Madras. In collaboration with Professor Balaraman Ravindran and his team at the Department of Computer Science and Engineering at IIT, I have initiated a Center for Computational Brain Research.

Work at the Center will be devoted to the analysis of neuroanatomical data sets from whole-brain light microscopy, including those from our projects. Computer scientists from IIT Madras and students will be visiting CSHL in summer 2015 to further this collaborative work. Continuing in our laboratory in 2014 are Sandra Michelsen (Administrator), Alex Tolpygo (Laboratory Manager), Kevin Weber (Laboratory Technician), Neil Franciotti (Laboratory Technician), Nicholas Baltera (Laboratory Technician), Stephen Savoia (Laboratory Technician), Angeliki Field-Pollatou (Visiting Graduate Student), Gregor Havkin (Computational Science Manager), and Ferenc Mechler (Computational Science Analyst, Project Manager). Amit Mukherjee (Postdoctoral Fellow) and Vadim Pinskiy (Computational Science Analyst) have recently left in pursuit of their careers in industry. Daniel Ferrante (Computational Science Manager) joined in 2015.

## Mouse Brain Architecture Project

V. Pinskiy, A. Tolpygo, K. Weber, N. Franciotti, N. Baltera, S. Savoia, G. Havkin, A. Mukherjee, D. Ferrante, F. Mechler

The MBA project, now in its fifth year, aims to construct a comprehensive mesoscale wiring map of the adult mouse brain using a classical neuroanatomical approach that we scaled to a high-throughput data acquisition pipeline. Using computer-guided stereotaxic targeting, we microinject neurotracer substances (two anterograde or two retrograde) in one of the 390 predetermined sites systematically mapped on a three-dimensional grid that covers the entire brain volume. The sample brains are sectioned, histologically processed, and digitally scanned. Labeled somata or axonal projections are detected and analyzed in registered, annotated, high-resolution whole-brain image stacks, and the mesoscale connectivity matrix is assembled from the analyzed sample brains. The MBA data enjoys critical advantages over the current alternative mouse brain connectivity mapping projects

**3D Viewer**

**GET STARTED:** Select your injection point below by clicking on the button to the right. Once an injection point is selected you will be brought back up to the 3D navigation. Click on the injection point to see a list of brains which will appear in the right hand column. Click on "Learn More" to see the full picture.

CLEAR INJECTION POINTS

SELECT YOUR INJECTION POINTS

FILTER BY TRACER

FILTER BY REGION

**Brains Featured in this Area**

**MouseBrain\_1309 F**  
 Inj Region: Primary somatosensory area  
 Coordinates: (2.77 mm, 0.34 mm, 1.30 mm)  
 Tracer: AAV2/1-CAG-eYFP/cre/PT6, TrkB

**MouseBrain\_1057 F**  
 Inj Region: Primary somatosensory area  
 Coordinates: (2.00 mm, -0.12 mm, 1.00 mm)  
 Tracer: RV-60P1813x ("Theragra")

**MouseBrain\_1319 HC**  
 Inj Region: Primary somatosensory area  
 Coordinates: (2.48 mm, 1.15 mm, 1.65 mm)  
 Tracer: CTB

**Injection Points**

**How it works:** Navigate below by using the dots to drill down and click on the name to select that point above in the 3D viewer

**Brain tree navigator**

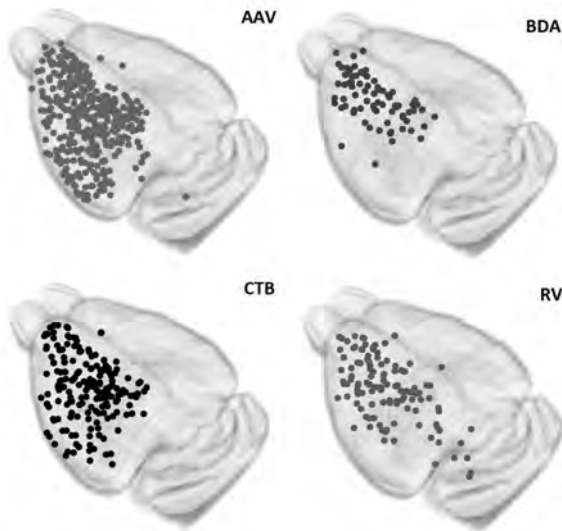
- Cortical plate (618)
- Cortical thalplate (7)
- Hippocampal formation (62)
- Isocortex (64)
- Olfactory areas (64)
- Aggessor insular area (2)
- Anterior cingulate area (1)
- Auxiliary areas (14)
- Subthalamic area (2)
- Frontal pole, medial cortex (1)
- Insular areas (2)
- Infralimbic area (1)
- Orbital areas (3)
- Postorbital area (2)
- Posterior parietal association areas (6)
- Prelimbic area (3)
- Prepyriform area (37)
- Subcallosal areas (17)
- Somatosensory areas (16)
- Visual areas (6)
- Primary somatosensory area (14)
- Region

Figure 1. The new three-dimensional injection viewer and brain tree navigator on the MBA portal.

(including the Allen Institute's), including higher spatial resolution and high-specificity retrograde tracing.

Registered project brains and metadata are released to the public on the MBA web portal (<http://mouse.brainarchitecture.org>). The portal interface and user experience itself have continued to improve with the new addition of a rotating three-dimensional injection

viewer (Fig. 1). To date, we have processed and released ~800 brains (Fig. 2). With continued support from the Mathers Foundation and with the current processing and publishing rate at approximately one to two brains per day, our expectation is to be able to complete data acquisition with at least one anterograde (AAV) and one retrograde (CTB) tracer by year end.



**Figure 2.** Injection data sets published to date, arranged in a common reference frame, separated by tracer. AAV  $n = 390$ ; CTB  $n = 190$ ; BDA  $n = 80$ ; RV  $n = 120$ .

**Outreach.** We continue to work to highlight the MBA project to the scientific public. I have published a perspectives paper on mesoscale circuit mapping in the brain (Mitra 2014) and have been involved in organizational activity with the National Science Foundation for BRAIN-I workshops. The MBA project has been represented in talks, posters, panel discussions, and a demonstration booth at the major annual scientific forums for the field (Max Planck/HHMI Connectomics Conference Berlin 2014; Annual Meeting of the Society for Neuroscience, Washington, DC).

### Alterations in Brain-Wide GABAergic Neuroanatomy in Autism Mouse Models

V. Pinskiy, A. Tolpygo, A. Mukherjee, F. Mechler

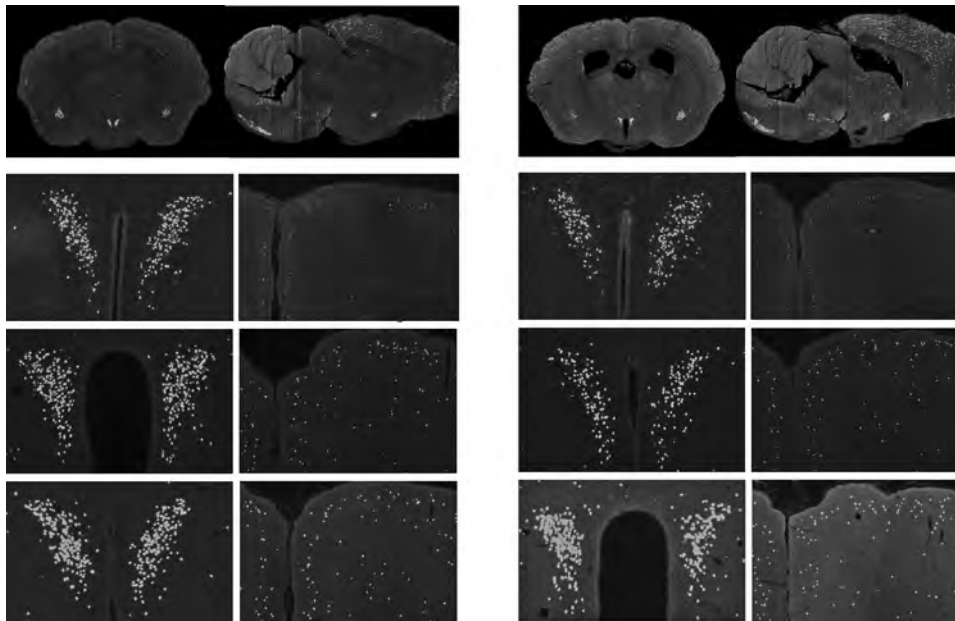
In our collaboration with Josh Huang, funded by the Simons Foundation Autism Research Initiative (SFARI), we address the developmental neuroanatomy of dysfunctional inhibition implicated in autism spectrum disorders (ASD). In a systematic brain-wide approach, we generate whole-brain maps of genetically targeted key subpopulations of inhibitory neurons that have specialized cellular morphology reflecting distinct physiological function. We quantify the cell distribution and long-range axonal projections of these GABAergic neurons and compare

them in an ASD mouse model (the *16p(df/+)* heterozygous copy-number-deficit mutant) and the “wild type” (C57BL/6). Last year, we reported on the distribution of somatostatin-expressing (SOM+) GABAergic bitufted neurons. Here, we report on data collected since then for a comparative analysis of the density distribution of the corticotropin-releasing hormone-positive (CRH+) subtype of GABAergic neurons during development (Fig. 3). The data acquisition is drawing to completion, and we have started the analysis phase. Software tools are being developed for the rigorous quantitative analysis of these data sets. In this process, aligned whole-brain image stacks are registered to a reference atlas, and the normalized cell counts obtained with automatic cell detection within annotated anatomical compartments are compared across genotypes. Data sets and analysis results will be released to the public on a dedicated portal.

### A Mesoscale Circuit Map of the Marmoset Brain

A. Tolpygo, N. Franciotti, D. Ferrante

The common marmoset (*Callithrix jacchus*) is a small New World primate with accelerated development and is the first primate with established stable transgenic lines. For neuroscientists, the species is important for being the simplest organism that shares many of the features that make primate brains special. Several years ago, Dr. Mitra joined world-renowned marmoset expert Dr. Marcello Rosa (Monash University, Australia) and several other international collaborators to create the first digital marmoset brain connectivity atlas. This project was seeded with existing cortical connectivity data originally collected by Dr. Rosa, and it currently focuses on new data acquisition and web portal development. More recently, Professor Mitra was invited to join a parallel international collaborative effort centered at RIKEN Brain Science Institute (Japan) to map the anatomical and functional connectivity of the marmoset brain, with Dr. Mitra's efforts contributing toward a comprehensive mesoscale circuit map of the marmoset brain. The project is modeled after the MBA project at CSHL: It will be using injections of neuronal tracer substances, including viral vectors, placed on a systematic grid designed to cover one hemisphere of the brain and will use a



**Figure 3.** The distribution of CRH+ subtype of GABAergic neurons in the ASD model *16p(df/+)* mutant mouse (*left*) and its wild-type littermate (*right*) at the postnatal age of p8, p21, and p56 (*top to bottom* rows, respectively). Overall distribution and local CRH+ densities are qualitatively similar in ASD model and wild type at all developmental stages.

newly built high-throughput neurohistology pipeline. The raw digitized image data will be made publicly available through a data portal, as well as subjected to computational neuroanatomical analysis and circuit reconstruction.

### The Cavity Method for Phase Transitions in Sparse Reconstruction Algorithms

P. Mitra [in collaboration with A. Sengupta, Rutgers University]

Compressed sensing relies on sparse reconstruction algorithms to retrieve high-dimensional sparse signals from a limited amount of measurements under suitable conditions. These algorithms exhibit sharp phase transition boundaries where signal retrieval breaks down as the number of variables goes to infinity. A common approach to predict these algorithmic breakdown boundaries uses mean field approximation of associated statistical mechanical model systems at a finite temperature and subsequently taking a zero-temperature limit. Despite the success of this approach, the nontrivial last step obscures the reasons for the algorithmic breakdown. We present an alternative approach for the derivation of the phase

transition boundaries that uses the “cavity method,” working directly in the zero-temperature limit. The method has the advantage of naturally generating a susceptibility that provides insight into different phases in this system, and it can be generalized for analysis of a broader class of sparse reconstruction algorithms.

### PUBLICATIONS

- Carmel JB, Mechler F, Mukherjee A, Wen T, Tolpygo A, Mitra PP. 2014. Whole rat brain analysis of anterograde tracers from motor cortex and retrograde tracers from cervical spinal cord: Validation of an efficient, automated method with stereology. Meeting of the Society for Neuroscience, Washington, DC.
- Cressy M, Valente D, Altick A, Kockenmeister E, Honegger K, Qin H, Mitra PP, Dubnau J. 2014. Laboratory evolution of adenylyl cyclase independent learning in *Drosophila* and missing heritability. *Genes, Brain Behavior* **13**: 565–577.
- Geyer S, Reimann K, Weiss M, Rose D, Silva AC, Rosa MGP, Pinskiy V, Tolpygo A, Giese G, Scherbarth A, Tolpygo A, Mechler F, Mitra PP. 2014. Combining brain clearing and light-sheet fluorescence microscopy with tape-transfer assisted sectioning and histochemistry for whole mouse brain and spinal cord imaging. Meeting of the Society for Neuroscience, Washington, DC.
- Grange P, Bohland JW, Okaty BW, Sugino K, Bokil H, Nelson SB, Ng L, Hawrylycz M, Mitra PP. 2014. Cell-type-based model explaining co-expression patterns of genes in the brain. *Proc Natl Acad Sci* **111**: 5397–5402.

- Lee J, Liu J, Lee H, Lin P, Fang P, Weitz Z, Fisher A, Pinskiy V, Mitra PP, Schiff N. 2014. Whole brain dissection of central thalamic circuit function with optogenetic fMRI. Meeting of the Society for Neuroscience, Washington, DC.
- Mechler F, Mitra PP. 2014. Mapping brain-wide connectivity in mouse. Max-Planck and HHMI Connectomics Conference, Berlin, Germany.
- Mitra PP. 2014. The circuit architecture of whole brains at the mesoscopic scale. *Neuron* **83**: 1273–1283.
- Mitra PP, Turner R. 2014. The microanatomical origin of the high-field magnetic resonance imaging signal in the marmoset cerebral cortex: myelin sheaths or cell bodies? Meeting of the Society for Neuroscience, Washington, DC.
- Ramezani M, Mitra PP, Sengupta AM. 2014. The cavity method for phase transitions in sparse reconstruction algorithms. arXiv1501.03194.
- Striedter GF, Belgard TG, Chen CC, Davis FP, Finlay BC, Güntürkün O, Hale ME, Harris JA, Hecht EE, Hof PR, et al. 2014. NSF workshop report: Discovering general principles of nervous system organization by comparing brain maps across species. *J Comp Neurol* **522**: 1445–1453.

*In Press*

- Okano H, Mitra PP. 2015. Brain-mapping projects using the common marmoset. *Neurosci Research* (in press).

# IDENTIFICATION OF DISRUPTED BRAIN CIRCUITS IN MOUSE MODELS OF AUTISM AND SCHIZOPHRENIA

P. Osten G. Fitzgerald C. Mende R. Palaniswamy J. Taranda  
Y. Kim A. Narasimhan E. Szelenyi K.U. Venkataraju

We are interested in studying how brain circuits assemble during development and give rise to behavior in the juvenile and adult mouse, both under normal conditions and in genetic mouse models of autism.

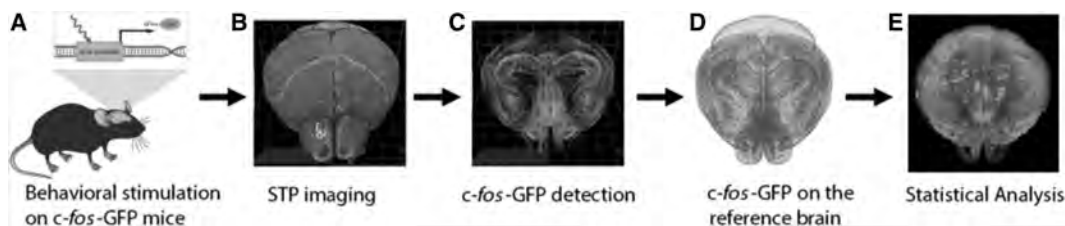
## Mapping Mouse Brain Activation

For the past several years, we have been developing a largely automated and highly quantitative approach to mapping neuronal activation in the mouse brain at cellular resolution. We have introduced a novel microscopy, called serial two-photon (STP) tomography, which achieves high-throughput fluorescence imaging of whole mouse brains by integrating two-photon microscopy and tissue sectioning (Ragan et al., *Nat Methods* 9: 255 [2012]; Osten and Margrie, *Nat Methods* 10: 515 [2013]). More recently, we have established a computational pipeline that allows the visualization of the immediate-early gene *c-fos*, a molecular marker of neuronal activation, by STP tomography in transgenic *c-fos*-GFP mice. The *c-fos*-GFP-positive neurons are computationally detected, their distribution is warped to a reference brain registered to the Allen Mouse Brain Atlas (ABA), and the activated brain regions are identified by rigorous statistical tests comparing *c-fos*-GFP cell counts in more than 600 anatomical ABA regions (Fig. 1). We have demonstrated the power of these methods by generating whole-brain activation

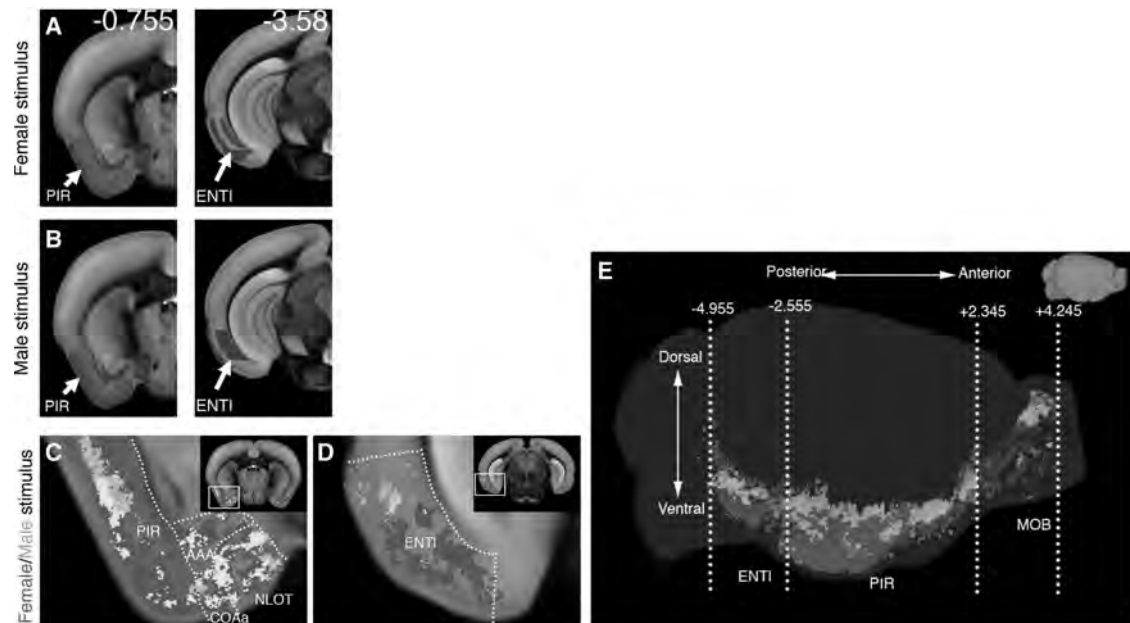
maps for two complex behaviors: a brief social interaction between a male and either a female mouse (sexual behavior) or another male mouse (agonistic behavior) (Kim et al. 2014). The result uncovered many novel findings including dorsoventral separation in activation patterns from olfactory cortices based on the two different social stimuli (Fig. 2). We anticipate that our methods will enable routine quantitative generation of whole-brain activation maps representing different complex behaviors, including other innate behaviors (e.g., sex, aggression, or fear) and higher cognitive behaviors (e.g., decision making). We predict that such maps will become a valuable resource for the systems neuroscience community. Furthermore, the application of our methods to the screening of brain activation in genetic mouse models of autism may reveal brain circuit deficits common to multiple susceptibility genes, which could serve as clinically relevant brain-circuit-based targets (or biomarkers) for the development of novel therapeutics.

## Toward Quantitative Cell-Type-Based Mapping of the Mouse Brain

The mouse brain comprises approximately 70 million neurons and approximately 30 million glia and other cells. In addition to the method for mapping brain activation described above, we have developed an assay for



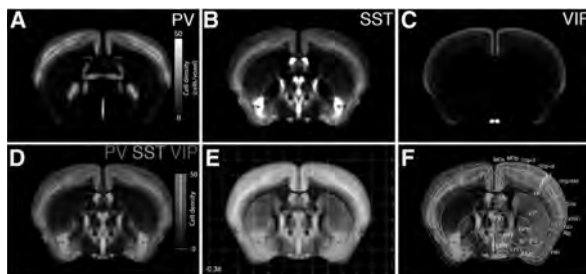
**Figure 1.** Mapping mouse brain activation by STP tomography. (A) Brain activation causes the induction of *c-fos*-GFP. (B) The brain is imaged by STP tomography. (C) *c-fos*-GFP cells are computationally detected. (D) The data sets are warped to a “reference” brain and Allen Mouse Brain Atlas. (E) The distribution of *c-fos*-GFP cells is compared by statistical analyses and the results are tabulated in a spreadsheet of brain regions.



**Figure 2.** Social brain circuit: Dorsoventral separation in the mouse olfactory bulb (MOB) downstream circuitry. (A,B) The male–female (A) and male–male groups (B) are compared to the handling group and significantly activated ROIs downstream from the MOB are displayed. Note similar activation in piriform (PIR) and entorhinal cortices (ENTI). Numbers in A represent bregma A/P coordinates. (C–E) Voxel-based analysis revealed activation patterns selective for the female stimulus, the male stimulus, and shared by both stimuli. Dorsoventral separation was detected between the male- and female-evoked activation in the PIR (C) and ENT (D). Such a pattern is clearly visualized by a three-dimensional side view (E).

quantitative mapping of the many neuronal and glial cell types present in the mouse brain. This work aims to provide a complete picture of cell type distribution in the mouse brain during development, in the adult, and during aging. The study of the developing brain will shed light on the temporal sequence of cell-type- and region-specific circuit assemblies that give rise to

emerging motor, sensory-perceptual, and cognitive specializations in the young brain. The experiments in the adult brain will focus primarily on the questions of gender differences in cell numbers that have been described for several brain areas, but have not been examined at a whole-brain level. As a first step toward this goal, we applied our mapping method to quantify expression of three major cortical interneurons throughout the whole brain (Fig. 3). In addition, we further developed a computational way to measure volumes of anatomical regions. The result provides complete cell counting and density measurement of genetically labeled cells. The established method will be applied to gain quantitative understanding of different cell types in the context of neurodevelopment and disease models.



**Figure 3.** Whole-brain PV, SST, and VIP+ interneuron distribution. (A–C) Quantitative spatial cell type mapping of PV-IRES-cre (A), SST-IRES-cre (B), and VIP-IRES-cre (C) crossed with LoxP-Stop-LoxP-H2B-GFP reporter mice. The heat map represents the number of cells per voxel. (D–F) Virtual overlay of signals from three different cell types in the same space (D) and over the reference STP brain (E,F).

## PUBLICATIONS

Kim Y, Venkataraju KU, Pradhan K, Mende C, Taranda J, Turaga SC, Arganda-Carreras I, Ng L, Hawrylycz MJ, Rockland KS, et al. 2014. Mapping social behavior–induced brain activation at cellular resolution in the mouse. *Cell Rep* 10: 292–305.



# NEURAL CIRCUITRY FOR SOCIAL COMMUNICATION

S. Shea   B. Cazakoff   D. Eckmeier  
M. Dinh   B. Lau

The overarching research goal in our laboratory is to understand how processing in specific brain circuits works to support natural communication behaviors. We aim to reveal neural mechanisms that allow organisms to detect and recognize familiar individuals, to gather information about their identity and social status, and to select appropriate behaviors. Mice are capable of acquiring detailed profiles on one another from the smells and sounds experienced during their social encounters. These dossiers may include information on a mouse's sex, genetic identity, reproductive state, levels of distress or sexual interest, or even recently consumed foods—details that are indispensable for survival and mating success. Initially, we are working to understand the neuronal activities and mechanisms in primary sensory brain areas that support these forms of communication. In the future, we anticipate moving deeper into the brain to ascertain where the sensory data from those regions are collected and integrated into hormonal and electrical signals that promote appropriate behavioral choices. We are particularly interested in how all stages of this decision arc are flexible according to experience and behavioral state.

The scientific benefit of this approach is twofold. First, we want to identify fundamental principles for how the circuitry of the brain adaptively controls complex behavior. In our pursuit of this goal, innate social behaviors are advantageous because they allow us to study neural circuits in light of problems they are exquisitely adapted to solve. Second, it is also our goal to pinpoint disturbances in the circuits that impair appropriate use of social information. Difficulty with social perception and cognition are core features of the autism spectrum disorders (ASD); for example, patients may have trouble perceiving and interpreting communication gestures such as speech, facial expressions, and “body language.” This broad feature is recapitulated in many mouse models of ASD that carry genetic variants identified in human ASD populations. Therefore, if we can ascertain the neural circuit substrates of social behavior in normal mice, we

can make and test predictions for how the circuitry is affected in the mouse models. The results are likely to tell us more about the synaptic modifications that occur in human autism.

## State-Dependent Neuronal Coding in Granule Cells of the Olfactory Bulb

B. Cazakoff, B. Lau, K. Crump, H. Demmer

---

Arguably, the most critical and finely tuned instrument providing a mouse with information about its social world is its sense of smell. Detection and interpretation of olfactory stimuli in the brain begins at the olfactory bulb. Odor signals passing through the olfactory bulb are subject to processing and modification by the intrinsic neuronal circuitry of the bulb. Among these intrinsic neurons, inhibitory granule cells are certainly the most numerous, arguably the most important, and yet somehow the most mysterious. These cells are the primary conduit for the olfactory bulb to integrate odor information with signals originating deep in the brain. Furthermore, granule cell anatomy along with a handful of functional studies suggests that they are crucial for olfactory discrimination and learning. Nonetheless, their activity patterns *in vivo* are poorly understood, and their electrophysiological properties in awake animals are completely unknown. This is because granule cells do not yield to conventional recording techniques. Former lab postdoc Dr. Heike Demmer therefore developed reliable methods for recording and labeling olfactory bulb granule cells. In a tour de force series of experiments, Watson School of Biological Sciences student Brittany Cazakoff, postdoctoral fellow Dr. Billy Lau, and CSHL Undergraduate Research Program student Kerensa Crump applied these methods to mice that were awake with their heads fixed, but running freely on a foam ball and receiving water rewards from a lick tube. Their recordings represent the first reported view of the activity of this important cell type during wakeful behavior. The team's results conclusively

show that granule cells become much more active during wakefulness and surprisingly uncouple from the animal's breathing and sniffing pattern. This seems to suggest that upon waking, granule cell activity becomes dominated by internal signals from the brain rather than the sensory properties of odors. If so, we have functionally identified a pathway central to the remarkably labile processing of odor stimuli. In April of this year, we published our first study of this system in *Nature Neuroscience* (Cazakoff et al. 2014).

Since then, Brittany Cazakoff has been extending her recordings of granule cells into mice that are not only awake, but actively learning new associations of odor stimuli with reward (water) and punishment (bitter taste). She has developed a behavioral paradigm in which mice learn that two new odors each signal either impending reward or the bitter taste, respectively. They are also asked to learn that the odor associations have switched. The mice show that they have learned by licking when reward is signaled and withholding licks when the odor portends bitter taste. This learning develops in a timeframe that allows us to monitor activity in individual granule cells throughout. Such recordings are ongoing.

### Noradrenaline Stores Olfactory Memories through Dynamic Regulation of Inhibition

H. Demmer, D. Eckmeier, B. Cazakoff [in collaboration with A. Koulakov, Cold Spring Harbor Laboratory]

How do we remember individuals that we have previously encountered? Substantial evidence indicates that many animals remember each other based on olfactory cues. Memories are especially strong for individuals encountered during key life events such as mating with a new partner or the birth of a litter of young. These important events typically evoke massive release of the neurochemical noradrenaline (NA), initiating a heightened state of emotion and arousal. This surge appears to cause long-lasting modifications to the responses to odorants in the olfactory bulb, which is the first processing station for scent in the mammalian brain. Indeed, it was hypothesized that the coincidence of an odor stimulus with a surge of NA is minimally sufficient to store a memory. We have been exploiting the intimate relationship between NA, olfactory bulb activity, and behavior to trigger and observe olfactory memories in mice. We discovered

that indeed, when NA release is evoked by stimulating locus coeruleus, the source of most NA, while the sleeping mouse sniffed an odorant, neural responses to that odor underwent specific long-term alterations. Remarkably, once awake, the mouse's subsequent behavior toward the odorant was also changed. In other words, the mouse seemed to remember the odor and treat it as though it were familiar. This finding has motivated several of our studies into how NA causes lasting changes to odor processing circuitry.

Using several experimental approaches, we are actively pursuing the following question: How are olfactory memories for individuals stored mechanistically among the specific synaptic connections of the various neuronal types in the olfactory bulb? Dr. Heike Demmer made recordings from granule cells to observe their activity during the induction of NA-dependent plasticity. She found that stimulation of NA release like that which occurs during social encounters suppresses granule cells. Suppression of granule cells transiently increases the excitability of mitral cells by relieving them from inhibition. This transient release of mitral cells from granule cell control is a key trigger for the synaptic plasticity that likely underlies memory. Indeed, this disinhibitory event triggers selective long-term changes to mitral cells that signal the presence of the learned odor. We have therefore pinpointed the synaptic connections between granule cells and mitral cells as an important synaptic substrate of individual recognition memories.

Dr. Dennis Eckmeier took a different approach to observe population mechanisms of NA-dependent memory formation with functional neural imaging. He used a technique in which the sensory neurons that provide input to the olfactory bulb are labeled with a fluorescent activity sensor that allows him to monitor the strength of activation in foci called "glomeruli." By comparing the response strength of each of these glomeruli before and after NA release, Dr. Eckmeier showed that NA weakens the response to paired odors. This effect is only observed in cases where NA is released during the odor by stimulation of a brainstem structure called locus coeruleus. These data establish the synaptic input to the mouse olfactory bulb as a target of noradrenergic modulation, thus revealing that the effects of learning and plasticity in the olfactory system extend to remarkably early sensory signaling events. We published these findings in November 2014 in *The Journal of Neuroscience* (Eckmeier and Shea 2014).

## Auditory Plasticity Is Impaired in a Mouse Model of Rett Syndrome

B. Lau, G. Ewall [in collaboration with J. Huang and K. Krishnan, Cold Spring Harbor Laboratory]

---

Far outside the range of our hearing, in the ultrasound range, mice are constantly holding conversations with one another in a language that is poorly understood at best. Many types of vocalizations are emitted by males and females, juveniles and adults, in a variety of behavioral contexts. We would like to better understand the perceptual significance of these calls to the mice and how they are used to guide behavioral choices. One form of vocalization that is reasonably well understood is the ultrasonic distress vocalization (USV). Young mice prior to vision and full mobility will occasionally become separated from the nest. This is stressful for them and they will therefore call out to their mother with a very high frequency peep. New mothers develop sensitivity to these cries and respond by moving toward their source (phonotaxis) to retrieve the pup. Moreover, females who have never given birth do not innately show approach responses to pups or their calls; however, they can learn to perform the behavior with experience.

Dr. Billy Lau is collaborating with CSHL professor Dr. Josh Huang and his postdoctoral fellow Dr. Keerthi Krishnan to examine how vocal perception of pup calls is affected in mice that are missing one copy of a gene called *MeCP2*. Impairments in the function of this gene are understood to cause the autism spectrum disorder Rett syndrome, and the Huang lab's findings regarding mutations in *MeCP2* suggest that they may affect inhibitory networks in the auditory cortex important for development of pup call sensitivity. Indeed, we find that females that possess only a single copy of *MeCP2* are not able to develop proficiency at gathering pups. This is an important finding because we describe the first robust behavioral phenotype to be observed in mice that most closely genetically resemble humans with Rett syndrome. It is also significant that this phenotype is a social communication deficit.

This year, we completed our initial study the effects of *MeCP2* mutation on auditory cortical circuitry and plasticity. Using a viral genetic strategy, we were able to knock down expression of *MeCP2* selectively in the auditory cortex of adult females and impaired retrieval behavior. This indicates that *MeCP2* has a critical role in maintenance and plasticity of the auditory cortex

in support of this maternal behavior. Further experiments revealed abnormalities in the response of auditory cortical inhibitory networks to maternal learning in mutant mice. In a particularly exciting set of experiments, we were able to repair inhibitory function with genetic and pharmacological manipulations, thereby restoring maternal gathering behavior. We are hopeful that these findings will open new therapeutic avenues for Rett syndrome.

## Neural Activity during Social Encounters

D. Eckmeier

---

We have begun experiments that are ultimately aimed at recording individual neurons during social encounters and other behavioral assays involving the perception of social and nonsocial information. There are two broad related goals to this approach. The first goal is to examine the encoding of social information such as body odors and vocalizations in primary sensory structures of awake animals. We hypothesize that activity in response to these signals may be labile to associative learning, attention, and arousal, which we may be able to manipulate in the context of social encounters. The second goal is to record from neurons in deep brain neuromodulatory centers during these encounters as well. Neurons that release noradrenaline and dopamine are likely responsive to social signals and may modulate encoding of sensory data and associative plasticity. Understanding the context-dependent activity patterns of these neurotransmitters is therefore critical to developing models for how they affect behavior.

## Olfactory and Auditory Interactions

M. Dinh, J. Shen

---

Social encounters are multisensory events that include smells and sounds. Increasingly, we are interested in neural mechanisms that integrate olfactory and auditory signals into a coherent perception of the social world. For example, how does the smell of a male mouse interact the sounds he makes in the brain of the female mouse? As a first step, this year, Michael Dinh and Jia Shen performed experiments in which they activated olfactory projections to the auditory cortex using optogenetic tools. The preliminary data from these experiments strongly suggest that regions

carrying social olfactory cues can modify responses to auditory stimuli in a complex manner.

#### PUBLICATIONS

- Cazakoff BN, Lau BY, Crump KC, Demmer HS, Shea SD. 2014. Broadly tuned and respiration-independent inhibition in the mouse olfactory bulb. *Nat Neurosci* **17**: 569–578.
- Eckmeier D, Shea SD. 2014. Noradrenergic plasticity of olfactory sensory neuron inputs to the main olfactory bulb. *J Neurosci* **34**: 15234–15248.
- In Press*
- Cazakoff BN, Shea SD. 2015. Simultaneous collection of *in vivo* functional and anatomical data from individual neurons in awake mice. In *Neural tracing methods* (ed. Arenkiel BR), Vol. 92, pp. 85–99. Springer, New York (in press).

# NEURAL CODING AND MEMORY FORMATION IN THE *DROSOPHILA* OLFACTORY SYSTEM

G. Turner M. Brill T. Hige M. Modi

The brain has a tremendous capacity to form different memories. These memories are, under normal circumstances, highly accurate for particular stimuli and situations. This is precisely the facility that is lost in diseases such as Alzheimer's and other dementias. Our overall goal is to understand how the brain forms such precise memories: How does the specificity arise, and how does the brain derive such a tremendous capacity for forming different memories?

We address these questions by studying olfactory memory formation in *Drosophila*. Just like Pavlov's dogs, *Drosophila* learn to form associations between smells and reward or punishment. A specific area of the fly brain, known as the mushroom body (MB), is essential for the flies to form olfactory memories. We are investigating how the neural activity patterns in this brain area are used to form specific olfactory memories. To achieve this, we monitor activity using both electrophysiological and functional imaging techniques. We have found that individual MB neurons exhibit highly odor-specific responses; at the population level, this translates into sparse activity patterns that are distinct for different odors. This specificity is thought to underlie the accuracy of memory; modifying the synapses of highly odor-specific neurons would lead to relatively precise memories. Sparse representations by highly stimulus-specific neurons are a general feature of brain areas involved in learning and memory, including the hippocampus and cerebellum in humans.

Using the simplicity and genetic manipulability of *Drosophila*, our goal is to understand several fundamental properties of neural responses in learning and memory-related brain areas: What mechanisms give rise to the stimulus specificity? What exactly is meaningful about the activity patterns? Is it simply which cells respond, or does the precise timing of activity matter? How are these response patterns modified by learning? And ultimately, what are the roles of the many genes implicated in learning on network level activity in the brain?

## Is the Timing of Spikes Important for the Neural Code in the Olfactory System?

M. Brill

---

Neuronal signals propagate from one cell to the next via electrical impulses, or spikes, that evoke neurotransmitter release. These trains of spikes are like the ones and zeros of the binary code of a computer. Spike trains have many different features that could potentially carry information: the rate of spikes, the timing of the onset, and the relative timing of firing across neurons. Although spike rate is widely recognized as important, other features of those spike trains more related to spike timing could also be important signaling factors. Understanding what features of these spike trains convey information is a fundamental aspect of understanding the brain. It is a little like finding the key to a computer's binary code.

We have begun to investigate this fundamental question by constructing artificial spike trains with these different features (spike rate vs. spike timing) under our experimental control. Using optogenetic techniques, we can manipulate the activity patterns at an early layer of the olfactory system. By then monitoring downstream layers of the circuit using *in vivo* imaging techniques, we can assess what features of those artificial spike trains are effectively transmitted to those downstream neurons, i.e., what features of the code are relevant. We are currently focusing on the importance of the onset of spiking activity. If downstream decoder neurons receive inhibition that is slow and takes time to develop, early-onset spikes could potentially be more effective at driving downstream circuitry than those arriving late. By combining this optogenetic control of spike trains with measures of the impact on behavior, we plan to reveal which spike train features are important for the perception of the animal, getting to the core features of the neural code in this system.

## Dopamine Controls the Signal-to-Noise Ratio of Transmission through the Mushroom Body

T. Hige, M. Modi

Neuromodulators such as dopamine and octopamine are essential for learning, but how they act and what their precise role is during the learning process have not been established. We examined the effects of these neuromodulators on activity in the MB. Surprisingly, we found that dopamine affects the signal-to-noise ratio of odor representations in the MB. Preliminary results suggest that dopamine increases the reliability of MB responses. Reliability is difficult to achieve in sparse representations because neurons typically respond with small numbers of spikes. Thus, it seems particularly useful for the circuit to use neuromodulators to increase signal-to-noise ratio depending on the behavioral context the animal is in; essentially dopamine could control the salience of a stimulus.

Additionally, we have shown that dopamine diminishes odor responses in the neurons that are downstream of the MB, the MB output neurons. Together, these results suggest a model where the dopamine acts to make the circuit a selective filter, increasing the signal-to-noise ratio of responses in the MB while decreasing signal transmission through the MB. These results may be an excellent entry point for studying the fundamental cognitive process of attention, which is essentially a process of filtering out distracting information to focus on pertinent inputs. We are now testing this possibility and investigating the precise mechanisms of the dopamine effects.

## Mushroom Body Output Neurons and Flexible Odor Categorization

T. Hige

Almost no studies have examined what happens to sensory information downstream from areas with sparse representations. So what happens to odor information after it leaves the MB? In collaboration with Gerry Rubin's lab at the Janelia Research Campus, we have functionally characterized odor coding in the complete set of 34 MB output neurons (MBONs). The following are our main findings.

1. Synaptic plasticity diversifies MBON tuning properties. We found that uniquely identifiable MBONs have strikingly different odor response properties

across different flies. To test whether these individual differences are experience-dependent, we examined odor tuning of the same neuron across the two hemispheres of the brain. Identical MBONs in the two hemispheres have identical odor tuning properties; however, the tuning of that neuron is very different across different individual flies. This result suggested that MBON tuning in each fly is partly a product of that individual's developmental and experimental history. We tested this by showing that a central gene in learning and memory, *rutabaga*, is required for these individual differences to arise. This experiment is one of the first to show that the odor tuning properties of MBONs are shaped by synaptic plasticity.

2. MBONs represent categories/groups of odors. We examined odor responses across the entire population of MBONs. We found that MBONs do a poor job of representing odor identity, especially in comparison to the upstream neurons that constitute the MB itself. This is because certain odors evoke extremely similar patterns of response across the MBON population. In fact, only three main groups of odors were identified based on the pattern of activity in the MBONs. Interestingly, one group was composed of odors that are repellent to flies, whereas another group contains food-based odors. This is an exciting result because few investigators have considered how different sensory inputs can be categorized into behaviorally meaningful classes. We may be seeing olfactory information shaped into behaviorally relevant neuronal signals at this layer, finding the connection between sensory input and motor output

These two results indicate that MBON odor responses are highly plastic and that the MBON response pattern signals the positive or negative quality of an odor. Together, these findings suggest that the MB represents odors categorically and that experience could potentially shift a neutral odor into either a positive or negative category, depending on the fly's past experience with that odor. This work has now been through two rounds of reviews at *Nature* and the outlook is positive, although it has not yet been accepted.

## PUBLICATIONS

Campbell RA, Eifert RW, Turner GC. 2014. OpenStage: A low-cost motorized microscope stage with sub-micron positioning accuracy. *PLoS ONE* 9: e88977.

# CIRCUITRY UNDERLYING CORTICAL PROCESSING AND DECISION MAKING

A. Zador   B. Burbach   S. Koh   A. Reid  
D. Gizatullina   U. Livneh   V. Vasily  
S. Ghosh   F. Marbach   Q. Xiong  
S. Jaramillo   H. Oyibo   P. Znamenskiy  
J. Kebschull   I. Peikon

My laboratory is interested in how neural circuits underlie cognition, with a focus on the cortical circuits underlying auditory processing and decision making. To address these questions, we use a combination of molecular, electrophysiological, imaging, and behavioral approaches. In addition, we are developing a method that will allow us to use high-throughput DNA sequencing to determine the wiring diagram of the brain at single-neuron resolution rapidly and efficiently, at low cost.

## Role of Corticostriatal Plasticity in Auditory Decisions

Q. Xiong, P. Znamenskiy

---

Corticostriatal plasticity has a key role in reinforcement learning, but how associations between stimuli and motor responses are established remains unclear. Recent work from our group demonstrated a causal role of corticostriatal neurons in driving choices during an auditory discrimination task, inspired by the classic random dot motion task used by Newsome and colleagues in macaques, in which subjects were required to choose the left or the right response port depending on the perceived frequency of a sound. We are testing the hypothesis that changes in the strength of corticostriatal synapses underlie the association between sound and action required to perform this task. Our results indicate that changes in the strength of a specific subset of corticostriatal synapses encode the arbitrary association between stimulus and motor response. Because all sensory cortical areas send projections to the striatum, our findings suggest a general mechanism for the formation of arbitrary sensorimotor transformations.

## In Vivo Two-Photon Imaging of Cortical Activity in Mice Performing an Auditory Discrimination Task

U. Livneh, F. Marbach

---

We are studying how activity in the auditory cortex is used to guide behavior. To address this, we have developed a suitable behavioral paradigm in mice that is compatible with two-photon imaging of behaviorally evoked neural activity. This task is modified from one we previously used in freely moving rats (Znamenskiy and Zador, *Nature* 497: 482 [2013]). This paradigm will allow us to identify the pathways important for the flexible association of sensory stimuli with desired actions, a critical function of higher nervous systems.

## Sequencing the Connectome

I. Peikon, D. Gizatullina, J. Kebschull, H. Oyibo, V. Vasily

---

We are developing an entirely novel approach to mapping neural connectivity that harnesses the high-throughput power of next-generation sequencing. We have engineered a suite of molecular tools to barcode individual neuronal connections (BOINC) for read-out via DNA sequencing. By mapping neural connectivity onto DNA sequencing, we have rendered neural circuit reconstruction tractable with current techniques.

## Sequencing the Projectome

J. Kebschull, I. Peikon, D. Gizatullina

---

Understanding the brain requires knowledge of its anatomy. Recently, much progress has been made in

mapping out “mesoscale” (area-to-area) connectivity in the mouse brain. However, most current approaches rely solely on bulk mapping of projections and cannot resolve the projection patterns of single neurons; techniques with the resolution needed to resolve single-neuron connectivity suffer from very low throughput. We are developing a high-throughput, multiplexable method to determine long-range projections in the mouse brain at single-cell resolution. To do so, we are leveraging the vast combinatorial space of nucleic acid barcodes to uniquely label individual neurons. We then use the power of modern high-throughput sequencing technology to reliably, quickly, and inexpensively read out these labels at projection target sites.

### Proximity Ligation Assay for Synapse Detection

J. Kebschull, D. Gizatullina

---

We are developing a novel method to label synapses between two genetically identified neural populations based on the proximity ligation assay (PLA). PLA uses antibodies to detect the close (<40-nm) apposition of two proteins. By targeting a pre- and postsynaptic protein, we can identify synapses with high reliability. This method has the potential to multiplex synapse detection to several pathways through labeling with orthogonal proximity probes.

### Role of Corticostriatal Plasticity in Stimulus-Motor Association

S. Ghosh

---

Previous work from our lab has shown that corticostriatal projections have a causal role in decision making in a complex auditory discrimination task (Znamenskiy and Zador, *Nature* 497: 482 [2013]; Xiong et al. 2015). Moreover, learning of this discriminatory task results in a specific pattern of plasticity in this pathway, which is determined by the frequency tuning of the neurons and the corresponding motor movement. Nevertheless, how this pattern of plasticity overlays onto the complex striatal circuitry remains unknown. We are therefore testing the hypothesis that differential patterns of plasticity in specific subtypes of corticostriatal synapses underlie the ability to associate a

particular auditory cue with a given motor action in the “tone cloud” task.

### Circuits Underlying Auditory Representations and Decisions

S. Jaramillo, U. Livneh, F. Marbach

---

We are studying the circuits underlying auditory representations and how these representations give rise to decisions. We train rats and mice to perform well-controlled auditory tasks and then use electrical and optogenetic methods to monitor and perturb neural activity. We use this approach to study the role of the auditory thalamus, cortex, and striatum in these behavioral processes. Using tasks in which the stimulus-action associations vary within a behavioral session, we have uncovered striking similarities in how thalamic and cortical neurons are modulated by the animals' choices.

### Interhemispheric Connections in the Auditory Cortex

S. Koh

---

The callosal projection is a long-range connection between the left and right auditory cortex, and it is one of the major corticocortical projections in the auditory cortex. However, its role in auditory coding is unclear. We are studying how manipulating the callosal projection affects auditory representations in the cortex. We found that silencing the auditory cortex in one hemisphere suppresses sound-evoked activity in the other hemisphere. The suppression effect was stronger for late response with latency longer than 40 msec and also for the response to sound coming from ipsilateral locations. Our results suggest that the callosal projection is well suited for slow interhemispheric integration of location-relevant information.

### Using Rci Recombinase to Generate Cellular Barcodes In Vivo

I. Peikon, D. Gizatullina

---

Heterogeneity is a ubiquitous feature of biological systems. A complete understanding of such systems



requires a method for uniquely identifying and tracking individual components and their interactions with each other. We have developed a novel method of uniquely tagging individual cells in vivo with a genetic “barcode” that can be recovered by DNA sequencing. Our method is a two-component system composed of a genetic barcode cassette whose fragments are shuffled by Rci, a site-specific DNA invertase. The system is highly scalable, with the potential to generate theoretical diversities in the billions. Currently, this method could be used to track the dynamics of populations of microbes through various bottlenecks. Advances of this method should prove useful in tracking interactions of cells within a network and/or heterogeneity within complex biological samples. The barcoding method developed here represents a major first step toward our goal of uncovering the connectivity of neural circuits using DNA sequencing.

## Role of Inhibitory Interneurons in Auditory Cortex Function

A. Reid

---

Fast synaptic inputs to neurons in the auditory cortex are either inhibitory or excitatory. Cortical interneurons are tremendously diverse. One inhibitory interneuron subclass, defined molecularly by the expression of parvalbumin (“PV+”), seems to be ideally positioned to mediate the fast component of the characteristic barrage of inhibition elicited by a sound. We are testing the hypothesis that PV+ inhibitory interneurons mediate fast sound-evoked

inhibitory synaptic currents in the auditory cortex. We seek to establish a causal link between a physiological property—the fast sound-evoked inhibition that contributes to receptive field dynamics—and a component of the underlying cortical circuitry. We approach the problem at three different levels, from brain slices through in vivo physiology to behavior. Although we are currently focusing on the role of one particular interneuron subclass (PV+), our approach combining electrophysiological and molecular tools can readily be generalized to other subclasses and can be extended to probe the circuitry underlying other behavior- and sensory-elicited neuronal responses.

## PUBLICATIONS

- Jaramillo S, Zador AM. 2014. Mice and rats achieve similar levels of performance in an adaptive decision-making task. *Front Syst Neurosci* **8**: 173.
- Jaramillo S, Borges K, Zador AM. 2014. Auditory thalamus and auditory cortex are equally modulated by context during flexible categorization of sounds. *J Neurosci* **34**: 5291–5301.
- Livneh U, Zador A. 2014. Sensory systems: Sound processing takes motor control. *Nature* **513**: 180–181.
- Oyibo HK, Znamenskiy P, Oviedo HV, Enquist LW, Zador AM. 2014. Long-term Cre-mediated retrograde tagging of neurons using a novel recombinant pseudorabies virus. *Front Neuroanat* **8**: 86.
- Peikon ID, Gizatullina DI, Zador AM. 2014. In vivo generation of DNA sequence diversity for cellular barcoding. *Nucleic Acids Res* **42**: e127.
- In Press*
- Xiong Q, Znamenskiy P, Zador A. 2015. Selective corticostriatal plasticity during acquisition of an auditory discrimination task. *Nature* (in press).

# NEURAL BASIS OF BEHAVIOR IN *DROSOPHILA*

Y. Zhong J. Beshel Y. Shuai M. Zhang

## Multiple Neural Pathways Mediate Active Forgetting in *Drosophila*

Y. Shuai

---

Evidence exists that short-term memories are removed by active forgetting processes, which involve the activation of signal transduction pathways independent of those for learning. To investigate the neural basis of such active forgetting, we performed a neural-silencing screen targeting extrinsic neurons connected to the *Drosophila* olfactory memory site, the mushroom body (MB). We identified two populations of neurons that are not involved in learning, but that function in parallel to drive forgetting (i.e., stimulating their activity induces memory loss), whereas silencing their functions preserves labile memory. These neurons include a cluster of dopamine neurons termed PAM- $\beta$ 1 and a pair of glutamate neurons termed MBON- $\gamma$ 4 >  $\gamma$ 1 $\gamma$ 2. The identified neurons are required for time-dependent memory decay specifically, but not for acute forgetting during reversal learning. The differential requirement suggests the presence of diversified forgetting circuits that may be recruited in a context-dependent manner.

## Mapping the Food-Odor Circuit Anatomically and Functionally

J. Beshel

---

We have previously reported that the range of attraction flies display to food odors can be accounted for by looking at the odor-evoked activity of neurons immunopositive for the *Drosophila* homolog of mammalian neuropeptide Y or *Drosophila* neuropeptide F (dNPF). These dNPF neurons are both necessary and sufficient to promote food-odor-attraction behavior. They reside within the central brain and demonstrate the possibility of one of perhaps several discrete brain sites that explicitly encodes odor value. We continue to map the food-odor circuit anatomically and functionally and have made significant progress in the isolation of single dNPF-positive neurons using multiple genetically based techniques. Single-cell isolation considered in combination with visualization of pre- and postsynaptic marker proteins has allowed us to further refine proposed dNPF-upstream and -downstream brain regions. We continue to study loss/gain-of-function phenotypes through a combination of RNA interference and acute genetic manipulation on neuronal subsets determined to be part of the dNPF circuitry.

# 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 critical for stem cell maintenance, and studies in this past year have identified additional factors that are required for transport of this important regulator. The lab also continues to identify additional 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 class of receptors completely different from those in animals. Their discovery helps to explain how signaling from diverse receptors is achieved in plants. This year, they also discovered a fundamentally new mechanism of stem cell control that involves signaling from differentiating cells back to the stem cell niche. Several studies in plants and animals have hinted at such a mechanism, but their work has for the first time identified the receptor and signals involved. They are applying this work to agriculture by finding weak mutations in some of these genes and are finding in some cases that they can enhance seed production in maize, which could lead to yield increases. Separately, the lab has characterized system-wide networks of gene expression, using “next-gen” profiling and chromatin immunoprecipitation methods that have revealed 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 focuses on the process of flowering and flower production, which is a major contributor to plant reproductive success and agricultural yield. By identifying genes that control how tomato plants produce their flowers in their characteristic repeated zigzag arrangement (e.g., tomatoes on a vine), Lippman's lab is addressing when and how flowering branches known as inflorescences develop on plants, particularly fruit-bearing plants. Of particular interest is how these “reproductive phase transitions” have contributed to the evolution of diverse inflorescence branching patterns in tomato's larger *Solanaceae* family, which includes plants that make just one flower, such as pepper and petunia, in each inflorescence, to plants whose inflorescences produce dozens of branches and hundreds of flowers, such as many wild species of tomato. Using a combination of genetic, genomic, and molecular approaches, Lippman is dissecting the gene networks that are responsible for the variation in inflorescence branching found in nature. He hopes to leverage these discoveries to improve crop yields. Already, his work on genes that are responsible for the production and activity of a universal flowering hormone known as florigen has resulted in novel approaches to fine-tune plant architecture and flower production, boosting yield beyond leading commercial varieties. To continue hunting for new genes, Lippman has adopted a systems-biology approach and next-generation sequencing technology to capture those genes that are active as stem cells mature from a vegetative to a reproductive state. Nearly 4000 genes were found to reflect the existence of a “maturation clock,” and one of the clock genes known as *Terminating Flower* acts as a key regulator to maintain a progressive pace to flowering—which in turn dictates how many flowers are produced on each tomato inflorescence. Finally, the Lippman

lab determined the genome sequence of the “currant tomato,” the wild ancestor of larger-fruited cultivated tomatoes, in order to better understand how flower and fruit production changed during the process of crop domestication.

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 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. With collaborators in Mexico, Martienssen has also coaxed *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 lab has also showed 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. This year, as part of a collaborative project to generate a high-quality full genome map of the oil palm plant, Martienssen and his colleagues identified a single gene that 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 growing tips of plants, called meristems, contain a population of stem cells that serve as a persistent source of daughter cells from which new organs, such as leaves, arise. **Marja Timmermans** and colleagues are studying the genetic networks that regulate plant stem cell activity. Using genomic approaches, they have defined gene expression signatures that distinguish indeterminate stem cells from their differentiating derivatives. They have also worked out the mechanism that suppresses stem-cell fate to allow cells to differentiate and have shown that this process requires a highly conserved epigenetic gene silencing mechanism. In particular, Timmermans’ group has shown that specific DNA-binding proteins mediate the recruitment of Polycomb repressive complexes to stem cell factors, an action that stably represses their expression in differentiating organs. This work addresses a major unresolved question in the field of epigenetics: how Polycomb proteins, which do not bind DNA themselves, recognize defined targets. Plant stem cells also produce signals important for the patterning of lateral organs. The lab has discovered that small RNAs can traffic from cell to cell and are among the stem-cell-derived signals. They have found that polarity in leaves is established via opposing gradients of mobile small RNAs that act as morphogen-like signals. Their most recent findings identified a third small RNA gradient involved in maintenance of organ polarity. These findings illustrate the complexity with which small RNAs generate developmental patterns. Currently, they are investigating parameters of small RNA mobility and the unique patterning properties of resulting small RNA gradients. Mathematical modeling predicts that gradients might serve to generate robustness during development.

# DEVELOPMENTAL BIOLOGY: STEM CELLS, SIGNALING, AND CROP PLANT ARCHITECTURE

D. Jackson	R. Balkunde	M. Fuchs	M. Shen
	H.T. Bui	B.I. Je	K. Vera
	H. Claeys	T. Lau	Q. Wu
	E. Demesa-Arevalo	I. Liao	T. Zadrozny
	A. Eveland	M. Rutigliano	

The ultimate goal of our research is to improve crop plant growth and yield. More specifically, we aim to identify genes, signals, and pathways that regulate plant architecture and development. All higher organisms develop by carefully controlling the flow of information that passes between cells and tissues. We are interested in discovering the signals that carry this information, in finding out how the signals are transmitted, and in determining how they function.

A major focus has been identification of genes that control development through effects on stem cell maintenance and identity. Examples from the past year include discovery of a redox enzyme that controls shoot architecture and its interacting partner, a DNA-binding transcription factor. We used “next-gen” profiling methods to identify system-wide networks of gene expression in inflorescence development, and the genes that function downstream from the redox enzyme transcription factor complex. We also continue to characterize the role of G proteins in stem cells, identifying candidate upstream receptors, as well as mutations in partner proteins, using CRISPR. Our discoveries help explain how signaling from diverse receptors is achieved in plants. This work could help improve crop yields, because we previously demonstrated that weak mutations in one of the receptor genes enhance seed production in maize.

In another project, we 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. We are using genetic screens to identify novel factors that control the transport of transcription factors between cells in the plant stem cell niche. Finally, we also made a major contribution to the maize genetics toolbox by creating a collection of lines that can drive expression of any

reporter or experimental gene in any tissue type. Such tools are of great interest to researchers and are being made available to the broader scientific community, enabling experiments never before possible in crop plants.

## Regulation of Cell-To-Cell Trafficking of the KNOTTED1 Transcription Factor

R. Balkunde, H. Bui, K. Vera, M. Shen

---

Cell-to-cell communication functions to specify cell fate and coordinate development in all multicellular organisms. However, unlike animal cells, plant cells have cell walls that could limit exchange of signal molecules between neighboring cells. Plants have therefore developed direct connections using special structures called plasmodesmata (PD). These channels allow the selective trafficking of signaling macromolecules, such as transcription factors and RNAs, between cells. In addition to cell fate specification, they are involved in viral movement, transport of metabolites, and cell-to-cell spread of RNA interference (RNAi), which points to their fundamental importance in coordinating plant defense, metabolism, and development. Despite the discovery of PD more than 100 years ago, the underlying components and mechanisms of PD trafficking remain poorly understood. Hence, we are implementing an unbiased genetic strategy to dissect these molecular mechanisms, using a reporter system involving *Arabidopsis* leaf trichomes as an easily scorable phenotype and the developmentally important mobile proteins KNOTTED1 (KN1) from maize and its *Arabidopsis* homolog SHOOTMERISTEMLESS (STM).

The trichome rescue system designed by our lab is widely used to characterize protein mobility. Briefly, GLABRA1 (GL1) protein is cell-autonomous and is required for trichome formation in the leaf epidermis.

Therefore, *GL1*, when expressed in mesophyll in *gl1* mutants, does not rescue trichomes. However, when fused to KN1 (GL1-KN1), it now results in trichome rescue, as the mobile KN1 brings GL1 into the epidermis. Ethylmethane sulfate (EMS) mutagenesis screening using this system, followed by Illumina high-throughput sequencing, resulted in identification of a mutation in a gene encoding CCT8 (chaperone containing tailless-complex 8). This led to the exciting proposal that chaperonins facilitate cell-to-cell trafficking and stem cell maintenance in *Arabidopsis*, supporting the functional relevance of chaperonin-mediated trafficking through PD. These results highlight the importance of protein conformational changes for PD trafficking.

Current efforts are focused on understanding the mechanism of CCT8 action and on identifying new trafficking regulators. Our efforts have identified mutants with interesting developmental defects, in addition to loss of trichome rescue. For example, mutant 1058 has leaf shape and floral organ defects, and mutant 1066 has narrow and elongated leaves, enlarged vegetative meristems, and fasciated shoots. Using a combination of marker-based mapping and Illumina high-throughput sequencing, we mapped these mutations to candidate genes. Next, using transformable bacteria artificial chromosomes (TACs) for complementation, we identified the causal mutations in *NUCLEOSTEMIN* (*NSNI*) and *ASPARTYL tRNA SYNTHETASE* (*AspRS*) genes for mutants 1058 and 1066, respectively. We further confirmed these genes by crossing to T-DNA knockout lines. *NSNI* encodes a GTPase that localizes to the nucleolus and nucleoplasm. The exact molecular function of *NSNI* is not clear; however, studies in mammalian cells suggest that *NSNI* is required for cell proliferation, rRNA biogenesis, and ribosome assembly, and its GTPase activity is required for trafficking from the nucleoplasm to the nucleolus. Interestingly, the *Drosophila* *NSNI* homolog was shown to function non-cell-autonomously. The second candidate, *AspRS*, is a nuclear gene, but the protein is targeted to the chloroplast and mitochondria. Aminoacyl-tRNA synthetases (AARSs) catalyze the attachment of an amino acid to its cognate transfer RNA molecule. AARSs have also been shown to be required for the gametogenesis and embryo development and for organ patterning. An example of a non-cell-autonomous function for tRNA synthetases was shown for the *SYCO* gene, which encodes a cysteinyl tRNA synthetase. During female gametophyte development

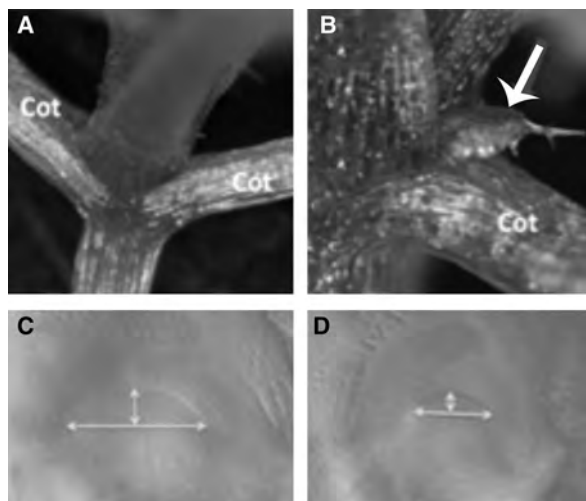
in *Arabidopsis*, *SYCO* is expressed and is localized to the mitochondria of central gametic cells, but it is absent in the antipodal cells. However, it controls the life span of antipodal cells, by nonautonomous signaling.

Interestingly, both *NSNI* and *AspRS* proteins have been reported in the PD proteome, suggesting that these proteins may have a direct role in PD regulation and protein trafficking. Currently, we are investigating protein localization to PD using fusions to mCherry or YPET. We will next perform a detailed characterization of the mechanism by which these candidates regulate protein trafficking. In an effort to identify additional regulators of PDs, we are also screening additional EMS-mutagenized M2 families.

## Role of SHOOTMERISTEMLESS Trafficking in the Shoot Apical Meristem

R. Balkunde

*STM* encodes a homeodomain transcription factor and is required for the maintenance of pluripotent stem cells in the shoot. *STM* is expressed throughout the meristem and is down-regulated in developing organ primordia. *STM* protein moves between cells, but the significance of this is not well understood. A common approach to prevent cell-to-cell movement is to add nuclear localization sequences (NLSs) to sequester the protein in the nucleus. We made an *STM* native expression construct with yellow fluorescent protein (YFP) fused at the amino terminus, and a second construct that also contained 2xNLS sequences at the amino terminus, to promote nuclear import. We found that the native pSTM::YFP-*STM* construct fully rescued *stm* phenotypes; however, the pSTM::2xNLS-YFP-*STM* construct did not. The 2xNLS-YFP-*STM* plants had axillary shoots in the axils of the cotyledons and smaller SAMs (shoot apical meristems) with fewer YFP-*STM* expressing cells (Fig. 1). We interpret the results as evidence that *STM* trafficking is important for proper SAM development. For example, having a smaller or “less potent” SAM could lead to axillary bud formation in the axils of cotyledons by a loss of apical dominance. This phenotype also suggested a possible defect in defining a correct organ boundary. It is known that *STM* and organ boundary genes *CUP SHAPED COTYLEDON1/2*/*(CUC1/2)* reinforce each other, and thus we are checking the expression of boundary genes. Subtle



**Figure 1.** Blocking STM movement causes SAM phenotypes. (A) *stm11* seedling complemented by pSTM::YFP-gSTM. (B) *stm11* seedling complemented by pSTM::2xNLS-YFP-gSTM, showing abnormal axillary shoot in the axils of cotyledon. These seedlings also have a smaller SAM (D, compared to control, C).

differences in the localization of STM and 2xNLS-STM at the organ boundary region may influence the levels of boundary gene expression, which may explain the developmental differences between the lines expressing STM and those expressing 2xNLS-STM.

### Generation of a pOp-LhG4 *trans*-Activation System for High-Throughput Protein Localization and Functional Studies in Maize

E. Demesa-Arevalo, Q. Wu, T. Zadrozny [in collaboration with A. Chan, J. Craig Venter Institute; A. Sylvester, University of Wyoming]

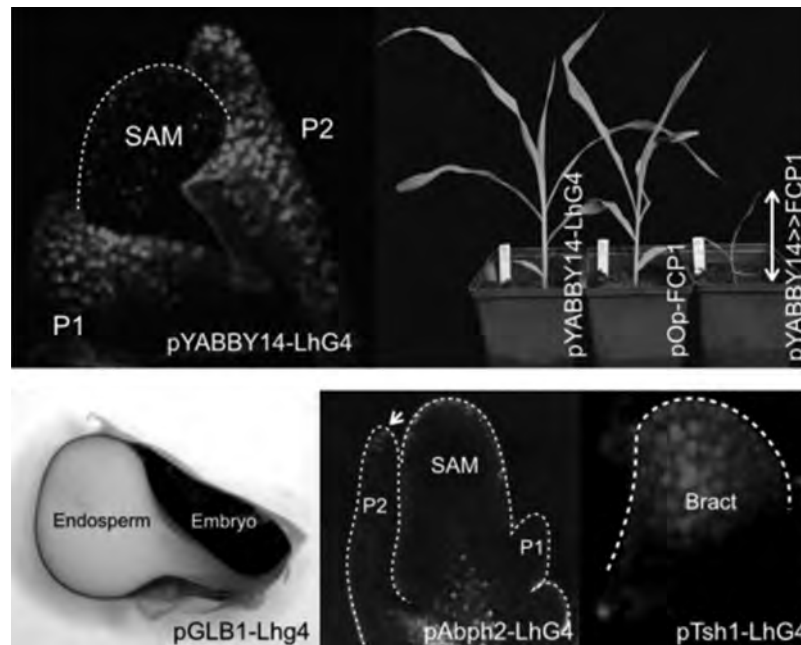
Despite the growing abundance of resources available for the study of various model organisms, the classical approach of characterizing a single gene mutant still provides invaluable information regarding gene function. We previously generated transgenic lines by tagging various genes with fluorescent markers, using their endogenous promoters to provide visual information with respect to time and location of expression as well as the ability to study the gene products at the protein level. Using our experience in promoter analysis, we are currently establishing tools for complex analysis at cellular, tissue, or organ level. We have developed a pOp-LhG4 *trans*-activation system in maize that allows us to express proteins in specific tissues and/or transient developmental stages. The

pOp-LhG4 system includes (1) a transcription activator, LhG4, which is a fusion between a high-affinity DNA-binding mutant of *lac* repressor, *LacI*<sup>His17</sup>, and transcription-activation domain II of GAL4 from *Saccharomyces cerevisiae*, and (2) a chimeric promoter, pOp, that consists of *lac* operators cloned upstream of a minimal CaMV (cauliflower mosaic virus) promoter, not activated in the reporter lines until crossed with the LhG4 activator line. We have already generated several LhG4 drivers, as well as four pOp responder lines, including *Zea mays* FON2-LIKE CLE PROTEIN1 (pOp::FCP1) and the FLOWERING LOCUS T-like *Zea mays* CENTRORADIALIS 8 (pOp::ZCN8). By crossing pOp::ZCN8 plants with a constitutive promoter line, pEF1A::LhG4, we found that neither the driver plants, pEF1A::LhG4, nor the responder plants, pOp::ZCN8-YFP, showed an early flowering phenotype; whereas the F<sub>1</sub> plants containing both pEF1A::LhG4 and pOp::ZCN8-YFP showed early flowering. We are now analyzing additional tissue-specific LhG4 driver lines expressing pOp::FCP1. pYABBY14-LhG4 is expressed in young leaf primordia, but not in the SAM. Overexpression of ZmFCP1 specifically in the leaf using this promoter reduces seedling size, possibly by affecting meristem size non-cell-autonomously (Fig. 2, upper panel). Currently, 20 LhG4 driver constructs have been generated, and our goal is to produce a comprehensive array of cell and tissue-specific lines (Fig 2, lower panel), which we will use to (1) isolate specific cell types by fluorescence-activated cell sorting (FACS) for transcriptional profiling and (2) drive the misexpression of developmentally regulated genes in order to better understand the function of stem cell niches during development. Data on the characterization of our FP and LhG4 transgenic lines, including confocal micrographs, movies, and recent publications, can be found on our website, <http://maize.jcvi.org/cellgenomics>.

### The Regulation of Meristem Size in Maize

B.I. Je, Q. Wu, M. Fuchs, A. Eveland [in collaboration with M. Komatsu and H. Sakai, Dupont Crop Genetics]

All plant organs derive from populations of stem cells called meristems. These stem cells have two purposes: to divide to replace themselves, and to give rise to daughter cells, which will differentiate into lateral organs. Consequently, meristems must precisely control the size of the



**Figure 2.** A maize *trans*-activation system. (*Upper panel*) pYABBY14-LhG4//pOp-TagRFpT is specifically expressed in leaf primordia (P, white dots = nuclei), plants containing both constructs pYABBY14-LhG4 and pOp-ZmFCP1 are reduced in size, but the seedlings containing only pYABBY14-LhG4 or pOp-FCP1 were normal. (*Lower panel*) pGLB1-LhG4//pOp-GUS, pAbph2-LhG4//pOp-TagRFpT and pTsh1-LhG4//pOp-TagRFpT are specific promoter drivers for embryo, young leaf (arrow) and young bracts, respectively.

stem cell niche via a network of positive and negative feedback signals. A loss of function in a negative regulator of stem cell fate can result in an enlarged or fasciated meristem phenotype and a dramatic alteration in the morphology of the maize ear and tassel.

Maize is an excellent genetic model system because of a large collection of developmental mutants and a sequenced reference genome. Our lab has undertaken a forward genetic approach to identify key regulators of stem cell homeostasis and meristem size. Two previously cloned mutants, *fasciated ear2* and *thick-tassel dwarf1*, encode orthologs of the *Arabidopsis thaliana* genes *CLAVATA1* and *CLAVATA2*, indicating that the well-known *CLAVATA*-*WUSCHEL* regulatory feedback loop is conserved from dicots to monocots. However, little else is known about the control of this important developmental process in maize. Here, we describe our progress in identifying additional genes contributing to stem cell niche homeostasis.

Recently, we have shown that the maize *COMPACT PLANT2* (*CT2*) gene, which encodes the predicted  $\alpha$  subunit of a heterotrimeric GTPase ( $G\alpha$ ), functions in the *CLAVATA* pathway to control meristem size through its interaction with *FEA2*. Genetic

data indicate that *ct2* and *fea2* are in a common pathway, and co-immunoprecipitation (Co-IP) and gel-filtration data further suggest that *FEA2* and *CT2* are in the same complex. Interestingly, our Co-IP experiments in *Nicotiana benthamiana* suggested that scrambling the intracellular tail of the *FEA2* receptor did not affect the *FEA2*-*CT2* interaction, suggesting their interaction may require additional mediators. Furthermore, fluorescence resonance energy transfer (FRET) and bimolecular fluorescence complementation (BiFC) data also suggested that *FEA2* and *CT2* do not directly associate with each other and that other proteins may bridge their interaction. To find the bridging protein(s), we used IP-mass spectrometry to find proteins that can be pulled down by both *FEA2* and *CT2*. We identified an uncharacterized leucine-rich repeat receptor-like kinase (LRR-RLK) that interacted with both *FEA2* and *CT2*, and these results were confirmed by Co-IP experiments using a tobacco transient expression system. Furthermore, BiFC results suggested that the LRR-RLK protein directly associates with *CT2*, but a kinase domain truncated version did not, suggesting that the kinase domain is required for this interaction. This LRR-RLK



is preferentially expressed in meristems, and we are now searching for phenotypes to illuminate its biological function in meristem development.

To further explore G-protein signaling, we generated a constitutively active form (CA-CT2), using a point mutation to abolish the GTPase activity, and tagged this protein with mTFP1. We are introgressing CA-CT2-mTFP1 into *ct2* and *fea2* mutant backgrounds to study how constitutive G-protein signaling affects maize development. We also knocked out the only G $\beta$  subunit of maize using CRISPR-Cas genome editing and will study the role of G $\beta$  in meristem regulation. All of these studies will facilitate our understanding of G-protein signaling in maize development.

Another fasciated ear mutant that we have cloned is *fasciated ear 3* (*fea3*), which was derived from irradiation mutagenesis. *fea3* shows an overproliferation of the inflorescence meristems, leading to enlarged or fasciated ears and thick tassels. Double mutants of *fea3* with *td1* or *fea2* have additive and synergistic fasciated phenotypes, indicating that they act in independent pathways that converge on the same downstream target to control meristem size. We cloned the *fea3* gene using a map-based cloning approach, and the mutant results from an insertion of a partial retrotransposon into an exon of the *fea3* locus. We confirmed this identity by isolation of three additional alleles of *fea3* from a targeted EMS mutagenesis. *FEA3* encodes a predicted leucine-rich repeat receptor-like protein related to *FEA2*. In-situ hybridization and imaging of red fluorescent protein (RFP)-tagged transgenic plants show that *FEA3* is expressed in the organizing center of the SAM and in leaf primordia (Fig. 3), as well as in the root apical meristem. Remarkably, expression of maize *WUSCHEL*, a marker for the stem cell niche

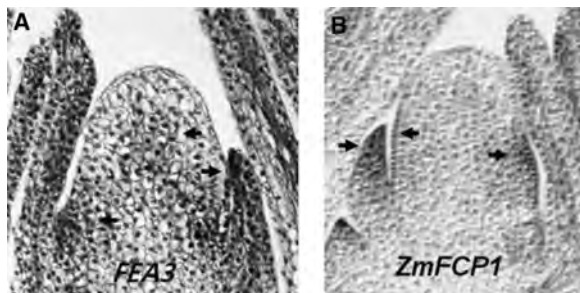
organizing cells, spreads downward in *fea3* mutants, which is strikingly different from its response in the known *CLAVATA* stem cell mutants.

*FEA3* is localized in the plasma membrane, and to see if it responds to a CLV3-related (CLE) peptide, we tested its sensitivity to different peptides. *fea3* mutants showed reduced peptide sensitivity, but interestingly, they responded to a different CLE peptide, ZmFCP1 (FLORAL ORGAN NUMBER2-like CLE protein 1), compared to *FEA2*. ZmFCP1 is expressed in leaf primordia (Fig. 3), suggesting that it serves as a signal from differentiating cells to repress meristem growth via the *FEA3* receptor. Our results suggest that the *FEA3* receptor function is in a new pathway for stem cell control that is spatially distinct from the known CLV receptors and uses a different peptide signal. This signaling system appears to be universal, as we have also found orthologs in *Arabidopsis* with similar properties.

In addition to these advances, we are in the process of mapping additional fasciated ear mutants from EMS mutagenesis screens. We have determined candidate map positions for several of these mutants and are proceeding with fine mapping and molecular cloning.

## Control of Branching and Determinacy in Plant Shoots

H. Claeys, E. Demesa-Arevalo, T. Zadrozny, A. Eveland  
[in collaboration with M. Komatsu and H. Sakai, Dupont  
Crop Genetics]



**Figure 3.** In-situ hybridization of *FEA3* and *ZmFCP1*. *FEA3* is expressed in the organizing center of the SAM and in leaf primordia (A). *ZmFCP1* is expressed in the leaf primordia (B).

The *RAMOSA* (*RA*) genes in maize function to impose determinacy on axillary meristem growth; consequently, *ra* loss-of-function mutants (*ra1*, *ra2*, and *ra3*) have more highly branched inflorescences. *RA3* encodes a trehalose phosphate phosphatase, which catalyzes the conversion of trehalose-6-phosphate (T6P) to trehalose. T6P was shown to be an important regulatory metabolite that connects sucrose levels, and thus the sugar status, to plant growth and development, but its mode of action is still unclear. *RA3* is expressed in a localized domain at the base of axillary inflorescence meristems, and it localizes to distinct puncta in both nuclear and cytoplasmic compartments, suggesting that its effect on development may not be simply metabolic. The *RAMOSA* genes define a boundary domain that surrounds the developing branch meristem, rather than being deployed in the meristem itself. These data support



**Figure 4.** SEM of young ears of B73, the *ra3* mutant in a B73 background, and *ra3* combined with a *rel3* mutant allele. B73 ears do not have branches, whereas *ra3* ears form branches at the base of the ear. In contrast, *ra3; rel3* ears have a significantly increased number of branches that are not confined to the base of the ear.

the hypothesis that *RA* genes may serve as mediators of signals, maybe a sugar signal, originated at the boundary domain and regulating determinacy. *RA3* itself may have a transcriptional regulatory function, since it affects the expression of specific genes.

We are taking genetic approaches to identify factors that act in the same pathway with *RA3* to control spikelet pair meristem determinacy by screening for enhancers of the *ra3* phenotype. Typically, *ra3* mutants in a B73 background have three to eight branches only at the base of the ear, and thus we mutagenized *ra3* mutants and looked for plants that have more branches and/or have branches in the upper part of the ear (Fig. 4). Notably, *ra3* in the Mo17 background was also mutagenized. In this background, *ra3* is suppressed, so the presence of branches also indicates that an enhancer mutation has been found. Combining both backgrounds, about 20 mutants have been isolated so far. Mapping of a subset of these mutants is currently under way. So far, two independent mutant alleles of *REL3* (*RAMOSA ENHANCER LOCUS 3*), an *RA3* paralog, were identified. Both contain amino acid substitutions, and we are currently trying to understand how these mutations affect protein activity.

As a third approach to better understand how *RA3* functions, we are looking for interacting proteins. To this end, a yeast two-hybrid screen was performed, and a number of predominantly nuclear proteins were found to interact with *RA3*, which fits with its partially nuclear localization. In parallel, meristematic ear tissue from plants expressing an HA-tagged *RA3* protein was used for IP-mass spectrometry, revealing a number of potential *in vivo* interactors. The biological roles of a number of these interactors in ear development are currently being studied using insertional mutants and mutant alleles generated using CRISPR-Cas9.

## Natural Variation in Inflorescence Architecture

H. Claeys, S. Vi, I. Liao, T. Lau, M. Rutigliano

Maize inflorescence architecture has been a target for extensive selection by breeders since domestication; hence, different maize inbreds vary greatly in these traits. The genetic basis underlying this diversity is largely unknown, but it is of great interest for both fundamental and applied science. Now with the wealth of genomic data, we are equipped to study them. To identify natural variation relevant to inflorescence traits, we looked for inbred backgrounds that can enhance or suppress the phenotypes of different mutants. We focused on the 25 NAM (nested association mapping) founder inbreds because they were selected to capture the diversity of maize germplasm and because of the genetic tools available for these inbreds. We have crossed these 25 inbreds to our collection of mutants (in a B73 background) and screened the F<sub>2</sub> for plants with suppressed or enhanced phenotypes compared to the mutant in the B73 background.

We identified a suppressor of *ramosa3* (*ra3*) from the Mo17 inbred, an enhancer from the Ki11 inbred, and an enhancer of *fea2* from NC350. Segregation ratios suggest one semidominant locus for the Mo17-derived suppressor, two or more loci for the Ki11-derived enhancer, and one recessive locus for the NC350-derived enhancer. We are now in the process of rough mapping by bulked segregant analysis. Additionally, because the natural modifiers are often the results of quantitative trait loci (QTL) rather than single-gene effects, we crossed the mutants to the corresponding NAM-founder/B73 recombinant inbred lines (RILs), in order to identify and map potential modifying QTL. Using this approach, we successfully mapped the *fea2*

enhancer from NC350 to a region that colocalizes with a previously identified kernel row number (KRN) QTL. This fits with our lab's recent finding that hypomorphic alleles of *fea2* lead to increased KRN, linking fasciation and KRN. Fine mapping of this locus to isolate one or more candidate genes is ongoing.

## PUBLICATIONS

Chen H, Jackson D, Kim JY. 2014. Identification of evolutionarily conserved amino acid residues in homeodomain of KNOX proteins for intercellular trafficking. *Plant Signal Behav* **9**: e28355.

Eveland A, Goldshmidt A, Pautler M, Morohashi K, Liseron-Monfils C, Lewis M, Kumari S, Hiraga S, Yang F, Unger-Wallace E, et al. 2014. Regulatory modules controlling maize inflorescence architecture. *Genome Res* **24**: 431–443.

*In Press*

Bui HT, Balkunde R, Jackson D. 2015. Plasmodesmata. In *The plant sciences: Cell biology* (ed. Assmann S, Liu B). Springer, New York. (In press.)

Jackson D. 2015. Plasmodesmata spread their influence. *F1000 Prime Rep* (in press).

Krishnakumar V, Choi Y, Beck E, Wu Q, Luo A, Sylvester A, Jackson D, Chan AP. 2015. A maize database resource that captures tissue-specific and subcellular-localized gene expression, via fluorescent tags and confocal imaging (Maize Cell Genomics Database). *Plant Cell Physiol* **56**: e12.

Pautler M, Eveland AL, LaRue T, Yang F, Weeks R, Lunde C, Je BI, Meeley R, Komatsu M, Vollbrecht E, Sakai H, Jackson D. 2015. *FASCIATED EAR4* encodes a bZIP transcription factor that regulates shoot meristem size in maize. *Plant Cell* (in press).

Yang F, Bui HT, Pautler M, Llaca V, Johnston R, Lee B, Kolbe A, Sakai H, Jackson D. 2015. A maize glutaredoxin gene, *Abphyl2*, regulates shoot meristem size and phyllotaxy. *Plant Cell* (in press).



B.I. Je

# PLANT DEVELOPMENTAL BIOLOGY, STEM CELLS, AND FLORAL BRANCHING SYSTEMS

Z. Lippman   C. Brooks   K.L. Liberatore   S.J. Park   S. Thomain  
K. Jiang   C.A. MacAlister   S. Soyk   C. Xu

Research in our laboratory is aimed at revealing, understanding, and manipulating the genetic and molecular networks that control when, where, and how flowers are produced on plants. Flowers form on inflorescences, the reproductive branches that originate from stem cell populations at the growing tips of shoots called meristems. The number of inflorescences produced on a plant, as well as how many branches and flowers form on each inflorescence, can vary dramatically both within and between species. At the center of this diversity lie two critical processes of meristem growth and regulation: (1) meristem maturation, during which stem cell populations transition from a vegetative growth program to a reproductive flower-promoting program, and (2) meristem maintenance, which modulates stem cell proliferation and meristem size. Our research takes advantage of natural and mutant variation in inflorescence production and architecture in tomato and related Solanaceae species to explore the hypothesis that differences in meristem maturation and size explain the remarkable diversity in inflorescence branching, flower production, and reproductive success observed in nature and agriculture.

Meristems typically start growing as small flat structures that give rise to leaves before maturing into large domes that develop into inflorescences. This generic description, however, fails to capture that species-specific variation in this flowering transition underlies the remarkable diversity in inflorescence architecture. For example, both within and between species, not all doming meristems become flowers, and for some meristems, maturing into a flower can be sudden, like a switch. We use the inflorescence of tomato as a model to investigate the mechanisms controlling meristem maturation and size in sympodial plants, which comprise half of all flowering plants and display some of the most impressive diversity in shoot complexity. This is because, compared to monopodial plants like *Arabidopsis*, in which meristems persist after the transition to reproductive growth, meristems in sympodial plants terminate in flowers, and new growth

continuously arises from specialized axillary (sympodial) meristems to produce compound shoots. The tomato inflorescence is an excellent system for studying meristem maturation and size, because meristems are easily accessed for morphological and molecular analyses, and genetic perturbations have provided a foundation to build a molecular understanding of how these two processes are regulated.

## Gene Editing in Tomato Using CRISPR/Cas9 Technology

C. Brooks

In the last decade, rapid advances in technology for genome editing have made it possible to engineer mutations in genes in a sequence-specific manner. Two of these strategies, zinc finger nucleases (ZFNs) and transcription activator-like nucleases (TALENs), are based on protein–DNA interactions, whereas a third technology, clustered regularly interspaced short palindromic repeats (CRISPR)/Cas9, is an RNA-guided DNA endonuclease system. CRISPR/Cas9 has seen a meteoric rise in the last 2 years, in large part due to a number of advantages of CRISPR/Cas9 over ZFNs and TALENs, including simple construct design and assembly and a high degree of success in multiple animal and plant systems.

This year, we tested the CRISPR/Cas9 system for gene editing in tomato. In CRISPR/Cas9, an engineered single guide RNA (sgRNA) targets the Cas9 DNA endonuclease to a specific chromosomal region to create double-strand breaks. Imprecise repair results in mutations, typically insertions and deletions (indels). Specific deletions can be created using two sgRNAs separated by a spacer region. We tested the efficacy of CRISPR/Cas9 in tomato by targeting the *ARGONAUTE7 (SLAGO7)* gene (Fig. 1). *SLAGO7* is required for the biogenesis of *trans*-acting short interfering RNAs (tasiRNAs), which regulate organ polarity through posttranscriptional silencing of *AUXIN*

*RESPONSE FACTOR (ARF)* genes. The first leaves of loss-of-function mutations in *SLAGO7* produce leaflets without petioles, and later-formed leaves lack laminae, providing distinctive recessive “wiry” leaf phenotypes that are easily recognized.

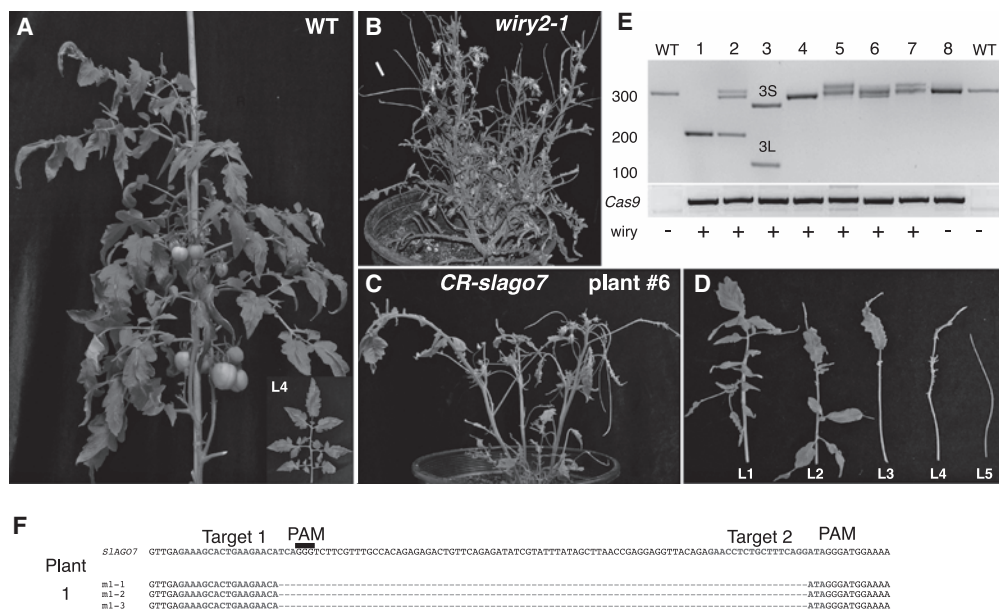
We designed a CRISPR/Cas9 construct containing two sgRNAs with the goal of creating large, defined deletions. Working with our collaborator Joyce Van Eck at the Boyce Thompson Institute Center for Plant Biotechnology, we generated 29 independent transgenic tomato plants and confirmed by polymerase chain reaction (PCR) that each carried an integrated CRISPR/Cas9 transfer DNA. Remarkably, nearly half of these plants were indistinguishable from a known strong *slago7* allele (*wiry2-1*), indicating a high efficiency of CRISPR/Cas9-mediated mutagenesis (Fig. 1). We evaluated the types of mutations generated using PCR and found one plant homozygous for the expected deletion. The remaining plants were chimeric, having the expected deletion and/or indels of various sizes. Importantly, we also demonstrated that an induced indel mutation could be transmitted through the germ line, and we targeted additional genes, including simultaneous induction

of homozygous deletions in two closely related homologs. Our findings demonstrated the power of the CRISPR/Cas9 system to create targeted mutations in tomato and likely many other dicot crops, providing a superior alternative to RNA interference (RNAi) for reverse genetics studies and crop genome engineering.

## Arabinosyltransferases Control Shoot Meristem Size in Tomato

K. Liberatore, C. Xu, C. MacAlister, S. Thomain

Despite its importance in agriculture, surprisingly little is known about the control of meristem size in plants beyond knowledge gained during the last two decades in *Arabidopsis* and the cereal crops maize and rice. Meristem size is controlled through a maintenance mechanism, in which stem cells lost to lateral organ formation are replenished through a feedback circuit involving the stem-cell-promoting *WUSCHEL* homeobox gene and its repression by the *CLAVATA* (*CLV*) signaling pathway. In *CLV* signaling, a cell surface receptor complex containing the leucine-rich repeat (LRR) receptor kinase *CLV1* and related



**Figure 1.** CRISPR/Cas9-mediated gene editing in stable transgenic tomato plants. (A) A wild-type (WT) tomato plant at 9 weeks of age, and the fourth produced leaf from the primary shoot (*inset*). (B) The strong *wiry2-1* allele of *slago7*. First-formed leaves have leaflets lacking petioles and later-formed leaves are radialized. (C,D) A representative CRISPR/Cas9-*slago7* (*CR-slago7*) plant (C) and its first five leaves showing the distinctive loss-of-function recessive wiry syndrome (D). Scale bar, 1 cm. (E,F) PCR genotyping of eight representative *CR-slago7* plants showing the type of DNA lesions generated (E), including homozygosity for the expected 90-base-pair deletion in plant #1 (F).

receptors is bound by CLV3, a secreted peptide that is modified with a chain of three arabinose sugars. However, the pathway responsible for modifying CLV3, and its importance in controlling meristem size, is unknown.

By screening a large tomato mutagenesis population, we identified novel mutants that develop branched inflorescences that produce fasciated flowers with extra floral organs. Molecular analyses showed that these phenotypes resulted from a broken CLV-WUS circuit, which causes stem cells to overproliferate and meristems to enlarge. The first mutant was weakly fasciated and was found to be defective in the homolog of CLV1. However, surprisingly, a much more extreme mutant, *fasciated inflorescence (fin)*, was defective in a hydroxyproline *O*-arabinosyltransferase (HPAT), which catalyzes the addition of arabinose to diverse plant proteins, including peptides related to CLV3. Chemically induced and CRISPR/Cas9-engineered mutations in the *FASCIATED AND BRANCHED 2 (FAB2/XEG113)* and *REDUCED RESIDUAL ARABINOSE 3 (RRA3)* genes, encoding arabinosyltransferases predicted to extend arabinose chains, resulted in plants resembling *fin* mutants, but they were less severely fasciated while having additional defects. We also used CRISPR/Cas9 to mutate all homologs of the *CLV* pathway in tomato, including *CLV3* and a second receptor gene, *CLV2*, that functions redundantly with *CLV1*. All of these mutants were also fasciated like the arabinosyltransferase mutants. Through a series of genetic, expression, and functional assays, we found that CLV3 and related CLE peptides must be fully arabinosylated for normal activity, demonstrating that a cascade of arabinosyltransferase activity is required to fully activate the CLV-WUS circuit. These findings have exposed arabinosyltransferase genes as critical new components of the *CLV* signaling pathway, revealing a new layer of complexity in the control of stem cell homeostasis and meristem size during plant development.

## Diverse Developmental Roles for Arabinosyltransferases in Other Plants

C. MacAlister

---

The *FIN/HPAT*, *FAB2/XEG113*, and *RRA3* arabinosyltransferase genes are all members of the plant and animal glycosyltransferase (GT) superfamily, several members of which have recently been reported to function in diverse developmental processes, including pollen

tube growth and root hair elongation in *Arabidopsis* and nodule formation in pea and clover. Yet, surprisingly, no meristem phenotypes for any of the available arabinosyltransferase mutants have been reported. That our analysis of extreme fasciated mutants in tomato uncovered a crucial role for arabinosyltransferase genes in stem cell homeostasis suggested that members of different arabinosyltransferase families might have been co-opted to target distinct sets of proteins and signaling peptides in different species and developmental contexts. To compare developmental roles of *FIN/HPAT* genes in diverse species, we analyzed mutations in the three *HPAT* family members in *Arabidopsis*. Surprisingly, we observed no phenotypes beyond pollen tube growth defects in any *hpat* mutant combination. In homozygous *hpat* double mutants, we found that pollen grains can germinate; however, tube growth arrests in the transmitting tract of the ovary, resulting in a failure to target ovules and poor seed set. Data from our lab and others' suggest that these phenotypes might be due to loss of arabinosylation on extensin proteins, which are critical for cell wall assembly.

To further explore the diverse, and potentially ancient, roles of the *FIN/HPAT* family, we created mutations in two homologs in the moss *Physcomitrella patens*. Moss development involves germination of a haploid spore followed by formation of a vegetative body (colony) composed of filamentous protonemal network containing two cell types: slower-growing chloronema for photosynthesis and faster-growing caulonema to colonize the surrounding substrate. We found that knocking out the more highly expressed *FIN* homolog (*PpFINLa*) results in larger and faster colonizing colonies, because of the enhanced caulonema production relative to chloronema. Notably, although pollen tubes and caulonema are both tip-growing cells, the process of tip growth itself is unaffected in both systems. We are currently investigating the molecular basis for the filament growth defects in moss *Ppfinla* mutants, and preliminary data suggest that a distinct protein family is targeted for arabinosylation compared to tomato and *Arabidopsis*.

## A New Mechanism Controlling Meristem Maturation in Tomato

C. Xu, C. MacAlister

---

Several years ago, we found that mutations in the genes *COMPOUND INFLORESCENCE (S)*, homolog

of *WUSCHEL HOMEODOMAIN 9*), *ANANTHA* (*AN*, homolog of *UNUSUAL FLORAL ORGANS*), and *FALSIFLORA* (*FA*, homolog of *LEAFY*) cause more branches and flowers to develop, because meristem termination is delayed (*s* mutants) or never achieved (*an*, *fa*). *S* is expressed late in maturation just before activation of *AN*, which encodes an F-box protein. From work in petunia, we understand how an ortholog of *AN* promotes the final stage of maturation; upon its late activation, the *AN* protein interacts with *FA* to form a flower specification complex, which activates floral identity genes. In tomato, the rate of maturation promoted by *S* and *AN-FA* permits only one lateral meristem to form in each sympodial cycle, resulting in a multiflowered, unbranched inflorescence.

To investigate whether a mechanism also exists to repress maturation, we studied the *tmf* mutant, which flowers early and produces a single flower inflorescence on the primary shoot. We found that the *TMF* gene encodes a member of the ALOG (*Arabidopsis* LSH1 and *Oryza* G<sub>1</sub>) protein family. ALOGs contain a DNA-binding domain, suggesting transcription factor function. Consistent with this, *TMF* is in the nucleus and can activate transcription in yeast. ALOGs have developmental roles in *Arabidopsis* and rice floral organ identity and, like *TMF*, rice inflorescence architecture.

In a yeast two-hybrid (Y2H) screen using *TMF* as bait, we identified several transcriptional regulators, the most interesting to us being homologs of the *Arabidopsis* BTB/POZ domain-containing BOP family and the AP2/ERF/ethylene responsive factor (AP2/ERF) family. These genes have diverse developmental functions in *Arabidopsis*, including in floral meristem determinacy; however, nothing is known about these genes in tomato. Tomato has three *BOP* genes (*SIBOP1/2/3*), and we used bimolecular fluorescence complementation (BiFC) to show that all three *SIBOP* proteins interact with *TMF* in tomato nuclei. Using RNAi, we knocked down expression of the *SIBOP* genes. RNAi of *SIBOP1* (*slbop1-RNAi*) resulted in increased leaf complexity and loss of floral organ abscission, similar to *Arabidopsis* *bop1/2* mutants. RNAi of *SIBOP2/3* together (*slbop2/3-RNAi*) causes not only the same phenotypes, but also inflorescences with architectural defects. To further study the *SIBOPs* and their roles in inflorescence architecture, we are taking advantage of CRISPR/Cas9 technology and have engineered deletion mutations in each *SIBOP* (*CR-slbop*). We have also engineered

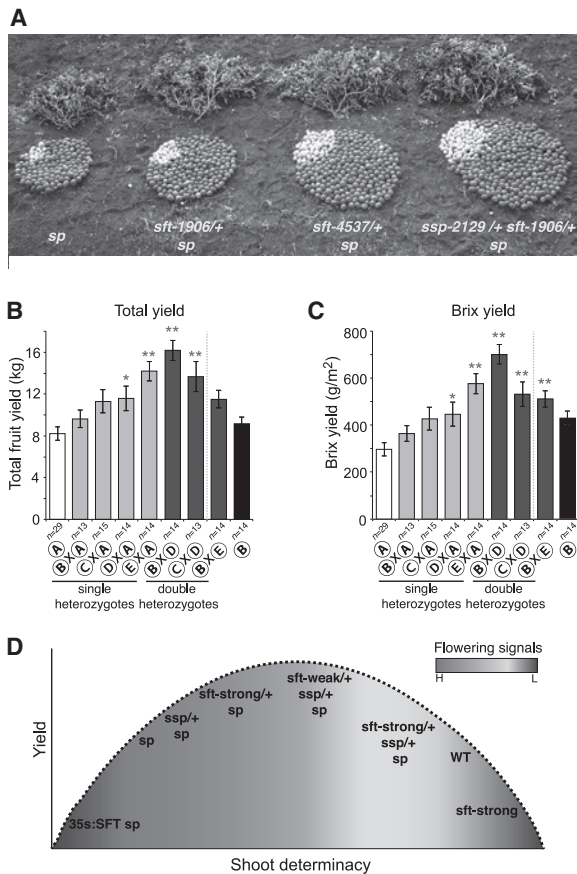
mutations in *AP2/ERF* family members. In the coming year, these mutants will be studied for their effects on flowering time and inflorescence development as it relates to meristem maturation. We have also developed biochemical tools to begin identifying transcriptional targets of *TMF* and *S*.

## Improving Tomato Productivity Using Mutations in the Florigen Pathway

S.J. Park, K. Jiang

By studying novel meristem maturation mutants of tomato, we uncovered a new way to customize and optimize plant architecture and yield using genes in the universal florigen flowering pathway. A major driver of crop domestication was the modification of wild-plant architectures into new architectural forms that improved flower, fruit, and seed production. In tomato, selection during domestication was for exceptionally large fruits, and shoot architecture underwent a radical change 85 years ago with the discovery of the *self pruning* (*sp*) mutation, which transformed “indeterminate” plants into a new “determinate” form that spawned the processing tomato industry. *SP* is the homolog of *Antirrhinum* *CENTRORADIALIS* (*CEN*) and *Arabidopsis* *TERMINAL FLOWER1* (*TFL1*), and it encodes a repressor of flowering in the CETS family of proteins, which includes the generic flowering hormone florigen. All wild tomatoes and classical cultivated varieties are indeterminate, a reflection of their sympodial growth program, which is regulated by the opposing activities of *SP* and florigen. The regularity and perpetuation of sympodial cycling are made possible by a balance of flower promoting and repressing signals: *SP* is produced in meristems to counterbalance florigen, encoded by *SINGLE FLOWER TRUSS* (*SFT*), but in *sp* mutants, florigen is no longer inhibited, causing shoots to terminate progressively faster until cycling eventually stops. The result is a burst of inflorescence production and nearly synchronous fruit ripening on compact bushy plants that can be grown at high density and harvested mechanically, providing major advantages for large-scale open field production.

Homologs of *SP* and *SFT* were repeatedly targeted for agricultural adaptations in many crops. Yet, the limited number and types of alleles provided by nature may not be providing optimal balances of



**Figure 2.** Exploiting dosage sensitivity of the florigen pathway leads to a new optimum for tomato yield. (A) Representative plant and yield from an M82 control plant, an *sft-1906/+ sp* single heterozygote, the previous highest-yielding *sft-4537/+ sp* genotype, and the new highest-yielding *sft-1906/+ ssp-2129/+ sp* double heterozygote. (B,C) Statistical comparisons of mean values ( $\pm$  s.e.m.) for total fruit yields (B), and Brix-yield (C) between *sp* plants (white bar), *ssp-2129 sp* indeterminate plants (black bar), and the single (light gray) and double (dark gray) heterozygous genotypes. Asterisks indicate significantly different yields from *sp* (Dunnett's "compare with control," \* $P < 0.05$ , \*\* $P < 0.01$ ). (D) Model of tomato yield with respect to shoot determinacy and the shifting ratios of flower promoting and repressing signals. The ratio in nature provides typical indeterminate growth. When flower-promoting signals are extremely high (e.g., *35S:SFT sp*) or extremely low (e.g., strong *sft* mutants), yields are low, because of rapid termination or failure to initiate sympodial growth, respectively. In *sp* mutants, flower-repressing signals are reduced, resulting in classical determinate growth. As flower-promoting signals are progressively reduced, determinacy is gradually relieved, and yields reach a new maximum before a further loss of flower-promoting signals results in a new balance that approaches WT, resulting in a switch to indeterminate growth and a drop in yield. H, high; L, low.

flowering signals to allow maximum productivity. In a telling example from our prior work, we found that induced mutations in *SFT* can increase yield in determinate plants. In *sft* mutants, a loss of florigen impedes initiation of sympodial growth, resulting in highly vegetative plants with few flowers and fruits. However, when *sft* mutations are heterozygous, a partial dose-dependent reduction of florigen modifies the balance of flowering signals, causing a weak suppression of *sp* that results in more sympodial shoots and inflorescences.

We reasoned that other mutants that suppress determinacy might provide new ways to modify tomato flowering, meristem maturation, and flower production. Several years ago, we identified three *suppressor of sp* (*ssp*) mutants that restored indeterminate growth. We cloned the *SSP* genes and found that they were defective in genes encoding components of a florigen activation complex, originally described in rice. Notably, *ssp-1906* was a weak mutation in *SFT*, whereas the other two *ssp* mutants were defective in a basic leucine zipper (bZIP) transcription factor component of the florigen complex. By combining these mutations in heterozygous states in hybrid plants, we pinpointed an optimal balance of flowering signals, which resulted in a new partially determinate architecture that doubled the yields of *sp* plants and surpassed the yield gain of *sft* heterozygotes by 15% (Fig. 2). These findings indicate that harnessing mutations in the highly conserved florigen pathway to customize plant architecture and flower production offers a broad tool kit to maximize productivity in many crops.

## PUBLICATIONS

- Brooks C, Nekrasov V, Lippman ZB, Van Eck J. 2014. Efficient gene editing in tomato in the first generation using the CRISPR/Cas9 system. *Plant Physiol* **166**: 1292–1297.
- Park SJ, Eshed Y, Lippman ZB. 2014. Meristem maturation and inflorescence architecture: Lessons from the Solanaceae. *Curr Opin Plant Biol* **17**: 70–77.
- Park SJ, Jiang K, Tal Y, Yichie Y, Gar O, Zamir D, Lippman ZB. 2014. Optimization of crop productivity in tomato using induced mutations in the florigen pathway. *Nat Genet* **46**: 1337–1342.



# EPIGENETIC INHERITANCE IN PLANTS AND FISSION YEAST

R. Martienssen	S. Bhattacharjee	R. Herridge	B. Roche
	S. Castel	Y. Jacob	A. Schorn
	A.-Y. Chang	H.S. Kim	A. Shimada
	K. Creasey	C. LeBlanc	J. Simorowski
	F. de Sousa Borges	S.C. Lee	R. Tirado-Magallanes
	M. Donoghue	A. Molla-Morales	R. Umamaheswari
	E. Ernst	J.-S. Parent	C. Underwood
	M. Gutbrod	M. Regulski	
	J.-J. Han	J. Ren	

Plants and fission yeast provide excellent models for epigenetic mechanisms of transposon regulation, heterochromatic silencing, and gene imprinting, important both for plant breeding and for human health. We are investigating the role of RNA interference (RNAi) in heterochromatic silencing in the fission yeast *Schizosaccharomyces pombe* and in the model plant *Arabidopsis thaliana*, as well as continuing our work on *Zea mays*. In fission yeast, we have found that RNAi promotes DNA replication and repair by releasing RNA polymerase II (Pol II). In plants, we have found that the genome undergoes limited reprogramming of DNA methylation in pollen, guided by DNA glycosylases and small RNA. Using mutants deficient in histone and DNA methylation, we have found that histone variants and microRNA (miRNA) contribute to reprogramming by targeting transposons and determining the specificity of epigenetic histone modifications. Our results suggest a model for imprinting, transposon control and the origin of epialleles. We continue to develop duckweeds as a source of biofuel and have developed an efficient transformation and gene control system in the clonally propagated aquatic macrophyte *Lemna minor*. Sequencing of the duckweed, oil palm, and sugarcane genomes has provided important clues as to how to increase biofuel yields and reduce tropical deforestation.

This year, we said good-bye to Roberto Tirado-Magallanes, who has started his Ph.D. at Ecole Normale Supérieure in Paris, and Stéphane Castel, who

started a postdoc at the New York Genome Center. We welcomed Watson School of Biological Sciences student Michael Gutbrod.

## Dicer Promotes Transcription Termination at Sites of Replication Stress to Maintain Genome Stability

S.E. Castel, J. Ren, S. Bhattacharjee, A.-Y. Chang, R. Martienssen [In collaboration with F. Antequera, CSIC/Universidad de Salamanca, Salamanca, Spain; and B. Arcangioli, Institut Pasteur, Paris France]

Nuclear RNAi is an important regulator of transcription and epigenetic modification, but the underlying mechanisms remain elusive. Using a genome-wide approach in the fission yeast *S. pombe*, we have found that Dcr1, but not other components of the canonical RNAi pathway, promotes the release of Pol II from the 3' end of highly transcribed genes, and, surprisingly, from antisense transcription of rRNA and tRNA genes, which are normally transcribed by Pol I and Pol III. These Dcr1-terminated loci correspond to sites of replication stress and DNA damage, likely resulting from transcription-replication collisions. At the rDNA loci, release of Pol II facilitates DNA replication and prevents homologous recombination, which would otherwise lead to loss of rDNA repeats especially during meiosis. Our results reveal a novel role for Dcr1-mediated transcription termination in genome maintenance and may account for widespread regulation of genome stability by nuclear RNAi in higher eukaryotes.

## Chromosome Conformation Maps in Fission Yeast Reveal Cell-Cycle-Dependent Subnuclear Structure

R. Martienssen [In collaboration with R.S. Grand, J. O'Sullivan Liggins Institute, University of Auckland, New Zealand; J.R. Allison, Massey University, Auckland, New Zealand]

Successful progression through the cell cycle requires spatial and temporal regulation of gene transcript levels and the number, positions, and condensation levels of chromosomes. We have performed a high-resolution survey of genome interactions in *S. pombe* using synchronized cells to investigate cell-cycle-dependent changes in genome organization and transcription. Cell-cycle-dependent interactions were captured between and within *S. pombe* chromosomes. The clustering of telomeres and retrotransposon long terminal repeats (LTRs) was observed throughout the cell cycle. There were clear correlations between transcript levels and chromosomal interactions between genes, consistent with a role for interactions in transcriptional regulation at specific stages of the cell cycle. In silico reconstructions of chromosome organization were made by polymer modeling. These models suggest that groups of genes with high and low or differentially regulated transcript levels have preferred positions within the nucleus. The observation that chromosomal interactions are maintained even when chromosomes are fully condensed in M phase implicates genome organization in epigenetic inheritance and bookmarking.

## RNAi Acts in Essential Novel Mechanisms in Quiescence in the Fission Yeast *S. pombe*

B. Roche, R. Martienssen [in collaboration with B. Arcangioli, Institute Pasteur, Paris]

Quiescence is defined as the state of a nondividing cell ( $G_0$ ), which is still metabolically active and is able to return to the cell cycle with full viability. Quiescent cells have a predominant role in unicellular organisms and in multicellular organisms, in which they are essential, for example, in humans (in stem cell niches and memory B cells). However, the mechanisms governing the transition between the cell cycle and quiescence, and quiescence maintenance, are not well understood. In *S. pombe*, cells can be induced into  $G_0$  by nitrogen starvation of a heterothallic strain and keep

full viability for a long period (>3 wk). We found that RNAi mutants ( $\Delta dcr1$ ,  $\Delta rdp1$ ,  $\Delta ago1$ ) are predominantly impaired specifically in quiescence and lose viability rapidly. Intriguingly, heterochromatin mutants ( $\Delta clr4$ ,  $\Delta swi6$ ,  $\Delta chp2$ ) do not show this phenotype, suggesting that RNAi functions in quiescence independently of its role in heterochromatin assembly, which we have previously shown to occur in S phase. In accordance with this model, we found that centromeric small RNAs are not detectable in  $G_0$  in wild type. However, the Dicer catalytic mutant *dcr1-5* has the  $\Delta dcr1$   $G_0$  phenotype, which suggests the involvement of cryptic new small RNAs. To uncover the novel pathways involved, we established a genetic screen for  $G_0$ -specific suppressors of RNAi mutants. This screen allowed us to isolate >30 independent suppressors of  $\Delta dcr1$  in the absence of a mutagen, potentially indicating a role in DNA repair.

## DNA Demethylation in *Arabidopsis* Pollen and Consequences for Transgenerational Epigenetic Inheritance

M. Donoghue, F. de Sousa Borges, Y. Jacob, C. LeBlanc, R. Martienssen [in collaboration with D.I. Bouyer and V. Colot, Ecole Normale Supérieure, Paris]

Genome reprogramming of DNA methylation is important for transposon regulation and epigenetic inheritance. In *Arabidopsis thaliana*, dynamic changes of DNA methylation occur in the male gametophyte, where two sperm cells (SCs) and a vegetative cell differentiate from a precursor microspore after postmeiotic cell divisions. One pathway toward epigenetic variation in plants comes from the activity of DNA glycosylases/demethylases in the pollen vegetative cell nucleus (VN). We have found that previously identified variable epialleles and imprinted genes are specifically demethylated in the VN, by combined activity of the DNA glycosylases ROS1 and DME. In contrast, heterochromatic CHH methylation, where H is A, C, or T, is almost completely lost during or after meiosis, and restored in the VN, but not in the gametes that remain hypomethylated at the mature pollen stage. The role of targeted demethylation in the VN is important for pollen development and fertility, as well as for stable repression of transposable elements in the germline and subsequent generations. We analyzed genome-wide methylomes of fluorescence-activated cell sorting (FACS)-purified

*dme/+*, *ros1* and *dme/+*, *ros1* mutant sperm and VN and showed that DME is responsible for demethylation of DNA transposons that are found in regulatory regions near imprinted genes and epialleles. Strikingly, the double-mutant *dme/+*, *ros1* showed a cumulative effect, indicating that DNA demethylation in the VN involves combined activity of DME and ROS1.

### Selective Methylation of Histone H3 Variant H3.1 Regulates Heterochromatin Replication

Y. Jacob, M. Donoghue, C. LeBlanc, C. Underwood, R. Martienssen [in collaboration with S. Michaels, Indiana University; D. Reinberg, New York University Medical School; and J.-F. Couture, University of Ottawa]

Histone variants have been proposed to act as determinants for posttranslational modifications (PTM) with widespread regulatory functions. We have found a histone-modifying enzyme that selectively methylates the replication-dependent histone H3 variant H3.1. The crystal structure of the SET domain of the histone H3 lysine 27 (H3K27) methyltransferase ARABIDOPSIS TRITHORAX-RELATED PROTEIN 5 (ATXR5) in complex with an H3.1 peptide shows that ATXR5 contains a bipartite catalytic domain that specifically “reads” alanine 31 of H3.1. Variation at position 31 between H3.1 and replication-independent H3.3 is conserved in plants and animals, and threonine 31 in H3.3 is responsible for inhibiting the activity of ATXR5 and its paralog ATXR6. Our results suggest a simple model for the mitotic inheritance of the heterochromatic mark H3K27me1 and the protection of H3.3-enriched genes

against heterochromatinization during DNA replication. Unmodified histone H3.1 may trigger overreplication of heterochromatin and contribute to reprogramming in pollen, which lacks H3.1 in the VN.

### miRNAs Trigger Widespread Epigenetically Activated siRNAs from Transposons in *Arabidopsis*

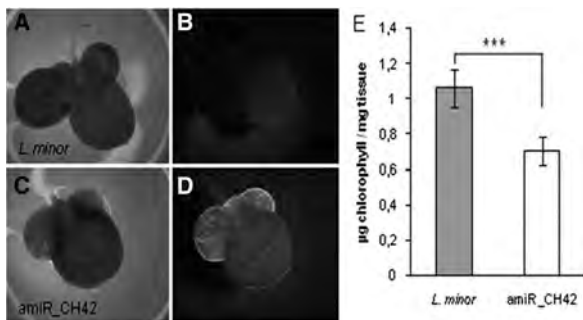
K. Creasey, M. Regulski, F. de Sousa Borges, A. Shimada, R. Martienssen [In collaboration with B.C. Meyers, University of Delaware]

Transposons in *Arabidopsis* give rise to abundant 21-nt “epigenetically activated” small interfering RNAs (easiRNAs) in *DECREASE IN DNA METHYLATION1* (*ddm1*) and *DNA METHYLTRANSFERASE1* (*met1*) mutants, as well as in pollen from wild-type plants, in which heterochromatin is lost during reprogramming. easiRNA biogenesis is dependent on *ARGONAUTE1* (*AGO1*), *DICER-LIKE4* (*DCL4*) and *RNA-DEPENDENT RNA POLYMERASE6* (*RDR6*), resembling 21-nt trans-acting (ta)siRNAs and other secondary siRNA in this respect. However, the factor that triggers easiRNA biogenesis from transposons has remained elusive. We have found that specificity is provided by targeted cleavage of thousands of transposon transcripts by at least fifty micro (mi)RNAs, some of which are themselves encoded by transposons, but most of which are highly conserved with well known roles in plant development. Interestingly, the loss of easiRNAs in *ddm1 rdr6* is compensated by the gain of 24-nt heterochromatic (het)siRNAs which can guide RNA-directed DNA methylation (RdDM). This suggests that *RDR6*-directed easiRNA production acts antagonistically to *RDR2*-directed hetsiRNA production, thereby inhibiting transcriptional gene silencing. Widespread targeting of transposons may reflect the evolutionary origin of microRNA in genome surveillance.

### Efficient Transformation and Artificial miRNA Gene Silencing in *Lemna minor*

A. Molla-Morales, E. Ernst, R. Martienssen [in collaboration with J. Shanklin, Brookhaven National Laboratory, New York; and B. Meyers, University of Delaware]

Despite rapid doubling time, simple architecture, and ease of metabolic labelling, a lack of genetic tools in the Lemnaceae (duckweed) has impeded the full implementation of this organism as a model for biological research. We



**Figure 1.** Phenotype of *L. minor* duckweed expressing an artificial miRNA targeting chlorophyll biosynthesis and green fluorescent protein (GFP). Wild-type *Lemna minor* (A,B) and transformed amiR\_CH42 (C,D) fronds observed under bright-field (A,C) and fluorescence (B,D) light. Quantification of the chlorophyll content in *L. minor* and amiR\_CH42 (E). The s.d. values are shown.  $n = 3$ . Stars indicate significance in two-tailed z-test,  $***P < 0.0001$ .

have developed technologies to facilitate high-throughput genetic studies in duckweed. We developed a fast and efficient method for producing *Lemna minor* stable transgenic fronds via *Agrobacterium*-mediated transformation and regeneration from tissue culture. Additionally, we engineered an artificial miRNA (amiRNA) gene silencing system (Fig. 1). We identified a *Lemna gibba* endogenous miR166 precursor and used it as a backbone to produce amiRNAs. As a proof of concept, we induced the silencing of CH42, a magnesium chelatase subunit, using our amiRNA platform. Expression of CH42 in transgenic *L. minor* fronds was significantly reduced, which resulted in reduction of chlorophyll pigmentation. These techniques will enable us to tackle future challenges in the biology and biotechnology of Lemnaceae.

### Duckweeds: Aquatic Macrophytes for Biofuel Production

E. Ernst, A. Molla-Morales, S.C. Lee [in collaboration with J. Shanklin and J. Schwender, Brookhaven National Laboratories; D. Pappin, Cold Spring Harbor Laboratory]

Petroleum availability and atmospheric carbon accumulation are some of the main concerns of our era: Currently, biofuels produced from corn grain and sugarcane are predominant alternatives, but they directly compete for land, with food production raising sustainability concerns. Lemnaceae species (aquatic duckweeds) include the smallest flowering plants, and they have considerable potential as biofuel feedstocks because of their extreme growth rates and clonal propagation. Our goal is to engineer duckweeds to increase their oil levels for biofuel production. To do so, we aim to increase the expression of genes related to the production of triacylglycerol (TAG), silence the genes that have a role in the oxidation of lipid bodies, or redirect the starch metabolism to oil production by silencing the key genes that lead to starch accumulation. We have sequenced the genome of *Lemna gibba* DWC131, as well as the tetraploid *L. minor*. These genome sequences will enable tackling future challenges in the biology and biotechnology of Lemnaceae. For example, nitrogen deficiency is an environmental stress that can enhance biosynthesis and storage of starch and TAG in vegetative tissues of model plant species such as tobacco, and TAG is a precursor of biodiesel. *L. gibba* grown on nitrate-lacking media had increased numbers of lipid droplets in the chloroplast and increases in saturated and monounsaturated fatty acids such as palmitic acid (16:0) and palmitoleic acid

(16:10). Transcriptome analysis is under way to identify the genes responsible, and we are using our transformation system to enhance oil yields further via CRISPR genome editing technology.

### Sugarcane Genome Sequencing by Methylation Filtration

M. Regulski, R. Martienssen [In collaboration with A.S. Hemerly and P.C. Ferreira, Universidade Federal do Rio de Janeiro, Brazil; and W. McCombie, Cold Spring Harbor Laboratory]

Large tracts of methylated repeats occur in plant genomes that are interspersed by hypomethylated gene-rich regions, reaching an extreme in the polyaneploid hybrid sugarcane, a relative of sorghum and one of the most efficient biofuel crops. Gene-enrichment strategies based on methylation profiles offer an alternative to sequencing repetitive genomes. We have applied methyl filtration with McrBC endonuclease digestion to enrich for euchromatic regions in the sugarcane genome. The use of methyl filtration filtered out 35% of the sugarcane genome and produced 1.5× more scaffolds and 1.7× more assembled sequences compared with unfiltered. The coverage of sorghum-coding sequences (CDS) by MF scaffolds was at least 36% higher, and we increased by 134× the coverage of gene regions of the monoploid sugarcane genome. The MF reads assembled into scaffolds that covered 100% of genes on existing sugarcane bacterial artificial chromosomes (BACs), 97.2% of sugarcane expressed sequence tags (ESTs), 92.7% of sugarcane RNA-Seq reads, and 98.4% of sorghum protein sequences. A large number of miRNA genes were also identified in the MF scaffolds. The information achieved by the MF data set provides a valuable tool for genomic research in the genus *Saccharum* and for improvement of sugarcane as a biofuel crop.

### The Oil Palm VIRESCENS Gene Controls Fruit Color and Encodes a R2R3-MYB Transcription Factor

R. Martienssen [In collaboration with R. Singh, E.T. Low, R. Nookiah, and R. Sambanthamurthi, Malaysian Palm Oil Board, Selangor, Malaysia; N. Lakey, S.W. Smith, J.M. Ordway, Orion Genomics, St. Louis, Missouri]

Oil palm, a plantation crop of major economic importance in Southeast Asia, is the predominant source of edible oil worldwide. We identified the virescens (VIR)

gene, which controls fruit exocarp color and is an indicator of ripeness. VIR is a R2R3-MYB transcription factor that has homology with Lily LhMYB12 and similarity with *Arabidopsis* production of anthocyanin pigment 1 (PAP1). We identified five independent mutant alleles of VIR in more than 400 accessions from sub-Saharan Africa that account for the dominant-negative VIR phenotype. Each mutation results in premature termination of the carboxy-terminal domain of VIR, resembling McClintock's C1-I allele in maize. The abundance of alleles likely reflects cultural practices, by which fruits were venerated for magical and medicinal properties. The identification of VIR will allow selection of the trait at the seed or early-nursery stage, 3–6 years before fruits are produced, greatly advancing introgression into elite breeding material.

## PUBLICATIONS

- Cantó-Pastor A, Mollá-Morales A, Ernst E, Dahl W, Zhai J, Yan Y, Meyers BC, Shanklin J, Martienssen R. 2014. Efficient transformation and artificial miRNA gene silencing in *Lemna minor*. *Plant Biol* **17**: (Suppl 1) 59–65.
- Castel SE, Ren J, Bhattacharjee S, Chang A-Y, Sanchez M, Valbuena A, Antequera F, Martienssen RA. 2014. Dicer promotes transcription termination at sites of replication stress to maintain genome stability. *Cell* **159**: 572–583.
- Creasey KM, Zhai J, Borges F, Van Ex F, Regulski M, Meyers BC, Martienssen RA. 2014. miRNAs trigger widespread epigenetically activated siRNAs from transposons in *Arabidopsis*. *Nature* **508**: 411–415.
- Grand RS, Martienssen R, O'Sullivan JM. 2014. Potential roles for interactions between the mitochondrial and nuclear DNA throughout the cell cycle of *Schizosaccharomyces pombe*. *Mitochondrion* **17**: 141–149.
- Grand RS, Pichugina T, Gehlen LR, Jones MB, Tsai P, Allison JR, Martienssen R, O'Sullivan JM. 2014. Chromosome conformation maps in fission yeast reveal cell cycle dependent sub nuclear structure. *Nucleic Acids Res* **42**: 12585–12599.
- Grativol C, Regulski M, Bertalan M, McCombie WR, da Silva FR, Zerlotini Neto A, Vicentini R, Farinelli L, Hemerly AS, Martienssen RA, Ferreira PC. 2014. Sugarcane genome sequencing by methylation filtration provides tools for genomic research in the genus *Saccharum*. *Plant J* **79**: 162–172.
- Heard E, Martienssen RA. 2014. Transgenerational epigenetic inheritance: Myths and mechanisms. *Cell* **157**: 95–109.
- Jacob Y, Bergamin E, Donoghue MTA, Mongeon V, LeBlanc C, Voigt P, Underwood CJ, Brunzelle JS, Michaels SD, Reinberg D, et al. 2014. Selective methylation of histone H3 variant H3.1 regulates heterochromatin replication. *Science* **343**: 1249–1253.
- Kuscu C, Zaratiegui M, Kim HS, Wah DA, Martienssen RA, Schalch T, Joshua-Tor L. 2014. CRL4-like Ctr4 complex in *Schizosaccharomyces pombe* depends on an exposed surface of Dos1 for heterochromatin silencing. *Proc Natl Acad Sci* **111**: 1795–1800.
- Singh R, Low ET, Ooi LC, Ong-Abdullah M, Nookiah R, Ting NC, Marjuni M, Chan PL, Ithnin M, Manaf MA., et al. 2014. The oil palm *VIRESCENS* gene controls fruit colour and encodes a R2R3-MYB. *Nat Commun* **5**: 4106.
- Thiebaut F, Rojas CA, Grativol C, Motta MR, Vieira T, Regulski M, Martienssen RA, Farinelli L, Hemerly AS, Ferreira PC. 2014. Genome-wide identification of microRNA and siRNA responsive to endophytic beneficial diazotrophic bacteria in maize. *BMC Genomics* **15**: 766.
- Thiebaut F, Grativol C, Tanurdzic M, Carnavale-Bottino M, Vieira T, Motta MR, Rojas C, Vincentini R, Chabregas SM, Hemerly AS, Martienssen RA, Ferreira PC. 2014. Differential sRNA regulation in leaves and roots of sugarcane under water depletion. *PLoS One* **9**: e93822.



S. Parent

# PLANT DEVELOPMENTAL GENETICS

M. Timmermans   A. Benkovics   C. Fernandez-Marco   S. Knauer   D. Skopelitis  
M. Dotto   A. Husbands   K. Petsch

Specification of adaxial–abaxial (upper/lower) polarity in the developing leaf drives the flattened outgrowth of the leaf blade and directs the differentiation of distinct cell types within the leaf’s adaxial/top and abaxial/bottom domains. Both are important innovations in the evolution of land plants that maximize photosynthesis while minimizing water loss to the environment. In addition to being a key developmental process, adaxial–abaxial patterning is of particular interest from a mechanistic point of view and has proven to be an excellent model to study small RNA-mediated gene regulation. We previously showed that patterning of this developmental axis involves a cascade of opposing small RNAs, in which microRNA (miRNA) miR390 triggers the biogenesis of the *TAS3*-derived 21-nucleotide tasiR-ARF on the adaxial side of developing leaves. These confine the accumulation of abaxial determinants, including AUXIN RESPONSE FACTOR3 (ARF3) and miR166, to the lower side of leaves. miR166, in turn, demarcates the abaxial side by repressing expression of class III homeodomain leucine zipper (HD-ZIPIII) transcription factors that specify adaxial fate.

Importantly, our analysis of the *TAS3* trans-acting short interfering RNA (tasiRNA) pathway in *Arabidopsis* demonstrated that tasiR-ARF functions as a mobile positional signal in adaxial–abaxial patterning. Movement of this small RNA from its defined source of biogenesis in the two uppermost cell layers of leaves creates a gradient of accumulation that dissipates abaxially. The tasiR-ARF gradient generates a sharply defined expression domain of the abaxial determinant ARF3 on the bottom side of leaf primordia, suggesting that small RNAs can establish pattern through a morphogen-like activity. This work provided the first direct evidence that small RNAs are mobile and can function as instructive signals in development, thereby revealing a novel patterning activity of small RNAs. We are currently studying the role of this specific small RNA pathway and the properties of mobile small RNAs in general in maize, *Arabidopsis*, and the moss *Physcomitrella patens*.

## Small RNA Gradients Create Stable Developmental Boundaries

A. Benkovics, D. Skopelitis, C. Fernandez-Marco

---

Mathematical modeling of the tasiR-ARF–ARF3 interaction predicts that small RNA gradients resulting from mobility are uniquely suited to generate sharply defined boundaries of target gene expression. To test this, we are taking advantage of the *pARF3:ARF3-GUS* reporter generated previously. In an *rdr6* mutant background, which blocks tasiR-ARF production, this reporter is expressed throughout the developing leaf. In this background, we express an artificial miRNA targeting *ARF3* (miRARF) from different leaf-specific promoters to modulate the position, direction, and steepness of this small RNA gradient. Target gene expression is being monitored by GUS histochemistry and miRARF expression patterns will be determined by in situ hybridizations. Our results strongly support the idea that small RNA gradients have the inherent ability to generate sharply defined domains of target gene expression.

Patterning via small RNA gradients is also being addressed using a reporter system that monitors the readout of a miR166 gradient. A miR166-insensitive HD-ZIPIII reporter (PHB\*-YFP) that is active throughout the leaf induces an adaxialized leaf phenotype. In this background, we express a modified version of miR166 (miR166\*) that specifically cleaves the PHB\*-YFP transcripts. Our results show that miR166\* expression from the abaxial epidermis is sufficient to suppress the PHB\*-YFP leaf defects, indicative of movement of this small RNA. Visualization of PHB\*-YFP expression patterns by confocal microscopy indicates that similar to results obtained via our miR-ARF3–ARF3-GUS system, miR166\* gradients generated in different directions in the leaf create a sharp on-off expression pattern of target expression, suggesting that this is a general patterning property of small RNA gradients.

Considering that patterning of the adaxial–abaxial axis involves two opposing small RNA gradients, these might serve to confer robustness onto the

leaf development program by stabilizing the adaxial–abaxial boundary throughout organogenesis and under a range of environmental conditions. We are comparing variability in leaf parameters under normal and mild stress conditions between wild-type plants and plants in which either small RNA gradient is perturbed. We are also analyzing the expression patterns of adaxial and abaxial reporters in these lines to monitor the effect of perturbing small RNA gradients on the position and uniformity of the adaxial–abaxial cell fate boundary under normal and stress conditions. The outcome of these experiments will reveal whether small RNA gradients provide robustness to the leaf developmental program under a range of environmental conditions.

### Dissecting Small RNA Mobility in Plants

D. Skopelitis, C. Fernandez-Marco

Given the scope of miRNA-regulated gene networks, the cell-to-cell movement of small RNAs has important implications with respect to their potential as instructive signals in development or in response to physiological and stress stimuli. We are using artificial miRNAs targeting easy-to-score reporter genes to study parameters of miRNA movement, such as tissue specificity, directionality, dose-dependence, and the kinetics of movement. To investigate miRNA mobility in multiple distinct developmental contexts, we are expressing an artificial miRNA targeting a cell-autonomous green fluorescent protein (GFP) reporter (miR-GFP) from a number of promoters with distinct spatiotemporal patterns of expression. Expression of miR-GFP from the epidermis-specific *ATML1* promoter, mesophyll-specific *RBCS* promoter, or vasculature companion cell-specific *SUC2* promoter showed that miRNAs move bidirectionally between all three cell layers and that the number of cells across which a small RNA moves is determined in part by its abundance. Moreover, analysis of GFP fluorescence in the vasculature supports the idea that miRNA-mediated gene regulation is a dose-dependent process and depends on the relative abundance of the miRNA versus the target mRNA. No GFP silencing was observed in the symplically isolated stomata, suggesting that miRNA movement occurs via plasmodesmata, small channels that connect most plant cells.

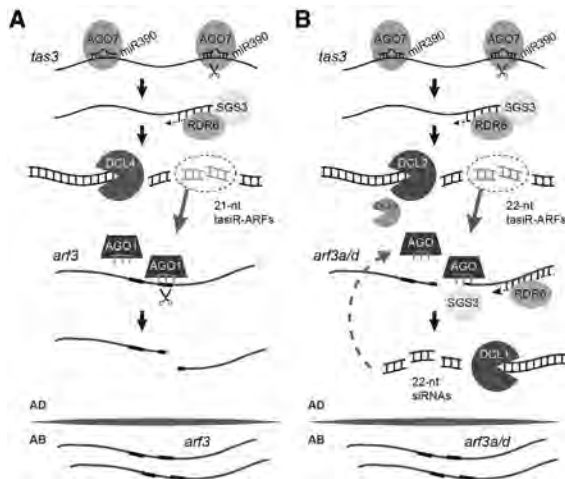
Interestingly, production of miR-GFP in young leaf primordia surrounding the shoot apical meristem

(SAM) resulted in silencing of GFP inside the SAM, demonstrating that miRNAs traffic from determinate organs into the shoot-stem-cell niche. In addition, miR-GFP was able to move from the vasculature below the SAM into the niche. These findings suggest that miRNAs may function as mobile signals between differentiated tissues and stem cells, possibly integrating environmental/physiological cues and plant development. Analyses of miRNA mobility in other developmental contexts are still ongoing, but the data obtained thus far suggest that small RNA mobility is developmentally regulated and follows rules that are distinct from those that govern protein trafficking. In addition, with the knowledge that miRNAs can traffic from the epidermis into underlying tissue layers, we are carrying out forward genetic screens to identify factors influencing this process.

### The START Domain Regulates HD-ZIPIII Activity and Organ Polarity

A. Husbands [in collaboration with V. Yong and H. Djaballah, Memorial Sloan-Kettering Cancer Center, New York]

On the basis of our previous observations regarding the expression and function of miR166 and tasi-ARF, it is evident that adaxial–abaxial patterning involves a cascade of positional signals. The mobile signals that pattern the newly formed leaf are distinct from those that maintain polarity during subsequent development. This project aims to identify additional signals in adaxial–abaxial patterning. The adaxial promoting HD-ZIPIII transcription factors contain a predicted START lipid-binding domain (Fig. 1). Modeling of the START domain of the HD-ZIPIII member PHABULOSA (PHB) suggests a structural similarity to human PC-TP, a phosphatidylcholine-binding START domain. This model and sequence conservation with other START domains was used to predict amino acids in PHB critical for START domain ligand binding. Analysis of *Arabidopsis* lines that express PHB-YFP fusion proteins with mutations in key residues lining the START ligand-binding pocket indicates that the START domain is required for proper HD-ZIPIII function. Such mutations block PHB-mediated target gene expression, indicating positive regulation of HD-ZIPIII activity by an unknown ligand.



**Figure 1.** Novel DCL1-dependent 22-nucleotide siRNAs bypass the loss of DCL4 in development. (A) In wild type, the 21-nucleotide tasiR-ARFs are generated via the two-hit *TAS3* tasiRNA pathway. Biogenesis of these small RNAs is confined to the adaxial/upper side of leaf primordia via the localized accumulation of miR390. tasiR-ARFs target *arf3* transcripts, limiting expression of these abaxial determinants to the lower side of developing leaves. (B) In *dcl4*, targeting of *TAS3* transcripts by miR390-AGO7 complexes proceeds on the adaxial side of leaf primordia as in wild type. However, upon loss of DCL4 activity, the double-stranded *TAS3* precursors are processed by DCL2/3 and this leads to the accumulation of 22-nucleotide tasiR-ARF variants rather than the normal 21-nucleotide species. The 22-nucleotide tasiR-ARFs trigger a unique one-hit mechanism of secondary siRNA biogenesis at the *arf3a* and *arf3d* transcripts specifically, leading to the processing of *arf3a/d* transcripts into 22-nucleotide siRNAs by DCL1. As a result, expression for both ARF3 abaxial determinants remains largely unaffected and the phenotype of *dcl4* mutants is suppressed in comparison to that of similar mutations in other *TAS3* tasiRNA pathway components.

Experiments to determine the mechanism through which the START domain controls HD-ZIPIII function are ongoing. Ligand binding does not appear to affect the subcellular localization of these transcription factors, as confocal imaging showed that the PHB-YFP derivatives correctly localize to the nucleus. Likewise, ligand binding is not essential for PHB dimerization. However, ligand binding does affect PHB's ability to bind DNA and possibly the interaction with protein partners. As part of these analyses, we developed a yeast-based assay and conducted a high-throughput screen for chemical compounds that bind the HD-ZIPIII START domain and modulate protein activity. Using yeast growth as readout, we identified compounds that reduced yeast growth (antagonists) and others that promoted growth (agonists).

Validation experiments in yeast and plants are under way to eliminate false positives and demonstrate that these chemicals act via modulation of HD-ZIPIII activity. In a parallel approach, we have generated plants expressing a biotin-tagged form of the START domain and will be performing immunoprecipitations and mass spectrometry to identify the endogenous ligand bound by HD-ZIPIII proteins. These latter experiments are under way in collaboration with the CSHL proteomics facility.

### tasiRNA Pathways in Maize and New Players in Leaf Polarity

M. Dotto [in collaboration with M. Hammell, Cold Spring Harbor Laboratory; M. Aukerman, M. Beatty, and R. Meeley, DuPont-Pioneer]

tasiRNAs are 21-nucleotide small RNAs that regulate the activity of key genes involved in plant growth and development. Mutations in maize *leafbladeless1 (lbl1)* that disrupt tasiRNA biogenesis give rise to plants with thread-like leaves that have lost top/bottom polarity. We used genomic approaches to identify *lbl1*-dependent small RNAs and their targets to determine the basis for these polarity defects. This revealed substantial diversity in small RNA pathways across plant species and identified unexpected roles for LBL1 in the regulation of repetitive elements within the maize genome. We further showed that only tasiRNA loci belonging to the *TAS3* family function in the maize vegetative apex. The *TAS3*-derived tasiR-ARFs are the main tasiRNA active in the apex, and misregulation of their *ARF3* targets emerges as the basis for the *lbl1* leaf polarity defects. Supporting this, we show that plants expressing *arf3a* transcripts insensitive to tasiR-ARF-directed cleavage recapitulate the phenotypes observed in *lbl1*. The *TAS3* tasiRNA pathway, including the regulation of *ARF3* genes, is conserved throughout land plant evolution, yet the phenotypes of plants defective for tasiRNA biogenesis are strikingly different. Our data lead us to propose that divergence in the processes regulated by the ARF3 transcription factors or the spatiotemporal pattern during development in which these proteins act underlies the diverse developmental contributions of this small RNA pathway across plants.

In a separate approach, we are working to identify additional small RNAs with a role in maize



development. Small RNAs associate with specific AGO proteins to form an RNA-induced silencing complex (RISC) that directs the cleavage and degradation of complementary mRNA targets. In *Arabidopsis*, AGO10 (AtAGO10) exclusively binds miR166, and this interaction is required for shoot apical meristem maintenance. However, interestingly, deep sequencing of immunoprecipitates of maize AGO10 (ZmAGO10), obtained using a peptide antibody that recognizes both AGO10 homologs, demonstrated that these AGO proteins bind both miR166 and miR394. In situ hybridization and laser microdissection–RNA-Seq analyses showed that miR394 accumulates in a gradient on the adaxial side of developing maize leaves, whereas its F-box targets accumulate in the complementary abaxial domain. This presents the possibility that organ polarity is regulated also by a fourth small RNA signal. Indeed, maize *ago10* double mutants display a downward-curved leaf phenotype, consistent with a role for this pathway in leaf development. We are currently characterizing single and double mutants for the remaining components of this pathway, *zma-miR394a* and *zma-miR394b*, as well as the two F-box protein targets. We are also analyzing how structural differences between these miRNA precursors and the AGO10 proteins in *Arabidopsis* and maize may account for the differential loading of miR394 in these species.

### Effects of Natural Variation Present in Maize Inbreds on the tasiRNA Pathway

K. Petsch [in collaboration with former postdoc Marie Javelle and M. Scanlon, Cornell University, Ithaca, New York; G. Muehlbauer, University of Minnesota, Minneapolis; J. Yu, Kansas State University, Manhattan; P. Schnable, Iowa State University, Ames]

Mutants that disrupt tasiRNA biogenesis exhibit defects in adaxial–abaxial patterning. Interestingly, the severity of these phenotypes is greatly dependent on the inbred background. tasiRNA pathway mutants, as well as other small RNA biogenesis mutants, typically exhibit weaker phenotypes when introgressed into B73 as opposed to other inbred backgrounds, such as W22, Mo17, and A619. To understand how inbred background influences expressivity of tasiRNA pathway mutants, we are using the severe *ragged1* allele of *leafbladeless1* (*lbl1-rgd1*). In the B73 inbred

background, *lbl1-rgd1* conditions a classical leaf polarity phenotype with most leaves developing as radial, abaxialized organs. In W22, however, *lbl1-rgd1* embryos lack a shoot meristem and germinate with just a root. We have generated an F<sub>2</sub> population to map loci that contribute to the variable expressivity of *lbl1-rgd1* in the two inbreds. This identified a major modifier locus on chromosome 8, which we have now fine-mapped to a region containing ~140 genes. In addition, we performed transcriptome analysis of *lbl1-rgd1* embryos and nonmutant siblings in B73 and W22. This identified 214 genes that are differentially expressed in *lbl1-rgd1* embryos in both backgrounds. These include known tasiRNA targets and genes showing differential expression across the adaxial–abaxial axis, as well as genes predicted to function in the embryonic shoot meristem. Most notably, ~30% of genes down-regulated in W22 mutant embryos have “meristem-associated” functions. Further investigation into specific genes is ongoing. Importantly, this study identified 40 genes within the chromosome 8 modifier region that are expressed in the developing embryo and show differential expression between the inbreds or between *lbl1-rgd1* and nonmutant siblings. And finally, this resource identified genes that are polymorphic between B73 and W22. By combining these approaches, we hope to gain a better understanding of the factors contributing to natural variation in tasiRNA pathway function.

### Novel DICER-LIKE1 siRNAs Bypass the Requirement for DICER-LIKE4 in Development

K. Petsch [in collaboration with O. Tam and M. Hammell, Cold Spring Harbor Laboratory; R. Meeley, DuPont-Pioneer; P. Manzotti and G. Consonni, University of Milan, Italy]

Dicer enzymes function at the core of RNA silencing to defend against exogenous RNA, or as an endogenous mechanism of gene regulation. Plant DICER-LIKE4 (DCL4) performs dual functions, acting in antiviral defense, as well as in development via the biogenesis of tasiR-ARFs. These small RNAs have an essential role in the grasses and act to spatially define the expression domain of AUXIN RESPONSE FACTOR3 (ARF3) transcription factors. However, contrary to tasiR-ARFs’ essential function in development, DCL4 proteins exhibit strong evidence of

recurrent adaptation typical of host factors involved in antiviral immunity. Here, we address how DCL4 balances its role in development with pressures to diversify in response to viral attack. We show that in contrast to other tasiR-ARF biogenesis mutants, *dcl4* null alleles condition an uncharacteristically mild phenotype, correlated with normal expression of select *arf3* targets. Loss of DCL4 activity yields a class of 22-nucleotide tasiR-ARF variants associated with the processing of *arf3* transcripts into 22-nucleotide secondary siRNAs by DCL1. Our findings uncover the presence of a novel DCL1-dependent siRNA pathway that bypasses the otherwise adverse developmental effects of DCL4 mutations (Fig. 1). This novel pathway is predicted to have important implications for DCL4's role in antiviral defense by reducing the selective constraints on DCL4 and allowing it to diversify in response to viral suppressors.

### High-Resolution Gene Expression Atlas for the Maize Shoot Apex

S. Knauer [in collaboration with former postdoc Marie Javelle, and M. Scanlon, Cornell University, Ithaca, New York; G. Muehlbauer, University of Minnesota, Minneapolis; J. Yu, Kansas State University, Manhattan; P. Schnable, Iowa State University, Ames]

Plants have the distinctive ability to form new organs throughout their lifetime, which can span hundreds or even thousands of years. The growing tip of a plant contains a population of stem cells that are located within a specialized niche, termed the shoot apical meristem (SAM). These stem cells divide to maintain the SAM and to generate daughter cells from which lateral organs, such as leaves and flowers, arise. The gene regulatory networks controlling meristem indeterminacy and organogenesis remain largely obscure. To gain insight into such networks, we have generated a high-resolution gene expression atlas for the maize shoot apex. Functional domains of the meristem were isolated by laser microdissection and analyzed by RNA deep sequencing. These include

the whole meristem, the stem cell-containing SAM tip, the newly initiating leaf (P0), the epidermal layer of the SAM (L1), and the subepidermal region (L2). Because gene networks controlling meristem maintenance and leaf development are largely interconnected, we further compared expression profiles in an ontogenic series of leaf primordia, P1, P2, and P3. Using this extensive data set, we were able to identify genes that mark specific cell types or are differentially expressed between distinct tissue types. Using CAST clustering, we further identified genes that show dynamic expression profiles during the transition from a stem cell to an organ primordium. For instance, we identified numerous genes that specifically mark the stem cell domain of the meristem and several genes that mark the presumptive organizing center. Most notably, our analysis identified unique transcription factor signatures that, next to the interplay of hormones, distinguish stem cells from their differentiating descendants within the SAM. Reverse genetic resources available for maize are being used to directly test the function of select genes in these processes. This has identified several new mutants affecting meristem maintenance and leaf morphology. Overall, our findings present a novel approach to modulating plant architecture by manipulating master regulators of cell identity.

### PUBLICATIONS

- Benkovics AH, Timmermans MCP. 2014. Developmental patterning by gradients of mobile small RNAs. *Curr Opin Genet Dev* **27**: 83–91.
- Dotto MC, Petsch KA, Aukerman MJ, Beatty M, Hammell M, Timmermans MCP. 2014. Genome-wide analysis of *leafbladeless1*-regulated and phased small RNAs underscores the importance of the *TAS3* ta-siRNA pathway to maize development. *PLoS Genet* **10**: e1004826.
- Li L, Eichten SR, Shimizu R, Petsch K, Yeh CT, Wu W, Chetoor AM, Givan SA, Cole RA, Fowler JE, Evans MM, et al. 2014. Genome-wide discovery and characterization of maize long non-coding RNAs. *Genome Biol* **15**: R40.
- Thompson AM, Crants JE, Schnable PS, Yu J, Timmermans MCP, Springer NM, Scanlon MJ, Muehlbauer GJ. 2014. Genetic control of maize shoot apical meristem architecture. *G3* **4**: 1327–1337.

There has been a growing appreciation in recent years that gene function is frequently context-dependent, with a large part of that context provided by the activities of other genes. But trying to understand how genes interact to produce function is a hugely complicated problem and one that appears likely to become more so as genomic information becomes more detailed. **Jesse Gillis** and colleagues are computational biologists who are presently challenging an oft-taken approach to the problem in which the functions of genes are interpreted in the context of networks derived from gene association data. Such networks consist of millions of interactions across thousands of genes, derived from protein-binding assays, RNA coexpression analysis, and other sources. Historically, many attempts to understand gene function through networks have leveraged a biological principle known as “guilt by association.” It suggests that genes with related functions tend to share properties (e.g., physical interactions). In the past decade, this approach has been scaled up for application to large gene networks, becoming a favored way to grapple with the complex interdependencies of gene functions in the face of floods of genomics and proteomics data. Gillis’ work centers on identifying the limits of the approach and making fundamental improvements to its operation, as well as applying those improvements to neuropsychiatric gene network data.

**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 (ENCyclopedia of DNA Elements) 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.

**Gholson Lyon’s** lab focuses on analyzing human genetic variation and its role in severe neuropsychiatric disorders and rare diseases, including intellectual disability, autism, and schizophrenia. By recruiting large groups of related individuals living in the same geographic location (e.g., Utah), Lyon’s lab can study the breadth and depth of genetic variants in a similar environmental background. Using the exome—the parts of the genome that code for protein—and whole-genome sequencing, his lab looks for mutations that segregate with syndromes in the various populations. A second focus of the Lyon lab is to study the mechanistic basis of a new rare disease that they described in 2011. This is the first human disease involving a defect in the amino-terminal acetylation of proteins, a common modification of eukaryotic proteins carried out by amino-terminal acetyltransferases (NATs). The team has been using several different cellular model systems to better understand the disease pathophysiology and the basic process of amino-terminal acetylation.

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 sequence), the fission yeast *Schizosaccharomyces pombe*, as well as the human genome and 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** lab 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 lab 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 demands of increasing population and a changing environment. The lab also has brought fully sequenced genomes into an integrated data framework, to enhance the power of their comparative studies. This past year, Ware was named as its 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.

# GENE NETWORK ANALYSIS

J. Gillis S. Ballouz W. Verleyen

Research in our lab centers on using computational methods to understand gene function. Computational biology has taken up the challenge of determining gene function mainly by attempting to interpret the activities of genes in the context of networks derived from gene association data. As data sets characterizing genes grow in size and complexity, it seems self-evident that computation can assist in inference as to gene function. Gene network analysis intended to provide insight into complex disorders is a dominant interest in the field. Gene associations (of various sorts) are believed to encode functional interaction, and this interaction is frequently shown to be able to substantially predict gene function across all functions. This approach, commonly called “Guilty by Association,” is embedded in everything from prioritization of *de novo* variants to uncovering novel molecular phenotypes or mechanisms of disease. Our research focuses on identifying limitations in gene network analysis and making fundamental improvements to its operation for the interpretation of neuropsychiatric genomics data. We are developing gene network resources that improve both the sophistication and breadth of data available for determining functional convergence among disease-causal genes. We develop carefully calibrated networks from data that are specific to age, species, sex, tissue, and disease state, as well as machine learning tools customized to their use for neuropsychiatric analysis. In addition to postdocs Wim Verleyen and Sara Ballouz, Maggie Crow joined us in 2015.

## Machine Learning Functional Properties of Genes

---

Increasingly, biologists have turned to computational methods to sift through the vast array of preexisting genomics data for validation that a candidate variant has a molecular role in the phenotype of interest or to prioritize a candidate as disease-causal. These

computational methods usually fit under the rubric of “machine learning,” using network data that represent the interaction of genes or their products.

The use of our burgeoning genomics data to characterize gene function has been one of the central research objectives of computational biology in the postgenomics era. Despite this intense focus, the abundance of public data, and methods development in machine learning and “big data” analysis, progress has been surprisingly uncertain. Attaching novel functions to genes based on prior data remains difficult, and a large number of genes still have comparatively little information attached to them. This is a critically important problem in evaluating disease candidate variants, many of which involve otherwise unstudied genes. To better understand the forces driving algorithmic and data performance, developers have focused on comparative assessment. Despite some important progress enabled by these assessments, the underlying mechanisms driving methods performance are poorly understood. Each developer may use separate data resources, different algorithms, and quite specific means for parsing the data, making it quite hard to explore factors affecting performance. Field-wide progress is hampered by our inability to know why a specific algorithm worked on a particular data set for a potentially narrow task.

To overcome this lack of detail in assessment, we developed a computational infrastructure that allows us to implement representative samples of cutting-edge machine learning algorithms. These algorithms can then be customized to specific biological problems and, when benchmarked against prior data, obtain very high levels of performance. Moreover, we learn more by combining them through meta-analysis than we do from their individual results, as is often the case in machine learning. Because we now have in-house versions of multiple algorithms, we can explore variance and data dependencies in a way not typically possible, leading to several useful heuristics for meta-analysis that we have recently published (Verleyen et al. 2014).

We have focused on data based on networks derived from protein–protein interaction, sequence similarity, aggregated coexpression, and semantic similarity to study underlying patterns of performance. Our infrastructure allows us to characterize in detail why aggregation improves performance, where results are robust and reproducible, and what artifacts are potentially problematic in data interpretation. We plan to make this replication across methods available as a public resource, allowing geneticists using function prediction methods to better assess reproducibility and data dependencies in target genes of interest.

### Condition-Specific RNA-Seq Coexpression Networks

---

The idea of condition-specific networks exists in a number of forms within the field of systems biology, most prominently as a way of capturing network dynamics; creating multiple condition-specific networks is a way of capturing how interactions between genes or their products change. Network construction depends on data type, but it usually takes the form of profiles of genes across a range of conditions. Our interest in such data is in determining not only that disease genes interact, but also under what conditions they do so. To assess this, we exploit RNA-Seq measures of expression activity from diverse prior data.

RNA-Seq offers profound biological and technical advantages over microarray technologies, such as detecting the whole transcriptome and an improved dynamic range. However, as was once the case for microarrays, RNA-Seq's utility is hobbled by our inability to determine sound consensus standards. One major reason is that RNA-Seq data have quite different statistical properties from those of microarray data, and thus the same principles do not apply for differential expression. Important questions about data quality, such as the role of biological noise in experimental design, are not readily addressed. A broader characterization of RNA-Seq data through meta-analysis offers one avenue to investigate problems of this nature.

Although differential expression analysis is a more common means for interpreting transcriptomic data, coexpression analysis is far more routine in the context of meta-analysis, with thousands of expression

profiles aggregated to generate robust signatures using repurposed data. One advantage of coexpression methods is that it already tends to require meta-analysis of disparate data sets with quite different properties, so approaches for aggregating across data are already common. We have developed a series of methods appropriate for RNA-Seq meta-analysis as well as producing reference networks for general use. RNA-Seq coexpression poses novel statistical and bioinformatics challenges; we have identified major confounds and developed appropriate control experiments necessary for network construction, laying the groundwork for functional analyses into RNA-Seq coexpression. Our analysis platform allows us to easily repurpose preexisting raw data, to construct large-scale aggregate coexpression networks from a variety of conditions, and to use both RNA-Seq and microarray data.

### Bias Tradeoffs in the Creation and Analysis of Protein–Protein Interaction Networks

---

Networks constructed from aggregated protein–protein interaction data are commonplace in biology. But the studies from which these data are derived were conducted with their own hypotheses and foci. Focusing on data from budding yeast present in BioGRID, we determined that many of the downstream signals present in network data are significantly impacted by biases in the original data. We determined the degree to which selection bias in favor of biologically interesting bait proteins goes down with study size, and we also find that promiscuity in prey contributes more substantially in larger studies. We analyze interaction studies over time with respect to data in the Gene Ontology and find that reproducibly observed interactions are less likely to favor multifunctional proteins. We find that strong alignment between coexpression and protein–protein interaction data occurs only for extreme coexpression values, and we use these data to suggest candidates for targets likely to reveal novel biology in follow-up studies (Gillis et al. 2014).

### Ligand-Binding Networks

---

With the development of large-scale target-ligand databases, and new cheminformatics metrics, it is

now feasible to directly correlate ligand similarities between proteins to their bioinformatic or proteomic similarities. We hoped that by investigating this alternate source of network data, new bridging ligands may be found for targets whose activity is perturbed in disease. In collaboration with the Shoichet lab at the University of California, San Francisco, we calculated similarities for 1131 proteins, each of which had known ligands, and compared ligand-based similarities to those derived from sequences, from protein–protein interaction networks, and from coexpression networks. Surprisingly, the proteomic associations were often orthogonal to the ligand-based associations; two targets that were intensely related by protein–protein interaction or by coexpression only rarely recognized similar ligands, and most pairs of targets that bound similar ligands were unrelated proteomically. Even sequence similarity was an unreliable predictor of shared ligand recognition among targets: Most related sequences had only a 55% chance of sharing similar ligands, and 29% of the targets that recognized similar ligands had no significant sequence similarity. Still, there were regions of high proteomic similarity that overlapped with high ligand similarity and that were coannotated for diseases. We sought single ligands that could bridge disease-linked targets, where no such bridging ligand was known. On experimental testing, we found such bridging ligands for seven pairs of targets, previously unknown to share ligands but linked proteomically in neuropsychiatric diseases.

### Improved Quality Control for RNA-Seq

---

Recent analyses of RNA-Seq have emphasized the value added by experimental designs with more samples, permitting better estimation of biological and technical variability. Unfortunately, this does not automatically translate into an experimental design decision for more replicates (biological or technical), largely because of the relatively high cost of RNA-Seq. Study design is further complicated by complex dependencies on the platform chosen, library preparation, and normalization methods. All of these factors affect the concordance of downstream differential expression analysis. Although researchers must be aware of the pitfalls that can arise when designing their experiment, they receive only modest feedback

in the experimental data itself as to whether their choices were successful. Many optimized protocols or best practices exist, but there are few means for assessing the quality of a given transcriptomic study, particularly outside of purely technical concerns and where focusing on novel biology.

In our recent work, we have demonstrated a straightforward but surprisingly powerful method to measure replicability by modeling the effects of noise within observed coexpression. Because of the breadth and extent of coexpression, it provides a sensitive, yet general means of performing a quality control. We begin our analysis by performing some sample and replicate-based analyses of RNA-Seq quality to establish general properties, then we show how common heuristics for quality control appear in a coexpression-based framework, and finally we provide a means of directly quantifying this by disruption of known coexpression pairs. By providing a direct quality control measure, we hope to make it straightforward to customize experiments to do better than a general heuristic—such as discarding all low-expressing genes—would normally permit. We plan to make our software available in a convenient form for bioinformatic use.

### De Novo Variant Network Analysis

---

One of the principal applications to which we put machine learning and coexpression networks is in looking for functional convergence among *de novo* variants. For the last several months, we have been relying on the recently published Simons simplex collection autism variant list to determine features that predict disease properties among genes. The key to our approach is to look for features that distinguish not only the disease candidates from other genes, but also high-quality disease candidates from low-quality ones. It is important to acknowledge that even these high-quality gene sets have false positives (perhaps as many as 50% of the total) with characteristic features; thus, it is not sufficient simply to look for points of convergence, because the false positives, too, have overlapping features. Instead, as our coexpression data have expanded, we have found that we can predict which candidates will recur in the population (true positives) based on their coexpression patterns.

## PUBLICATIONS

- Gillis J, Ballouz S, Pavlidis P. 2014. Bias tradeoffs in the creation and analysis of protein-protein interaction networks. *J Proteomics* **100**: 44–54.
- Mazurek A, Park Y, Miething C, Wilkinson JE, Gillis J, Lowe SW, Vakoc CR, Stillman B. 2014. Acquired dependence of acute myeloid leukemia on the DEAD-box RNA helicase DDX5. *Cell Rep* **7**: 1887–1899.
- McCarthy SE, Gillis J, Kramer M, Lihm J, Yoon S, Berstein Y, Mistry M, Pavlidis P, Solomon R, Ghiban E, et al. 2014. De novo mutations in schizophrenia implicate chromatin remodeling and support a genetic overlap with autism and intellectual disability. *Mol Psychiatry* **19**: 652–658.
- In Press*
- Verleyen W, Ballouz S, Gillis J. 2014. Measuring the wisdom of the crowds in network-based gene function inference. *Bioinformatics* (in press).



# GENOME ORGANIZATION AND REGULATION AND FUNCTIONAL ROLES OF NONCODING RNAs

T.R. Gingeras    P. Batut                    J. Drenkow                    F. Schlesinger  
                         S. Chakraborty    G. Nechooshtan    L.-H. See  
                         C. Davis            A. Prakash            H. Xue  
                         A. Dobin            A. Scavelli            C. Zaleski

## Human ENCODE Project

P. Batut, C. Davis, A. Dobin, J. Drenkow, A. Scavelli, L.-H. See, C. Zaleski

A total of 173 human primary cell line and tissue samples (done in replica) have been analyzed for long RNA (>200 nucleotides), short RNA (<200 nucleotides), and RAMPAGE (5' cap sites) this year. These and all generated for this phase of ENCODE data are available at (<https://www.encodeproject.org/search/?type=experiment&lab.title=Thomas%20Gingeras,%20CSHL>). As part of the model genomes ENCODE (mod-ENCODE), we and our collaborators from the University of Connecticut, University of California, Berkeley, University of Indiana, RIKEN, Harvard University, and the Rockefeller University have identified new genes, transcripts, and proteins using poly(A)<sup>+</sup> isolated from cultured cell lines and dissected organ systems from *Drosophila melanogaster*. We found that a small set of mostly neuron-specific genes has the potential to encode thousands of transcripts each through extensive alternative promoter usage and RNA splicing. The magnitudes of splicing changes were seen to be larger between tissues than between developmental stages, and most sex-specific splicing is gonad-specific. Gonads expressed hundreds of previously unknown coding and long noncoding RNAs (lncRNAs), some of which are antisense to protein-coding genes and produce short regulatory RNAs. Furthermore, previously identified pervasive intergenic transcription occurs primarily within newly identified introns. The fly transcriptome is substantially more complex than previously recognized, with this complexity arising from combinatorial usage of promoters, splice sites, and polyadenylation sites. Using these and other data consisting of matched RNA-sequencing data from human, worm, and fly that we and our

collaborators in the ENCODE and mod-ENCODE consortia have generated has allowed comparison across metazoan phyla, extending beyond earlier within-phylum transcriptome comparisons and revealing ancient conserved features. Specifically, we discovered coexpression modules shared across animals, many of which are enriched in developmental genes. Moreover, we used expression patterns to align the stages in worm and fly development and find a novel pairing between worm embryo and fly pupae, in addition to the embryo-to-embryo and larvae-to-larvae pairings. Furthermore, we found that the extent of noncanonical, noncoding transcription is similar in each organism, per base pair. Finally, we find in all three organisms that the gene-expression levels, both coding and noncoding, can be quantitatively predicted from chromatin features at the promoter using a “universal model” based on a single set of organism-independent parameters.

## Comparisons of Transcriptomes of Human and Mouse

C. Davis, A. Dobin, C. Zaleski

The laboratory mouse shares the majority of its protein-coding genes with humans, making it the premier model organism in biomedical research—yet the two mammals differ in significant ways. To gain greater insight into both shared and species-specific transcriptional and cellular regulatory programs in the mouse, the Mouse ENCODE Consortium has mapped transcription, DNase I hypersensitivity, transcription factor binding, chromatin modifications, and replication domains throughout the mouse genome in diverse cell and tissue types. By comparing with the human genome, we not only confirmed substantial conservation in the newly annotated potential

functional sequences, but also found a large degree of divergence of sequences involved in transcriptional regulation, chromatin state, and higher-order chromatin organization. To better understand these two species on a molecular level, we and our Mouse ENCODE collaborators performed a comparison of the expression profiles of 15 tissues by deep RNA sequencing and examined the similarities and differences in the transcriptome for both protein-coding and -noncoding transcripts. Although commonalities are evident in the expression of tissue-specific genes between the two species, the expression for many sets of genes was found to be more similar in different tissues within the same species than between species. However, comparison with transcriptome profiles in human cell lines also revealed substantial conservation of transcriptional programs and uncovered a distinct class of genes with levels of expression that have been constrained early in vertebrate evolution. This core set of genes captures a substantial fraction of the transcriptional output of mammalian cells and participates in basic functional and structural housekeeping processes common to all cell types. Perturbation of these constrained genes was associated with significant phenotypes including embryonic lethality and cancer. Evolutionary constraint in gene expression levels was not reflected in the conservation of the genomic sequences, but it was associated with conserved epigenetic marking, as well as with the characteristic posttranscriptional regulatory program, in which subcellular localization and alternative splicing have comparatively large roles.

## Database and Software Development

A. Dobin, C. Zaleski

The infrastructure of our laboratory's analysis programs is used for the continued development and maintenance of the large databases collected as part of the human, mouse, and model genome (mod) ENCODE projects. The notable achievements in this area include the following:

### Database Development

- Continued management and organization of sequencing data consisting of raw fastq files and genome-mapped data files. These include bam, bigwig, and other formats in a growing repository almost 30 TB in size.

- Continued execution of a sequencing production pipeline. This consists of sequence mapping into multiple formats, multiple QC steps, and file transfer to outside locations.
- Began implementation of new software tools for annotation-agnostic analysis of sequencing data. Completed part 1, consisting of a program that uses a recursive algorithm to generate "intervals" or "regions" of expression from RNA-Seq data.
- Managed our primary storage hardware to accommodate >40 TB (terabytes) of data that were migrated to a new hardware system and connected to the shared computing cluster as well as our own lab's servers.
- Established an archival storage tier that organized and migrated almost 25 TB of data to the archive tier. Also implemented a shared storage tracking system to monitor usage of live and archived data.

### Software Development

- Developed standardized RNA-Seq pipelines for the ENCODE consortium. The long and small RNA-Seq, as well as RAMPAGE, pipelines have been transferred to the ENCODE DCC and are ready for uniform processing of the RNA-Seq data.
- Developed and supported the STAR RNA-Seq aligner. STAR's popularity is rising: The paper has been cited ~230 times in the 2 years since publication in 2013, with hundreds of users using the software (~700 unique users in GitHub repository).
- Developed comprehensive RNA-Seq analysis software STARtools. Although numerous tools have been developed for performing separate analysis tasks, linking them into a processing pipeline is a demanding job due to peculiarities of the file formats, necessity of maintaining an assortment of executables, and the wide range of required computational resources. To resolve these difficulties, we have developed STARtools, a software suite for comprehensive analysis of RNA-Seq data, which can perform a variety of common analysis tasks in a single pass. STARtools is based on our RNA-Seq mapper STAR, which is gaining popularity because of its high mapping speed and precision. Conversion tool converts alignments into various output formats, which can be visualized in genome browsers. Quantification tool quantifies gene and isoform expression using highly efficient and accurate

algorithms, allowing for differential expression assessment. In addition, it outputs the alignments in the transcriptomic coordinates, which can be input into popular quantification software such as RSEM or eXpress. Chimeric detection tool discovers and quantifies chimeric (fusion) transcripts, and it is also capable of discovering circular RNAs and shuffled exons. Long read tool is used for extracting full-length transcripts from Pacific Biosciences and Oxford Nanopore data. We will show that combining mapping of the RNA-Seq data with the downstream processing results in a substantial reduction of required computational resources as compared to the consecutive implementation, and at the same, it yields same or better accuracy as the existing tools.

- Developed a pipeline for mapping and annotating PacBio Circular Consensus reads. The full-length transcript reconstruction from data obtained from Pacific Biosciences and Oxford Nanopore technologies has been a key goal of this year's efforts. Because of the high error rate, Oxford Nanopore reads require a significantly revised mapping strategy, which is currently under development. Using long reads allows for improvements of existing annotations and predicting novel genes and isoforms in hundreds of human genome loci.

## Evolutionary Dynamics of the Genomic Elements That Regulate Gene Expression

P. Batut, A. Dobin

Animal development is orchestrated by the unfolding of regulatory programs that dynamically specify the expression patterns of thousands of protein-coding genes and noncoding transcripts. It has long been thought that changes to these regulatory programs drove the evolution of metazoans, and yet little is known about the evolutionary dynamics of the genomic elements that regulate gene expression. To experimentally study the evolution of transcriptional promoters, we generated genome-wide profiles of promoter activity throughout embryonic development in five *Drosophila* species spanning ~25 million years of divergence (*D. melanogaster*, *D. simulans*, *D. erecta*, *D. ananassae*, *D. pseudoobscura*). We found that promoter gain and loss have been very active processes throughout the evolutionary history of these species. Developmental gene expression profiles tend to be

tightly conserved, and regulatory divergence is shaped by systems-level constraints on gene function and developmental stages. We detected extremely prevalent noncoding transcription throughout embryogenesis: 4050 lncRNA promoters, most of which have never been described before, are dynamically expressed during that critical period. We identified a *Drosophila* core set of more than 1000 lncRNA promoters conserved over 25 million years and showed that they are under substantial purifying selection at the levels of promoter sequence and expression specificity. This suggests that noncoding transcription in *Drosophila* not only is prevalent, but also fulfills essential functions during embryonic development. lncRNA promoters do, however, evolve substantially faster than their protein-coding gene counterparts. The observation that lineage-specific lncRNA promoters have recently come under strong purifying selection suggests that noncoding transcription is a major driving force behind the adaptive evolution of *Drosophila* species. We are currently conducting functional studies on the FBgn0264479 locus, which harbors an lncRNA that is expressed extremely highly, and in a spatially restricted fashion, during a 3–4-h period around the onset of gastrulation.

## Extracellular Vesicles Containing Processed Forms of RNY5 Potentially Contribute to Tumor Microenvironment

S. Chakraborty, A. Prakash

Tumor microenvironments are critical to the establishment and growth of tumors. Extracellular vesicles (EVs) have been hypothesized as a means to establish and shape such environments. Explorations of the diversity of extracellular vesicles were conducted to evaluate their biological functionalities. EVs can be loosely grouped based on their size distributions into exosomes, microvesicles, and apoptotic bodies. To improve our ability to isolate and enrich for various types of EVs, we tested and developed more viable methods of exosome isolation for small RNA-sequencing experiments. Through small RNA sequencing of their respective RNA cargo, we explored the activity of cells sending out significantly conserved messages as RNA using any of these types of vesicles as a carrier. Evidence for the biological relevance of the EV-mediated communication of RNAs was observed when cultured human primary cells were exposed to EVs from each

of five human cancer cell lines of distinct lineages. Rapid apoptosis of the primary cells is seen as a result, whereas cancer cells treated with EVs from primary or cancer cells do not trigger this response. Analysis of the RNA cargos highlighted the prominence of RNY5 and its fragments. Clearly distinct expression patterns within and outside cancer and primary cells of the RNY5 RNAs and their fragments were observed. Overrepresentation and pathway analysis were developed to understand the pathways triggered leading to the apoptotic phenotype in primary cells. Results from these studies are being assembled for publication.

## In Vitro Processing of RNY5

G. Nechooshtan

Full-length fragments of the RNY5 transcript are seen in EVs. Verification of the in vivo steady-state levels of the full-length form of RNY5 found in cells and EVs was established, as was the relative abundance of processed forms of the transcript by northern hybridization analysis. EVs were found to be enriched for the processed forms, whereas whole cells harbor exclusively the full-length transcript. Using whole-cell and EV protein extracts for in vitro processing of a synthetic full transcript revealed that although the whole-cell extract is able to protect the transcript from degradation by ribonucleases, EV protein extracts efficiently process the full-length RNY5 into typical shorter species found in vivo. These shorter species then remain protected from further degradation. To identify the cellular factors that have a role in the phenotype exerted by RNY5 in recipient cells and in its processing and targeting to EVs, proteomic analysis of binding partners of RNY5 were carried out by RNA-protein pulldowns with complete and processed forms of RNY5 acting as bait. In addition to known binding counterparts of Y5 RNA, this analysis revealed several other candidate factors that will be further explored.

## PUBLICATIONS

Abdelhamid RF, Plessy C, Yamauchi Y, Taoka M, de Hoon M, Gingeras TR, Isobe T, Carninci P. 2014. Multiplicity of 5' cap structures present on short RNAs. *PLoS ONE* **9**: e102895.

Andersson R, Gebhard C, Miguel-Escalada I, Hoof I, Bornholdt J, Boyd M, Chen Y, Zhao X, Schmid C, Suzuki T, et al. 2014. An atlas of active enhancers across human cell types and tissues. *Nature* **507**: 455–461.

Arner E, Forrest AR, Ehlund A, Mejhert N, Itoh M, Kawaji H, Lassmann T, Laurencikiene J, Rydén M, Arner P, FANTOM Consortium. 2014. Ceruloplasmin is a novel adipokine which is overexpressed in adipose tissue of obese subjects and in obesity-associated cancer cells. *PLoS ONE* **9**: e80274.

Bassett AR, Akhtar A, Barlow DP, Bird AP, Brockdorff N, Duboule D, Ephrussi A, Ferguson-Smith AC, Gingeras TR, Haerty W, Higgs DR, et al. 2014. Considerations when investigating lncRNA function in vivo. *Elife* **3**: e03058.

Brown JB, Boley N, Eisman R, May GE, Stoiber MH, Duff MO, Booth BW, Wen J, Park S, Suzuki AM, et al. 2014. Diversity and dynamics of the *Drosophila* transcriptome. *Nature* **512**: 393–399.

Cheng Y, Ma Z, Kim BH, Wu W, Cayting P, Boyle AP, Sundaram V, Xing X, Dogan N, Li J, et al. 2014. Principles of regulatory information conservation between mouse and human. *Nature* **15**: 371–375.

Fagegaltier D, König A, Gordon A, Lai EC, Gingeras TR, Hannon GJ, Shcherbata HR. 2014. A genome-wide survey of sexually dimorphic expression of *Drosophila* miRNAs identifies the steroid hormone-induced miRNA let-7 as a regulator of sexual identity. *Genetics* **98**: 647–668.

FANTOM Consortium and the RIKEN PMI and CLST (DGT), Forrest AR, Kawaji H, Rehli M, Baillie JK, de Hoon MJ, Haberer V, Lassmann T, Kulakovskiy IV, Lizio M, Itoh M, et al. 2014. A promoter-level mammalian expression atlas. *Nature* **507**: 462–470.

Gerstein MB, Rozowsky J, Yan KK, Wang D, Cheng C, Brown JB, Davis CA, Hillier L, Sisu C, Li JJ, et al. 2014. Comparative analysis of the transcriptome across distant species. *Nature* **512**: 445–448.

Hasegawa Y, Tang D, Takahashi N, Hayashizaki Y, Forrest AR; FANTOM Consortium, Suzuki H. 2014. CCL2 enhances pluripotency of human induced pluripotent stem cells by activating hypoxia related genes. *Sci Rep* **4**: 5228.

Kellis M, Wold B, Snyder MP, Bernstein BE, Kundaje A, Marinov GK, Ward LD, Birney E, Crawford GE, Dekker J, et al. 2014. Reply to Brunet and Doolittle: Both selected effect and causal role elements can influence human biology and disease. *Proc Natl Acad Sci* **111**: E3366.

Kellis M, Wold B, Snyder MP, Bernstein BE, Kundaje A, Marinov GK, Ward LD, Birney E, Crawford GE, Dekker J, et al. 2014. Defining functional DNA elements in the human genome. *Proc Natl Acad Sci* **111**: 6131–6138.

Lin S, Lin Y, Nery JR, Urich MA, Breschi A, Davis CA, Dobin A, Zaleski C, Beer MA, Chapman WC, Gingeras TR, et al. 2014. Comparison of the transcriptional landscapes between human and mouse tissues. *Proc Natl Acad Sci* **111**: 17224–17229.

Morikawa H, Ohkura N, Vandenbon A, Itoh M, Nagao-Sato S, Kawaji H, Lassmann T, Carninci P, Hayashizaki Y, Forrest AR, et al. 2014. Differential roles of epigenetic changes and Foxp3 expression in regulatory T cell-specific transcriptional regulation. *Proc Natl Acad Sci* **111**: 5289–5294.

Yue F, Cheng Y, Breschi A, Vierstra J, Wu W, Ryba T, Sandstrom R, Ma Z, Davis C, Pope BD, et al. 2014. A comparative encyclopedia of DNA elements in the mouse genome. *Nature* **515**: 355–364.

*In Press*

Pervouchine DD, Djebali S, Breschi A, Davis CA, Barja PP, Dobin A, Tanzer A, Lagarde J, Zaleski C, See LH, et al. 2015. Enhanced transcriptome maps from multiple mouse tissues reveal evolutionary constraint in gene expression. *Nat Commun* **6**: 5903.

# HUMAN GENETICS AND GENOMIC MEDICINE

G. Lyon M. Doerfel J. O’Rawe  
P. Kota Y. Wu

Our lab focuses on analyzing human genetic variation and its role in severe neuropsychiatric disorders and rare diseases, including intellectual disability, autism, and schizophrenia. By recruiting large groups of related individuals living in the same geographic location (e.g., Utah), we can study the breadth and depth of genetic variants in a similar environmental background. Using the exome—the parts of the genome that code for protein—and whole-genome sequencing, we look for mutations that segregate with syndromes in the various populations and undertake comprehensive functional studies of many of the newly identified mutations.

## Ogden Syndrome and the Amino-Terminal Acetylation of Proteins

M. Doerfel, Y. Wu [in collaboration with R. Marmorstein, Philadelphia, Pennsylvania; T. Arnesen and N. Reuter, Norway; and P. van Damme, Belgium]

---

More than 85% of human proteins are acetylated at their amino-terminal amino group; hence, amino-terminal acetylation (NTA) is one of the most abundant modifications of eukaryotic proteins. Despite its discovery more than 30 years ago, very little is known about the cellular effects/functions of this modification. In humans, six distinct amino-terminal amino-acetyltransferases (NATs) catalyze the transfer of an acetyl group from acetyl-CoA to the amino-terminal amino group of their specific target proteins. The major human acetyltransferase, NatA, consists of an auxiliary subunit, Naa15, and a catalytic active subunit, Naa10. We have previously described two families with a lethal X-linked disorder of infancy called Ogden syndrome. This disorder comprises a distinct combination of an aged appearance, craniofacial anomalies, hypotonia, global developmental delays, cryptorchidism, and cardiac arrhythmias. Using X-chromosome exon sequencing, we identified a c.109T > C (p.Ser37Pro) variant in Naa10 as contributing to this disease. Biochemical analysis and immunoprecipitation assays in combination with

liquid chromatography/mass spectrometry (LC/MS) demonstrated a reduced catalytic capacity and revealed an impaired binding of the S37P mutant toward specific interaction partners, including Naa15 and Naa50. Analysis of the amino-terminal acetylation of patient cells revealed a decreased acetylation of a subset of NatA substrates, indicating that a reduced binding capability and an affected enzymatic activity of the Naa10 S37P mutation is a prominent feature in Ogden syndrome. Characterization of *NAA10/NAA15* knockout yeast strains revealed various phenotypes, including growth defects at elevated temperatures and altered sensitivity toward cytotoxic stresses. These effects could be rescued by overexpressing human wild-type Naa15/Naa10 from plasmids; however, overexpressing mutant Naa15/Naa10 S37P only partially rescues these effects. Interestingly, introduction of both a human Naa15/Naa10 wild type and S37P mutant into the endogenous locus of the corresponding yeast genes failed to reverse the effects. We also continued our efforts with establishing induced pluripotent stem cells (iPSCs) from skin fibroblasts from one of the boys with Ogden syndrome, and we are also establishing knock-in mice containing the mutation of interest in *NAA10*. Ongoing work will focus on characterizing the cells and mice.

## RykDax Syndrome: Characterization and Analysis of an Idiopathic Intellectual Disability Syndrome

J. O’Rawe, Y. Wu [in collaboration with H. Fang, L. Jimenez Barron, Cold Spring Harbor Laboratory; E. Yang, Boston; A. Rope, Oregon; J. Swensen, Arizona; R. Robison, Utah; and K. Wang, California]

---

We worked on the discovery of a new genetic syndrome, RykDax syndrome, driven by a whole-genome sequencing (WGS) study for one family from Utah. This family had two affected brothers presenting with severe intellectual disability (ID), a characteristic intergluteal crease, and very distinctive facial features, including a broad,

upturned nose, sagging cheeks, downward sloping palpebral fissures, prominent periorbital ridges, deep-set eyes, relative hypertelorism, thin upper lip, a high-arched palate, prominent ears with thickened helices, and a pointed chin. Illumina-based WGS was performed on 10 members of this Caucasian family, with additional complete genomics-based WGS performed on the nuclear portion of the family (mother, father, and the two affected males). Using WGS data sets from 10 members of this family, we could increase the reliability of the biological inferences with an integrative bioinformatic pipeline. In combination with insights from clinical evaluations and medical diagnostic analyses, the DNA sequencing data were used in the study of three plausible genetic disease models that might uncover genetic contribution to the syndrome. We found a two- to fivefold difference in the number of variants detected as being relevant for various disease models when using different sets of sequencing data and analysis pipelines. We derived greater accuracy when more pipelines were used in conjunction with data encompassing a larger portion of the family, with the number of putative *de novo* mutations being reduced by 80%, because of false-negative calls in the parents. The boys carry a maternally inherited missense variant in the X-chromosomal gene *TAF1*, which we consider as disease relevant. We took a “genotype-first” approach to find other families with variants in *TAF1* and containing individuals having a remarkably similar clinical presentation. *TAF1* is the largest subunit of the general transcription factor IID (TFIID) multiprotein transcription complex, and our results have implicated mutations in *TAF1* as having a critical role in the development of this new intellectual disability syndrome. We published the mutation as part of a paper developing SeqHBase, described below, and we also posted a preprint with much more detailed clinical description to the BioRxiv preprint server while we continue searching for mutations in *TAF1* in additional families.

### Development of Comprehensive Whole-Genome Sequencing Analysis Pipelines

J. O’Rawe, Y. Wu [in collaboration with H. Fang, L. Jimenez Barron, M. Schatz, and G. Narzisi, Cold Spring Harbor Laboratory; K. Wang, California; and M. He, Wisconsin]

We continued developing various bioinformatics approaches for the analysis of exome and whole-genome

sequencing data. For example, in one project, we showed that the accuracy of detection of small insertions and deletions (INDELs) is greater when using WGS versus exon capture and sequencing. We also calculated that 60× WGS depth of coverage from the Illumina HiSeq platform is needed to recover 95% of INDELs detected by Scalpel. Although this is higher than current sequencing practices, we proposed that the deeper coverage may save total project costs because of the greater accuracy and sensitivity. Finally, we investigated sources of INDEL errors (e.g., capture deficiency, polymerase chain reaction [PCR] amplification, homopolymers). During the past 12 months, we reported the results of several other ongoing bioinformatics projects as well, as shown in the below publications. For example, we developed SeqHBase, a big data-based tool set for analyzing family-based sequencing data to detect *de novo*, inherited homozygous or compound heterozygous mutations that may be disease contributory. We demonstrated SeqHBase’s high efficiency and scalability, which is necessary, as WGS and WES (whole exome sequencing) are rapidly becoming standard methods to study the genetics of familial disorders. We also recently published an opinion piece regarding the current state of uncertainty quantification in DNA-sequencing applications, and we proposed methods that can be used to account for propagation of errors and their uncertainties through subsequent calculations.

### Summarizing the State of Human Genetics, Including the Genetic Architecture of Human Disease

J. O’Rawe

We prepared a comprehensive book chapter summarizing the current state of human genetics, and we continue to expand upon this work. In brief, there are ~12 billion nucleotides in every cell of the human body, and there are ~25–100 trillion cells in each human body. Given somatic mosaicism, epigenetic changes and environmental differences, no two human beings are the same, particularly as there are only ~7 billion people on the planet. One of the next great challenges for studying human genetics will be to acknowledge and embrace complexity. Every human *is* unique, and the study of human disease phenotypes (and phenotypes in general) will be greatly enriched by moving

from a deterministic to a more stochastic/probabilistic model. The dichotomous distinction between “simple” and “complex” diseases is completely artificial, and we argue instead for a model that considers a spectrum of diseases that are variably manifesting in each person. The rapid adoption of WGS and the Internet-mediated networking of people promise to yield more insight into this century-old debate. Comprehensive ancestry tracking and detailed family history data, when combined with WGS or at least cascade-carrier screening, might eventually facilitate a degree of genetic prediction for some diseases in the context of their familial and ancestral etiologies. However, it is important to remain humble, as our current state of knowledge is not yet sufficient, and in principle, any number of nucleotides in the genome, if mutated or modified in a certain way and at a certain time and place, might influence some phenotype during embryogenesis or postnatal life.

### Expanding Collection and Sequencing of Other Rare Genetic Syndromes

J. O’Rawe, Y. Wu [in collaboration with H. Fang, L. Jimenez Barron, M. Yoon, I. Iossifov, Cold Spring Harbor Laboratory; R. Robison, Utah; K. Wang, California; A. Rope, Oregon; and others]

We continue to meet and collect many families in Utah and elsewhere with very rare, idiopathic genetic syndromes. The total number of DNA samples collected to date is approaching 2000, and this includes detailed phenotyping information. Some of these samples have undergone exome or WGS, and we are currently analyzing 18 whole genomes and 35 exomes generated as part of this project. This includes the ongoing analysis of whole genomes from three families with singleton cases of autism, being performed in collaboration with Ivan Iossifov at CSHL, and an analysis of nine whole genomes from a pedigree with Prader–Willi syndrome (PWS), hereditary hemochromatosis, familial dysautonomia (FD), and Tourette syndrome. We did publish about one case in China where two siblings both began to develop idiopathic progressive cognitive decline starting from age six and were suspected of having an undiagnosed neurological disease. Initial clinical assessments included review of medical history, comprehensive physical examination, genetic testing for metabolic diseases, blood tests, and brain imaging. We performed exome sequencing with

Agilent SureSelect exon capture and the Illumina HiSeq2000 platform, followed by variant annotation and selection of rare, shared mutations that fit a recessive model of inheritance. To assess the functional impact of candidate variants, we performed extensive biochemical tests in blood and urine, and examined their possible roles by protein structure modeling. Exome sequencing identified *NAGLU* as the most likely candidate gene with compound heterozygous mutations (chr17:40695717C > T and chr17:40693129A > G in hg19 coordinate). Sanger sequencing confirmed the recessive patterns of inheritance, leading to a genetic diagnosis of Sanfilippo syndrome (mucopolysaccharidosis IIIB). Biochemical tests confirmed the complete loss of activity of  $\alpha$ -*N*-acetylglucosaminidase (encoded by *NAGLU*) in blood, as well as significantly elevated dermatan sulfate and heparan sulfate in urine. Structure modeling revealed the mechanism whereby the two variants affect protein structural stability. This successful diagnosis of a rare genetic disorder with an atypical phenotypic presentation confirmed that such “genotype-first” approaches can particularly succeed in areas of the world with insufficient medical genetics expertise and with cost-prohibitive in-depth phenotyping.

### Writing Policy and Opinion Pieces in Genomic Medicine

C. Barash, Cold Spring Harbor Laboratory; and others

The PI is an ongoing advocate for open access and data sharing, along with assisting with efforts at CSHL Press to roll out new initiatives, such as the BioRxiv preprint server and the new journal, *Molecular Case Studies*, which is designed to make it easier to publish and share the underlying data from single cases, families, or small cohorts of human conditions.

### Collaborating on Genetics of Tourette Syndrome

This work was done in collaboration with the Tourette Syndrome Association International Consortium for Genetics.

The PI continues to collaborate on this international effort to understand the genetics of Tourette syndrome (TS). Psychiatric comorbidity is common in TS; when present, these conditions typically cause

more distress and impairment than do tics. High rates of attention-deficit/hyperactivity disorder (ADHD) and obsessive-compulsive disorder (OCD) are well documented and thought to be core components of the TS phenotype; however, few studies have fully characterized other comorbidities. We therefore characterized the prevalence and impact of psychiatric comorbidity in a large sample of individuals with TS and their family members. The lifetime prevalence of comorbid psychiatric disorders, their heritabilities, and ages of risk was determined in participants with TS ( $N = 1374$ ). The lifetime prevalence of any psychiatric comorbidity was 85.7%; 57.7% of participants had at least two psychiatric disorders. The mean number of lifetime comorbid diagnoses was 2.1. Prevalence of mood, anxiety, and disruptive behavior disorders was ~30%. The age of greatest risk for most psychiatric disorders was 4 to 10 years, except eating and substance use disorders, which begin in adolescence. TS was associated with increased risk of anxiety and decreased risk of substance use disorders; high rates of mood disorders among participants with TS may be accounted for by comorbid OCD. Parental history of ADHD was associated with a higher burden of non-OCD, non-ADHD psychiatric disorders. Genetic correlations between TS and mood, anxiety, and disruptive behavior disorders were better accounted for by ADHD and, for mood disorders, by OCD. This study confirmed that psychiatric comorbidities begin early in life and are extremely common among individuals with TS and demonstrates that some may be mediated by the presence of comorbid OCD or ADHD.

## PUBLICATIONS

- Brownstein C, Beggs A, Homer N, Merriman B, Yu TW, Flannery KC, DeChene ET, Towne MC, Savage SK, Price EN, et al. 2014. An international effort towards developing standards for best practices in analysis, interpretation and reporting of clinical genome sequencing results in the CLARITY Challenge. *Genome Biol* **15**: R53.
- Fang H, Wu Y, Narzisi G, O'Rawe JA, Jimenez Barron LT, Rosenbaum J, Ronemus M, Iossifov I, Schatz MC, Lyon GJ. 2014. Reducing INDEL calling errors in whole-genome and exome sequencing. *Genome Med* **6**: 89.
- McGrath LM, Yu D, Marshall C, Davis LK, Thiruvahindrapuram B, Li B, Cappi C, Gerber G, Wolf A, Schroeder FA, et al. 2014. Copy number variation in obsessive-compulsive disorder and Tourette syndrome: A cross-disorder study. *J Am Acad Child Adolesc Psychiatry* **53**: 910–919.
- Narzisi G, O'Rawe JA, Iossifov I, Fang H, Lee YH, Wang Z, Wu Y, Lyon GJ, Wigler MH, Schatz MC. 2014. Accurate de novo and transmitted indel detection in exome-capture data using microassembly. *Nat Methods* **11**: 1033–1036.
- Shi L, Li B, Huang Y, Liu T, Lyon GJ, Xu A, Wang K. 2014. "Genotype-first" approaches on a curious case of idiopathic progressive cognitive decline. *BMC Med Genomics* **7**: 66.
- In Press*
- Barash CI and Lyon GJ. 2015. Open access and data sharing: Easier said than done. *Appl Transl Genomics* (in press).
- He M, Person TN, Hebring SJ, Heinzen E, Ye Z, Schrodi SJ, McPherson EW, Lin SM, Peissig PL, Brilliant MH, O'Rawe J, et al. 2015. SeqHBase: A big data toolset for family-based sequencing data analysis. *J Med Genet* (in press).
- Hirschtritt ME, Lee PC, Pauls DL, Dion Y, Grados MA, Illmann C, King RA, Sandor P, McMahon WM, Lyon GJ, et al. 2015. Age of risk, and etiology of comorbid psychiatric disorders in Tourette syndrome. Tourette Syndrome Association International Consortium for Genetics. *JAMA Psychiatry* (in press).
- Lyon GJ, O'Rawe J. 2015. Clinical genetics of neurodevelopmental disorders. In *The genetics of neurodevelopmental disorders* (ed. Mitchell K). Wiley, New York (in press). Preprint also posted on BioRxiv preprint server: doi: 10.1101/000687.
- Lyon GJ, Bird L, Rope A. 2015. X-linked malformations and infantile lethality. In *Epstein's inborn errors of development* (ed. Zimmerman C). Oxford University Press, Oxford (in press).
- Myklebust LM, Van Damme P, Støve SI, Dörfel M, Kalvik TV, Abboud A, Grauffel C, Jonckheere V, Kaasa H, Liszczak G, et al. 2014. Biochemical and cellular analysis of Ogden syndrome reveals downstream Nt-acetylation defects. *Hum Mol Genet* (in press).
- O'Rawe JA, Ferson S, Lyon GJ. 2015. Accounting for uncertainty in DNA sequencing data. *Trends Genet* (in press).
- O'Rawe J, Rope A, Wu Y, Swensen J, Robison R, Wang K, Lyon GJ. 2015. A variant in TAF1 is associated with a new syndrome with severe intellectual disability and characteristic dysmorphic features. *BioRxiv* (in press).



# DEVELOPMENT OF NEXT-GENERATION SEQUENCING TECHNOLOGY

W.R. McCombie   E. Antoniou   S.M. Eskipehlivan   J. Lihm   S. Teng  
N. Azamy   E. Ghiban   S. McCarthy  
Y. Berstein   S. Goodwin   O. Mendivil Ramos  
P. Deshpande   M. Kramer   S. Muller

One of the great challenges remaining in biology is to understand how variation in the genome of an organism dictates its phenotype. This is of great theoretical interest as well and is also central to understanding the impact of genetic variation on susceptibility to disease. We are studying this phenomenon in a variety of contexts ranging from analyzing the variants that contribute to significant human diseases such as psychiatric disorders and cancer, to understanding how variation contributes to the evolution of plant species. We are also looking at the impact of epigenetic variation (particularly in cancer) as well as identifying the de novo mutations contributing to psychiatric disorders. In this past year, we have found de novo variants that appear to be contributing to schizophrenia, which also overlap with genes involved in autism. To better understand these variants, we have developed, with the Hammell lab here at CSHL, a high-throughput testing system to identify potential functions of genes that may be contributing to psychiatric illnesses. We have also made substantial improvements to our understanding of the genes and the variants within them that contribute to evolution among plants.

## Investigating Bipolar Disorder Using Exome Sequencing of Case, Control, and Family Samples

M. Kramer, S. Muller, E. Ghiban, W.R. McCombie [in collaboration with J. Potash, V.L. Willour, and E. Monson, University of Iowa; P. Zandi, M. Pirooznia, F. Goes, and A. Chakravarti, Johns Hopkins School of Medicine]

Bipolar disorder is a debilitating psychiatric disorder with complex genetic architecture. In 2010, we began a collaboration with Johns Hopkins University to apply exome sequencing to discover genetic variants associated with bipolar disorder. Our early work has been detailed in previous reports, including our efforts last year to dramatically increase throughput to double our study's sample size. Building on our 2013 scale-up of sequencing

efforts and the establishment of our informatics pipeline (details of which were published in the journal *Human Genomics*), as well as the previous analysis of the family samples, in 2014, we turned our focus to the case-control analysis of our enlarged data set. In an effort to generate more accurate calling across the entire data set, we used the updated version of GATK (genome analysis toolkit) with multisample joint genotyping to reprocess the 1135 bipolar cases and 1142 control subjects. After stringent data quality assurance (QC), we called more than 883,000 variants, which were then annotated using ANNOVAR. The rare coding variants (with a minor allele frequency <5%) that were predicted to be damaging (nonsense, missense, splice, and frameshift variants) were then subjected to statistical analysis at the variant, gene, gene set, and pathway levels. Although several findings rose to a suggestive level of significance, no particular variant, gene, gene set, or pathway has risen to a genome-wide significant value thus far. This is perhaps not unexpected considering that much larger sample sizes may be required to overcome the genetic heterogeneity in these disorders. The expanded capture target set we used in 2013 included more than 1400 genes that encode proteins localized in the synapses of the brain, as well as promoter regions of those genes, and we will include these noncoding regions in our analysis. In addition, we will be performing a meta-analysis of our data to look for replication in other large data sets from the PGC. We plan to finalize all of the case-control and family analyses in 2015, and publish our findings.

## Comparative Genomics of Seed Evolution

E. Antoniou, M. Kramer, W.R. McCombie [in collaboration with R. Martienssen, Cold Spring Harbor Laboratory; D. Sasha, K. Varala, and G. Coruzzi, New York University; D. Stevenson, New York Botanical Gardens; R. DeSalle, American Museum National History]

The goal of this project is to exploit plant genome diversity to discover new genes involved in the development

of seeds. Data (transcriptome, orthologs), resources (phylogenetic matrix and tree), informatics tools, and pipelines for phylogenomics and machine learning will be generated. It is envisioned that the data and software resources created will empower comparative genomic researchers to exploit plant diversity to identify genes associated with any trait of interest or economic value. For example, a major event in seed crops was the shift from prefertilization storage tissue in gymnosperms to the postfertilization endosperm in angiosperms. Other seed traits of economic interest are onset dormancy, shift of fleshy seeds to dry winged seeds, and evolution of de novo appendages such as arils, each of which have multiple origins within seed plants. By studying the sequence and expression of genes supporting nodes where these independent origins occur, we can test whether the same or different gene(s) give rise to the trait.

The first aim is to generate RNA-Seq data for nodal species plus outgroups for analysis of leaf, ovule, and seed transcriptomes to fill in gaps in depth/breadth of taxa/ortholog coverage in phylogenetic analysis. The second aim is to generate deep transcriptome data to be used for unigene assembly, ortholog identification, and expression data. Chip-Seq data will also be generated to elucidate relationships between gene expression and evolution of regulatory genes.

We have sequenced 306 RNA-Seq libraries and 68 Chip-Seq libraries prepared at New York University in 2014. Analysis of the data is ongoing.

### Functional Screening in the PVD Neuron of *Caenorhabditis elegans* for Psychiatric Candidate Genes

O. Mendivil Ramos, C. Aguirre-Chen, S. McCarthy, M. Kramer, C. Hammell, W.R. McCombie

Despite the complex genetic heterogeneity of neuropsychiatric disorders, the development of large-scale genomic techniques has led to important advances in the identification of genes and loci associated with schizophrenia, bipolar disorder, and major depressive disorder. However, the functional interpretation of genetic variation is the rate-limiting step in determining the impact of mutations on biological function and disease pathogenesis. Functional analyses have not kept up with our ability to detect variants, and interpretation relies on inconsistent computational

methods. This speaks to the need for high-throughput functional validation of novel genes and variants. We propose a new strategy in which we would use a tiered approach in model organisms, using an inexpensive, high-throughput initial test in *C. elegans* that would feed into more expensive and time-consuming downstream analyses.

Late in 2013, we established a collaboration with the Hammell lab at CSHL. Using bacterial ingestion-mediated RNA interference (RNAi) to knock down candidate gene orthologs, it was possible to detect altered branching of the PVD neuron in a subset of psychiatric candidate genes we had identified. Building on that pilot work, we expanded this research line in 2014 by extending this list of homologs to additional candidates, as well as doing positive control studies to better estimate the efficacy of this model in assessing the large numbers of genetic variants being identified for major psychiatric disorders. The genes assayed in this extended experiment cover the gamut of candidates identified from large-scale genetic studies on schizophrenia and autism as well as genes harboring damaging mutations present in healthy individuals. From previously published data of candidate genes for autism and schizophrenia, we identified 34 homologs of human-*C. elegans*, four of which altered the PVD neuron—a success rate of ~12%. We compared our results to an unbiased PVD neuron assay on 2635 genes of chromosome IV of *C. elegans*. From these genes, 338 are homologs in humans and five of them had a disrupted PVD neuron—a success rate of 1.5%. Thus, we show that the enrichment of this procedure was approximately eightfold. As arborization and neuronal migration are likely part of the molecular pathogenesis of psychiatric disorders in some patients, this assay is a useful tool to further validate at a functional level the genes contributing to these important illnesses.

### The Epigenome of Fear

S. McCarthy, N. Azamy, P. Deshpande, W.R. McCombie [in collaboration with S. Ahrens, L. Sabin, and B. Li, Cold Spring Harbor Laboratory]

In 2014, the Li and McCombie labs initiated a collaboration to explore the hypothesis that critical epigenetic changes controlling gene expression in the amygdala underlie long-term fear memory formation and contribute to fear dysregulation in mental disorders. This collaboration unites discoveries made by the

McCombie and Li groups which showed that damaging mutations in genes important for chromatin organization may be strong risk factors for schizophrenia and that fear conditioning in mice induces robust memory formation through specific plasticity in the SOM<sup>+</sup> and SOM<sup>-</sup> inhibitory neurons of central amygdala, respectively. Fear dysregulation brought about by traumatic life experiences is implicated in many psychiatric disorders including depression, schizophrenia, and, in particular, posttraumatic stress disorder (PTSD). Therefore, both studies suggest that genetic and activity-dependent changes in gene expression converge on epigenetic regulation as a potentially important biological mechanism in the etiology of psychiatric disorders.

We are combining the powers of genetic and behavioral techniques with cutting-edge genomic technologies to profile the “*Epigenome of Fear*”: changes in structural features of the genome in SOM<sup>+</sup> and SOM<sup>-</sup> neurons that regulate the molecular response of long-term fear memory formation. Using “assay for transposase-accessible chromatin” (ATAC-Seq), we are first identifying differences in the open chromatin profiles of the murine cortex and blood as a proof of principle to develop bioinformatics tools that broadly distinguish non-specific tissue types. In parallel we are also gathering data on SOM<sup>+</sup> nuclei to fine-tune our analysis pipeline to distinguish cell-type- and function-specific differences.

During 2014, we optimized nuclei extraction to overcome the problems caused by mitochondrial contamination. Specifically, we used fluorescence-activated cell sorting (FACS) to fluorescently label nuclei from the cortex and the amygdala, thereby removing neuronal mitochondria and increasing the specificity of neuronal nuclei for analysis by ATAC-Seq. Furthermore, we streamlined our analysis pipeline to identify tissue-specific open-chromatin profiles. First, ATAC-Seq data are aligned to the mouse genome. Uniquely mapping reads are filtered to remove polymerase chain reaction (PCR) duplicates. Second, peaks of stacked reads representing regions of open chromatin tagged by the Tn5 transposase are called using *MACS2*. ATAC-Seq peaks that overlap the transcription start site of Ref-Seq genes are then identified and filtered for uniqueness in either the cortex or blood. The clustering of genes into biological pathways and ontologies is then carried out using a number of different analysis packages in R. Preliminary data analysis

suggests that we can clearly identify patterns of open chromatin at transcriptional start sites of genes unique to the cortex and blood. For example, unique open chromatin profiles in the cortex are enriched in genes that function in neuronal projection, at the synapse, at the cell junction, in cell differentiation, and in synaptic transmission. Conversely, unique open chromatin profiles in blood are enriched in leukocyte activation, T-cell proliferation, and immunity. In addition, we have evidence of isoform-specific open chromatin profiles in the brain and blood at genes such as *TIMELESS*, which is critical in circadian rhythm, further exemplifying the power of ATAC-Seq to distinguish cell functions and high resolution.

Going forward in 2015, we will further refine our analysis to profile open chromatin in the amygdala, laying the foundation for experience-dependent and genetic-dependent models of fear conditioning.

### Targeted Resequencing of 264 *DISC1* Interactome Genes in 1543 Samples

S. Teng, S. McCarthy, M. Kramer, W.R. McCombie [in collaboration with D. Blackwood, D. Porteous, P. Thompson, and I. Deary, University of Edinburgh; A. McRae, University of Queensland]

Continuing our growing collaboration with the University of Edinburgh, to enhance our understanding of *DISC1* genetics and biology, we strengthened our analysis of targeted and resequenced exons, promoters, and conserved regions of 264 genes, which we call the *DISC1* Interactome, that directly interact with *DISC1* or are correlated with *DISC1* expression and form functional biological pathways related to *DISC1* activity. We previously completed sequencing of 1543 individuals comprising 241 schizophrenia patients, 221 bipolar patients, 192 recurrent major depressive disorder patients, and 889 healthy controls.

In a quality filtered cohort of 1464 (95%) samples that had greater than 20× coverage for more than 80% of the Interactome target, we discovered 226,953 single-nucleotide variants (SNVs). We identified 3146 missense variants and reported 71 validated disruptive variants, including 47 nonsense and 24 splice site variants. In 2014, we focused our attention on the patterns of rare (MAF < 1%) but also common variants (MAF ≥ 1%) for their associations with psychiatric and cognitive conditions. We found an enrichment of rare disruptive SNVs in schizophrenia cases versus

controls ( $P = 0.0038$ ), which is supported by previous findings from other schizophrenia exome sequence studies demonstrating that the rare damaging mutations were enriched in the genes involved in neural development or synaptic function. Damaging mutations in schizophrenia and recurrent major depressive disorder patients were greatly enriched in a gene that indirectly interacts with *DISC1*. Furthermore, we found an increased burden of damaging singleton variants associated with reduced cognitive ability measures such as the National Adult Reading Test and Moray House Test at age 11 and 70.

In our analysis of common variants, we identified an association between an SNP in a strong copy number variant (CNV) candidate gene and the combined set of psychiatric disorders ( $OR = 0.56$ ,  $P = 3E-6$ , Adjusted  $P = 0.047$ ). In contrast to previous associations with CNV analysis, the effect we observed is protective, suggesting complex roles for this gene in disease susceptibility.

Going forward in 2015, our objective is to prioritize variants and genes for functional follow-up in model systems such as *Caenorhabditis elegans*. In addition, we aim to use our data and findings in the *DISC1* Interactome, the whole-genome sequencing of members of the *DISC1* translocation family, as well as gene expression data to analyze epistatic interactions between genes that could lead to variable phenotype and incomplete penetrance in the *DISC1*  $\tau(1:11)$  family.

## Colorectal Cancer Genomics

E. Antoniou, W.R. McCombie [in collaboration with X. Wang, X. Yu, E. Li, and J. Williams, Stony Brook University]

Colorectal cancer (CRC) is the third most commonly diagnosed cancer and is the second leading cause of cancer-related deaths in the Western world. The mortality from colorectal cancer in Caucasian Americans (CAs) has been declining, but, in contrast, it has been rising in African Americans (AAs). The response to chemopreventive agents differs, as do the biological/genetic heterogeneity of the tumors of AAs and CAs. Because chemopreventive agents exert their effect through a molecular target, this disparity would suggest differences at the genetic level. This project aims to determine the attribution of genetic and epigenetic differences to the colon cancer racial/ethnic health disparity.

In previous years, we used RNA-Seq to obtain gene expression measurements on 18 tumor/normal paired AA tissue samples and as many samples from CA patients. We also prepared 57 reduced representational bisulfite sequencing (RRBS) libraries, 52 of them from paired tumor/normal tissues.

In 2014, we worked with Dr. X. Wang from Stony Brook University to develop a new way to analyze RRBS sequences produced by the Illumina Next-Seq500 sequencer. The new bioinformatic pipeline almost doubled the number of mapped CpG sites. Preliminary analysis of the RRBS data obtained in 2013 identified differential methylation of genes such as *BMP3* and *GPR75*. However, it was also observed that our mean base coverage ( $3.5\times$ ) was too low to confidently identify most differentially methylated sites. Consequently, we generated additional sequences from eight AA and seven CA patients and increased the base coverage to  $10\times$ . Analysis of these data is ongoing.

## De Novo Assembly of Yeast and Rice Genomes Using Long Insert Libraries and the Pacific Biosciences RS II Instrument

P. Deshpande, S.M. Eskihehliyan, M. Kramer, S. Goodwin, E. Antoniou, W.R. McCombie [in collaboration with J. Gurtowski, H. Lee, M.C. Schatz, and J. Hicks, Cold Spring Harbor Laboratory; C. Heiner and G. Khitrov, Pacific Bioscience; S. McCouch, Cornell University]

Next-generation sequencing is an invaluable tool for the understanding of genetics across many organisms. However, the relatively short reads generated by instruments such as the Illumina HiSeq pose a limitation to the de novo assembly of large and complex genomes. The Pacific Biosciences instrument allows for the sequencing of fragments several tens of thousands of base pairs in length. Such long reads provide superior structural information for de novo assembly, and they retain structural elements important to the understanding of genetic variation.

As a continuation of our yeast work from 2013, we used the Pacific Biosciences instrument to sequence three separate strains of yeast: W303 (as we have previously used), ethanol red, and a sake strain (kyokai #7). We compared the various purification, library prep, and loading approaches to generate a quality sequence. Overall, we generated more than  $50\times$  coverage of each strain with sequence read N50s in excess of 10 kb. These strains will be assembled using HGAP and

the Celera Assembler. Once the genomes are assembled, they will be probed for variations contributing to their industrial phenotypes, in particular their ethanol production capacity. These assemblies will also be compared to prior assemblies of these strains using a long read assembly to identify features not detected in the short read assembly approaches.

Building on our yeast work, we began sequencing a more complex genome, the IR64 strain of rice. The rice sequences were generated using the Pacific Biosciences RS II instrument with P5/C3 chemistry. We generated greater than 115× coverage across the genome with 48× coverage in reads longer than 10 kb. The longest read was 53,652 bp long. A de novo assembly was built using HGAP and the Celera Assembler. The N50 contig size is 4.02 Mbp, up from 20 kb in the previous Illumina-only assembly. The longest contig is 16 Mbp long, approaching the size of a complete chromosome arm. The assembly has a high sequence identity (>99%) with known IR64 genes and produced complete gene models and promoter regions for nearly every gene. This dramatic increase in contig length will likely have a significant positive impact on our ability to annotate the features of these genes. Both the yeast and rice work are under preparation for publication of a manuscript.

### Comparative Genomics of Three Rice Strains Using Whole-Genome Illumina Sequencing

E. Antoniou, M. Kramer, E. Ghiban, W.R. McCombie [in collaboration with M.C. Schatz, J.C. Stein, and D. Ware, Cold Spring Harbor Laboratory; L.G. Maron and S.R. McCouch, Cornell University]

In collaboration with the Ware lab at CSHL and the McCouch lab at Cornell University, we previously sequenced three strains of rice on Illumina HiSeq sequencers, representing the two major varietal groups and three of the five major subgroups (*indica* [IR64], *aus* [DJ123], and *temperate japonica* [Nipponbare]). In 2014, we constructed de novo assemblies of the three genomes using the ALLPATHS-LG software with the recommended combination of a 180-bp overlap library with 2-kb and 5-kb jumping libraries. We then analyzed the genomes by comparing them to the published reference sequences for Nipponbare and Indica, as well as comparing the three genomes to one another. We published our findings in the journal

*Genome Biology*. We documented several megabases of each genome absent in the other two strains (4.8 Mbp to 8.2 Mbp [423 kbp to 930 kbp exonic]) and analyzed the detailed genomic structures of several loci associated with important agricultural traits such as hybrid sterility, submergence tolerance, improved yield, and phosphorus deficiency. These three genome assemblies will support future breeding and functional studies of this important food staple.

### Technology Development and Assessment of Nanopore Sequencing

S. Goodwin, P. Deshpande, W.R. McCombie [in collaboration with J. Gurtowski, and M.C. Schatz, Cold Spring Harbor Laboratory]

The Oxford Nanopore device is a new instrument for next-generation sequencing that is currently undergoing a beta testing phase. This instrument relies on the translocation of a long DNA molecule through a protein pore in order to perform sequencing. Unlike most sequencing-by-synthesis methods, this approach does not rely on an optical or fluorescent signal. Rather, as the DNA is translocated through the pore, a characteristic change in voltage is used to identify the composition of a 5-mer as it passes through the channel. This approach allows for the sequencing of very long contiguous molecules with minimal reagent costs.

The goal of this project is to evaluate the use of the Oxford Nanopore device for various sequencing routines. We began sequencing the yeast W303 strain in mid-2014. This strain was selected because we have previously worked with it for long read sequencing on the Pacific Biosciences instrument. Over the course of this project, we used 48 flow cells from three separate versions of the chemistry: R6.0, R7.0, and R7.3. We also saw a large increase in the performance of the instrument. Early on, error rates were ~30% for the R6.0 chemistry. Those error rates have decreased to between 10% and 15% with the R7.3 flow cells, and we hope to see continued decreases. The typical N50 has also improved, ranging from ~5K for an R6.0 flow cell to ~10K for an R7.3 flow cell (with a 10K input fragment size). We have generated contiguous sequences in excess of 150K bp. However, because of the high error rate and limited tools available for such long reads, we have been unable to align fragments longer than 60K. Finally, to overcome the high error rate present in the instrument, we have developed a

pipeline, called nanocorr, that can select the best sequences from our data set, correct them with MiSeq data, and assemble them with a greater than 99.8% accuracy. This work is currently under review.

In 2015, we will continue to evaluate the instrument for hybrid and de novo assembly uses using the yeast model. We will also be evaluating the use of the instrument for full-length isoform sequencing from human and other cDNA samples.

### Exome-Based Sequencing of Parent-Offspring Trios with Schizophrenia

S. McCarthy, M. Kramer, J. Lihm, Y. Berstein, R. Solomon, E. Ghiban, E. Antoniou, W.R. McCombie [in collaboration with J. Gillis and S. Yoon, Cold Spring Harbor Laboratory; M. Mistry, and P. Pavlidis, The University of British Columbia; E. Kelleher, C. O'Brien, G. Donohoe, M. Gill, D.W. Morris, and A. Corvin, Trinity College Dublin]

In 2014, in collaboration with Trinity College Dublin, we published findings from our study of de novo variants in 57 parent-offspring trios with schizophrenia (42 trios with no family history of psychosis [*sporadic*] and 15 trios with a family history of psychosis [*familial*]), in the journal *Molecular Psychiatry*. Despite a small sample, our work shed further light on the contribution of de novo variation to the genetic risk of schizophrenia (first found at CSHL) and also highlighted a common genetic pathology across neurodevelopmental disorders of possible therapeutic potential.

In summary, exome sequencing analysis of 57 trios with schizophrenia revealed a significant increase in the rate of nonsense mutations in sporadic cases compared to what was expected based on publicly available data of healthy unaffected trios ( $P = 0.01$ ), suggesting

that some of the de novo mutations identified could be private, penetrant risk factors. Expanding our analysis to include missense mutations (de novo mutations likely to disrupt or damage protein function that were enriched in genes less tolerant to new or rare mutations) increased the strength of the result ( $P = 2 \times 10^{-5}$ ). Notably, many potentially functional de novo mutations were identified in genes with prior genetic evidence for a role in autism and intellectual disability such as *CHD8*, *MECP2*, *AUTS2*, *MLL2* ( $P = 0.02$ ). Furthermore, de novo mutations significantly overlapped chromatin modifiers important for transcription regulation, suggesting that perturbing the components controlling the epigenetic regulation of gene expression could be a molecular pathogenetic mechanism shared between schizophrenia and autism. Since we published our work, other available exome studies of neurodevelopmental disorders support our hypothesis and open exciting avenues to explore novel mechanisms for potential therapeutics for schizophrenia and autism.

### PUBLICATIONS

- McCarthy SE, Gillis J, Kramer M, Lihm J, Yoon S, Berstein Y, Mistry M, Pavlidis P, Solomon R, Ghiban E, et al. 2014. De novo mutations in schizophrenia implicate chromatin remodeling and support a genetic overlap with autism and intellectual disability. *Mol Psychiatry* **19**: 652–658.
- Pirooznia M, Kramer M, Parla J, Goes FS, Potash JB, McCombie WR, Zandi PP. 2014. Validation and assessment of variant calling pipelines for next-generation sequencing. *Hum Genomics* **8**: 14.
- Schatz MC, Maron LG, Stein JC, Hernandez Wences A, Gurtowski J, Biggers E, Lee H, Kramer M, Antoniou E, Ghiban E, et al. 2014. Whole genome de novo assemblies of three divergent strains of rice, *Oryza sativa*, document novel gene space of *aus* and *indica*. *Genome Biol* **15**: 506.

# PLANT GENOME ORGANIZATION, EVOLUTION, AND FUNCTION

D. Ware	K. Chougule	Z. Lu	M. Regulski	W. Ware
	Y. Jiao	M.K. Monaco	J. Stein	S. Wei
	V. Kumar	J. Mulvaney	J. Thomason	L. Zhang
	S. Kumari	M. Neves Dos Santos Leite	P. Van Buren	
	Y.K. Lee	C. Noutsos	B. Wang	
	C. Liseron-Monfils	A. Olson	L. Wang	

Our lab has two primary goals: (1) research into plant genomics, with a focus on functional and comparative genomics of model plant systems and agriculturally important crop plants, and (2) development of tools, data sources, and resources for the genomics research community. We also contribute actively to three large-scale, multi-institutional, cyber-infrastructure collaborations designed to serve broad research and educational communities: Gramene, the National Science Foundation iPlant collaborative, and the Department of Energy (DOE) Systems Biology Knowledgebase (KBase).

Our lab is a dynamic research group, and its composition has changed during the past year. During this period, we have been joined by Vivek Kumar, Kapeel Chougule, Mike Regulski, and Joseph Mulvaney. Andrea Eveland and Ken Youens-Clark have advanced their careers, taking positions as an assistant professor and a scientific developer, respectively. Visiting scientists Zhiming Zhang and Fazhan Qiu have returned to China to lead their own research groups.

## PLANT GENOME RESEARCH

---

In the last decade, the sequencing and annotation of complete plant genomes have helped us to understand plant function and evolution, as well as how to alter economically important traits. Efforts in many disparate disciplines are required to generate reference genomes. The work starts with laboratory scientists, who generate the raw sequence data. Next, computational biologists and bioinformaticians, such as those in our lab, kick off a series of steps to interpret the raw data. The process of interpretation involves the assembly of raw sequence reads into overlapping segments (“contigs”), which are used to create a scaffold. This

scaffold, in turn, discerns the position, relative order, and orientation of contigs within the chromosomes. The next step is annotation, the discovery and description of genes and other functional elements, and homologies (evolutionary relationships) with other genomes. This information must be faithfully communicated and visualized in Web-based platforms, such as Gramene.

All of these activities are rapidly evolving with the advent of new sequencing technologies, algorithms, and data-handling requirements. For example, high-depth and low-cost sequencing of RNA transcripts is providing a vast stream of new evidence that informs genome annotation; this in turn has spurred the development of new software for modeling and performing genome annotation. Low-cost sequencing has also made it possible to ask whole new classes of questions, moving beyond the generation of single references for individual species and supporting the development of multispecies representation as “pan-genome.” Ongoing projects within the maize, rice, and *Arabidopsis* research communities are now sequencing hundreds or thousands of genotypic backgrounds, chosen from carefully constructed populations, wild populations, and breeding germplasms in each species. Information about genetic variation is helping scientists to understand the genetic basis of phenotypic traits and the origins of domesticated crop species. New technologies are also driving research in epigenetics, the study of heritable variation not caused by changes in the underlying genome sequence. Epigenetic mechanisms include modification of DNA by methylation and various forms of histone modification that can cause changes in gene expression and other phenotypes. Both types of modifications can be studied using new sequencing technologies and analytical methods.

## Updating the Maize B73 Reference Genome and Developing Pan-Genome Variation Pipelines

A. Olson, J. Stein, Y. Jiao, B. Wang, M. Regulski [in collaboration with The Maize Genome Sequencing Consortium; J. Glaubitz and E. Buckler, Cornell University; R. Fulton and R. Wilson, Washington University; K. Guill and M. McMullen, Missouri University, USDA ARS; D. Rank, P. Peluso, and T. Clark, Pacific Biosciences; M.Y. Law; and M. Yandell, University of Utah, Salt Lake City]

Work continues on refining the assembly and annotation of the Maize B73 reference sequence. The third version (“RefGen\_v3”) of the B73 assembly and its annotations are hosted on Gramene and have been deposited in GenBank along with contigs assembled from Roche/454 sequencing of a whole-genome shotgun library. RefGen\_v3 incorporates some of these contigs in order to increase coverage of missing gene space. The contigs were selected based on alignment to FLcDNA sequences and then inserted into gaps in the RefGen\_v2 assembly, guided by a genetic map as well as synteny with rice and sorghum. Approximately 500 genes were added or improved using this method, and many unplaced physical map contigs could also be anchored based on these maps. The RefGen\_v3 assembly was released following acceptance by GenBank.

Gene structural annotations were originally called using the Gramene/Ensembl Gene Builder, which takes into consideration *ab initio* gene predictions as well as aligned protein, cDNA, and EST evidence. To improve these gene models, MAKER-P software was run on RNA-Seq data from 96 public data sets, in collaboration with Mark Yandell (Law et al. 2014). Untranslated regions of 1393 existing gene models were extended, and an additional 4466 protein-coding and 2522 noncoding genes were predicted.

We also reviewed the current status of long-read sequencing from PacBio Systems to support improvement of the B73 reference assembly. In 2014, these efforts met with moderate success, as we combined the PacBio Systems technologies with short-read sequencing from the more established Illumina technology for the purposes of error correction. In both sequencing and assembly, the integration of multiple approaches improved the length and accuracy of the genome sequences we obtained, relative to what we could have achieved using either technology alone. For the maize reference B73 genome, we are using PacBio long reads to order contigs and fill in gaps. Based on the complexity of the maize genome, our goal is to achieve 50×

coverage; to date, we have generated 40 Gb of PacBio long reads data, representing 17.8× coverage. We have used these data to improve the order and orientation of sequences in the existing reference assembly and extend existing contigs; in the process, we have recovered 160 Mb of missing sequences in the current maize B73 genome. In addition to our work on genome assembly, we are using the same long-read sequencing technology to generate full-length transcripts, characterization, and annotation of alternatively spliced transcripts.

In *Zea mays*, only ~50% of genomic content is shared between lines because of the tremendous haplotype diversity in this species. Therefore, a single reference assembly for maize cannot serve as a sufficient backbone for capturing and describing this variation, and this constraint poses an obstacle to our understanding of the genetic variation that controls interesting traits. To overcome this challenge, we need a suitably detailed and flexible species-level representation (i.e., a maize “pan-genome”). Our initial efforts are focused on developing a maize pan-genome comprising the B73 reference assembly and novel sequences from other inbred maize lines. There are three steps involved: (1) assembly of the additional accessions, (2) characterization of novel haplotypes, and (3) genetic anchoring of the novel sequence. As proof of concept, we have generated 30× Illumina reads of B97. At this depth, high-repeat content in maize challenges existing methods for *de novo* assembly; based on simulations of 30× reads sampled from the B73 reference, we expect to achieve ~80% coverage of annotated genes. Because different assembly algorithms have different advantages, we have adopted a meta-assembly strategy. A meta-assembly of B97 was generated, and after performing an alignment to the bacterial artificial chromosome (BAC)-based B73 reference, we identified ~19.78 Mb of novel B97 sequences. We are now using population-based mapping approaches to support anchoring of these scaffolds, relative to the reference B73 physical map. This genetic mapping method was sufficiently accurate to place 77.6% of 1000 B73 gene sequences on the 10 maize chromosomes.

## The Rice Pan-Genome

J. Stein [in collaboration with R. McCombie and M. Schatz, Cold Spring Harbor Laboratory; S. McCouch, Cornell University]

This year marks the 10th anniversary of the completion of the rice genome, *Oryza sativa*. The Nipponbare reference sequence stands out as the highest-quality assembly



available for any crop plant, and thus serves as a cornerstone of functional genomics research. However, to fully probe the genetic diversity of rice, researchers need additional high-quality assemblies that are more representative of the species as a whole. The complex history of rice domestication gave rise to five subpopulations, each with specialized traits and cultivation geographies. Many large-scale structural differences among the different varieties have resulted in substantial variation in gene content. The full complement of genes that exist in diverse varieties is currently poorly understood in rice.

To begin to characterize the rice pan-genome, this NSF-funded research project produced new reference assemblies for two additional rice varieties belonging to the *indica* and *aus* subpopulations (Schatz et al. 2014). Using Nipponbare (a *japonica* strain) as a positive control, we demonstrated that low-cost sequencing combined with advanced computational methods can yield gene-enriched de novo assemblies with high accuracy and completeness. Examination of several known quantitative trait loci (QTL) revealed differences in haplotype structure between the varieties. These differences correlated with known phenotypic differences that influence traits such as crop yield, submergence tolerance, phosphorus uptake, and hybrid sterility. Remarkably, each of the three varieties possesses several megabases of genomic sequence that are absent from the other two, and these unique sequences harbor hundreds of genes. We do not yet know how much these differences in gene content contribute to adaptive traits (such as tolerance to abiotic stress) or agronomic features (such as grain quality) that distinguish these rice varieties. Notably, however, an unusually large proportion of strain-specific genes belong to the R-gene family, which confers resistance to pathogenic microbes. Therefore, it is reasonable to speculate that differential expansion or contraction of the R-gene complement accompanied the geographical radiation of rice varieties, allowing them to develop resistance to disease-causing agents endemic to particular geographical regions.

### Discovery and Application of Epigenetic Variation in Plant Genomes

Z. Lu, A. Olson, M. Regulski [in collaboration with R. Martienssen and R. McCombie, Cold Spring Harbor Laboratory; R. Klein and P. Klein, USDA/Texas A&M University; and S. Brady, University of California, Davis]

DNA methylation plays important roles in regulating gene expression and controlling transposable elements

(TEs). Patterns of DNA methylation, referred to as the “methylome,” must be faithfully propagated in order for plants and animals to undergo proper development. To study methylation, genomic DNA is treated with sodium bisulfite (Methyl-Seq), which converts unmethylated cytosine to thymine; methylation can then be detected at single-base resolution by sequencing the treated DNA and comparing the reads to the sequence of the unmodified genome. In previous work, we achieved 20× – 30× coverage over the mappable portion of the maize genome and observed methylation patterns correlated with gene expression, small RNA genes, and alternative splicing. Cytosine methylation patterns were diverse in TEs, and especially in genes. These patterns were largely heritable in recombinant inbred lines (RILs); however, we observed significant deviations from heritability, many of which were conserved in different RILs. These findings will help us to reveal the roles of DNA methylation in gene regulation and other biological functions in the future.

In *Sorghum*, we are participating in a pilot project using laser capture microdissection (LCM) to extract specific tissues (rather than larger structures, which have more heterogeneous gene expression and methylation). In these tissue-specific RNA-Seq and Methyl-Seq experiments, we will identify gene expression and DNA-methylation patterns associated with responses to abiotic stresses, such as drought and salinity, in the reference strain BTx623 and widely cultivated varieties.

### Gramene: Comparative Genomic Resource for Plants

K. Chougule, Y. Jiao, V. Kumar, S. Kumari, M. Karey Monaco, J. Mulvaney, A. Olson, J. Stein, J. Thomason, S. Wei, K. Youens-Clark [in collaboration with P. Jaiswal, Oregon State University; P. Kersey and H. Parkinson, EMBL-European Bioinformatics Institute; L. Stein, Ontario Institute of Cancer Research; C. Taylor, American Society of Plant Biologists; and R. Wing, University of Arizona, Tucson]

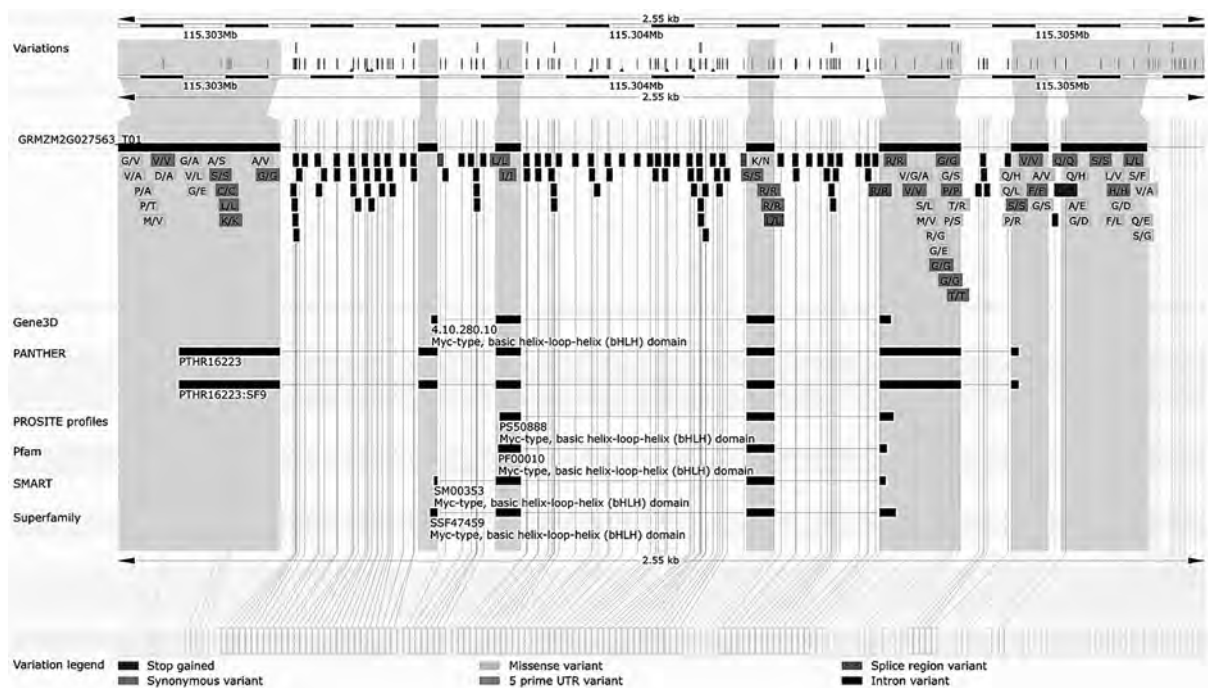
The Gramene project provides online reference resources for plant genomes and curated pathways to aid functional genomics research in crops and model plant species. Our website, [www.gramene.org](http://www.gramene.org), facilitates studies of gene function by combining plant genome and pathway annotation 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



**Figure 1.** Complete reference genomes in Gramene as of data release 44 (January 2015).

to predict gene function in other species. This ability should advance our fundamental understanding of plant physiology, including economically important traits such as hybrid vigor, grain development, seed dormancy, drought tolerance, and resistance to diseases. Last year, we significantly expanded the Website and streamlined the user interface and backend functions. The project accomplished several major milestones, culminating in our 43rd release (December 2014) since the inception of the project (Fig. 1). Highlights include the addition of 11 new species to our genome browsers: peach (*Prunus persica*), cocoa (*Theobroma cacao*), six rice-related species in the *Oryza* and *Leersia* genera, the multifaceted crop *Brassica oleracea*, the primitive flowering plant *Amborella trichopoda*, and the unicellular alga *Ostreococcus lucimarinus*, which brought the total to 39 reference genomes. We also incorporated new genetic variation data (Fig. 2) for six crops including maize, sorghum, barley, tomato, and two species of rice, bringing the total to 10. Furthermore, we expanded our Plant Reactome database, which now includes 200 curated rice pathways and pathway projections in 17 species.

The Gramene team develops comparative genomics databases in collaboration with the Ensembl Plants project at the European Bioinformatics Institute (EBI).



**Figure 2.** Single-nucleotide genetic diversity displayed in the context of functional protein domains.

Perhaps the most significant achievement from this collaboration was the addition of four Triticeae cereal genomes, including barley (*Hordeum vulgare*) and bread wheat (*Triticum aestivum*), whose whole-genome shotgun sequence assemblies had until recently been severely compromised due to their size and highly repetitive DNA composition. The barley genome is currently a “gene-ome” (i.e., mostly an assembly of gene regions anchored to a physical map).

The wheat genome is the largest (~17 Gbp) and most complex angiosperm genome ever sequenced, and it is more than five times larger than the human genome. Its hexaploid genome (AABBDD) comprises three nearly identical ancestral subgenomes (AA, BB, and DD) and is estimated to contain about 124,000 genes. The genomes of two of those diploid progenitors, *Triticum urartu* (AA) and *Aegilops tauschii* (DD), were recently added to the Gramene/Ensembl Plants collection. Distinguishing among genes in the three subgenomes has been a challenge for plant geneticists. Consequently, the first wheat draft, although gene-rich, is still highly fragmented. Its assembly and gene annotation have been improved by integrating complementary and heterogeneous sequence-based genomic and genetic data sets; by performing alignments to ancestrally related DNA, RNA-Seq, expressed sequence tags (ESTs) and UniGene clusters; and by identifying homeologous genes across the subgenome components. The addition of the wheat genome to Gramene completed the triumvirate of major cereal grains (wheat, rice, and maize) in our collection. These three crops constitute 90% of worldwide grain production and are the keystones of future agricultural sustainability.

Gramene continues to build a genus-level phylogenomics resource for rice and related *Oryza* species. Rice emerged ~15 million years ago. Genus *Oryza* includes two cultivated species (Asian and African rice) and 21 wild species adapted to a broad range of tropical and subtropical habitats around the world. Genomics research aimed at exploiting such diversity for the purpose of crop improvement is also establishing *Oryza* as a model system for genus-level study of trait evolution, speciation, and domestication. Working with the NSF-funded *Oryza* Genome Evolution project (Principal Investigator [PI]: Rod Wing, University of Arizona) and the international consortium known as I-OMAP, Gramene currently hosts high-quality genome assemblies for more than half of those species.

Last year, Gramene added complete reference assemblies for six of these rice-related genomes, bringing the total to 12. One of the challenges of comparative genomics research is that different genome projects use different protocols to identify genes, leading to methodological biases that confound comparative analysis. To remedy this, Gramene staff worked with the NSF *Oryza* Genome Evolution (OGE) project to develop a consistent annotation protocol (using the MAKER-P pipeline) that was applied to both the new genomes and the *Oryza* genomes previously sequenced and annotated by other projects. To compare both sets of annotations, the Gramene team created a new Website ([oge.gramene.org](http://oge.gramene.org)) for visualization and analysis of I-OMAP genomes and phylogenetic trees.

## PLANT SYSTEMS BIOLOGY

### Exploring *Arabidopsis* Gene Regulatory Networks

L. Zhang, Y.K. Lee, C. Liseron-Monfils, C. Noutsos, B. Wang [in collaboration with S. Brady, University of California, Davis; DuPont-Pioneer, Inc.]

The goal of this work is to characterize gene regulatory networks associated with development and stress response. microRNA (miRNA) and transcription factors (TFs) have important roles in plant development and responses to the environment. To understand the important regulators of miRNA expression, we are using a gene-centered yeast one-hybrid (Y1H) experimental system to screen for TFs that bind to promoters of miRNA genes and their targets. Over the years, we have developed a library resource of 950 *Arabidopsis* TFs, representing >95% of the TFs expressed in roots. Roots are important plant organs responsible for uptake of water and nutrients as well as the physical support of the plant. Understanding the genes responsible for root system architecture is critical to addressing key traits, including drought tolerance, cycling of nutrients such as phosphorus, and nitrogen uptake.

Expanding upon previous years' work, the new interaction data enabled us to generate a gene regulatory network (GRN) comprising more than 5000 protein–DNA interactions (PDIs) involving 645 TFs and 180 promoters. To this network, we have also incorporated known posttranscriptional regulatory pathways mediated by miRNAs upon their targets.

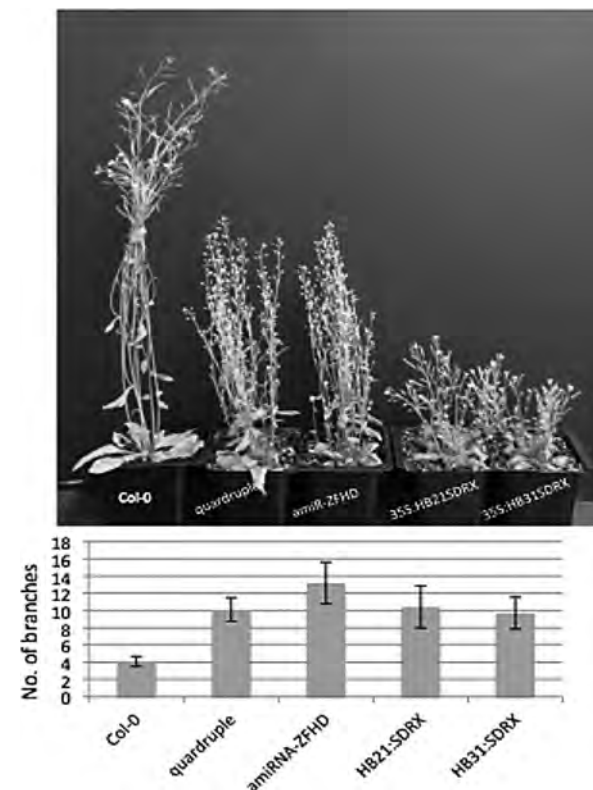
Although the network is enriched for genes expressed in root, it also serves as a model for other tissues in which those genes are expressed. Using comparative genomic projections, we are also beginning to model these networks in other species, including maize and sorghum. Ultimately, we plan to use these gene regulatory networks to generate testable hypotheses about cell-type-specific phenotypes.

To further validate the GRN and to dissect the phenotypic consequences of network perturbation, we are characterizing mutant TF lines. One T-DNA insertion mutant, *hat22*, exhibits a short root phenotype. Overexpression of AtHAT22, a homeobox domain transcription factor, results in longer primary roots, suggesting that this protein has a positive role in root development. Cellular-level investigations revealed that the short root length in the *hat22* mutant was the result of reduced cell number in meristem zone. To obtain further insight into the role of AtHAT22 in the regulation of root development, we characterized the spatiotemporal expression pattern of AtHAT22. Specifically, we performed RNA-Seq and ChIP-Seq in order to directly investigate the genes regulated by AtHAT22: 3754 genes showed significant ( $p < 0.01$ ) differential expression between the *hat22* mutant and the wild type. Furthermore, most of the HAT22-binding peaks were close to TSSs (transcription start sites), and many present in intergenic regions. We also found a consensus HAT22-binding motif using MEME. The Chip-Seq data revealed that HAT22 bound the promoter region of HB21, a zinc finger homeobox (ZF-HD) gene. Like the *hat22* mutant, the *hb21* mutant also exhibited a short root phenotype; moreover, the *hat22/hb21* double mutant had much shorter roots than either single mutant, suggesting their genetic interaction within a common pathway. Further comparison of the transcriptome revealed a 30% overlap between the sets of genes differentially expressed in the *hb21* and *hat22* mutants.

GRNs are characterized by highly connected genes as “hubs” in the network. The ZF-HD TFs were identified as hub genes within the miRNA network. An initial survey of six single loss-of-function mutants identified from existing T-DNA collections revealed no observable root phenotypes, suggesting that the ZF-HD TFs are functionally redundant. To test this hypothesis, we generated multiple loss-of-function mutants using different approaches, including genetic

crosses to a single loss-of-function mutant, generation of artificial miRNA, and independent repressor lines (Fig. 3). Combining multiple loss-of-function mutants within a single *Arabidopsis* line, we observed several phenotypes, including altered flower structures and an increased vegetative branching. In addition, we performed transcriptome profiling of ZF-HD TF mutants in order to characterize differentially expressed genes. The gene-expression profiles and phenotypic properties of the mutants suggested a role for ZF-HD TFs as regulators of developmental transitions. Moreover, the results of this work demonstrate that the miRNA GRN can be applied more generally, beyond the root system.

The basic leucine zipper (bZIP) TFs are another highly connected TF family. One member of the family, GBF2, binds many miRNA and miRNA target promoters. As with the ZF-HD TFs, single-gene loss-of-function mutants do not have clear phenotypes. We are in the process of generating multiple



**Figure 3.** Phenotype of the ZF-HD TF mutant. (Top) Branching phenotype of quadruple mutant, artificial miRNA, and two independent repressor lines compared to wild-type plant Col-0. (Bottom) Lateral branching number of wild type and mutants.

loss-of-function mutants by performing genetic crosses between the available single-gene mutant lines. Initial screens for homozygous mutants from multiple loci suggests that these mutants may be lethal or have large fitness disadvantages.

In addition to our miRNA targeted GRN, as part of a collaboration with Siobhan Brady's lab, the Y1H library has been used to mapped networks focused on the secondary cell wall (Taylor-Teeple et al. 2015) and nitrogen use efficiency (NUE). The plant cell wall is an important factor for determining cell shape, function, and response to the environment. Secondary cell walls, such as those found in xylem, are composed of cellulose, hemicelluloses and lignin and account for the bulk of plant biomass. The coordination between transcriptional regulation of synthesis for each polymer is complex and vital to cell function. The resulting GRN allows us to develop and validate new hypotheses about secondary wall gene regulation under abiotic stress.

As sessile organisms, plants have limited means to explore and acquire nutrients from the environment, but they have evolved genetic mechanisms that control uptake and other responses to varying levels of mineral nutrients. Nitrogen is required by plants in greater amounts than any other mineral and is necessary for basic cellular function. Higher plants most commonly take up nitrogen in the form of nitrate, but nitrate sources are often insufficient and heterogeneous in the soil. Nitrogen metabolism is regulated at the transcriptional, translational, and posttranslational levels. We are mapping transcriptional control of genes involved in nitrogen uptake, assimilation, and downstream metabolism.

To support cross-species comparisons, in the last year, we obtained cloned sequences of 2017 unique maize TFs from the NSF-funded maize TFome project awarded to Drs. Grotewold, Ohio State University, and Gray, University of Toledo (Burdo et al., *Plant J* 80: 356 [2014]), and utilized them to develop a maize TF resource. The current collection consists of 1846 unique TFs, arranged in 96-well format. We have generated projections of the *Arabidopsis* network in maize and sorghum and are currently testing selected specific hypotheses.

As part of this research, we are evaluating methods to combine expression data with gene network models. To support these activities, we have developed the NECorr tool. In contrast to other network inference

tools that generate networks from expression data, NECorr analyzes an existing network in conjunction with expression data to define the most important part of the network. To achieve this goal, NECorr uses a linear model to combine topological properties of the network with coexpression and differential expression data. To validate NECorr, we used a receiver operative characteristic (ROC) to determine the relationship between sensitivity and specificity. Preliminary results indicate that NECorr identifies candidate genes more accurately than methods relying on differential expression alone ( $AUC_{NECCorr} = 0.75-0.9$ ;  $AUC_{Differential\ Expression} = 0.663$ ). Using NECorr and our miRNA GRN, and making use of publicly available spatial-temporal expression data from *Arabidopsis* root, we determined that miR-398 mediates significant cross-talk between stress responses and development within transition zones of roots. These results were subsequently confirmed by mutant analysis.

## Developmental Networks Controlling Inflorescence Architecture in Grasses

A.L. Eveland, S. Kumari, C. Liseron, Y. Jiao [in collaboration with D. Jackson, Cold Spring Harbor Laboratory; and Z. Xin and J. Burke, USDA ARS]

The goal of this work is to integrate genetics and genomics data sets to find molecular networks that influence the morphology (architecture) of grass inflorescences (flowers). Because inflorescences bear the fruits and grains that we eat, the genetic and regulatory factors that govern their formation are clearly relevant to important agronomic traits such as grain yield and harvesting ability. Our work in maize focuses on inflorescence primordia sampled during key developmental transitions and in perturbed genetic backgrounds. The latter includes loss-of-function mutants in three important regulators of the *RAMOSA* (*RA*) pathway, which controls stem-cell fate decisions and ultimately the decision to branch. We established a robust system for investigating the networks that modulate branching, including characterization of precise timing of developmental events and the associated spatio-temporal changes in gene expression. We integrated genome-wide mRNA-Seq data to resolve coexpression networks during key stages of maize inflorescence development, and we are currently working to expand these networks by incorporating additional data sets,

including genome-wide TF occupancy profiles and *cis*-regulatory information.

The following were the primary objectives for this project.

1. Establish a comprehensive pipeline for mRNA-Seq and ChIP-Seq data analysis in maize by evaluating and optimizing available software for mapping and quantitation. This also included testing various statistical methods to extract biological relevance.
2. Characterize genome-wide expression signatures specific to a given developmental event or branching phenotype. To this end, we made use of known developmental marker genes and their spatiotemporal transcriptional responses to genetic perturbation in order to test our experimental system and establish a proxy for developmental staging.
3. Evaluate and implement clustering approaches to identify candidate genes that are coexpressed with key regulators and/or coincide with specific developmental events. These efforts resulted in identification of candidate genes, in particular developmentally regulated TFs and novel genes of unknown function that are potentially involved in stem-cell maintenance and determinacy. We continue to examine coexpression clusters for enrichment of functional processes and *cis*-regulatory motifs near the transcriptional start sites (TSSs) of coexpressed genes.
4. Identify targets of the *RAI* TF using ChIP-Seq and integrate the results with data from parallel mRNA-Seq experiments. On the basis of this approach, we showed that one-third of genes with altered expression levels in the *rai* mutant are also bound by *RAI* (Eveland et al. 2014). As they become available, we are incorporating additional ChIP-Seq data sets to investigate combinatorial binding of TFs associated with the branching pathway. The ChIP-Seq data also provide *in vivo* confirmation of binding sites of developmental regulators in maize, and the resultant information is being used in efforts to resolve *cis*-regulatory modules across the maize genome.

In addition to maize, we have begun work on sorghum, an important emergent bioenergy crop that is also used for human consumption in sub-Saharan Africa. In particular, we are using next-generation sequencing approaches to identify single-nucleotide mutations associated with an increase in seed number.

Using this approach, we have identified two genes that can change the structure of flowers and generate more seeds. We intend to apply this strategy to a large-scale mutant study aimed at identifying candidate genes associated with developmental traits in the roots, shoots, and flowers of sorghum.

In the next phase, we will prioritize candidates identified in this stage of the project by overlaying *Arabidopsis* regulatory network information (see previous section). The resulting hypotheses (e.g., those related to stress responses) can be tested in *Arabidopsis* and ultimately translated into agronomic systems. Candidate genes that are maize- and/or grass-specific are also of high priority, because they may contribute to the unique morphology of maize inflorescences and/or features shared among other grasses. In addition, we will use comparative genomics approaches, including both cross-species comparisons and integration of analogous RNA-Seq data sets from closely related grasses such as sorghum, to identify candidate genes that may contribute to grass-specific aspects of inflorescence architecture.

## CYBERINFRASTRUCTURE PROJECTS

### The iPlant Collaborative

L. Wang, Z. Lu, C. Noutsos, K. Chougule, P. Van Buren, J. Stein

This project is a collaboration with Cold Spring Harbor Laboratory, employs more than 100 staff, and is headquartered at the University of Arizona (PIs: Stephen Goff, Doreen Ware, Nirav Merchant, Matthew Vaughn, and Eric Lyons). Dozens of collaborators are located at more than 20 institutions.

Our world is changing rapidly. The human population is increasing while arable land and fisheries are decreasing and food cultivation is being diverted for fuel production. Climate instability and energy sustainability are impacting agricultural and ecological systems while concomitant changes in land-use patterns affect global biodiversity. To successfully address these issues, we need to understand how organisms' appearance, physiology, and behavior are shaped by the interactions between their genetic makeup and the environment. Although these global challenges are sobering, the efforts to respond productively will lead to exciting science—provided that the computational infrastructure is in place to handle the necessary data sets, analyses, interpretation of results, and dissemination of knowledge.

Advances in biological research technology have enabled scientists to amass unprecedented amounts of data, and many researchers find themselves drowning in this sea of data. Foreseeing this major bottleneck in biological research, the NSF established the iPlant Collaborative (<http://iplantcollaborative.org>) in 2008 to develop cyberinfrastructure for life sciences research and democratize access to U.S. supercomputing capabilities.

Having completed its first 5-year grant, the iPlant Collaborative has made extensive progress toward meeting these goals, and last year, it was renewed for another 5 years. Work in the last year culminated in the development of multiple platforms that are now being used by the research and educational communities. Staff at CSHL contributed directly to some of these cyberinfrastructure platforms or built upon existing platforms to provide scientists and educators with ready access to needed software and analysis tools. Within the Ware lab, these platforms include the iPlant Data Store, Discovery Environment (DE), and Atmosphere.

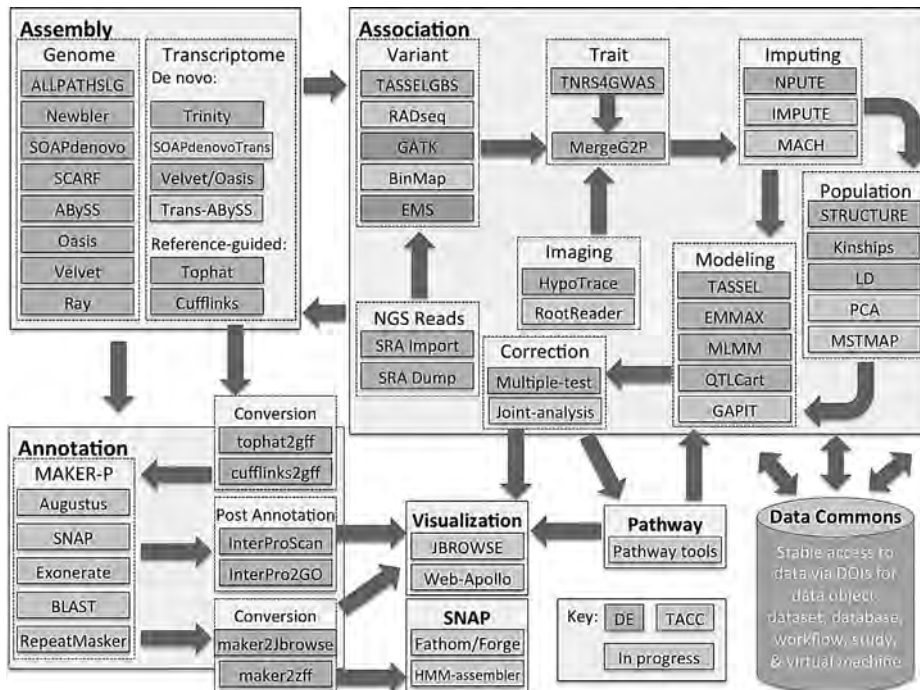
The iPlant Data Store was designed as a distributed storage system for hosting data on the Cloud, providing convenient access. The distributed system allows us to “bring the infrastructure to the data” to enable fast computing and, at the same time, to reduce the bandwidth required to transfer large amounts of next-generation sequencing data. Currently, the iPlant Data Store consists of two major storage systems with fast Internet connections, one at the University of Arizona (UA) and another one at the Texas Advanced Computing Center (TACC). CSHL itself has relatively limited bandwidth relative to TACC and UA, making our institution an ideal case for validation of the iPlant model of bringing the infrastructure to the data. Because it makes more sense to utilize local computing resources with iPlant pre-built analysis workflows via the iPlant AGAVE API than to copy the data to outside clusters (e.g., it takes >10 min to move 1 GB data from CSHL to the TACC clusters), the Ware lab at CSHL decided to migrate the important iPlant platforms at CSHL. The first target is the data store, and a local resource server (WildCat) has been successfully added into the iPlant storage system, allowing data transfer that is 80 times faster than with outside servers. For the next step, the Ware lab is going to synchronize the resource server with iPlant’s AGAVE servers at TACC, which will allow iPlant’s pre-built workflows to be run locally on CSHL computing

servers. The migration of these important platforms will provide a proof-of-concept demonstration of the portability of iPlant Platforms for enrolling more institutes for efficient data management and possibly utilization of national computing resources provided with good bandwidth.

The DE is the most visible portal for iPlant tools and services. This Web-based platform supports an “app store” model of user-extensible tools, automated workflows, and data storage. Users can take advantage of existing tools, integrated by iPlant staff and the user community, or add their own tools to use privately or share. Although users may not be aware of this, the underlying infrastructure provides access to iPlant’s massive data store at UA and TACC. Computationally intensive tasks are handled by supercomputers located at TACC and other centers within the Extreme Science and Engineering Discovery Environment (XSEDE). So far, more than 300 tools have been integrated into the DE; these tools enable a broad range of research activities, including genome/transcriptome assembly, annotation, RNA-Seq quantitation, variant detection, GWAS (genome-wide association study), and phylogenetics. Members of the Ware lab have played important roles in contributing to workflow design, tool integration, validation, science tutorials, and documentation.

Atmosphere is iPlant’s configurable and Cloud-enabled computational resource for the plant research community. From Atmosphere’s Web interface, users can launch a virtual machine (VM) with pre-configured working environments and precustomized, ready-to-use software. Users can also create their own applications and environments as VMs and share them with others via Atmosphere. As with the DE, Atmosphere is a gateway to iPlant’s core infrastructure resources, including the high-performance grid computing environment and big-data storage system. Using the Atmosphere platform, we created a VM to be used for image analysis based on high-throughput phenotyping.

The success of genome research depends on our ability to accurately assemble, annotate, and derive meaning from sequence data; however, the extremes of genome size, polyploidy, diversity, and repeat content push the limits on the algorithms, expertise, and computational power currently available to researchers. In response, iPlant is fostering a community effort to identify best practices and state-of-the-art tools,



**Figure 4.** iPlant Workflow: Summary of iPlant apps and their relationships. The flowchart is divided into Assembly, Annotation, Association, Visualization, and other components. A genome (or transcriptome) can be assembled by various assemblers and annotated with MAKER-P. The assembled genome will serve as reference genome for variant callers, and variant can be combined with trait data for association analysis. Annotation and association results can be visualized with various browsers and combined with pathway databases. All data can be accessed from the iPlant Data Commons.

install them, optimize their performance on the nation's most powerful supercomputers, and make them available as free online resources. During the last 3 years, the iPlant Discovery Environment has matured to provide a comprehensive set of tools and services for sequence handling, read alignments, RNA-Seq profiling, and de novo genome and transcriptome assembly (Fig. 4). To extend these capabilities, we have incorporated MAKER-P, a standardized, portable, and easy-to-use plant-genome annotation engine with built-in methods for quality control. As part of this effort, MAKER-P was specifically optimized to take advantage of the parallel computing environment of the TACC Lonestar cluster and is now a supported module. Performance testing showed that MAKER-P provides high-quality, full-fledged annotation pipelines on even the largest plant genomes in a matter of hours. MAKER-P is currently available for use as an Atmosphere image.

A major mission of iPlant is to promote adoption of the cyberinfrastructure through training workshops and outreach at academic institutions and scientific

meetings. In 2014, members of our lab participated as instructors in several "Big Data" workshops focused on transcriptomics and population genomics using the DE and Atmosphere platforms.

## KBase: The Department of Energy Systems Biology Knowledgebase

V. Kumar, S. Kumari, J. Thomason

This project is a collaboration with the DOE National laboratories and is led by PI Adam Arkin of Lawrence Berkeley National Laboratory (LBNL), with co-PIs Rick Stevens of Argonne National Laboratory (ANL), Robert Cottingham of Oak Ridge National Laboratory (ORNL), and Sergei Maslov of Brookhaven National Laboratory (BNL). I serve as the Plants Science Team Lead for KBase. Other PIs include Mike Schatz of Cold Spring Harbor Laboratory, Dave Weston of ORNL, and Pamela Ronald of the University of California, Davis.

The Systems Biology Knowledgebase (KBase, [www.kbase.us](http://www.kbase.us)) has two primary goals. The scientific goal



is to produce predictive models, reference data sets, and analytical tools and to demonstrate their utility in DOE biological research relating to bioenergy, carbon cycle, and the study of subsurface microbial communities. The operational goal is to create the integrated software and hardware infrastructure needed to support the creation, maintenance, and use of predictive models and methods for the study of microbes, microbial communities, and plants.

KBase's computational infrastructure is supported by a distributed, high-performance, cloud-based system that includes >3 petabytes of storage, more than 12,000 cores for data processing, and 90 GBit/sec bandwidth over DOE's Energy Sciences Network (ESnet). The KBase data model supports more than 900 data types, including sequence reads, assemblies, genomes, annotations, expression data, quantitative phenotype data, and metagenomic profiles. The KBase data repository has integrated various data sets from public resources for thousands of microbes, hundreds of microbial communities, and tens of plants.

In addition, KBase supports more than 1000 analysis and access functions grouped into more than 30 integrated service modules, including sequence assembly, variation analysis, genome annotation, expression analysis, and various types of modeling (e.g., metabolic, regulatory, and flux balance analysis [FBA]). To capture complex, multistep analyses, users can build persistent and provenance-based workflows that can capture complex, multistep analyses and mobilize rich annotations, visualization widgets, reusable workflows, and custom scripts. This capability empowers users to create science that is transparent, reproducible, and reusable.

During 2014, the plants team implemented multiple workflows, also known as narratives. These narratives supported functionality such as (1) incorporating functional and network information (ontology, coexpression, protein-protein interaction, and pathway) in genome-wide association studies (GWAS) to identify genetic variants associated with traits of interest; (2) using RNA-Seq data to quantify expression and identify differentially expressed genes, constructing a coexpression network from gene expression data, and conducting an annotation enrichment analysis of densely interconnected clusters to identify functional modules and draw biological inferences; (3) generating a genome from plant transcript sequences provided by the user, based on a set of gene models in a plant genome or an assembly of transcripts

from next-generation sequencing (NGS) of a transcriptome, and annotating the genome with metabolic functions using PlantSEED Subsystems; and (4) generating a metabolic model using the annotated genome, followed by gap-filling and FBA (flux balance analysis) to simulate growth and identify the set of metabolic reactions that are active in plant biomass production.

Another focus within the project is the creation of a library of Web interface components ("widgets") for rapid development of reusable visualizations that can be dropped into disparate interfaces to present biological data with new insights. Such visualizations help scientists interpret complex data sets. Building upon standard Web libraries, such as jQuery, Bootstrap, and D3, has allowed the creation of a flexible and very familiar widget standard. This standard provides dynamic visualizations that can immediately update as the underlying data changes, as well as interactive visualizations that allow graphs to operate as user interfaces such that image manipulations translate into actual data modifications in real time.

## PUBLICATIONS

- Campbell MS, Law M, Holt C, Stein JC, Moghe GD, Hufnagel DE, Lei J, Achawanantakun R, Jiao D, Lawrence CJ, et al. 2014. MAK-ER-P: A tool kit for the rapid creation, management, and quality control of plant genome annotations. *Plant Physiol* **164**: 513–524.
- Eveland AL, Goldshmidt A, Pautler M, Morohashi K, Liseron-Monfils C, Lewis MW, Kumari S, Hiraga S, Yang F, Unger-Wallace E, et al. 2014. Regulatory modules controlling maize inflorescence architecture. *Genome Res* **24**: 431–443.
- Kersey PJ, Allen JE, Christensen M, Davis P, Falin LJ, Grabmuller C, Hughes DS, Humphrey J, Kerhornou A, Khobova J, et al. 2014. Ensembl Genomes 2013: Scaling up access to genome-wide data. *Nucleic Acids Res* **42**: D546–D552.
- Monaco MK, Stein J, Naithani S, Wei S, Dharmawardhana P, Kumari S, Amarasinghe V, Youens-Clark K, Thomason J, Preece J, et al. 2014. Gramene 2013: Comparative plant genomics resources. *Nucleic Acids Res* **42**: D1193–D1199.
- Schatz MC, Maron LG, Stein JC, Wences A, Gurtowski J, Biggers E, Lee H, Kramer M, Antoniou E, Ghiban E, et al. 2014. Whole genome de novo assemblies of three divergent strains of rice, *Oryza sativa*, document novel gene space of aus and indica. *Genome Biol* **15**: 506.
- Seaver SM, Gerdes S, Frelin O, Lerma-Ortiz C, Bradbury LM, Zallot R, Hasnain G, Niehaus TD, El Yacoubi B, Pasternak S, et al. 2014. High-throughput comparison, functional annotation, and metabolic modeling of plant genomes using the PlantSEED resource. *Proc Natl Acad Sci* **111**: 9645–9650.
- Zwickl DJ, Stein JC, Wing RA, Ware D, Sanderson MJ. 2014. Disentangling methodological and biological sources of gene tree discordance on *Oryza* (Poaceae) chromosome 3. *Syst Biol* **63**: 645–659.

*In Press*

- Law M, Childs KL, Campbell MS, Stein JC, Olson AJ, Holt C, Panchy N, Lei J, Jiao D, Andorf CM, et al. 2015. Automated update, revision, and quality control of the maize genome annotations using MAKER-P improves the B73 RefGen\_v3 gene models and identifies new genes. *Plant Physiol* **167**: 25–39.
- Monaco M K, Stein JC, Wei S, Youens-Clark K, Jaiswal P, Ware D. 2015. Gramene: A resource for comparative analysis of plants genomes and pathways. *Plant bioinformatics: Methods and protocols. Methods Mol Biol.*
- Taylor-Teeple M, Lin L, de Lucas M, Turco G, Toal TW, Gaudinier A, Young NF, Trabucco GM, Veling MT, Lamothe R, et al. 2015. An *Arabidopsis* gene regulatory network for secondary cell wall synthesis. *Nature* **517**: 571–575 (in press).
- Wang L, Ware D, Lushbough C, Merchant N, Stein L. 2015. A GWAS platform built on iPlant cyber-infrastructure. *Concurrency and computation: Practice and experience* **27**: 420–432.

# QUANTITATIVE BIOLOGY

---

**Gurinder S. “Mickey” Atwal** and colleagues are applying insights from the physical and computational sciences to the study of population genetics, machine learning, and human disease. The Atwal lab has modeled the process by which genetic variants, or alleles, have evolved in the last 100,000 years of human history. This has recently led to surprising insights about the role of *TP53*, a master tumor suppressor gene, in female fertility and furthered our understanding of how complex gene networks evolve. The lab has analyzed the comparative genomics and physical organization of cancer-related genes and their role in mediating tumorigenesis across numerous tissue types. Recently, they have begun to focus efforts on understanding cancer genome evolution on shorter timescales by analyzing nucleotide sequences from single cells.

**Ivan Iossifov** focuses on the development of new methods and tools for genomic sequence analysis and for building and using molecular networks, and he applies them to specific biomedical problems. He studies 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.

**Alexander Krasnitz** and colleagues use mathematical and statistical tools to discover key genetic elements involved in cancer and to understand how cancer cells evolve. Array-based comparative genome hybridization, a technique honed in the Wigler lab, and, more recently, sequencing experiments have revealed subtle patterns of frequent and widespread aberration in cancer genomes. Krasnitz hypothesizes that recurrent, aberrant genomic loci observed in a range of cancer types are under selection and therefore are enriched in important cancer genes. He has developed a novel, comprehensive methodology to discover such “cores” and has used it to analyze multiple genome data sets in breast, liver, ovarian, and prostate cancer. The results have been shared with cancer biology labs across CSHL, and they have been a key enabling agent of functional studies using mouse models and RNA interference. Krasnitz has begun to apply these novel statistical tools to the latest generation of experimental data, which have characterized tumor samples down to the level of single cells. By interpreting single-cell genomes, he and colleagues seek to learn how specific tumors evolve and how cancer cells migrate to invade adjacent tissues and metastasize.

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 encompassing thousands of nuclear families. After earlier working with high-resolution CGH arrays, Levy’s group now uses targeted sequence data. Levy 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 lab include analysis of single-cell RNA, phylogenetic reconstruction from sparse data sets, and disentangling haplotypes from sperm and subgenomic sequence data.

**Michael Schatz** is a computational biologist and an expert at large-scale computational examination of DNA sequencing data, including the alignment, assembly, and analysis of next-generation

sequencing reads. These methods have been used to reconstruct the genomes of previously unsequenced organisms, probe sequence variations, and explore a host of biological features across the tree of life. Recent improvements in sequencing technologies are challenging our capacity to store and analyze the huge volume of DNA sequence data being generated. Consequently, Schatz is particularly interested in capitalizing on the latest advances in distributed and parallel computing, especially cloud computing technologies, to advance the state of the art in bioinformatics and genomics. In a recent breakthrough, Schatz was able to create a hybrid software-based solution to eliminate errors in so-called third-generation sequencing. This makes it remarkably easier to compile, align, and analyze full-genome sequences.

# QUANTITATIVE BIOLOGY

G.S. Atwal R. Aboukhalil A. Gerwitz

Fueled by data generated from recent technological developments in DNA sequencing, our lab is primarily focused on population genetics, cancer biology, and high-performance computing. We often tackle scientific questions analytically and computationally by invoking theoretical concepts from statistical physics and machine learning. A long-running thread in our work is the analysis of cancer genomes from an evolutionary perspective in order to determine how this informs on clinical outcome and therapy.

A major new research direction in our lab is the investigation of ectopic expression of the germline set of genes in various tumors. Previous work in our lab has established the pleiotropic and ancient role of the TP53 pathway in germline development. Preliminary results have also identified ectopic expression of germline *piwi* genes in samples of glioblastoma multiforme extracted from the Cancer Genome Atlas. However, the landscape of ectopically expressed germline genes is unknown and their functional impact on cancer development remains elusive. We began an integrative study that sought to determine the ectopic expression of germline transcripts in glioblastoma multiforme and elucidate the predictive value of germline molecular programs as diagnostic markers and immunogenic therapeutic targets.

Previous work in our lab, in collaboration with colleagues at the Cancer Institute of New Jersey and Weill Cornell Medical College, has established the association between single-nucleotide polymorphisms (SNPs) in the p53 tumor suppressor pathway and female infertility in mice and humans. This hypothesis was first generated through computational investigations of haplotypic diversity and positive selection in genes in the TP53 pathway, where we detected signatures of recent positive selection in SNPs present at high frequencies in various human populations. Selected alleles in SNPs in the *LIF*, *Mdm2*, *Mdm4*, and *Hausp* genes, each of which lies in the p53 molecular network, were found to be enriched in women undergoing in vitro fertilization (IVF) treatment. Moreover, some of these SNPs have been demonstrated to be associated with estrogen-driven cancer risk, highlighting the pleiotropic character of genetic variants in the p53 pathway. We have also sought association of

SNPs in p63 and p73, and initial results are encouraging, although we await an increase in sample numbers before the results can be deemed to be statistically significant.

Following up on these earlier investigations, we have begun exploring the contribution of multiple alleles from both a population genetics and a human disease perspective. In addition, we have begun refocusing efforts applying the tools of population genetics to understand single-cell evolution from tumors. In particular, we have been addressing the ubiquitous problem of how many cells and how much read coverage we will need before the inferred cell phylogeny accurately reflects the evolutionary history of the tumor. In addition, our lab has continued research in the use of information theory and other machine learning tools in addressing the tsunami of data generated by next-generation sequencing.

Bud Mishra, a faculty member of the Courant Institute, New York University, is an ongoing collaborator. Bud and I co-organized the RECOMB-Comparative Genomics conference, held for the first time at Cold Spring Harbor Laboratory in October. Ricki Korff, an URP student during the summer of 2013, continues to intermittently work with us remotely on the project concerning the ectopic expression of germline genes in tumors. Aspects of this project were spearheaded by 2014 URP student Ariel Gerwitz. Robert Aboukhalil, a graduate student from the Watson School, has continued to make progress developing software to analyze copy-number data obtained from single tumor cells. In addition, two first-year students from the Watson School graduate program, Maria Nattestad and Daniel Kepple, joined the lab for rotation projects.

## Responses to “Equitability, Mutual Information, and the Maximal Information Coefficient”

This work was done in collaboration with J. Kinney (Cold Spring Harbor Laboratory).

We received two published responses to our widely publicized manuscript on equitability, to which we

responded in return. The original paper dealt with the following question: How should one quantify the strength of association between two random variables without bias for relationships of a specific form? Despite its conceptual simplicity, this notion of statistical “equitability” had yet to receive a definitive mathematical formalization. We argued that equitability is properly formalized by a self-consistency condition closely related to Data Processing Inequality. Mutual information, a fundamental quantity in information theory, was shown to satisfy this equitability criterion. These findings were at odds with the recent work of Reshef et al. (*Science* 334: 1518[2011]), which proposed an alternative definition of equitability and introduced a new statistic, the “maximal information coefficient” (MIC), said to satisfy equitability in contradistinction to mutual information. These conclusions, however, were supported only with limited simulation evidence, not with mathematical arguments. Upon revisiting these claims, we proved that the mathematical definition of equitability proposed by Reshef et al. cannot be satisfied by any (nontrivial) dependence measure. We also identified artifacts in the reported simulation evidence. When these artifacts were removed, estimates of mutual information were found to be more equitable than estimates of MIC. Mutual information was also observed to have consistently higher statistical power than MIC. We concluded that estimating mutual information provides a natural (and often practical) way to equitably quantify statistical associations in large data sets. The responses did not negate any of our original results and conclusions.

### Ectopic Germline Gene Expression in Cancer

This work was done in collaboration with R. Darnell (Rockefeller University) and A. Haase (Cold Spring Harbor Laboratory).

We initiated an integrative study that sought to determine the ectopic expression of germline transcripts in glioblastoma multiforme and elucidate the predictive value of germline molecular programs as diagnostic markers and therapeutic targets.

To determine the genome-wide landscape of restricted germline expression, we developed an empirical Bayesian computational approach to analyze normal

human soma and germline tissues represented across 1236 heterogeneous expression data sets. Germline tissues were found to express a significantly large, but tightly knit, network of genes ( $n = 1294$ ,  $q$ -value =  $10^{-2}$ ) that we inferred to be tissue-restricted; the functionality of these genes is enriched in molecular pathway processes such as piwi-interacting RNA (piRNA) metabolism ( $p = 10^{-6}$ ) and cell cycle processes ( $p = 10^{-36}$ ). By analyzing RNA-Seq data from the Cancer Genome Atlas, we identified 385 high-confidence germline genes ectopically expressed in glioblastoma multiforme samples, that have potential to serve as sources of neoantigens. Analyses of clinical survival data revealed a stratification of patients with high and low number of germline gene transcripts, demonstrating that the germline genetic program can inform on clinical outcome.

To elucidate the role of the germline piRNA processing pathway in cancer, our collaborators purified ectopic piwi-complexes in cultured cancer cells and investigated the identity of associated nucleic acids. Next-generation sequencing revealed that the piwi proteins are not loaded with piRNAs, suggesting a piRNA-independent mechanism for their reactivation. Current experimental efforts are geared toward investigating piRNA pathways in patient samples.

### Optimal Inference and Multistable Perception

The stochastic dynamics of multistable perception pose an enduring challenge to our understanding of neural signal processing in the brain. We formulated a mathematical theory to show that the emergence of perception switching and stability can be understood using principles of probabilistic Bayesian inference where the prior temporal expectations are matched to a scale-free power spectrum, characteristic of fluctuations in the natural environment. The optimal percept dynamics are inferred by an exact mapping of the statistical estimation problem to the motion of a dissipative quantum particle in a multiwell potential. In the bistable case, the problem is further mapped to a long-ranged Ising model. Optimal inference in the presence of a  $1/f$  noise prior leads to critical dynamics, exhibiting a dynamical phase transition from unstable perception to stable perception, as demonstrated in recent experiments. The effect of stimulus fluctuations and perception bias is currently being discussed.

PUBLICATIONS

- Atwal G. 2014. Statistical mechanics of multistable perception. [bioRxiv/10.1101/008177](https://doi.org/10.1101/008177)
- Garvin T, Aboukhalil R, Baslan T, Kendall J, Hicks J, Atwal G, Wigler , Schatz M. 2014. Interactive analysis and quality assessment of single-cell copy-number variations. [bioRxiv/10.1101/011346](https://doi.org/10.1101/011346)
- Kinney J, Atwal GS. 2014. Parametric inference in the large data limits using maximally informative models. *Neural Computation* **26**: 637–653.
- Kinney J, Atwal GS. 2014. Equitability, mutual information, and the maximal information coefficient. *Proc Natl Acad Sci* **111**: 3354–3359.
- Kinney J, Atwal GS. 2014. Reply to Reshef et al.: Falsifiability or bust. *Proc Natl Acad Sci* **111**: E3304.
- Kinney J, Atwal GS. 2014. Reply to Murrell et al.: Noise matters. *Proc Natl Acad Sci* **111**: E2161.
- Parida L, Atwal GS, Mishra B. 2014. Proceedings of the Twelfth Annual RECOMB Comparative Genomics Conference. *BMC Genomics* **15(Suppl 6)**: I1.

# DE NOVO MUTATION IN AUTISM

## I. Iossifov

I study the genetics of common diseases in humans, using two main tools: next-generation sequencing and molecular networks representing functional relationship among genetic loci. These tools in combination enable the large-scale studies necessary for furthering our understanding of the complex etiology of disorders such as autism, bipolar disorder, and cancer. I focus on both developing new methods (for sequence analysis and for building and using molecular networks) and applying them to specific biomedical problems.

### Exome Sequencing Reveals De Novo Gene Disruptions in Children on the Autistic Spectrum

---

We found that children on the autistic spectrum (ASD) have increased incidence of de novo (DN) missense, “likely gene-disrupting” (LGD), and copy-number variant (CNV) mutations compared to unaffected siblings. We estimated that 70% of DN CNV, 42% of DN LGD, and 13% of DN missense mutations contributed to 6%, 9%, and 12% of diagnoses in simplex families, respectively. DN mutation in the coding sequence contributes to nearly 30% of all simplex diagnoses and 45% of female diagnoses; 27 genes were recurrently hit by DN LGD mutations in affected children. CHD8 was hit by nine DN LGDs (in nine affected children), DYRK1A by four, and each of ANK2, GRIN2B, DSCAM, and CHD2 was hit by three DN LGDs. Males with DN LGDs or DN CNVs have lower IQ. The gene targets of DN LGDs in ASD males with lower IQ overlap with targets in ASD females and with individuals having intellectual disability or schizophrenia, but not significantly with targets in ASD males having higher IQ. We estimate the number of vulnerable genes in which an LGD mutation can cause ASD in females or lower IQ ASD in males to be ~400, with a similar number of genes vulnerable to missense mutation. LGD targets are enriched for chromatin modifiers and fragile X mental retardation protein (FMRP)-associated genes in both

affected males of lower IQ and females. Embryonically expressed genes are significantly enriched in DN targets, LGD, and missense only in affected females.

### Ongoing Projects

---

**Whole-genome sequencing of Simons Simplex Collection.** Whole-exome sequencing (WES) will not exhaust the information available in the Simons Simplex Collection (SSC); more can be uncovered by whole-genome sequencing (WGS). As with WES, the main goal of WGS is gene target discovery in simplex families through the identification of disruptive de novo events. Certainly, all the data that one can derive from exome capture can also be extracted from whole-genome data, hopefully with more even coverage of the exons. By that alone, we anticipate extracting on the order of 30% more events. Additionally, we will observe de novo events in promoter regions, in evolutionarily conserved regions near target genes, and in noncoding genes like microRNAs and long noncoding RNAs. These genomic regions are not targeted by WES, yet they may add to the discovery of recurrent targets and to the list of causative mechanisms for genetic disruption. Although the interpretation of any individual event is likely to be difficult, we can nevertheless calculate a differential signal between affected and unaffected siblings for various regions of suspected functionality. Finally, WGS will certainly facilitate the discovery of small copy-number changes and copy number–neutral rearrangements.

### PUBLICATIONS

- Cokol M, Weinstein ZB, Yilancioglu K, Tasan M, Doak A, Cansever D, Mutlu B, Li S, Rodriguez-Esteban R, Akhmedov M, et al. 2014. Large-scale identification and analysis of suppressive drug interactions. *Chem Biol* **21**: 541.
- Glessner J, Bick AG, Ito K, Homsy J, Rodriguez-Murillo L, Fromer M, Mazaika EJ, Vardarajan B, Italia MJ, Leipzig J, et al. 2014. Increased frequency of de novo copy number variations in congenital heart disease by integrative analysis of SNP array and exome sequence data. *Circulation Res* **115**: 884.



- Iossifov I, O’Roak BJ, Sanders SJ, Ronemus M, Krumm N, Levy D, Stessman HA, Witherspoon KT, Vives L, Patterson KE, et al. 2014. The contribution of de novo coding mutations in autism spectrum disorder. *Nature* **515**: 216.
- Narzisi G, O’Rawe JA, Iossifov I, Fang H, Lee YH, Wang Z, Wu Y, Lyon GJ, Wigler M, Schatz MC. 2014. Accurate de novo and transmitted indel detection in exome-capture data using microassembly. *Nat Methods* **11**: 1033.
- Yilancioglu K, Weinstein ZB, Meydan C, Akhmetov A, Toprak I, Durmaz A, Iossifov I, Kazan H, Roth FP, Cokol M. 2014. Target-independent prediction of drug synergies using only drug lipophilicity. *J Chem Inf Model* **54**: 2286.

# IN SILICO CANCER GENOMICS

A. Krasnitz M. McGillivray G. Sun

Research in our group is focused on in silico cancer genomics. In the last 5 years, there has been explosive growth in the volume as well as quality and detail of cancer-related genomic data available for analysis. This development came about through rapid accumulation of public data sets generated by individual laboratories worldwide, through large-scale cancer genome projects such as TCGA (the Cancer Genome Atlas) and ICGC (International Cancer Genome Consortium), and with the advent of new experimental methodologies, especially next-generation sequencing and single-cell genomics. We see our goal as channeling this flood of data into a number of clinically relevant applications, including pinpointing and prioritizing targets for functional analysis; discovering genomic markers for clinical outcome and molecular classification of cancer; and elucidating the clonal structure of tumors and its relation to progression, invasion, metastasis, and response to treatment. All our work is done in close coordination with experimental studies performed by the Wigler, Hicks, Powers, Lowe (presently at Memorial Sloan-Kettering Cancer Center), and Stillman laboratories here at CSHL.

## Recurrence Analysis of Genomic Interval Data

---

Large collections of intervals commonly arise in high-throughput genomics. For example, DNA copy-number analysis yields intervals of the genome corresponding to gains or losses of DNA segments. Likewise, chromatin structure is often reported as intervals of the genome. In such cases, a common goal is inference of contiguous genomic target regions, which under certain model assumptions generates the observed patterns in the data. We call such target regions “cores.” The typical evidence for cores is the presence of “recurrent” observations, suitably defined.

In cancer, genomes display complex patterns of DNA copy-number alteration, but recurrent aberrations are observed within a given cancer type. Detection and quantitative characterization of the cores in a given malignancy are potentially beneficial in two

ways. First, it is plausible that at least some of the recurrence owes to selective pressure on regions harboring cancer-related genes. Focusing on these regions may therefore facilitate cancer gene discovery. Second, once the cores are known, the genome of a tumor can be described in a drastically simplified fashion by indicating the presence or absence of copy-number alteration in any given core. This simplified form of the data is better suited for further analysis, such as finding subtypes in a given type of cancer or discovering associations between genomic properties and clinical parameters. Finally, differences in core patterns among individual tumor cells may be used to explore genomic heterogeneity of cancer and examine genealogical relationships among tumor cell populations.

We designed and implemented a method for identifying cores in large collections of genomic interval data. A descriptive name for the method is cores of recurrent events (CORE). Central to CORE is the notion of explanatory power. We say that a core explains an event and quantify the explanation provided as a number between 0 and 1. The explanation is a measure of how closely the event is matched by the core. We then seek a set of cores that jointly provide best possible explanation of the data, subject to additional criteria of statistical significance. CORE is now available to the community via the Comprehensive R Archive Network (CRAN).

## Tumor Cell Population Structure

---

Study of genomic and transcriptional properties of individual cells is now the focus of multiple research laboratories around the world. In application to cancer, this line of research has revealed the genomic complexity of the disease and the presence of multiple genealogically related cell populations in a tumor. Detailed knowledge of the clonal structure of a cancer potentially is of high clinical value: Multiplicity of clones or of lesions in most advanced clones is a possible measure of progression; spatial pattern of clone dispersal in a tumor may signal elevated propensity to

invade; lesions observed in individual clones but not in the bulk tissue may point to targets for therapy. DNA copy-number profiling of cells from low-coverage sequencing is an accurate, economically feasible technological approach to the study of cancer subpopulation structure. Novel multiplex sequencing techniques, developed by the Wigler lab at CSHL, permit simultaneous sequencing of dozens of single-cell DNA specimens and their subsequent copy-number profiling at up to 50-kb resolution. Optimal use of these data for robust reconstruction of cancer cell phylogenies is a challenging computational problem calling for new and robust informatic and statistical tools.

In response, we used CORE to transform copy-number profiles of single cells into a form suitable for phylogeny. As a result of this transformation, each profile is characterized in terms of presence or absence in it of archetypal copy-number events recurrent in the single-cell populations. The entire set of copy-number profiles of single cells sampled from a tumor is summarized as an incidence table, essentially a matrix with cores as columns, single cells as rows and elements quantifying, on a scale between 0 and 1, how well a given core is matched by a copy-number event in a given cell. Distances among cells as rows of this matrix can be readily computed and used for distance-based phylogeny. This methodology has been used to reconstruct the cell population structure in a number of samples from prostate cancers. Importantly, this study shows that well-defined clones of tumor cells are likely to exist in prostates of patients with high, but not with low, Gleason scores. Furthermore, we see evidence that some of the clones spread to multiple anatomical sites within the affected prostate, including sites with less malignant histopathological characteristics.

### **Prediction of Response to Anthracycline Therapy in Breast Cancer**

---

Anthracyclines are a class of chemotherapeutic drugs that are effective and widely used in treatment of a number of cancers, in particular those of breast. Sensitivity to these agents varies widely across the entire population of patients and likely depends on the pattern of genomic alterations in the tumor. Anthracycline administration also causes significant side effects, most importantly cardiotoxicity, and should

therefore be restricted to patient populations for which the expected benefits outweigh these risks. Yet despite more than 50 years of use, the anthracycline mechanism of action is not firmly established, and there are as yet no reliable predictive markers. In response, we applied CORE to DNA copy-number profiles derived by the Hicks laboratory at CSHL from breast tumors of approximately 250 patients enrolled in the NEAT/BR9601 clinical trial in Britain. The study was conducted in collaboration with the Bartlett group at the Ontario Institute for Cancer Research. To discover markers of sensitivity to anthracyclines, we performed CORE analysis of the set of profiles and computed the incidence table. Next, we evaluated each core as a prospective marker of sensitivity and found dramatically higher benefit from treatment in patients positive for a narrow amplified region of chromosome 8. None of the marker-positive patients treated by anthracyclines have suffered a relapse in 5 years past treatment, whereas the marker-positive patients who were not treated displayed a markedly higher rate of relapse than the remainder of the cohort. Given these results, a validation study has been initiated, with the goal to similarly analyze hundreds of breast tumor samples from the MA.5 clinical trial in Canada.

### **Significantly Distinct Branches of Hierarchical Trees**

---

Hierarchical clustering (HC) is widely used as a method of partitioning data and of identifying meaningful data subsets. Most commonly an application consists of visual examination of the dendrogram and intuitive identification of subtrees that appear clearly distinct from the rest of the tree. Obviously, results of such qualitative analysis and conclusions from it may be observer-dependent. Quantifying the interpretation of hierarchical trees and introducing mathematically and statistically well-defined criteria for distinctness of subtrees would therefore be highly beneficial.

We formulated and implemented a computational method termed tree branches evaluated statistically for tightness (TBEST) for identifying significantly distinct tree branches in hierarchical clusters. For each branch of the tree, a measure of distinctness, or tightness, is defined as a rational function of heights, both of a given branch and of its parent. A statistical

procedure is then developed to determine the significance of the observed values of tightness. We tested TBEST as a tool for tree-based data partitioning by applying it to four benchmark data sets. These represent four distinct types of biological data: mRNA expression, DNA copy-number variation, proteomics, and flow cytometry. In each of the four cases there is a well-defined and independently known partition of the data into classes. In all cases considered, TBEST reproduced these classes on par with or better than

the existing techniques. The R language implementation of the method is now available, as an eponymous package, from the Comprehensive R Active Network (CRAN).

#### PUBLICATIONS

Sun G, Krasnitz A. 2014. Significantly distinct branches of hierarchical trees: A framework for statistical analysis and applications to biological data. *bioRxiv* doi: 10.1101/002188.

# COMPUTATIONAL GENETICS

D. Levy V. Kumar

The data-rich environment at CSHL generates a wealth of opportunities for the application of mathematics and computation to further our understanding of biology and genetics. The primary activities of our lab are algorithm and protocol development, data analysis, and genetic theory, with a focus on sporadic human diseases such as autism, congenital heart disease, and cancer.

## Autism Genetics, Mutation, and Computational Methods

---

Together with Michael Wigler, Michael Ronemus, and Ivan Iossifov here at CSHL, our efforts in autism attempt to bridge the gap between genetic theory and large-scale data analysis. Together with groups at the University of Washington and the University of San Francisco, we have completed the whole-exome analysis on the entire Simons Simplex Collection (SSC), composed of 3000 families with exactly one child affected with autism. Analyzing more than 10,000 exomes, we established a causal role for deleterious coding mutations in sensitive genes. The study uncovered many candidate genes whose disruption is likely to result in an autism spectrum disorder, establishing a high-quality set of starting points for uncovering the functional relationships that underlie the disorder.

Although the study addressed the types of mutations best observed by exome sequencing, namely, single-nucleotide mutations and small insertion and deletion events, we are building tools to extract other mutational modalities such as copy-number variations (CNVs) and genomic rearrangements.

## Copy-Number Variation

---

The primary challenge in identifying CNVs from exome capture data is that the capture process distorts coverage in a manner that varies from sample to sample. Fortunately, the variability expressed in an individual sample is often recapitulated on a population level, and we can use our library of samples to identify the major

elements of systemic noise. Our normalization procedure is coupled to a hidden Markov model (HMM) that integrates over all the copy-number possibilities for a given sample and returns a probability for each copy state for each region of the genome. Joining our methods with a complementary approach developed by Chris Yoon, we have established a set of high-quality CNV calls from the noisy SSC exome data. Preliminary results support earlier microarray studies, establishing a similar rate for de novo CNVs and suggesting a role for rare transmitted variants that disrupt genes.

## Genomic Rearrangements

---

Together with Peter Andrews here at CSHL, we have developed an algorithm for identifying large genomic rearrangements—deletions, insertions, translocations, and inversions—by cataloging discontinuities in the mapping of reads against the reference genome. Taking advantage of memory-intensive computational methods, we are able to rapidly identify all maximal exact matches (MEMs) between a read and a reference genome. This provides a scaffold for cataloging and indexing all of the sequence data such that we can quickly identify discontinuities: areas where placing the reads on the scaffold bends or breaks the reference genome. By looking across all samples, we can identify common events, spurious rearrangements, and global discrepancies within the reference. With all reads indexed and anchored to the reference genome, we reassemble rare and de novo discontinuities within an individual or family. Analyzing large data sets requires an efficient and robust data format and processing algorithms, which we are currently deploying on both the 10,000 SSC exomes and a smaller whole-genome sequence data set.

## Genetic Models of Autism

---

Autism is often described as the most highly heritable neurological disorder. Concordance for autism among

identical twins is as high as 90%. There are several other interesting statistics in the epidemiology of autism that are less often noted. If a family has a single child with autism, the risk to the next born male child is about 20%. This elevated risk has led to an explanation that autism is a multifactorial disorder and that several gene disruptions are needed to produce an autism phenotype. That claim, however, is belied by another statistic: given two children in the family with autism, the risk to a third-born male child is 50%.

One model that fits this data is to postulate two risk classes for autism—nearly all families have a very low risk of having a child with autism, whereas a small percentage of families have a very high risk. Together with Michael Wigler and Swagatam Mukhopadhyay here at CSHL, we have been exploring simple genetic models that recapitulate the observed risk statistics, both for boys and girls, as well as observations that stem from genetic analyses of autism spectrum disorder (ASD) families. Our model includes parameters such as de novo mutation rates, variability of penetrance, assortative mating, and the number and type of genetic targets.

Combinatorics allow for fast simulations of infinite populations, providing a means for quickly exploring the space of possible models consistent with the observable measures. We find that a simple model with two gene classes and different selection profiles for males and females is sufficient to match observable risk and mutation rates. This model also makes predictions about gender ratios as a function of severity, the role of transmitted mutation, and the genetic load in the population.

## Mutational Sequencing

---

Despite vast improvements in DNA sequencing, many problems of interpretation arise when trying to count or assemble molecules (templates) that are largely identical. In a recent paper, we show that by randomly mutating DNA templates before amplification, many of these problems are resolved. We can accurately count the number of templates by counting the number of unique patterns. By introducing distinctive patterns onto otherwise identical spans, we enhance our ability to correctly assemble sequences. This idea can be implemented with currently available mutagenesis protocols, and such techniques can have applications in RNA expression analysis, haplotype phasing, copy-number determination, and genome assembly. Mutational sequencing (muSeq) solves counting problems and effectively generates long reads from short-read sequence data.

We are developing new protocols to generate random, uncorrelated mutations at a tunable rate. We are also developing algorithms for counting and assembling mutational patterns in the noisy context of real sequence data. Having established working protocols and algorithms, we are now exploring applications ranging from the sensitive detection of rare mutation to single-cell isoform profiling to de novo genome assembly.

## PUBLICATIONS

Levy D, Wigler M. 2014. Facilitated sequence counting and assembly by template mutagenesis. *Proc Natl Acad Sci* 111: E4632–E4637.

## COMPUTATIONAL SEQUENCE ANALYSIS

M. Schatz   H. Fang   S. Marcus   F. Sedlazeck  
T. Garvin   G. Narzisi   R. Sherman  
J. Gurtowski   M. Nattestad   A.H. Wences  
H. Lee   S. Ramakrishnan

Our lab develops novel computational analysis methods to study the structure and function of genomes, especially genomes with medical or agricultural significance. These include methods for assembling sequence data into complete genomes, for aligning sequences to discover variations or measure transcription levels, and for mining these data to discover relationships between an organism's genome and its traits. We then apply these methods to study several biological systems, for example, to discover mutations associated with human diseases such as cancer and autism or to reconstruct the genomes of important agricultural crops and biofuels.

In the past year, our lab has grown by several new members. Han Fang is a new Ph.D. student in Applied Math at Stony Brook University who is jointly supervised by myself and Dr. Gholson Lyon. Han is particularly interested in transcriptional and translational regulation, using RNA-Seq and Ribo-Seq profiling in yeast to study how genetic mutations can influence protein production. Maria Nattestad is a new student from the Watson School of Biological Sciences. Maria completed her undergraduate education in Biology at the University of the Pacific and is currently researching the role of structural variations in cancer using single-molecule sequencing from Pacific Biosciences (PacBio). Rachel Sherman joined us from Harvey Mudd College for the summer undergraduate research program and developed a new algorithm for multispecies comparative genomics. At the very end of the year, Fritz Sedlazeck joined us as a new postdoctoral researcher from Vienna, Austria, where he completed his Ph.D. in genomics and computer science. He will be researching novel methods for DNA sequence assembly, using new long-read sequencing technology from PacBio, Oxford Nanopore, and the new 10X Genomics platform. We were also proud of Giuseppe Narzisi for his new position as Bioinformatics Scientist at the New York Genome Center and of Shoshana Marcus for her new role as Assistant

Professor of Computer Science at the City University of New York at Kingsborough.

In the past year, our lab has made significant contributions to the study of the genetics of autism in collaboration with the Wigler, Iossifov, Levy, and Lyon labs here at CSHL. Most significantly, we published our state-of-the-art algorithm Scalpel for finding insertion and deletion mutations from short sequencing reads and applied it to identify several new autism candidate genes. We also contributed to several genome-sequencing projects in collaboration with the McCombie and Ware labs, including publishing the rice, wheat, and the parrot genomes. We also developed new algorithmic advances toward the assembly of genomes and the assembly of "pan-genomes." Finally, I am proud to say that I was awarded an NSF CAREER award to develop new algorithms for studying genomes using new single-molecule sequencing technologies. This prestigious award recognizes the significant research advances I have made since joining the CSHL faculty in 2010 as well as my future research plans.

### The Contribution of De Novo Coding Mutations to Autism Spectrum Disorder

---

Whole-exome sequencing has proven to be a powerful tool for understanding the genetic architecture of human disease. Here, we apply it to more than 2500 simplex families, each having a child with an autistic spectrum disorder. By comparing affected to unaffected siblings, we show that 13% of de novo missense mutations and 43% of de novo likely gene-disrupting (LGD) mutations contribute to 12% and 9% of diagnoses, respectively. Including copy-number variants, coding de novo mutations contribute to ~30% of all simplex and 45% of female diagnoses. Almost all LGD mutations occur opposite wild-type alleles. LGD targets in affected females significantly overlap

the targets in males of lower intelligence quotient (IQ), but neither overlaps significantly with targets in males of higher IQ. We estimate that LGD mutation in ~400 genes can contribute to the joint class of affected females and males of lower IQ, with an overlapping and similar number of genes vulnerable to contributory missense mutation. LGD targets in the joint class overlap with published targets for intellectual disability and schizophrenia and are enriched for chromatin modifiers, FMRP-associated genes, and embryonically expressed genes. Most of the significance for the latter comes from affected females.

### Reducing INDEL Calling Errors in Whole-Genome and Exome Sequencing Data

---

Genome insertions and deletions (INDELs), especially those disrupting protein-coding regions of the genome, have been strongly associated with human diseases. However, there are still many errors with INDEL variant calling, driven by library preparation, sequencing biases, and algorithm artifacts. We characterized whole-genome sequencing (WGS), whole-exome sequencing (WES), and polymerase chain reaction (PCR)-free sequencing data from the same samples to investigate the sources of INDEL errors. We also developed a classification scheme based on the coverage and composition to rank high- and low-quality INDEL calls. We performed a large-scale validation experiment on 600 loci and found that high-quality INDELs have a substantially lower error rate than low-quality INDELs (7% vs. 51%). Simulation and experimental data show that assembly-based callers are significantly more sensitive and robust for detecting large INDELs (>5 bp) than alignment-based callers, consistent with published data. The concordance of INDEL detection between WGS and WES is low (53%), and WGS data uniquely identifies 10.8-fold more high-quality INDELs. The validation rate for WGS-specific INDELs is also much higher than that for WES-specific INDELs (84% vs. 57%), and WES misses many large INDELs. In addition, the concordance for INDEL detection between standard WGS and PCR-free sequencing is 71%, and standard WGS data uniquely identify 6.3-fold more low-quality INDELs. Furthermore, accurate detection with Scalpel of heterozygous INDELs requires

1.2-fold higher coverage than that for homozygous INDELs. Finally, homopolymer A/T INDELs are a major source of low-quality INDEL calls, and they are highly enriched in the WES data. Overall, we show that accuracy of INDEL detection with WGS is much greater than that with WES even in the targeted region. We calculated that 60× WGS depth of coverage from the Hi-Seq platform is needed to recover 95% of INDELs detected by Scalpel. Although this is higher than the current sequencing practice, the deeper coverage may save total project costs because of the greater accuracy and sensitivity. We also investigate sources of INDEL errors (e.g., capture deficiency, PCR amplification, and homopolymers) with various data that will serve as a guideline to effectively reduce INDEL errors in genome sequencing.

### Accurate De Novo and Transmitted INDEL Detection in Exome-Capture Data Using Microassembly

---

We present here an open-source algorithm, Scalpel, that combines mapping and assembly for sensitive and specific discovery of insertions and deletions (INDELs) in exome-capture data. A detailed repeat analysis coupled with a self-tuning k-mer strategy allows Scalpel to outperform other state-of-the-art approaches for INDEL discovery, particularly in regions containing near-perfect repeats. We analyzed 593 families from the Simons Simplex Collection and demonstrated Scalpel's power to detect long (≥30 bp) transmitted events and enrichment for de novo likely gene-disrupting INDELs in autistic children.

### Whole-Genome De Novo Assemblies of Three Divergent Strains of Rice, *Oryza sativa*, Document Novel Gene Space of *aus* and *indica*

---

The use of high-throughput genome-sequencing technologies has uncovered a large extent of structural variation in eukaryotic genomes that makes important contributions to genomic diversity and phenotypic variation. When the genomes of different strains of a given organism are compared, whole-genome resequencing data are typically aligned to an established reference sequence. However, when the



reference differs in significant structural ways from the individuals under study, the analysis is often incomplete or inaccurate. Here, we use rice as a model to demonstrate how improvements in sequencing and assembly technology allow rapid and inexpensive de novo assembly of next-generation sequence data into high-quality assemblies that can be directly compared using whole-genome alignment to provide an unbiased assessment. Using this approach, we are able to accurately assess the “pan-genome” of three divergent rice varieties and document several megabases of each genome absent in the other two. Many of the genome-specific loci are annotated to contain genes, reflecting the potential for new biological properties that would be missed by standard reference-mapping approaches. We further provide a detailed analysis of several loci associated with agriculturally important traits, including the S5 hybrid sterility locus, the Sub1 submergence tolerance locus, the LRK gene cluster associated with improved yield, and the Pup1 cluster associated with phosphorus deficiency, illustrating the utility of our approach for biological discovery. All of the data and software are openly available to support further breeding and functional studies of rice and other species.

### High-Coverage Sequencing and Annotated Assemblies of the Budgerigar Genome

---

Parrots belong to a group of behaviorally advanced vertebrates and have an advanced ability of vocal learning relative to other vocal-learning birds. They can imitate human speech, synchronize their body movements to a rhythmic beat, and understand complex concepts of referential meaning to sounds. However, little is known about the genetics of these traits. Elucidating the genetic bases would require whole-genome sequencing and a robust assembly of a parrot genome. We present a genomic resource for the budgerigar, an Australian Parakeet (*Melopsittacus undulatus*), the most widely studied parrot species in neuroscience and behavior. We present genomic sequence data that includes more than 300× raw read coverage from multiple sequencing technologies and chromosome optical maps from a single male animal. The reads and optical maps were used to create three hybrid assemblies representing some of the largest genomic scaffolds to date for a bird, two of which

were annotated based on similarities to reference sets of nonredundant human, zebra finch, and chicken proteins and budgerigar transcriptome sequence assemblies. The sequence reads for this project were in part generated and used for both the Assemblathon 2 competition and the first de novo assembly of a gigascale vertebrate genome utilizing PacBio single-molecule sequencing. Across several quality metrics, these budgerigar assemblies are comparable to or better than the chicken and zebra finch genome assemblies built from traditional Sanger sequencing reads, and they are sufficient to analyze regions that are difficult to sequence and assemble, including those not yet assembled in prior bird genomes and promoter regions of genes differentially regulated in vocal-learning brain regions. This work provides valuable data and material for genome technology development and for investigating the genomics of complex behavioral traits.

### Large-Scale Sequencing and Assembly of Cereal Genomes Using Blacklight

---

Wheat, corn, and rice provide 60% of the world’s food intake every day, and just 15 plant species make up 90% of the world’s food intake. As such, there is tremendous agricultural and scientific interest in sequencing and studying plant genomes, especially to develop a reference sequence to direct plant breeding or to identify functional elements. DNA sequencing technologies can now generate sequence data for large genomes at low cost; however, it remains a substantial computational challenge to assemble the short sequencing reads into their complete genome sequences. Even one of the simpler ancestral species of wheat, *Aegilops tauschii*, has a genome size of 4.36 gigabase-pairs (Gbp), nearly 50% larger than the human genome. Assembling a genome this size requires computational resources, especially RAM to store the large assembly graph, which is out of reach for most institutions. Here, we describe a collaborative effort between Cold Spring Harbor Laboratory and the Pittsburgh Supercomputing Center to assemble large, complex cereal genomes starting with *Ae. tauschii*, using the XSEDE shared memory supercomputer Blacklight. We expect these experiences using Blacklight to provide a case study and computational protocol for other genomics communities to leverage this

or similar resources for assembly of other significant genomes of interest.

## On Algorithmic Complexity of Biomolecular Sequence Assembly Problem

---

Because of its connection to the well-known NP-complete shortest superstring combinatorial optimization problem, the sequence assembly problem (SAP) has been formulated in simple and sometimes unrealistic string-and-graph theoretic frameworks. We revisit this problem by reexamining the relationship between the most common formulations of the SAP and their computational tractability under different theoretical frameworks. For each formulation, we show examples of logically consistent candidate solutions that are nevertheless unfeasible in the context of the underlying biological problem. We hope that this material will be valuable to theoreticians as they develop new formulations of SAP as well as be of guidance to developers of new pipelines and algorithms for sequence assembly and variant detection.

## SplitMEM: A Graphical Algorithm for Pan-Genome Analysis with Suffix Skips

---

Genomics is expanding from a single reference per species paradigm into a more comprehensive pan-genome approach that analyzes multiple individuals together. A compressed de Bruijn graph is a sophisticated data structure for representing the genomes of entire populations. It robustly encodes shared segments, simple single-nucleotide polymorphisms, and complex structural variations far beyond what can be represented in a collection of linear sequences alone. We explore deep topological relationships between suffix trees

and compressed de Bruijn graphs and introduce an algorithm, SplitMEM, that directly constructs the compressed de Bruijn graph in time and space linear to the total number of genomes for a given maximum genome size. We introduce suffix skips to traverse several suffix links simultaneously and use them to efficiently decompose maximal exact matches into graph nodes. We demonstrate the utility of SplitMEM by analyzing the nine-strain pan-genome of *Bacillus anthracis* and up to 62 strains of *Escherichia coli*, revealing their core-genome properties.

## PUBLICATIONS

- Blood P, Marcus S, Schatz MC. 2014. Large-scale sequencing and assembly of cereal genomes using Blacklight. *Proc XSEDE14*. Atlanta, Georgia.
- Fang H, Narzisi G, O'Rawe J, Wu Y, Rosenbaum J, Ronemus M, Iossifov I, Schatz MC, Lyon GJ. 2014. Reducing INDEL calling errors in whole-genome and exome sequencing data. *Genome Med* **6**: 89.
- Ganapathy G, Howard JT, Ward JM, Li J, Li B, Li Y, Xiong Y, Zhang Y, Zhou S, Schwartz DC, Schatz MC, et al. 2014. High-coverage sequencing and annotated assemblies of the budgerigar genome. *GigaScience* **3**: 11.
- Iossifov I, Ronemus M, Levy D, Yamron B, Lee YH, Grabowska E, Narzisi G, et al. 2014. The contribution of de novo coding mutations to autism spectrum disorder. *Nature* **515**: 216–221.
- Marcus S, Lee HS, Schatz MC. 2014. SplitMEM: A graphical algorithm for pan-genome analysis with suffix skips. *Bioinformatics* **30**: 3476–3483.
- Narzisi G, Mishra B, Schatz MC. 2014. On algorithmic complexity of biomolecular sequence assembly problem. Algorithms for computational biology. Lecture Notes In: Computer science, vol. 8542, pp. 183–195.
- Narzisi G, O'Rawe JA, Iossifov I, Fang H, Lee YH, Wang Z, Wu Y, Lyon G, Wigler M, Schatz MC. 2014. Accurate de novo and transmitted indel detection in exome-capture data using micro-assembly. *Nat Methods* **11**: 1033–1036.
- Schatz MC, Maron LG, Stein JC, Hernandez Wences AS, Gurtowski JS, Biggers ES, Lee HS, Kramer M, Antoniou E, Ghiban E, et al. 2014. Whole genome de novo assemblies of three divergent strains of rice (*O. sativa*) documents novel gene space of *aus* and *indica*. *Genome Biol* **15**: 506.

# QUANTITATIVE BIOLOGY FELLOW

---

Quantitative Biology Fellows are independent researchers who enter this position soon after receiving a Ph.D. They come from the fields of mathematics, physics, engineering, or computer science and spend 3–5 years at Cold Spring Harbor Laboratory applying techniques from these disciplines to important questions in biology.

## Transcriptional Regulation, Biophysics, and Machine Learning

J. Kinney

---

My research combines theory, computation, and experiment in an effort to better understand sequence-function relationships in molecular biology.

My experimental work focuses primarily on the biophysics of transcriptional regulation in the bacterium *Escherichia coli*. In graduate school and as a postdoc, I developed an experimental method called “Sort-Seq” that allows the transcriptional activities of hundreds of thousands of different *E. coli* promoters to be measured in a single experiment. Sort-Seq data can be used to build quantitative biophysical models for how elementary protein–DNA and protein–protein interactions combine to regulate gene expression. I continue to use Sort-Seq to study transcriptional regulation in *E. coli*, but I am also using similar techniques to study antibody–antigen interactions.

The analysis of Sort-Seq data also highlights interesting problems in statistics and machine learning that have yet to be fully addressed. About half of my research effort is devoted to these theoretical and computational questions.

## Parametric Inference in the Large Data Limit Using Maximally Informative Models

---

Kinney and Atwal (2014a) describe a basic problem in statistical inference. Our work was motivated by my analysis of Sort-Seq data, but it is also relevant to the measurement of receptive fields in sensory neuroscience.

Consider an experiment that produces an unlimited supply of data (e.g., activity measurements for many different variants of a bacterial promoter). Assume, however, that we do not know the precise quantitative form of the experimental noise that contaminates

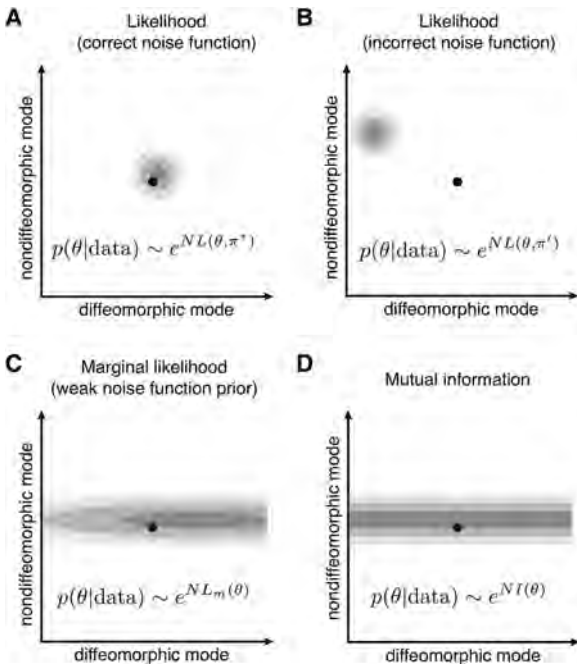
these measurements. Such situations are common in Sort-Seq experiments and in other high-throughput biology assays. Nevertheless, not knowing how to model experimental noise causes problems for the statistical methods typically used to fit models to such data. The reason is that standard statistics requires that one compute a quantity called “likelihood,” and computing this requires an explicit quantitative model for experimental noise.

Atwal and I show that using a quantity from information theory called “mutual information” in place of likelihood circumvents the problem of not knowing the quantitative details of the experimental noise (Fig. 1). We also introduce the concept of “diffeomorphic modes”—directions in parameter space that mutual information cannot pin down. Importantly, this work shows that different model parameters can respond in qualitatively different ways to data. This is a general mathematical finding applicable to a wide range of statistical regression problems, including many “Big Data” problems in biology.

## Equitability, Mutual Information, and the Maximal Information Coefficient

---

In Kinney and Atwal (2014b), we introduce a rigorous formulation of a statistical concept called “equitability.” Equitability is a heuristic criterion, introduced and advocated by Reshef et al. (*Science* 334: 1518 [2011]), that general measures of statistical dependence should satisfy. We show that this criterion is naturally formalized as a weakened form of the “Data Processing Inequality,” a core concept in information theory. We also emphasize that this formal equitability criterion is satisfied by the well-known mutual information measure and that estimates of mutual information made on finite data can adhere well to this notion of equitability in practice.



**Figure 1.** Schematic illustration of constraints placed on diffeomorphic and nondiffeomorphic modes in parameter space by different objective functions. The filled circle in each panel represents the correct model  $\theta^*$ ; shades of gray represent the posterior distribution  $p(\theta|\text{data})$ . (A,B) Likelihood places tight constraints along both diffeomorphic and nondiffeomorphic directions in parameter space. (A)  $\theta^*$  will typically lie within error bars if the correct noise function  $\pi^*$  is used to compute likelihood  $L$ . (B) However, if an incorrect noise function  $\pi'$  is used,  $\theta^*$  will generally violate inferred constraints along both diffeomorphic and nondiffeomorphic modes. (C) Alternatively, if uncertainties in the noise functions are explicitly accounted for in the computation of  $p(\theta|\text{data})$ , one will find tight constraints on nondiffeomorphic modes but only weak constraints along diffeomorphic modes. (D) Mutual information places tight constraints on nondiffeomorphic modes but provides no constraints whatsoever on diffeomorphic modes. (Figure from Kinney and Atwal 2014a.)

Although the heuristic concept of equitability is valuable, we also find that the primary claims of Reshef et al. regarding equitability were incorrect. In particular, we show that the “maximal information coefficient” (MIC), a statistic that Reshef et al. introduced and that was said to satisfy equitability, does not actually have this mathematical property. This finding is important because the paper by Reshef et al. has become one of the most highly cited statistics papers of the last few years, and MIC appears to have gained widespread use. We hope that our paper will motivate scientists in need of an equitable measure of

dependence to seriously consider using mutual information instead of MIC.

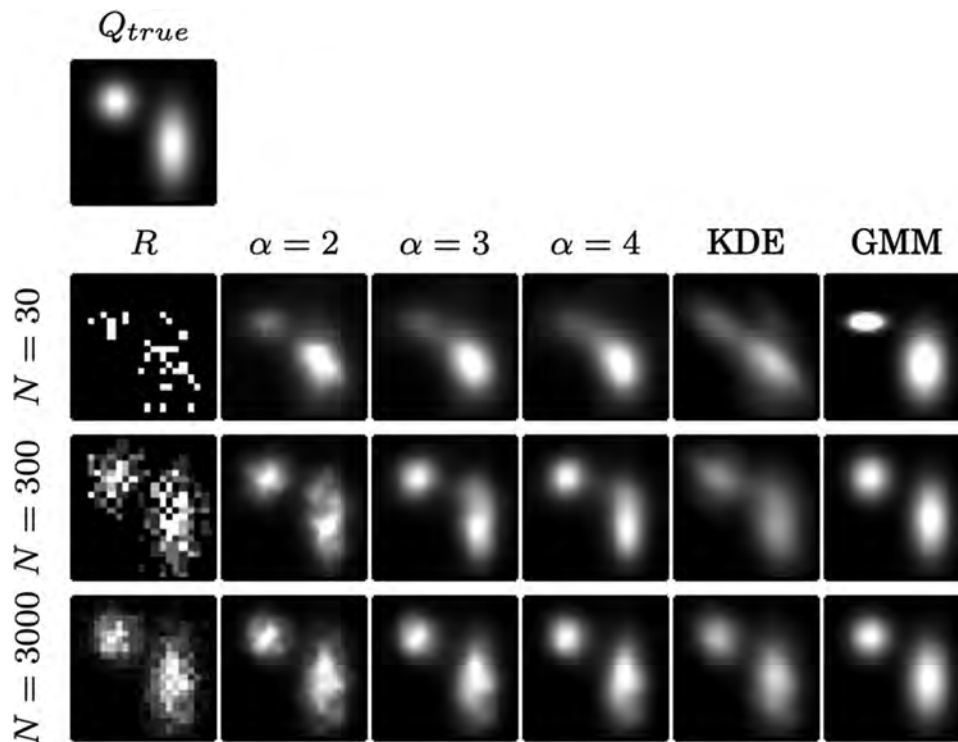
## Field Theoretic Methods for Estimating Low-Dimensional Probability Densities

As with many efforts to understand experimental data, analyzing Sort-Seq data requires the accurate estimation of continuous probability distributions (also known as probability densities). Despite the ubiquity and seeming simplicity of this problem, the best way to estimate a continuous probability distribution from data—even in one dimension—remains unresolved. I believe, however, that field theory methods from theoretical physics can ultimately provide a definitive practical solution to this important statistical problem.

Field theory methods for estimating probability densities date back to the mid 1990s. These methods are computationally intensive, however, and until Kinney (2014), no practical implementation of these methods was available. My work (Kinney 2014) shows that field-theory-based density estimates can be computed rapidly and deterministically using a homotopy algorithm. This paper also demonstrates, for the first time, that such density estimation methods can perform excellently on simulated data.

In a more recent work, I show that the boundary conditions that were placed on estimated densities in previous field-theoretic work are unnecessary. By removing these boundary conditions, one recovers the well-known “maximum entropy” approach to density estimation in the infinite smoothness limit. This unification of field theory and maximum entropy methods for density estimation further suggests a novel way of testing the hypothesis that one’s data comes from a probability distribution having a specified functional form.

By establishing the ability to rapidly compute field theory density estimates and by showing how to remove prior assumptions about the boundary behavior of these densities, my work provides a solid foundation upon which to build a definitive practical solution to the density estimation problem in low dimensions. In the near future, I hope to show how such density estimation methods can be used to accurately estimate mutual information and other important statistical quantities. I also hope to provide



**Figure 2.** Field-theoretic estimation of two-dimensional probability densities. Shown is a simulated probability density  $Q_{\text{true}}$  (composed of two Gaussian distributions), raw data sets  $R$  (of various size  $N$ ) drawn from  $Q_{\text{true}}$  and density estimates computed from  $R$  using either kernel density estimation, Gaussian mixture modeling or the field theory method of Kinney (2014) with various values for the parameter  $\alpha$ . (Figure from Kinney 2014.)

easy-to-use software that makes these statistical methods available for general use.

## Ongoing Projects

*A biophysical code for transcriptional regulation in E. coli.* I have established a long-term collaboration with Rob Phillips (California Institute of Technology), the goal of which is to build an accurate biophysical model of the transcriptional regulatory code of *E. coli*. We are focusing on this bacterium for two reasons. First, the underlying molecular biology of transcriptional regulation in *E. coli* is understood well at a qualitative level. This qualitative understanding is essential for the correct quantitative interpretation of Sort-Seq data. Second, transcription in *E. coli* is regulated in a hierarchical manner, with only a small number of proteins controlling a large number of regulatory sequences. Using in vitro single-molecule techniques together with in vivo Sort-Seq assays, we aim to systematically build a full sequence-dependent

biophysical model for how regulatory proteins bind DNA, interact with one another, and modulate the rate of mRNA transcription.

Initial work on this project has been pursued by Daniel Jones, a graduate student who did his Ph.D. work with Rob Phillips followed by a short stint as a postdoc in my lab. We are currently wrapping up this phase of the project.

*Sequence-function relationships in antibody–antigen recognition.* Despite the central role that antibodies have in the adaptive immune system and in biotechnology, surprisingly little is known about how antibody–antigen affinity depends on antibody sequence. For instance, how many possible antibodies can specifically bind an antigen of interest with 1-nM affinity, or with 1-pM affinity? Does the sequence-affinity landscape have a single optimum, or is it “glassy”? Do the sequence-affinity landscapes for different antibody–antigen pairs exhibit any general features, for example, a universal “density of states”? In collaboration with Aleksandra Walczak (ENS, Paris) and Thierry Mora (ENS, Paris), I am pursuing a combination

of theory, computation, and experiment aimed at making progress on these important problems.

During the last 2 years, Rhys Adams, a postdoc working with Walczak, Mora, and myself, has been performing and analyzing Sort-Seq experiments designed to address these questions. Using yeast display technology, he is able to measure the binding affinities of thousands of variant antibodies to a specific antigen of interest. Our hope is that these data will allow us to build quantitative models of the sequence-affinity landscape of this system, explore this landscape computationally, and look for features that might generalize to other antibody–antigen pairs.

*Formal diagrammatic methods for biochemical systems.* My work on transcriptional regulation requires building quantitative biophysical models of molecular interactions. However, the only systematic way to currently represent models of such interactions is to explicitly list all the possible states of a system, a task that becomes exponentially more difficult as system size increases. For this reason, I am developing a formal diagrammatic method for concisely and rigorously defining thermodynamic models of biochemical systems. This method allows equilibrium partition functions to be calculated systematically and should

eventually facilitate nonequilibrium calculations as well. This approach can be applied to a wide range of processes in systems biology. My hope is that such methods will serve a purpose much like Feynman diagrams do in physics, bridging the gap between one's intuitive visual understanding of a system and rigorous quantitative models thereof.

## PUBLICATIONS

- Boedicker J, Razo-Mejia M, Jones D, DeLuna A, Kinney JB, Phillips R. 2014. Comparison of the theoretical and real-world evolutionary potential of a genetic circuit. *Phys Biol* **11**: 026005.
- Kinney JB. 2014. Estimation of probability densities using scale-free field theories. *Phys Rev E* **90**: 011301(R).
- Kinney JB, Atwal GS. 2014a. Parametric inference in the large data limit using maximally informative models. *Neural Comput* **26**: 637–653.
- Kinney JB, Atwal GS. 2014b. Equitability, mutual information, and the maximal information coefficient. *Proc Natl Acad Sci* **111**: 3354–3359.
- Kinney JB, Atwal GS. 2014c. Reply to Reshef et al.: Falsifiability or bust. *Proc Natl Acad Sci* **111**: E3364.
- Kinney JB, Atwal GS. 2014d. Reply to Murrell et al.: Noise matters. *Proc Natl Acad Sci* **111**: E2161.
- Sheu Y-J, Kinney JB, Lengronne A, Pasero P, Stillman B. 2014. Domain within the helicase subunit Mcm4 integrates multiple kinase signals to control DNA replication initiation and fork progression. *Proc Natl Acad Sci* **111**: E1899–E1908.

## COLD SPRING HARBOR LABORATORY FELLOW

---

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 3 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 and the program of courses and meetings on diverse topics in biology contribute to a research environment that is ideal for innovative science 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. Most notably, 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 fellowship, Carol was on the CSHL faculty for 9 years, and she is currently the Daniel Nathans Professor and Director of Molecular Biology and Genetics at Johns Hopkins University School of Medicine.

Previous CSHL Fellows Adrian Krainer (1986) and Marja Timmermans (1998) are now Professors at the Laboratory, and Chris Vakoc (2008) and Florin Albeanu (2008) are currently holding Assistant Professor positions at CSHL. Scott Lowe (1995) is a Howard Hughes Medical Institute (HHMI) Investigator. After nearly 15 years on the CSHL faculty, he recently took on a Professorship at Memorial Sloan-Kettering Cancer Center in New York City. 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; and TERENCE Strick (2000) left at the end of his fellowship to become a Group Leader at the Institute 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 an Associate Professor and Associate Director of the Genome Institute at Washington University. 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.

Lingbo Zhang has been a Fellow at the Laboratory since 2013. He joined us from Harvey Lodish's laboratory at the Whitehead Institute of the Massachusetts Institute of Technology, where he studied the regulation of erythroid progenitor cell self-renewal. As a CSHL Fellow, Lingbo is conducting genetic and small-molecule screens to discover novel regulators of normal and aberrant stem cell biology.

# UNDERSTANDING AND TARGETING SELF-RENEWAL OF NORMAL AND MALIGNANT ADULT STEM AND PROGENITOR CELLS

L. Zhang S. Chiappone G. Trivedi

Self-renewal is a unique property of stem and progenitor cells, allowing them to undergo several cycles of cell division while maintaining an undifferentiated status. During normal development, self-renewal allows stem and progenitor cells in adult tissues and organs to expand, a phenomenon that leads to tissue regeneration and can serve as a key target for regenerative medicine. However, compared to the well-known embryonic stem cell, our understanding of self-renewal in adult stem and progenitor cells is extremely limited, thus preventing us from therapeutically utilizing these cells. In addition to the crucial role of self-renewal in maintaining tissue homeostasis, malignant cells can hijack this property under pathological conditions to form cancer stem cells (or tumor-initiating cells) and contribute to carcinogenesis. Therefore, understanding the molecular mechanisms underlying self-renewal of cancer stem cells is critical for the eradication of these cells and prevention of relapse. Furthermore, we are probing the mechanistic divergence of self-renewal in normal versus malignant stem and progenitor cells, a crucial strategy for the development of targeted therapy that selectively disrupts carcinogenesis but not normal tissue homeostasis.

The research in my laboratory focuses on the hematopoietic, or blood-forming, system, using this lineage as a model system to uncover mysteries behind adult stem and progenitor self-renewal. Currently, we are mainly studying two important cell types, the early erythroid progenitor and the leukemia stem cell, in an effort to develop novel small chemical-compound-based therapeutic strategies for a wide range of blood diseases, including severe anemias and leukemias that are resistant to existing treatments. To better recapitulate the genetic heterogeneity of these diseases, we are also using the CRISPR/Cas9 genome editing technology to develop novel preclinical mouse models that possess combinatorial genetic lesions.

At the moment, bone marrow transplantation (BMT) is the only curative treatment for severe anemias and leukemias. However, BMT is a difficult and

risky procedure, and it is usually very hard or impossible to find a suitable bone marrow donor. Therefore, the long-term goal of our research is to better understand the molecular mechanisms governing self-renewal of patients' normal and malignant hematopoietic stem and progenitor cells in order to develop novel chemical approaches for these life-threatening blood diseases that bypass the standard complex BMT procedure.

## Understanding and Targeting Early Erythroid Progenitor Self-Renewal as Novel Therapeutic Strategies for a Broad Spectrum of Resistant Anemias

---

Anemia is a global health problem, with more than 1.6 billion people suffering from this disease. Anemic patients have low numbers of red blood cells and hemoglobin, and severe life-threatening cases are associated with a wide variety of chronic diseases such as bone marrow failure disorders, myelodysplastic syndrome (MDS), many major types of cancer, and cancer chemotherapy/radiotherapy. A large portion of these severe anemias are resistant to existing hormone therapies, highlighting the need for novel treatment strategies. We therefore focus our research on the highly self-renewing early erythroid progenitor in order to harness its expansion and regeneration potential for the formation of thousands of erythrocytes in the treatment of resistant anemias.

To therapeutically use this cell type, we are currently dissecting molecular pathways specifically required for self-renewal of the early erythroid progenitor, as well as identifying druggable proteins that can be effectively targeted. We have conducted a short hairpin RNA (shRNA)-based functional screening, focusing on genes that are enriched in early erythroid progenitors. As a complementary approach, we have also performed a small chemical compound-based forward genetics screening for chemical compounds that are able to boost early erythroid progenitor self-renewal.



Strikingly, several proteins emerged as positive hits in both the shRNA and chemical compound screening. When these proteins are blocked by either shRNA or chemical compound approaches, expansion of primary early erythroid progenitors is triggered. Importantly, in a preclinical anemia mouse model, we have demonstrated that delivery of these small chemical compounds facilitates recovery from anemia by triggering self-renewal of the early erythroid progenitor. This physiological efficacy confirms these proteins as potential bona fide targets for anemia treatment.

### **Generation of Blood Disease Mouse Models with Combinatorial Genetic Lesions Using CRISPR/Cas9 Genome Editing Technology**

---

Blood malignancies are genetically heterogeneous and complex diseases, and current mouse models do not faithfully recapitulate this heterogeneity and multiplicity. Therefore, through collaboration with the Gene Targeting Facility here at CSHL, we are using

CRISPR/Cas9 genome editing technology to develop novel mouse models of blood diseases.

CRISPR/Cas9 technology allows us to concurrently modify multiple genomic loci by introducing point mutations and deletions with high efficiency. We have designed single-guide RNA (sgRNA) and single-stranded DNA (ssDNA) corresponding to genes that are mutated in blood malignancies. Together with the gene targeting facility, we are injecting sgRNA and ssDNA along with Cas9 mRNA directly into fertilized zygotes. After transferring blastocysts into the uteruses of pseudopregnant females, we will obtain and test newborn mice to determine whether they carry the targeted mutations. With these novel blood disease mouse models, we will be able to better evaluate the efficacies of both existing therapies and novel therapies developed in our lab, as well as to predict their efficacies for patients carrying these different combinatorial genetic mutations. Additionally, these mouse models will help us better understand the pathophysiology of these genetically complex diseases.

# AUTHOR INDEX

---

- Aboukhalil R., 78, 226  
Aguirre-Chen C., 54  
Ahrens S., 149  
Akerman M., 37  
Akshinthala S., 94  
Alagesan B., 100  
Albeanu D., 113  
Alexander J., 78  
Alexandre Garcia Da  
  Silva P., 113  
Allen J., 78  
Almeida A., 85  
Ambrico A., 98  
An X., 136  
Anaparthi N., 70  
Anczuków O., 37  
Andrews P., 78  
Anselmi F., 113  
Antoniou E., 206  
Arun G., 40  
Atwal G.S., 226  
Azamy N., 206  
Aznarez I., 37  
  
Baker L., 100  
Balkunde R., 170  
Ballouz S., 194  
Baltera N., 151  
Banerjee A., 113  
Baslan T., 70  
Battistoni G., 58  
Batut P., 198  
Bekritsky M., 78  
Bencze S., 88  
Bencze G., 94  
Benkovics A., 187  
Bergmann J., 40  
Berstein Y., 206  
Beshel J., 167  
Bezirci F., 37  
Bhagwat A., 48  
Bhattacharjee S., 182  
Biffi G., 100  
Bobrowski-Khoury N.,  
  139  
Bodnar M.S., 40  
Boettcher S., 58  
Borges Monroy R., 126  
Bressan D., 58  
  
Brill M., 162  
Brooks C., 177  
Brown A., 123  
Bui H.T., 170  
Burbach B., 164  
  
Cappellani J., 85  
Carlston C., 54  
Castel S., 182  
Cazakoff B., 158  
Chae H.G., 113  
Chakraborty S., 198  
Chang A.-Y., 182  
Chang K., 58  
Chang K.-C., 40  
Chang Y., 73  
Chartarifsky L., 123  
Chaudhary F., 94  
Chen W.-J., 37  
Cheng D., 100  
Chiappone S., 245  
Chio I.I., 100  
Cho H., 98  
Chougule K., 212  
Churchland A., 123  
Claeys H., 170  
Corbo V., 100  
Cox H., 70  
Creasey K., 182  
Czech B., 58  
  
Das S., 37  
Davis C., 198  
Davis M., 113  
de Sousa Borges F., 182  
Delas Vives M.J., 58  
Delevich K., 149  
Demesa-Arevalo E., 170  
Demir E., 139  
Deshpande P., 206  
Diermeier S., 40  
Dinh M., 158  
Dobin A., 198  
Doerfel K., 54  
Doerfel M., 202  
Donoghue M., 182  
dos Santos C., 58  
Dotto M., 187  
Drenkow J., 198  
  
Dubnau J., 126  
  
Eckersley-Maslin M., 40  
Eckmeier D., 158  
Egeblad M., 85  
El Demerdash O., 58  
Elkayam E., 32  
Elmer B., 102  
Elyada E., 100  
Engle D., 100  
Enikolopov G., 128  
Epstein A., 32  
Erard N., 58  
Ernst E., 182  
Eskipehlivan S.M., 206  
Esposito D., 78  
Eveland A., 170  
  
Faehnle C., 32  
Fagegaltier D., 58  
Falcatori I., 58  
Fang G., 94  
Fang H., 236  
Feigin M., 100  
Fein M., 85  
Fenoglio S., 58  
Fernandez-Marco C., 187  
Ferrante D., 146, 151  
Field-Pollatou A., 151  
Filippini D., 100  
Fitzgerald G., 156  
Forcier T., 78  
Franciotti N., 151  
Fuchs M., 170  
Furukawa H., 132  
  
Gao W., 45  
Garvin T., 236  
Gauss C.M., 88  
Gerwitz A., 226  
Ghiban E., 206  
Ghosh S., 164  
Gillis J., 194  
Gingeras T.R., 198  
Gizatullina D., 164  
Goh W.S., 58  
Goldsmith S., 32  
Goodwin S., 206  
Grabowska E., 78  
  
Gupta P., 113  
Gurtowski J., 236  
Gutbrod M., 182  
Guzzardo P., 58  
  
Haase A., 58  
Hakker I., 78  
Hammell C.M., 54  
Hammell M., 28  
Han J.-J., 182  
Handly-Santana A., 100  
Hangya B., 139  
Hannon G., 58  
Hao Y., 28  
Harrison E., 58  
Hassan T., 88  
Havkin G., 151  
Hearn S., 40  
Herridge R., 182  
Herzka T., 98  
Hicks J., 70  
Hige T., 162  
Hirokawa J., 139  
Ho Y.-J., 28  
Hodges E., 58  
Hohmann A., 48  
Hossain M., 45  
Hua Y., 37  
Huang Z.J., 136  
Hübner M., 40  
Huilgol D., 136  
Husbands A., 187  
Hwang C.-I., 100  
Hwang D.-W., 73  
Hwang G., 149  
  
Iossifov I., 229  
Ipsaro J., 32  
  
Jackson D., 170  
Jacob Y., 182  
Jaganathan A., 73  
Jansen J., 45  
Jaramillo S., 164  
Je B.I., 170  
Jeon H.-Y., 37  
Jia Y., 85  
Jiang K., 177  
Jiao Y., 212

- Jin Y., 28  
 Joshua-Tor L., 32  
  
 Kandasamy R., 78  
 Kara N., 45  
 Karakas E., 132  
 Kaufman M., 123  
 Kebschull J., 164  
 Kelly S., 136  
 Kendall J., 78  
 Kepecs A., 139  
 Kim H.S., 182  
 Kim Y., 156  
 Kinney J., 240  
 Klingler S., 105  
 Knauer S., 187  
 Knott S., 58  
 Koh M., 113  
 Koh S., 164  
 Kota P., 202  
 Koulakov A., 146  
 Krainer A.R., 37  
 Kramer M., 206  
 Krasnitz A., 231  
 Krishnan K., 136  
 Krishnan N., 94  
 Krug L., 126  
 Kudla M., 58  
 Kuettner V., 85  
 Kuhn C., 32  
 Kumar V., 212, 234  
 Kumaran R.I., 40  
 Kumari S., 212  
 Kvitsiani D., 139  
  
 Lau B., 158  
 Lau T., 170  
 LeBlanc C., 182  
 Lee E., 58, 100  
 Lee H., 236  
 Lee S.C., 182  
 Lee S.-J., 32  
 Lee Y.-H., 78  
 Lee Y.K., 212  
 Leotta A., 78  
 Levine J., 136  
 Levy D., 234  
 Li B., 149  
 Li F., 105  
 Li J., 40, 75  
 Li L., 94  
 Li S., 70  
 Li W., 126  
 Liao I., 170  
  
 Liberatore K.L., 177  
 Licata A., 123  
 Lihm J., 206  
 Lippman Z., 177  
 Liseron-Monfils C., 212  
 Liu Y.-H., 37  
 Livneh U., 164  
 Lorenc M., 139  
 Lu J., 136  
 Lu Z., 212  
 Ludwig M., 100  
 Lyon G., 202  
  
 Ma B., 78  
 MacAlister C.A., 177  
 Maiorino L., 85  
 Malinauskas T., 132  
 Manche L., 37  
 Marbach F., 113, 164  
 Marcus S., 236  
 Marks S., 78  
 Martienssen R., 182  
 Masset P., 139  
 Mazurek A., 45  
 McCarthy S., 206  
 McCombie W.R., 206  
 McGillivray M., 231  
 McIndoo J., 78  
 Mechler F., 151  
 Mejia L., 149  
 Mende C., 156  
 Mendivil Ramos O., 206  
 Mercan F., 48  
 Michelsen S., 151  
 Michurina T., 128  
 Milazzo J., 48  
 Mills A., 73  
 Minder J., 48  
 Mitra P., 151  
 Modi M., 162  
 Mofunanya A., 75  
 Moissidis M., 136  
 Molik D., 28  
 Molla-Morales A., 182  
 Monaco M.K., 212  
 Muerdter F., 58  
 Mukherjee A., 151  
 Mukhopadhyay S., 78  
 Muller S., 206  
 Mulvaney J., 212  
  
 Naguib A., 98  
 Najafi F., 123  
  
 Nandi S., 45  
 Narasimhan A., 156  
 Narzisi G., 236  
 Nattestad M., 236  
 Nechooshtan G., 198  
 Neves Dos Santos Leite M., 212  
 Nguyen H., 123  
 Ni S., 45  
 Nomakuchi T., 37  
 Noutsos C., 212  
 Novatt J., 37  
 Nowak D., 98  
 Nunes Raposo D., 123  
  
 O'Rawe J., 202  
 Odoemene O., 123  
 Olson A., 212  
 Olund E., 100  
 On K., 32  
 Oni T., 100  
 Osten P., 156  
 Otazu G., 113  
 Oyibo H., 164  
  
 Pal D., 91  
 Palaniswamy R., 156  
 Panagot E., 37  
 Paniagua E., 28  
 Pappin D., 88  
 Parent J.-S., 182  
 Park J., 85  
 Park J.H., 128  
 Park S.J., 177  
 Park Y., 100  
 Patel A., 28  
 Paul A., 136  
 Peacock S., 88  
 Peikon I., 164  
 Penzo M., 149  
 Petsch K., 187  
 Peunova N., 128  
 Pi H.J., 139  
 Pie J.L., 139  
 Pinskiy V., 151  
 Ponz-Sarvise M., 100  
 Powers S., 75  
 Prakash A., 198  
 Prazak L., 126  
 Preall J., 58  
  
 Qi H., 70  
  
 Ramakrishnan S., 236  
 Ramesh M., 94  
 Ranade S., 139  
 Rao M., 75  
 Rebbeck C., 58  
 Rebouillat B., 113  
 Regalado Perez J., 28  
 Regan M., 132  
 Regulski M., 182, 212  
 Reid A., 164  
 Ren J., 182  
 Riggs M., 70  
 Rivera K., 88  
 Roche B., 182  
 Rodgers L., 78  
 Roe J., 48  
 Romero Hernandez A., 132  
 Ronemus M., 78  
 Rosenbaum J., 78  
 Rozhkov N., 58  
 Rozhkova E., 58  
 Rus I., 94  
 Rutigliano M., 170  
 Ryan M., 123  
  
 Sabin L., 58  
 Sanders J., 139  
 Savoia S., 151  
 Scavelli A., 198  
 Scharner J., 37  
 Schatz M., 236  
 Schiff H., 149  
 Schlesinger F., 198  
 Schorn A., 182  
 Schuck S., 43  
 Sedlazeck F., 236  
 See L.-H., 198  
 Senturk S., 91  
 Shea S., 158  
 Shen C., 48  
 Shen M., 170  
 Sherman R., 236  
 Sheu Y.-J., 45  
 Shi J., 48  
 Shields M., 85  
 Shih M.F., 126  
 Shimada A., 182  
 Shirole N., 91  
 Shuai Y., 167  
 Sikkens T., 139  
 Simorowski N., 132  
 Simorowski J., 182  
 Sivaramakrishnan M., 37

- Skopelitis D., 187  
 Snider J., 88  
 Son Y.E., 73  
 Song D., 94  
 Sordella R., 91  
 Soyk S., 177  
 Spector D.L., 40  
 Stahlhut C., 98  
 Stein J., 212  
 Stenlund A., 43  
 Stepansky A., 70  
 Stephenson-Jones M.,  
     149  
 Stillman B., 45  
 Sun G., 231  
 Szelenyi E., 156  
  
 Tai Y., 102  
 Tajima N., 132  
 Tam O., 28  
 Tan W., 37  
 Taranda J., 156  
 Teng S., 206  
 Thomain S., 177  
 Thomason J., 212  
 Timmermans M., 187  
 Tirado-Magallanes R.,  
     182  
  
 Tiriatic H., 100  
 Tocilj A., 32  
 Tolpygo A., 151  
 Tonks N.K., 94  
 Trivedi G., 245  
 Troge J., 78  
 Trotman L., 98  
 Tucciarone J., 136  
 Turna N.S., 88  
 Turner G., 162  
 Tuveson D., 100  
  
 Umamaheswari R., 182  
 Underwood C., 182  
  
 Vagin V., 58  
 Vakoc C.R., 48  
 Van Aelst L., 102  
 Van Buren P., 212  
 Vargas Landín D., 37  
 Vasily V., 164  
 Vaughan A., 139  
 Venkataraju K.U., 156  
 Vera K., 170  
 Verleyen W., 194  
  
 Wagenblast E., 58  
 Walleshauser J., 32  
  
 Wang B., 212  
 Wang C.-L., 102  
 Wang E., 48  
 Wang L., 212  
 Wang M., 102  
 Wang Z., 78  
 Wappel R., 37  
 Ware D., 212  
 Ware W., 212  
 Wasik K., 58  
 Watrud K., 98  
 Weber K., 151  
 Wei S., 212  
 Wei Y., 146  
 Wences A.H., 236  
 Wendel P., 45  
 Wigler M., 78  
 Wilson J., 88  
 Wong S., 37  
 Wright K., 100  
 Wu C., 105  
 Wu P., 136  
 Wu Q., 170  
 Wu Y., 202  
 Wysocki R., 85  
  
 Xiong Q., 164  
 Xu C., 177  
  
 Xu Y., 48  
 Xue H., 198  
  
 Yamrom B., 78  
 Yang M., 94  
 Yang Y.-T., 102  
 Yao Z., 91  
 Yoon C., 78  
 Yordanov G., 100  
 Yu A., 40  
 Yu J.-R., 102  
 Yu K., 100, 149  
 Yu Y., 58  
  
 Zador A., 164  
 Zadrozny T., 170  
 Zaleski C., 198  
 Zepeda-Mendoza C., 40  
 Zhang L., 78, 105, 212,  
     245  
 Zhang M., 167  
 Zhang S., 94  
 Zhang X., 149  
 Zheng W., 94  
 Zheng H., 105  
 Zhong Y., 167  
 Zhou X., 58  
 Znamenskiy P., 164





WATSON SCHOOL OF  
BIOLOGICAL SCIENCES

## ADMINISTRATION

**Alexander Gann, Ph.D.**, WSBS Professor and Dean

**Carrie Cowan, Ph.D.**, Associate Dean

**Alyson Kass-Eisler, Ph.D.**, Postdoctoral Program Officer and Curriculum Director

**Kimberly Creteur, M.Ed., M.SEd.**, Admissions and Recruitment Manager

**Kimberley Geer**, Administrative Assistant

## EXECUTIVE COMMITTEE

### Chair

Alexander Gann

### Members

Mickey Atwal

Mikala Egeblad

Adrian R. Krainer

Bo Li

David L. Spector (Director of Research)

David Stewart

### Student Representatives

Brittany Cazakoff, WSBS

Debjani Pal, SBU

### Secretary

Carrie Cowan

## ADMISSIONS COMMITTEE

### Chair

Adrian Krainer

### Members

Mickey Atwal

Anne Churchland

Alexander Gann (ex officio)

Zachary Lippman

W. Richard McCombie

Stephen Shea

Nicholas Tonks

Christopher Vakoc

Linda Van Aelst

### Secretary

Kimberly Creteur

## QUALIFYING EXAM COMMITTEE

### Chair

Linda Van Aelst

### Members

W. Richard McCombie

Marja Timmermans

Glenn Turner

Christopher Vakoc

## CURRICULUM DEVELOPMENT AND INTEGRATION COMMITTEE

### Chair

Carrie Cowan

### Members

Alexander Gann

David Jackson

Michael Schatz

Nicholas Tonks

Glenn Turner

### Secretary

Alyson Kass-Eisler

## EXTERNAL ADVISORY COMMITTEE

### Keith Yamamoto (Chair)

Executive Vice Dean, School of Medicine

University of California, San Francisco

### Victor Corces

Professor, Department of Biology

Emory University

Professor, Howard Hughes Medical Institute

### Gail Mandel

Senior Scientist, Vollum Institute

Oregon Health and Science University

Investigator, Howard Hughes Medical Institute

### Marguerite Mangin

Academic Programs Director and Senior Research Associate

The Rockefeller University

### Barbara Meyer

Professor of Genetics and Development

University of California, Berkeley

Investigator, Howard Hughes Medical Institute

### Frank Solomon

Professor, Department of Biology and Center

for Cancer Research

Massachusetts Institute of Technology

# WATSON SCHOOL OF BIOLOGICAL SCIENCES DEAN'S REPORT

---

Perhaps the most significant event for the Watson School of Biological Sciences (WSBS) in 2014 was its successful completion of reaccreditation by the New York State Board of Regents and the New York State Education Department (NYSED). In consequence, the School is now accredited through September 16, 2021.

The reaccreditation process began in earnest in 2013 with the submission of a self-study by the school, followed by a site visit by state officials, scientists, and educators from institutions across New York State. The site visit was held from October 31 to November 1, 2013, after which the review committee submitted a report on the School's compliance with the Standards to NYSED and the Board of Regents. The School's response, together with the original review committee report, was sent to the Board of Regents and Commissioner of Education for evaluation. On September 16, 2014, the Board of Regents voted, as recommended by the Regents Advisory Council, and authorized renewal of the Laboratory's accreditation for a period of 7 years.

## Faculty Changes

Five new faculty members joined the Watson School in 2014: Douglas Fearon, Justin Kinney, Je Hyuk (Jay) Lee, Adam Siepel, and Jessica Tollkuhn. Professor Douglas Fearon studies how the immune system can be activated to fight cancer. His lab uses mouse models of pancreatic cancer to understand immune regulation in tumors and to develop therapies that use the immune system to combat disease. He joins the Watson School from his previous position at Cancer Research UK and the University of Cambridge. Justin Kinney, an Assistant Professor and former Quantitative Biology Fellow, uses next-generation sequencing to understand protein–protein and protein–DNA interactions in a quantitative way. He takes a combination of theoretical, computational, and experimental approaches to dissect open problems in transcriptional regulation, DNA replication, and immunology. Assistant Professor Jay Lee develops sophisticated imaging and molecular sequencing techniques to understand how cancer cells arise and evolve. He is particularly interested in how cells interact with their local environment and how signals from the environment lead to changes in gene expression. Jay comes to the Watson School from his postdoctoral tenure in the lab of George Church at Harvard University. Professor Adam Siepel, the Chair of the Simons Center for Quantitative Biology, uses approaches from statistics, computer science, and evolutionary genetics to understand various complex problems in genetics. His lab is especially interested in how genomes change over time and uses this information to inform human health. Adam remains a professor of computational biology at Weill Cornell Medical College. Assistant Professor Jessica Tollkuhn investigates the regulation of sex-specific behaviors. Her group studies how transient hormone spikes during development lead to changes in gene expression and chromatin that, in turn, change the neural circuits that underlie male- or female-specific behaviors. Jessica comes to the Watson School from postdoctoral research in the lab of Nirao Shah at the University of California, San Francisco. These new faculty members have already participated in WSBS activities. Each of them has given a Research Topics talk to the first-year students. They are also serving as members of thesis committees and expert qualifying examiners and have lectured in WSBS core courses. We look forward to their growing participation as members of the faculty.

Two faculty members also departed the School this year. Scott Lowe was one of the original instructors of the Scientific Reasoning and Logic Course. He also served on the Executive and Qualifying Exam Committees. He was on a number of thesis committees and was the thesis



research mentor to seven WSBS students. Scott Powers was also an instructor in the Scientific Reasoning and Logic Course and served as Chair on three thesis committees. We will miss them both and wish them every success in their future endeavors.

### The 11th WSBS Graduation

On May 4, 2014, we celebrated the Watson School's 11th graduation ceremony. Eleven students were awarded Ph.D. degrees: Marek Kudla, Hassana Oyibo, Michael Pautler, Yevgeniy Plavskin, and Joshua Sanders from the Entering Class of 2007; Philippe Batut, Dario Bressan, Melanie Eckersley-Maslin, Katie Liberatore, and Zinaida Perova from the Entering Class of 2008; and Kaja Wasik from the Entering Class of 2009. William Donovan from the Entering Class of 2012 was awarded a master's degree. An honorary degree was bestowed upon Dr. Richard Burgess, the James D. Watson Emeritus Professor of Oncology at the McArdle Laboratory for Cancer Research, University of Wisconsin, Madison. Dick has published more than 250 papers in the fields of biochemistry and molecular biology and has made important scientific discoveries, including the identification of the *Escherichia coli* RNA polymerase specificity factor, sigma. From 1992 to 2013, he taught the CSHL Protein Purification Course, making him one of the longest serving course instructors in the Laboratory's history. Additionally, he was part of the team that generated the 1996 CSHL Press manual based on the course, *Strategies for Protein Purification and Characterization: A Laboratory Course Manual*.

As with each graduation, we extended a special welcome to the family members and friends of our students who attended the ceremony.



2014 WSBS Graduates: (Left to right) WSBS Dean Alexander Gann, Chancellor Emeritus James Watson, Marek Kudla, Hassana Oyibo, CSHL Chairman Jamie Nicholls, CSHL President Bruce Stillman, Melanie Eckersley-Maslin, Dario Bressan, Kaja Wasik, Katie Liberatore, Michael Pautler, Philippe Batut, Yevgeniy Plavskin, Joshua Sanders

## 2014 WSBS DOCTORAL RECIPIENTS

Student	Thesis advisor	Academic mentor	Current position
Ozlem (Mert) Aksoy	Scott Lowe	John Inglis	Postdoctoral Fellow, University of California, San Francisco (Advisor: William Weiss)
Philippe Batut	Thomas Gingeras	Zachary Lippman	Graduate Student Post, CSHL (Advisor: Thomas Gingeras)
Mitchell Bekritsky	Michael Wigler	W. Richard McCombie	Bioinformatic Scientist, Illumina, Inc., Cambridge, United Kingdom
Dario Bressan	Gregory Hannon	Z. Josh Huang	Postdoctoral Fellow, University of California, San Francisco (Advisor: William Weiss)
Stephane Castel	Robert Martienssen	Lloyd Trotman	Postdoctoral Fellow, New York Genome Center (Advisor: Tuuli Lappalainen)
Kristen Delevich	Bo Li	Stephen Shea	Postdoctoral Fellow, Stanford University (Advisor: Linda Wilbrecht)
Melanie Eckersley-Maslin	David Spector	Gregory Hannon	Postdoctoral Fellow, Cambridge University (Advisor: Wolf Reik)
Sang-Geol Koh	Anthony Zador	Glenn Turner	Scientist, Start-up Company, Korea
Katie Liberatore	Zachary Lippman	Adrian Krainer	Postdoctoral Fellow, U.S. Department of Agriculture, University of Minnesota (Advisor: Shahryar Kianian)
Ian Peikon	Anthony Zador	Mickey Atwal	Scientist, Google X, California
Zinaida Perova	Bo Li	Linda Van Aelst	Postdoctoral Fellow, Cambridge University (Advisor: Tiago Branco)
Nilgun Tasdemir	Scott Lowe	Josh Dubnau	Postdoctoral Fellow, University of Pittsburgh (Advisor: Nancy Davidson)
Elvin Wagenblast	Gregory Hannon	Jan Witkowski	Graduate Student Post, CSHL (Advisor: Gregory Hannon)
Kaja Wasik	Gregory Hannon	Jan Witkowski	Graduate Student Post, CSHL (Advisor: Gregory Hannon)
Susann Weissmueller	Scott Lowe	Raffaella Sordella	Equity Research Associate, MLV & Co., Connecticut
Cinhyia Zepeda Mendoza	David Spector	Thomas Gingeras	Postdoctoral Fellow, Brigham and Women's Hospital, Harvard Medical School (Advisor: Cynthia Morton)

## 2014 THESIS DISSERTATION DEFENSES

## ENTERING CLASS OF 2008

Ozlem (Mert) Aksoy, May 7, 2014

*The role of E2F7 in senescence and development.*

## Thesis Examining Committee

Chair: Lloyd Trotman  
 Research Mentor: Scott Lowe  
 Academic Mentor: John Inglis  
 Committee Member: David Spector

(continued)

2014 THESIS DISSERTATION DEFENSES (*continued*)

Committee Member: **Emily Bernstein**, *Mount Sinai School of Medicine*  
 External Examiner: **Eva Hernando-Monge**,  
*New York University*

**Philippe Batut**, January 13, 2014

*Promoter evolution in Drosophila: Noncoding transcription and transposon-driven innovation.*

**Thesis Examining Committee**

Chair: **Gregory Hannon**  
 Research Mentor: **Thomas Gingeras**  
 Academic Mentor: **Zachary Lippman**  
 Committee Member: **Adrian Kraine**  
 Committee Member: **Josh Dubnau**  
 External Examiner: **Brenton Gravely**, *University of Connecticut*

**Mitchell Bekritsky**, June 25, 2014

*Detecting de novo microsatellite mutations in a population of families with sporadic autism.*

**Thesis Examining Committee**

Chair: **Mickey Atwal**  
 Research Mentor: **Michael Wigler**  
 Academic Mentor: **W. Richard McCombie**  
 Committee Member: **Raffaella Sordella**  
 Committee Member: **Michael Schatz**  
 External Examiner: **Bud Mishra**, *New York University*

**Dario Bressan**, January 27, 2014

*A novel technology for the space-specific recovery of biological molecules.*

**Thesis Examining Committee**

Chair: **Linda Van Aelst**  
 Research Mentor: **Gregory Hannon**  
 Academic Mentor: **Z. Josh Huang**  
 Committee Member: **Anthony Zador**  
 Committee Member: **Josh Dubnau**  
 External Examiner: **Robert Darnell**, *The Rockefeller University*

**Sang-Geol Koh**, June 26, 2014

*The role of callosal projections in auditory coding.*

**Thesis Examining Committee**

Chair: **Pavel Osten**  
 Research Mentor: **Anthony Zador**  
 Academic Mentor: **Glenn Turner**  
 Committee Member: **Adam Kepecs**  
 Committee Member: **Alexei Koulikov**  
 External Examiner: **Richard Mooney**, *Duke University School of Medicine*

**Nilgun Tasdemir**, May 21, 2014

*Investigating the role of Brd4 in cellular senescence and normal tissue homeostasis.*

**Thesis Examining Committee**

Chair: **Linda Van Aelst**  
 Research Mentor: **Scott Lowe**  
 Academic Mentor: **Josh Dubnau**  
 Committee Member: **Alea Mills**  
 Committee Member: **Raffaella Sordella**  
 External Examiner: **Johanna Joyce**, *Memorial Sloan-Kettering Cancer Center*

**Elvin Wagenblast**, July 23, 2014

*A molecular characterization of breast cancer progression, including onset, tumor growth, and metastasis.*

**Thesis Examining Committee**

Chair: **Marja Timmermans**  
 Research Mentor: **Gregory Hannon**  
 Academic Mentor: **Jan Witkowski**  
 Committee Member: **Mikala Egeblad**  
 External Examiner: **Charles Sawyers**, *Memorial Sloan-Kettering Cancer Center*

**Susann Weissmueller**, May 5, 2014

*Molecular mechanisms and therapeutic implications of chromosome 17p alterations.*

**Thesis Examining Committee**

Chair: **David Tuveson**  
 Research Mentor: **Scott Lowe**  
 Academic Mentor: **Gregory Hannon**  
 Committee Member: **Linda Van Aelst**  
 Committee Member: **Christopher Vakoc**  
 External Examiner: **Carol Prives**, *Columbia University*

**ENTERING CLASS OF 2009**

**Stephane Castel**, June 19, 2014

*Transcriptional regulation in the context of DNA replication by RNA interference in fission yeast.*

**Thesis Examining Committee**

Chair: **James Hicks**  
 Research Mentor: **Robert Martienssen**  
 Academic Mentor: **Lloyd Trotman**  
 Committee Member: **Matthew O'Connell**, *Mount Sinai School of Medicine*  
 External Examiner: **Benoit Arcangioli**, *Institut Pasteur*

**Kristen Delevich**, July 29, 2014

*Synaptic effects of Disrupted-in-Schizophrenia 1 loss-of-function in the medial prefrontal cortex and thalamofrontal feedforward inhibitory circuit.*

**Thesis Examining Committee**

Chair: **Z. Josh Huang**  
 Research Mentor: **Bo Li**

(*continued*)

2014 THESIS DISSERTATION DEFENSES (*continued*)

Academic Mentor: **Stephen Shea**  
 Committee Member: **Anthony Zador**  
 External Examiner: **Barry Connors, Brown University**

**Kaja Wasik, March 21, 2014**  
*Unusual aspects of piRNA pathways in mice and flatworms.*

**Wee Siong Goh, August 29, 2014**

*Biogenesis and Function of piRNAs.*

**Thesis Examining Committee**  
 Chair: **Marja Timmermans**  
 Research Mentor: **Gregory Hannon**  
 Academic Mentor: **Hiro Furukawa**  
 Committee Member: **Christopher Hammell**  
 Committee Member: **Adrian Krainer**  
 External Examiner: **Victor Ambros, University of Massachusetts Medical School**

**Thesis Examining Committee**  
 Chair: **Zachary Lippman**  
 Research Mentor: **Gregory Hannon**  
 Academic Mentor: **Jan Witkowski**  
 Committee Member: **Christopher Hammell**  
 External Examiner: **Ruth Lehmann, New York University**

**Ian Peikon, September 26, 2014**

*Encoding of neural circuit information into DNA.*

**Thesis Examining Committee**  
 Chair: **Josh Dubnau**  
 Research Mentor: **Anthony Zador**  
 Academic Mentor: **Mickey Atwal**  
 Committee Member: **Bruce Stillman**  
 External Examiner: **Cori Bargmann, The Rockefeller University**

**Cinthya Zepeda Mendoza, September 24, 2014**  
*Impact of copy-number variation on chromatin interactions at the mouse 4E2 chromosome region.*

**Thesis Examining Committee**  
 Chair: **Robert Martienssen**  
 Research Mentor: **David Spector**  
 Academic Mentor: **Thomas Gingeras**  
 Committee Member: **Alea Mills**  
 Committee Member: **Thomas Reid, National Institutes of Health**  
 External Examiner: **James Lupski, Baylor College of Medicine**

## Teaching Award

At the graduation ceremony this year, WSBS Associate Professor Zachary Lippman was presented with the ninth annual Winship Herr Faculty Teaching Award, named in honor of the School's founding Dean. Zach is the instructor of the Specialized Disciplines Course in Genetics and Genomics and was chosen by the students for this award, based on his enthusiasm and creativity in teaching. In addition to receiving this honor, Zach, who earned his Ph.D. degree from the Watson School in 2005, graduated his first student during this ceremony—that is, the School's first second-generation student. Here is some of what the students said about Zach's teaching in their nominations:



Zachary Lippman

*“Zach was one of the instructors who really managed to infect everyone with his enthusiasm about the subject. His preparation for the lectures was excellent—minimizing confusion and being able to explain important and complex concepts in an understandable manner. Moreover, he was dedicated to the course, available for feedback and questions at any time, and succeeded in understanding everyone's needs and adapting his teaching to them.”*

*“The critical and analytical thinking skills I learned from Zach will carry over into any area of research in which I find myself in the future.”*

## DOCTORAL THESIS RESEARCH

Student	Academic mentor	Research mentor	Thesis research
<b>ENTERING CLASS OF 2009</b>			
<b>Silvia Fenoglio</b> <i>Elisabeth Sloan Livingston Fellow</i>	Linda Van Aelst	Gregory Hannon	RNAi screening to identify putative therapeutic targets for the treatment of pancreatic cancer.
<b>ENTERING CLASS OF 2010</b>			
<b>Arkarup Bandyopadhyay</b> <i>Goldberg Lindsay Fellow</i>	Zachary Lippman	Florin Albeanu	Identity and intensity encoding of odors in rodents.
<b>Colleen Carlston</b> <i>John and Amy Phelan Student National Science Foundation Fellow</i>	Hiro Furukawa	Christopher Hammell	Identification and characterization of noise-suppressor genes that act via microRNAs in <i>Caenorhabditis elegans</i> larval development.
<b>Matthew Koh</b> <i>George A and Marjorie Anderson Fellow</i>	Bo Li	Florin Albeanu	Roles of olfactory bulb inhibitory microcircuits in shaping the temporal response properties of mitral cells.
<b>Lisa Krug</b> <i>NIH Predoctoral Trainee</i>	Stephen Shea	Josh Dubnau	Mechanisms of transposon regulation in the central nervous system.
<b>Jack Walleshauser</b> <i>Barbara McClintock/NIH Predoctoral Trainee</i>	Christopher Hammell	Leemor Joshua-Tor	Structural basis for TUT4 uridylation of pre-let-7/lin28 complex.
<b>ENTERING CLASS OF 2011</b>			
<b>Robert Aboukhalil</b> <i>NIH Predoctoral Trainee Starr Centennial Scholar</i>	Josh Dubnau	Mickey Atwal Michael Wigler	Using single-cell RNA-Seq to investigate tumor heterogeneity and evolution.
<b>Brittany Cazakoff</b> <i>Edward and Martha Gerry Fellow NSERC Scholar</i>	Christopher Hammell	Stephen Shea	Dynamic granule cell processing of odor information.
<b>Joaquina Delas Vives</b> <i>La Caixa Fellow Boehringer Ingelheim Fonds Fellow</i>	Nicholas Tonks	Gregory Hannon	Functional role of long non-coding RNAs in hematopoiesis.
<b>Anja Hohmann</b> <i>David H. Koch Fellow Boehringer Ingelheim Fonds Fellow</i>	John Inglis	Christopher Vakoc	Exploring the role of bromodomain-containing protein 9 (Brd9) in the maintenance of acute myeloid leukemia.
<b>Justus Kobschull</b> <i>David and Fanny Luke Fellow Genentech Foundation Fellow Boehringer Ingelheim Fonds Fellow</i>	Marja Timmermans	Anthony Zador	Grasping the brain.
<b>Fred Marbach</b> <i>Farish-Gerry Fellow</i>	Josh Dubnau	Anthony Zador	A study of auditory corticostriatal cells in the behaving mouse.
<b>Onyekachi Odoemene</b> <i>NIH Individual Fellowship William Randolph Hearst Scholar</i>	Stephen Shea	Anne Churchland	The role of neural inhibition in perceptual decision-making.
<b>Sophie Thomain</b> <i>George A. and Marjorie H. Anderson Fellow</i>	Josh Dubnau	Zachary Lippman	Characterization of a new meristem maintenance pathway in tomato and <i>Arabidopsis thaliana</i> and its relation to pollen tube growth.
<b>Charles Underwood</b> <i>William R. Miller Fellow</i>	Michael Schatz	Robert Martienssen	Epigenetic inheritance through mitosis and meiosis in <i>Arabidopsis thaliana</i> .
<b>ENTERING CLASS OF 2012</b>			
<b>Talitha Forcier</b> <i>NIH Predoctoral Trainee William Randolph Hearst Scholar</i>	Nicholas Tonks	Michael Wigler	Determining the change of transcriptional events from single-cell RNA analysis.
<b>Tyler Garvin</b> <i>NIH Predoctoral Trainee</i>	Zachary Lippman	Michael Schatz	Structural variants and gene networks underlying complex human disease.

(continued)

DOCTORAL THESIS RESEARCH <i>(continued)</i>			
Student	Academic mentor	Research mentor	Thesis research
<b>Yu-Jui (Ray) Ho</b> <i>David and Fanny Luke Fellow</i>	Michael Schatz	Molly Hammell	Methods development for low-input RNA-Seq analysis and application to breast cancer heterogeneity.
<b>Paul Masset</b> <i>Florence Gould Fellow</i>	Jan A. Witkowski	Adam Kepecs	Representations of decision confidence in the brain: From Bayes' rule to channelrhodospin.
<b>Annabel Romero Hernandez</b> <i>Genentech Foundation Fellow</i>	Adrian R. Krainer	Hiro Furukawa	Molecular mechanisms of inhibition in NMDA receptors.
<b>Abram Santana</b> <i>Robert and Theresa Lindsay Fellow</i> <i>William Randolph Hearst Scholar</i>	Lloyd Trotman	David Tuveson	Pancreatic tumor biology and therapy.
<b>ENTERING CLASS OF 2013</b>			
<b>Giorgia Battistoni</b> <i>Starr Centennial Scholar</i>	Christopher Hammell	Gregory Hannon	One-carbon metabolism and DNA methylome dynamics in pancreatic cancer.
<b>Lital Charatfisky</b> <i>Israeli WSBS Fellow</i>	John Inglis	Anne Churchland	Neural circuits for multisensory integration in normal and disease states.
<b>Sanchari Ghosh</b> <i>Charles A. Dana Fellow</i>	Josh Dubnau	Anthony Zador	Role of corticostriatal plasticity in learning an auditory discrimination task.
<b>Michael Gutbrod</b> <i>Bristol-Myers Squibb Fellow</i>	Zachary Lippman	Robert Martienssen	Small RNA and the RNAi pathway in transposable element regulation and differentiation in the preimplantation embryo.
<b>Daniel Kepple</b> <i>Crick-Clay Fellow</i> <i>NIH Predoctoral Trainee</i>	David Stewart	Alexei Koulakov	The human olfactory space.
<b>Laura Maiorino</b> <i>George A and Marjorie Anderson Fellow</i>	Nicholas Tonks	Mikala Egeblad	Understanding the role of the epithelial–mesenchymal plasticity in pancreatic cancer metastasis.
<b>Maria Nattestad</b> <i>John and Amy Phelan Student</i>	Linda Van Aelst	Michael Schatz	Algorithms for analyzing and assembling complex genomes.
<b>Georgi Yordanov</b> <i>Leslie C. Quick, Jr. Fellow</i>	Leemor Joshua-Tor	David Tuveson	Role of c-Myc in pancreatic cancer.

## Admissions 2014

The Watson School received 276 applications for the Entering Class of 2014 and is indebted to its Admissions Committee, which reviewed, interviewed, and selected candidates for our doctoral program. The Admissions Committee for the 2014 entering class comprised Adrian Krainer (chair), Mickey Atwal, Anne Churchland, Zachary Lippman, W. Richard McCombie, Stephen Shea, Nicholas Tonks, Christopher Vakoc, Linda Van Aelst, and myself.

## Entering Class of 2014

On August 18, 2014, the WSBS welcomed the 16th incoming class. The eight students—Emilis Bruzas, Hamza Giffar, Jacqueline Giovanniello, Devinn Lambert, Sashank Pisupati, Colin Stoneking, Jue Xiang Wang, and Anqi Zhang—hail from the United States, Canada, Germany, India, Lithuania, and the United Kingdom.

## Academic Mentoring

The Watson School takes great pride in the level of mentoring that it offers its students. One example is our two-tiered mentoring approach, whereby each student chooses both an

### ENTERING CLASS OF 2014

**Emilis Bruzas**, University of Cambridge: B.A. in Biochemistry (2014); Natural Sciences Tripos Pt IB College Scholarship (Thomas Hobbes 1631) (2012, 2013); Hardy Award for Biological Science (2013)

**Academic Mentor:** Alea Mills

**Hamza Giaffar**, Imperial College London: M.Sci. in Chemistry (2012), Associate of the Royal College of Science (ARCS)

**Academic Mentor:** Jan Witkowski

**Jacqueline Giovanniello**, Brown University: B.S. in Human Biology (2012); Distinction of Honors in the Department of Biology (2012); Dean's Discretionary Grant for Undergraduate Research (2010–2012); Cardiovascular Research Center Grant (2010–2012)

**Academic Mentor:** Bruce Stillman

**Elizabeth Hutton**, Princeton University: B.S. in Molecular Biology (2014); Quantitative and Computational Biology Certificate (2014); Office of the Dean, College Senior Thesis Funding Award (2013); Princeton University German Summer Work Program Award (2012)

**Academic Mentor:** Molly Hammell

**Devinn Lambert**, Rutgers University: B.S. in Biotechnology (2013); M.Phil. Plant Sciences, University of Cambridge (2014); Gates Foundation Fellowship (2014); National Science Foundation Graduate Research Fellowship Program (2013); Marshall Finalist (2012); Goldwater Scholar (2012); Singapore International PreGraduate Award (2011); National Science Foundation Biotechnology Scholar (2009–2013); Cap & Skull Senior Honors

Society (2012); Dunbar Fund for Excellence Grant (2011); Award for Academic Excellence, Rutgers VP of Undergraduate Education (2011); Rutgers Aresty Research Fellow (2010)

**Academic Mentor:** David Stewart

**Sashank Pisupati**, Indian Institute of Technology, Kanpur: B.Tech. in Biological Sciences and Bioengineering (2014); Baljit and Nirmal Dhindsa UG Scholarship (Fall 2010, 2011); KVPY Fellowship (2009); Runner up in "Mimamsa 2012," a national intercollegiate science challenge (2012); Second in Robocamp, a Robotics Workshop held by IIIT Hyderabad (2011)

**Academic Mentor:** Stephen Shea

**Colin Stoneking**, University of Cambridge: B.A. in Neuroscience (2012); M.S. in Statistics, ETH Zurich (2014); The Undergraduate Prize for Physiology, The Physiological Society (2012); Robert Dobson Memorial Prize, Emmanuel College, Cambridge (2010)

**Academic Mentor:** Zachary Lippman

**Jue Xiang Wang**, University of Cambridge: B.A./M.Sci. in Natural Sciences/Biochemistry (2013); First Class Grade (2013)

**Academic Mentor:** Mikala Egeblad

**Anqi Zhang**, McGill University: B.Sci. in Neuroscience, with a minor in Mathematics (2014); NSERC Undergraduate Student Research Award, University of Waterloo (Summer 2013); McGill RE Powell Scholarship (2010–2013); Dean's Honour List (2011)

**Academic Mentor:** Michael Schatz



2014 WSBS Entering Class: (Left to right) Sashank Pisupati, Anqi Zhang, Elizabeth Hutton, Emilis Bruzas, Jue Xiang Wang, Hamza Giaffar, Devinn Lambert, Jacqueline Giovanniello, Colin Stoneking

academic and a research mentor. The academic mentor is critical for monitoring students—and offering advice—during the intensive coursework of the first term, during their rotations, and when identifying a suitable research mentor. Furthermore, the academic mentor continues to follow 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 an 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 following are the Academic Mentors for the Entering Class of 2014.

STUDENT	MENTOR	STUDENT	MENTOR
Emilis Bruzas	Alea Mills	Sashank Pisupati	Stephen Shea
Hamza Giaffar	Jan Witkowski	Colin Stoneking	Zachary Lippman
Jacqueline Giovannello	Bruce Stillman	Jue Xiang Wang	Mikala Egeblad
Elizabeth Hutton	Molly Hammell	Anqi Zhang	Michael Schatz
Devinn Lambert	David Stewart	Emilis Bruzas	Alea Mills

### The Fall Term Curriculum

Our faculty continues to do an outstanding job developing and delivering the curriculum, and we are extremely grateful for the time and effort they spend on maintaining the high quality of the courses. The Curriculum Development and Integration Committee (CDIC)—Carrie Cowan (chair), David Jackson, Michael Schatz, Nicholas Tonks, and Glenn Turner—oversees development of the curriculum. In addition to the course instructors and guest lecturers from within the Laboratory, our courses attract an impressive array of guest lecturers from other institutions.

### Recruiting Efforts

This year, we once again focused primarily on targeted visits to graduate fairs and minority conferences to identify prospective students for the Watson School. In addition to these visits, the WSBS generated new mailing lists for the distribution of information to top undergraduate science departments around the world. A multiprogram booklet was designed for this recruitment season. Additionally, emails were sent to personalized contacts and an electronic mailing list of more than 50,000 individuals who receive information from the Cold Spring Harbor Laboratory Press or have attended CSHL Meetings and Courses or Banbury meetings. We are grateful to these departments for sharing this contact list. The targeting efforts appear to be successful in that the number of applications received—310—for the Entering Class of 2015 is the largest number to date.

#### WATSON SCHOOL OF BIOLOGICAL SCIENCES 2014 RECRUITMENT SCHEDULE

Event	Location	Date
Rutgers University STEPed UP for Success Program, Visit and Information Session	Cold Spring Harbor Laboratory	April 4
American Association of Cancer Research, Annual Meeting	San Diego, California	April 5–9
Molloy College, Visit and Information Session	Cold Spring Harbor Laboratory	April 25
Vassar College, Visit and Information Session	Cold Spring Harbor Laboratory	May 2
UMDNJ–NJ Medical School, PHRI, Visit and Information Session	Cold Spring Harbor Laboratory	August 1
Brookhaven National Laboratory, Graduate School Fair and Undergraduate Research Program Symposium	Brookhaven, New York	August 6

(continued)



**WATSON SCHOOL OF BIOLOGICAL SCIENCES 2014 RECRUITMENT SCHEDULE** *(continued)*

<b>Event</b>	<b>Location</b>	<b>Date</b>
Big 10+ Graduate School Expo (Purdue University), Graduate School Fair	West Lafayette, Indiana	September 21–22
University of Maryland, Baltimore County, Meyerhoff Scholars Program, Information Session	Baltimore, Maryland	October 1
Johns Hopkins University Information Session	Baltimore, Maryland	October 1
Hunter College MARC and RISE/MBRS, Information Session	New York, New York	October 7
University of Cambridge, Information Session	Cambridge, England	October 16
University of Southern California, Information Session	Los Angeles, California	October 16
Society for Advancement of Chicanos and Native Americans in Science (SACNAS), National Conference	Los Angeles, California	October 16–18
American Society for Human Genetics, Annual Meeting	San Diego, California	October 18–22
Columbia University, Information Session	New York, New York	October 20
Stanford University, Graduate School Fair	Stanford, California	October 21
University of California, Berkeley, Information Session	Berkeley, California	October 22
Harvard University, Information Sessions	Cambridge, Massachusetts	November 6
California Forum for Diversity in Graduate Education, Graduate School Fair	San Diego, California	November 8
Annual Biomedical Research Conference for Minority Students (ABRCMS), National Conference	San Antonio, Texas	November 12–15
Society for Neuroscience Annual Meeting, Graduate School Fair	Washington, D.C.	November 15–19
American Society for Cell Biology, Annual Meeting	Philadelphia, Pennsylvania	December 6–10

**Interinstitutional Academic Interactions**

WSBS students account for approximately half of the total graduate student population here at CSHL. The other half comprise 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 students are from Stony Brook University (SBU), via a program established between CSHL and SBU more than 30 years ago. Over the years, we have established relationships with other institutions around the world, enabling their students to conduct research here at CSHL. Currently, we have visiting students from institutions in France, India, Portugal, and the United States. The Watson 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 this year.

<b>STUDENT</b>	<b>CSHL RESEARCH MENTOR</b>	<b>STONY BROOK UNIVERSITY PROGRAM</b>
Derek Cheng	David Tuveson	Molecular Genetics and Microbiology
Hsiang-Chen Chou	Bruce Stillman	Molecular and Cellular Biology
Han Fang	Michael Schatz	Applied Mathematics and Statistics
Shanshan Li	James Hicks	Applied Mathematics and Statistics
Jason O'Rawe	Gholson Lyon	Molecular Genetics and Microbiology
Dongyan Song	Nicholas Tonks	Molecular and Cellular Biology
Allen Yu	David Spector	Molecular Genetics and Microbiology

**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 May, the other in October. Each Symposium

consists of senior students giving short talks. Coffee breaks and lunch provide opportunities for more informal interactions. The prize for best talk for the May session was awarded to WSBS student Brittany Cazakoff (Shea lab), and for the October session, it was shared by Annabel Romero Hernandez (WSBS, Furukawa lab) and Bobby Wysocki (SBU, Egeblad lab). We are grateful to the two student chairs, Debjani Pal (SBU) and Arkarup Bandyopadhyay (WSBS), for their hard work and to WSBS's Kim Geer for helping make the Symposium a great success.

## Graduate Student and Postdoctoral Fellow Departures

With each year come not only new arrivals, but also departures. The following graduate students and postdoctoral fellows left the Laboratory during 2014:

### Postdoctoral Fellows

Christine Ardito	Balazs Hangya	Haohong Li	Roberto Perales
Huyen Bui	Miao He	Jinyu Li	Lisa Prazak
Alexander Campbell	Michael Huebner	Li Li	Jonathan Preall
Marcela Dotto	Ke Jiang	Cora MacAlister	Andrea Schorn
Cheryl Eifert	Shimpei Kawaoka	Shoshana Marcus	Naoki Takada
Andrea Eveland	Sukant Khurana	Adaobi Mofunanya	Yi Wei
Pascal Grange	Claus-Dieter Kuhn	Amit Mukherjee	Huiqing Zhan
Baofeng Guo	Duda Kvitsiani	June Hee Park	Ling Zhang

### Graduate Students

Ozlem (Mert) Aksoy	Stephane Castel	Katie Liberatore	John Sheppard
Deanna Bahel	Hyejin Cho	Felix Muerdter	Guoli Sun
Taimour Balan	Shipra Das	Hassana Oyibo	Nilgun Tasdemir
Mitchell Bekritsky	William Donovan	Zinaida Perova	Assaf Vestin
Fatma Bezirci	Melanie Eckersley-Maslin	Yevgeniy Plavskin	Susann Weissmueller
Megan Bodnar	Paloma Guzzardo	Huan Qi	Chunsu Xu
Dario Bressan	Nihan Kara	Joshua Sanders	

## 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. I wish to thank faculty members Mickey Atwal, Mikala Egeblad, Adrian Krainer, Bo Li, David Spector, David Stewart, and Linda Van Aelst. I would also like to thank the student representatives Debjani Pal (SBU) and Brittany Cazakoff (WSBS), who contributed to discussions and provided useful suggestions and feedback from their colleagues.

## The Watson School Continues to Benefit from Generous Benefactors

We are extremely grateful for the generous donors whose one-time gifts or continued support made our 2014 programs possible, including the Annenberg Foundation, Lita Annenberg Hazen Foundation, Bristol-Myers Squibb Company, Mr. and Mrs. Landon Clay, Lester Crown, the Dana Foundation, Henriette and Norris Darrell, the Samuel Freeman Charitable Trust, the William Stamps Farish Fund, the Genentech Foundation, Mr. and Mrs. Alan Goldberg, Florence Gould Foundation, Gonzalo Río Arronte Foundation, William Randolph Hearst Foundation, Dr. and Mrs. Mark Hoffman, Annette Kade Charitable Trust, Mr. David H. Koch, Mr. and Mrs. Robert D. Lindsay and Family, Mr. and Mrs. David Luke III, Marjorie A. Matheson, Mr. and Mrs. William R. Miller, OSI Pharmaceuticals Foundation, Estate of Edward L. Palmer, Mr. and Mrs. John C. Phelan, the Quick Family, Estate of Elisabeth Sloan Livingston, the Starr Foundation, the Roy J. Zuckerberg Family Foundation, and anonymous donors.

We are also continuously grateful for our endowed lectureships: the John P. and Rita M. Cleary Visiting Lectureship, the George W. Cutting Lectureship, the William Stamps Farish Lectureship, the Martha F. Gerry Visiting Lectureship, the Edward H. Gerry Visiting Lectureship, the Edward H. and Martha F. Gerry Lectureship, the Susan T. and Charles E. Harris Visiting Lectureship, the Klingenstein Lectureship, the Mary D. Lindsay Lectureship, the Pfizer Lectureship, the George B. Rathmann Lectureship, the Seraph Foundation Visiting Lectureship, and the Sigi Ziering Lectureship.

We are also very fortunate to hold a prestigious National Research Service Award Predoctoral Training Grant from the National Institutes of Health, National Institute of General Medical Sciences.

### **Student and Alumni Achievements**

To date, 81 students have received their Ph.D. degrees from the WSBS. Nineteen graduates currently hold tenure-track faculty positions. In 2014, Monica Dus took an assistant professor position at the University of Michigan, Yaniv Erlich joined the New York Genome Center and Columbia University as an assistant professor, and Ira Hall moved to Washington University, St. Louis, as an associate professor. Our graduates have also moved into influential positions in administration, publishing, consulting, and industry. In 2014, Mitchell Bekritsky joined Illumina as a Bioinformatic Scientist, Rebecca Bish joined D.E. Shaw Research as a Senior Scientific Editor, Paloma Guzzardo took a scientist position at Haplogen, Fred Rollins joined LEK Consulting as a consultant, and Susann Weissmueller became an Equity Research Associate at MLV & Co.

The WSBS students continue to impress us with their accomplishments. They publish their research findings in prestigious international journals—more than 300 papers to date—and obtain fellowships to pursue their research interests.

In 2014, our current students and alumni were successful in receiving the following prestigious awards and fellowships:

- WSBS student Yu-Jui (Ray) Ho was awarded a Cutting-Edge Talent and Training Scholarship from the Ministry of Education of Taiwan.
- WSBS student Onyekachi Odoemene was awarded a Ruth L. Kirschstein National Research Service Award Individual Fellowship.
- WSBS graduate Shraddha Pai received a Canadian Institutes of Health Research Postdoctoral Fellowship.
- WSBS graduate Niraj Tolia received a Burroughs Wellcome Fund Investigators in the Pathogenesis of Infectious Disease Award.
- WSBS graduate Michelle Cilia received a Herbert L. Rothbart Outstanding Early Career Research Scientist Award.
- WSBS graduate Ira Hall was promoted to Associate Professor.

Prizes for the best posters by a postdoctoral fellow and by a graduate student were awarded at the Laboratory's annual In-House Symposium held in November 2014. The poster session provides a forum for the postdoctoral fellows and students to show off their research and gives the entire scientific community a chance to come together and share ideas beyond the walls of their individual laboratories. This year, Jack Walleshauser, a Watson School student from Leemor Joshua-Tor's laboratory, won the graduate student prize. The postdoctoral prize was shared between Keerthi Krishnan from Josh Huang's laboratory and Billy Lau from Stephen Shea's laboratory, for their collaborative project.

## 2014 WSBS STUDENT (CURRENT OR PREVIOUS) PUBLICATIONS

- Bolden JE, **Tasdemir N**, Dow LE, van Es JH, Wilkinson JE, Zhao Z, Clevers H, Lowe SW. 2014. Inducible in vivo silencing of Brd4 identifies potential toxicities of sustained BET protein inhibition. *Cell Rep* **8**: 1919–1929.
- Bots M, Verbrugge I, Martin BP, Salmon JM, Ghisi M, Baker A, Stanley K, Shortt J, Ossenkoppele GJ, Zuber J, **Rappaport AR**, Atadja P, Lowe, SW, Johnstone RW. 2014. Differentiation therapy for the treatment of t(8;21) acute myeloid leukemia using histone deacetylase inhibitors. *Blood* **123**: 1341–1352.
- Castel SE**, Ren J, Bhattacharjee S, Chang AY, Sánchez M, Valbuena A, Antequera F, Martienssen RA. 2014. Dicer promotes transcription termination at sites of replication stress to maintain genome stability. *Cell* **159**: 572–583.
- Chen C, Liu Y, **Rappaport AR**, Kitzing T, Schultz N, Zhao Z, Shroff AS, Dickins RA, Vakoc CR, Bradner JE, Stock W, Lebeau MM, Shannon KM, Kogan S, Zuber J, Lowe SW. 2014. MLL3 is a haploinsufficient 7q tumor suppressor in acute myeloid leukemia. *Cancer Cell* **25**: 652–665.
- Cazakoff BN**, Lau BYB, Crump KL, Demmer HS, Shea SD. 2014. Broadly tuned and respiration-independent inhibition in the olfactory bulb of awake mice. *Nat Neurosci* **17**: 569–576.
- Cressy M, Valente D, Altick A, Kockenmesiter E, **Honegger K**, Qin H, Mitra PP, Dubnau J. 2014. Laboratory evolution of adenyl cyclase independent learning in *Drosophila* and missing heritability. *Genes Brain Behav* **13**: 565–577.
- Dow LE, Nasr Z, Saborowski M, **Ebbesen SH**, Machado E, **Tasdemir N**, Lee T, Pelletier J, Lowe SW. 2014. Conditional reverse Tet-transactivator mouse strains for the efficient induction of TRE-regulated transgenes in mice. *PLoS ONE* **9**: e95236.
- Eckersley-Maslin MA**, Thybert D, Bergmann JH, Marioni JC, Flicke P, Spector DL. 2014. Random monoallelic gene expression increases upon embryonic stem cell differentiation. *Dev Cell* **4**: 351–365.
- Eckersley-Maslin MA**, Spector DL. 2014. Random monoallelic expression: Regulating gene expression one allele at a time. *Trends Genet* **30**: 237–244. Review
- Faehnle CR\*, **Walleshauser J\***, Joshua-Tor L. 2014. Mechanism of Dis3l2 substrate recognition in the Lin28-let-7 pathway. *Nature* **514**: 252–256.
- Goh WS**, Seah JW, Harrison EJ, Chen C, Hammell CM, Hannon GJ. 2014. A genome-wide RNAi screen identifies factors required for distinct stages of *C. elegans* piRNA biogenesis. *Genes Dev* **28**: 797–807.
- Ganapathy G, Howard JT, Ward JM, Li J, Li B, Li Y, Xiong Y, Zhang Y, Zhou S, Schwartz DC, Schatz M, **Aboukhalil R**, Fedrigo O, Bukovnik L, Wang T, Wray G, Rasolonjatovo I, Winer R, Knight JR, Koren S, Warren WC, Zhang G, Phillippy AM, Jarvis ED. 2014. High-coverage sequencing and annotated assemblies of the budgerigar genome. *Gigascience* **3**: 11.
- Gerstein MB, Rozowsky J, Yan KK, Wang D, Cheng C, Brown JB, Davis CA, Hillier L, Sisu C, Li JJ, . . . , **Schlesinger F**, et al. 2014. Comparative analysis of the transcriptome across distant species. *Nature* **512**: 445–448.
- Hohmann AF**, Vakoc CR. 2014. A rationale to target the SWI/SNF complex for cancer therapy. *Trends Genet* **30**: 356–363. Review.
- Jacob Y, Bergamin E, Donoghue MT, Mongeon V, LeBlanc C, Voigt P, **Underwood CJ**, Brunzelle JS, Michaels SD, Reinberg D, Couture J-F, Martienssen RA. 2014. Selective methylation of histone H3 variant H3.1 regulates heterochromatin replication. *Science* **343**: 1249–1253.
- Knott SR, Maceli AR, Erard N, Chang K, Marran K, Zhou X, Gordon A, El Demerdash O, **Wagenblast E**, Kim S, Fellmann C, Hannon GJ. 2014. A computational algorithm to predict shRNA potency. *Mol Cell* **56**: 796–807.
- Meikar O, Vagin VV, Chalmel F, Sostar K, Lardenois A, Hammell M, Jin Y, Da Ros M, **Wasik KA**, Toppari J, . . . , **Scuoppo C**, et al. 2014. PTEN action in leukaemia dictated by the tissue microenvironment. *Nature* **509**: 402–406.
- Odoemene O**, Churchland AK. 2014. Listening for the right sounds. *Neuron* **82**: 944–945. Review.
- Oyibo HK**, **Znamenskiy P**, Oviedo HV, Enquist LW, Zador AM. 2014. Long-term cre-mediated retrograde tagging of neurons using a novel recombinant pseudorabies virus. *Front Neuroanat* **8**: 86.
- Peikon ID**, Gizatullina DI, Zador AM. 2014. In vivo generation of DNA sequence diversity for cellular barcoding. *Nucleic Acids Res* **42**: e127.
- Rudalska R, Dauch D, Longrich T, **McJunkin K**, Wuestefeld T, Kang TW, Hohmeyer Rechavi O, Hourri-Ze'evi L, Anava S, **Goh WS**, et al. 2014. Starvation-induced transgenerational inheritance of small RNAs in *C. elegans*. *Cell* **158**: 277–287.
- Sanders JI**, Kepecs A. 2014. A low-cost programmable pulse generator for physiology and behavior. *Front Neuroeng* **7**: 43.
- Sun J, Ramos A, Chapman B, Johnnidis JB, Le L, **Ho YJ**, Klein A, Hofmann O, Camargo FD. 2014. Clonal dynamics of native haematopoiesis. *Nature* **514**: 322–327.
- Tschaharganeh DF, Xue W, Calvisi DF, Evert M, Michurina TV, Dow LE, Banito A, Katz SF, Kastenhuber ER, **Weissmueller S**, et al. 2014. p53-Dependent nestin regulation links tumor suppression to cellular plasticity in liver cancer. *Cell* **158**: 579–592.
- Wang M, **Perova Z**, Arenkiel BR, Li B. 2014. Synaptic modifications in the medial prefrontal cortex in susceptibility and resilience to stress. *J Neurosci* **34**: 7485–7492.
- Weissmueller S\***, Machado E\*, Saborowski M, Morris JP 4th, **Wagenblast E**, Davis CA, Moon SH, Pfister NT, Tschaharganeh DF, Kitzing T, et al. 2014. Mutant p53 drives pancreatic cancer metastasis through cell-autonomous PDGF receptor  $\beta$  signaling. *Cell* **157**: 382–394.

\* Authors contributed equally to the work.

Watson School student is designated in boldface.

**WSBS GRADUATES IN FACULTY AND SENIOR POSITIONS**

Name	Current position
Rebecca Bish	Senior Scientific Editor, D.E. Shaw Research
François Bolduc	Assistant Professor, University of Alberta, Canada
Darren Burgess	Associate Editor, <i>Nature Reviews Cancer</i> and <i>Nature Reviews Genetics</i> , United Kingdom
Amy Caudy	Assistant Professor, University of Toronto, Canada
Daniel Chitwood	Assistant Professor, Donald Danforth Plant Science Center, St. Louis
Michelle Cilia	Research Molecular Biologist, U.S. Department of Agriculture, The Boyce Thompson Institute; Adjunct Assistant Professor, Cornell University
Catherine (Cormier) Seiler	Scientific Liaison, Arizona State University
Monica Dus	Assistant Professor, University of Michigan
Yaniv Erlich	Assistant Professor, N.Y. Genome Center, Columbia University
Rebecca Ewald	Manager, Business Development, Ventana Medical Systems
Elena Ezhkova	Assistant Professor, Mount Sinai School of Medicine
Patrick Finigan	Associate II, Regulatory Affairs CMC, Genzyme
Ira Hall	Associate Professor, Washington University School of Medicine, St. Louis
Christopher Harvey	Assistant Professor, Harvard University
Tomáš Hromádka	Assistant Professor, Slovak Academy of Sciences
Keisha John	Director of Diversity Programs, Office of Graduate and Postdoctoral Affairs at University of Virginia
Zachary Lippman	Associate Professor, Cold Spring Harbor Laboratory
Marco Mangone	Assistant Professor, Arizona State University
Masafumi Muratani	Associate Professor, University of Tsukuba, Japan
Elizabeth Murchison	Reader, Cambridge University; Fellow, Kings College, United Kingdom
Patrick Paddison	Assistant Member, Fred Hutchinson Cancer Research Center
Emiliano Rial-Verde	Associate Principal, McKinsey & Co., Inc., Geneva, Switzerland
Frederick Rollins	Consultant, LEK Consulting, Boston, Massachusetts
Ji-Joon Song	Associate Professor, Korea Advanced Institute of Science and Technology (KAIST), South Korea
Niraj Tolia	Assistant Professor, School of Medicine, Washington University, St. Louis
Wei Wei	Assistant Professor, University of Chicago
Susann Weissmueller	Equity Research Associate, MLV & Co., Connecticut
Jeremy Wilusz	Assistant Professor, University of Pennsylvania

**Alexander Gann**  
*WSBS Professor and Dean*

# SPRING CURRICULUM

---

## Topics in Biology

---

ARRANGED BY Alyson Kass-Eisler and Jan A. Witkowski

Each year, invited instructors offer 7-day courses at the Banbury Conference Center to explore 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. In Spring 2014 there were two such courses: Immunology and Microbial Pathogenesis.

## Immunology

---

April 27–May 3 Attended by the entering classes of 2011 and 2012

INSTRUCTOR Hidde Ploegh, Whitehead Institute

VISITING LECTURERS Yasmine Belkaid, National Institutes of Health

TEACHING FELLOWS  
Vinidhra Mani, Harvard University  
Marisella Panduro Sicheva, Harvard University  
Matthew Woodruff, Harvard University



(Left to right; seated) Kachi Odoemene, Hidde Ploegh. (Left to right; standing) Tyler Garvin, Abram Santana, Giorgia Battistoni, Fred Marbach, Charles Underwood, Robert Wysocki, Paul Masset, Justus Kebschull, Yasmine Belkaid, Robert Aboukhalil, Vinidhra Mani, Annabel Romero Hernandez, Brittany Cazakoff, Anja Hohman, Matt Woodruff, Talitha Forcier, Marisella Panduro Sicheva, Joaquina Delas Vives

Immunology focuses on understanding the mechanisms by which multicellular organisms defend themselves against external threats of microbial aggression and internal threats associated with genetic instability and cellular transformation. The course focused on the innate immune system and the adaptive immune system. Innate immunity defends against microbes by recognizing evolutionarily conserved molecular patterns; the adaptive immune system has enormous flexibility in molecular recognition, but it can also target self to cause autoimmune diseases.

## Microbial Pathogenesis

---

May 4–10

Attended by the entering classes of 2010 and 2013

### INSTRUCTORS

Stanley Maloy, Center for Microbial Sciences, San Diego  
Ronald K. Taylor, Dartmouth Medical School

### VISITING LECTURERS

Nancy Freitag, University of Illinois, Chicago  
Karla Satchell, Northwestern University Medical School  
Michele Swanson, University of Michigan Medical School

This course focused on the mechanisms of microbial pathogenesis and the host response and the scientific approaches that are used to investigate these mechanisms. How do microbes adhere to host cells? How do environmental cues direct the response of microbial pathogens? How do microbial pathogens modulate host cells to expedite virulence? How do host cells respond to microbial pathogens? How does the host immune system react to microbial pathogens? What does genomics tell us about how microbial pathogens evolve? How do emerging pathogens take advantage of new ecological niches? Although there are numerous microbial pathogens, many use similar approaches to solve these common problems.



(Left to right) Ronald Taylor, Jack Walleshauser, Anja Hohmann, Stanley Maloy, Arkarup Bandyopadhyay, Giorgia Battistoni, Sanchari Ghosh, Michael Giangrasso, Maria Nattestad, Lital Chartarirfsky, Georgi Yordanov, Daniel Kepple, Laura Maiorino, Matt Koh, Colleen Carlston, Lisa Krug, Michael Gutbrod

## Teaching Experience at the Dolan DNA Learning Center

---

DIRECTOR	David A. Micklos
INSTRUCTORS	Amanda McBrien (Lead) Jennifer Galasso Elna Gottlieb Erin McKechnie Bruce Nash Brooke Roeper Jermel Watkins

Science has an increasing role in society, and there is an increasing need for biologists to educate nonscientists of all ages about biology. The Watson School of Biological Sciences doctoral program offers its students the opportunity to teach in the Laboratory's Dolan DNA Learning Center, where laboratory courses are taught to high school and middle school students. In so doing, they learn how to communicate with nonbiologists and to inspire and educate creative young minds. The teaching module entailed pairs of students teaching one morning or afternoon a week for 12 weeks. In the initial weeks, the Dolan DNA Learning Center instructors taught the Watson School students the didactic process; it was not until the fifth week that the graduate students taught on their own. At the end of the 12 weeks, the students were very excited about their teaching experience.

## Laboratory Rotations

---

The most important element of a doctoral education is learning to perform independent research. After the fall term courses, students participate in laboratory rotations. These rotations provide 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 make short presentations of their studies to the other students and their rotation advisors. These talks give students an opportunity to share their laboratory experiences and to practice giving scientific presentations. This year, 20 WSBS faculty members served as rotation mentors, some mentoring more than one student.

ROTATION MENTORS	Mickey Atwal Anne Churchland Mikala Egeblad Hiro Furukawa Jesse Gillis Thomas Gingeras Christopher Hammell	Adam Kepecs Alexei Koulakov Bo Li Robert Martienssen Michael Schatz Stephen Shea David Spector	Glenn Turner David Tuveson Christopher Vakoc Linda Van Aelst Michael Wigler Anthony Zador
------------------	--	--	--



# FALL CURRICULUM

---

The students began the semester by attending boot camps in Molecular, Cellular, and Quantitative Biology. These boot camps introduced the students to the techniques and terminology that they will encounter in subsequent courses. The Molecular and Cellular Biology boot camp featured seven lectures from faculty members Hiro Furukawa, Dick McCombie, Darryl Pappin, David Spector, Lloyd Trotman, Christopher Vakoc, and Linda Van Aelst. The Quantitative Biology boot camp lectures were given by Mickey Atwal, Justin Kinney, and Michael Schatz.

## CORE COURSES

### The Leslie C. Quick, Jr. Core Course on Scientific Reasoning and Logic

---

INSTRUCTORS	Linda Van Aelst (Lead) Christopher Hammell Leemor Joshua-Tor	Alexander Gann Bo Li David Tuveson
-------------	--	--

GUEST LECTURERS	Douglas Fearon Hiro Furukawa Justin Kinney Adrian Krainer Robert Martienssen
-----------------	--

This core course forms the heart of the curriculum, at which students (1) acquire a broad base of knowledge about the biological sciences, (2) learn the scientific method, and (3) learn how to think critically. The initial four to five modules are on a different general theme. For each module, students read an assigned set of research articles, and at the end of the module, they provide written answers to a problem set that guides them through several of the articles.

Twice weekly, students attend lectures related to the module's topic that include concepts and fundamental information as well as experimental methods. The students meet among themselves to discuss the assigned papers not covered by the problem set. Each week, students spend an evening discussing the assigned articles with faculty. In the final module of the course, students participate in a mock study-section in which real National Institutes of Health R01 grants are reviewed and critiqued. This allows students to evaluate the questions before the answers are known, assess routes toward discovery before knowing where they will end, and make critical judgments about how to proceed in the face of an uncertain outcome.

The following were the module topics for this course.

Topic	Instructor(s)
Mechanism and Structure of Gene Regulation	Alex Gann Leemor Joshua-Tor
Gene Regulatory Logic and the Construction of Multicellular Organisms: Insights from Flies, Plants, and Worms	Christopher Hammell
The Brain: Wiring, Plasticity, and Maladaptation	Bo Li
Cancer	David Tuveson
Study Section	Linda Van Aelst

## The Darrell Core Course on Scientific Exposition and Ethics

---

INSTRUCTORS	David J. Stewart (Lead) Mikala Egeblad Arne Stenlund
GUEST LECTURERS	Lisa Bianco Carrie Cowan Diane Esposito Alyson Kass-Eisler
VISITING LECTURERS	Olga Akselrod, The Innocence Project Michael Bielski, The Center for Biotechnology Keith Baggerly, M.D. Anderson Cancer Center Robert Charrow, Greenberg Traurig, LLC Stephen Hall, New York University Avner Hershlag, North Shore University Hospital Dennis Meredith, Research Communication Consultant Elizabeth Phimister, New England Journal of Medicine Yvette Seger, FASEB Tung-Tien Sun, New York University

The 2014 Scientific Exposition and Ethics (SEE) core course was led by David Stewart, Arne Stenlund, and Mikala Egeblad. The course hosted expert guest lecturers who covered topics including scientific presentations, the editorial review process, the ethical and legal responsibilities of scientists, DNA profiling and postconviction appeals, and reproductive genetics.

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 on 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

---

ORGANIZERS	Kimberley Geer Alyson Kass-Eisler
------------	--------------------------------------

This core course provided students with an in-depth introduction to the fields of research that the Laboratory scientists investigate. Students and faculty attended a weekly Research Topics seminar, at which faculty members and CSHL fellows presented their current research topics and methods of investigation each Wednesday evening over dinner. The students learned how to approach important problems in biology. These seminars, together with the annual fall Laboratory In-House Symposium, provided students with a basis for selecting laboratories in which to do rotations.

## Specialized Disciplines Courses

---

The students in the Entering Class of 2014 took a total of four Specialized Disciplines courses this fall: *Genetics and Genomics*, *Cellular Structure and Function*, *Systems Neuroscience*, and *Quantitative Biology*.

### Genetics and Genomics

---

INSTRUCTORS	Zachary Lippman (Lead) Michael Schatz
CSHL GUEST LECTURERS	Jesse Gillis David Jackson Gholson Lyon
EXTERNAL GUEST LECTURER	Kirsten Bomblied, Harvard University

This course placed modern genetics and genomics into the context of classical genetics. History, technique, and perspective of genetic inference were described around four levels of analysis: forward genetics, natural genetic variants, gene interaction, and genomics. Emphasis was on integrating classical with modern questions of genetic analysis: How are genes mapped and “cloned”? How do gene mutations help to define biological processes? How are more complex traits genetically dissected into simpler (underlying) components? What concepts and techniques are used to organize genes into pathways and networks? What defines a gene and what gene variation exists in natural populations? What are the functional consequences of gene variation, and how is it detected? How are genomes organized and coordinately regulated? How can genomic information be catalogued, organized, and mined? These questions and concepts were fleshed out using examples from the literature.

### Cellular Structure and Function

---

INSTRUCTORS	Carrie Cowan Linda Van Aelst
CSHL GUEST LECTURERS	Raffaella Sordella David Spector Bruce Stillman
EXTERNAL GUEST LECTURERS	Francesca Bartolini, Columbia University Aron Jaffe, Novartis Alexey Khodjakov, Wadsworth Laboratory Sandy Simon, The Rockefeller University Orion Weiner, University of California, San Francisco

With the complete genome sequence available for many organisms, it becomes imperative to understand the function of the gene products. This understanding requires an increasing appreciation of the structure and function of the cell, as well as dynamic associations within the cell. This

course provided a basic overview of the structural and functional organization of cells with particular emphasis on cellular compartmentalization and communication. Topics of focus included the cytoskeleton, cell adhesion and signaling, membrane transport, gene expression, and nuclear organization. In addition, the course provided insight into the basic toolbox of the cell biologist of the 21st century.

## Systems Neuroscience

---

**INSTRUCTORS** Adam Kepecs (Lead)  
Glenn Turner

**GUEST INSTRUCTORS** Anne Churchland  
Alexei Koulakov  
Stephen Shea  
Anthony Zador

This course introduced students to 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. The course started with the basics of electrical signaling in neurons: ion channels, action potentials, and synaptic transmission. The cellular basis of learning, including Hebb's postulates, and LTP (long-term synaptic potentiation), were discussed. The course explored the consequences of synaptic learning rules by examining how experience shapes the wiring of the nervous system during development and investigated how such building blocks translate into whole-organism behavior. The course then examined classical conditioning and asked how changes in synaptic transmission could underlie such behavior. Computational associative learning models of the learning process were also discussed.

From behaviors that focus on simple memories, the course turned to behaviors that require making perceptual decisions. This was accomplished by covering some basic concepts of perceptual neuroscience, such as neuronal "receptive fields," and using these to discuss current results and models of perceptual decision making. Finally, the course turned to the learning of behaviors through reward and punishment, what is known as reinforcement learning. The course concluded with a discussion of the role of dopamine in reward and learning, the theory of reinforcement learning, and pathologies of reward-seeking behavior.

## Quantitative Biology

---

**INSTRUCTORS** Mickey Atwal (Lead)  
Michael Schatz

**GUEST LECTURERS** Justin Kinney  
Alexander Krasnitz

With the advent of high-throughput technologies in biology, it has become necessary for biological researchers to be able to analyze and interpret high-dimensional data and frame hypotheses mathematically. To this end, this course aimed to equip students with a basic training in modern statistical methods. By the end of the course, students not only were able to answer many of the simple statistical questions that arise in data analyses, but also became familiar with the more

complex techniques used by computational biologists. Topics covered included error fluctuations, calculating the significance of an experimental result, Bayesian inference, information theory, power calculations, dimensional reduction, and DNA sequence analyses. In addition, this course introduced mathematical modeling, motivated by classic examples in quantitative biology such as the Delbrück–Luria experiment, Hopfield’s kinetic proofreading, and Kimura’s neutral theory of population genetics.

# POSTDOCTORAL PROGRAM

---

PROGRAM DIRECTOR

Nicholas Tonks

PROGRAM ADMINISTRATOR

Alyson Kass-Eisler

Cold Spring Harbor Laboratory (CSHL) is proud of its rich tradition in postdoctoral education and has a deep commitment to the postdoctoral experience. CSHL has long been recognized as a place for nurturing young scientists, with postdoctoral researchers being an integral part of the discovery process. Current fields of research expertise include genetics; molecular, cellular, and structural biology; neuroscience; cancer; plant biology; genomics; and quantitative biology. The postdoctoral community at CSHL is diverse and international, arriving at the Laboratory with many different backgrounds and with needs that change during the course of the postdoctoral years. The Postdoctoral Program Office works closely with the fellows and the Laboratory's administration to coordinate and organize educational and career development activities. Alyson Kass-Eisler, the Postdoctoral Program Officer, and Nicholas Tonks, the Scientific Director of the Postdoctoral Program, provide the day-to-day support for the postdoctoral program, which is overseen by the Dean of the Watson School of Biological Sciences (WSBS), Alexander Gann.

The Postdoctoral Liaison Committee (PDLC)—an elected group of postdocs who help share information and ideas between the administration and the postdoctoral community—continues to enhance the postdoctoral experience at CSHL. The PDLC is the voice of the community and has regular meetings with CSHL President Bruce Stillman. In 2014, new members Sara Ballouz, Daniel Ferrante, and Serif Senturk joined returning members Jonathan Ipsaro, Saikat Nandi, and Benjamin Roche. The PDLC organized two successful retreats this year aimed at fostering networking and collaboration. The day-long retreats took place at the Banbury Conference Center and included research talks, chalk talks, and discussions sessions. The PDLC also oversees and distributes funds to two career development groups as described below. These groups are primarily composed of postdoctoral fellows but also include graduate students. Today's postdocs face a number of challenges, including a very difficult job market. CSHL works hard to give its postdocs the upper hand when it comes to competing for the scarce jobs that are available. More and more, it has become our responsibility to introduce a diversity of career opportunities that are available and to provide the tools to our postdocs to prepare for these positions. Thanks to the assistance from PDLC and these career development groups, a number of fantastic events took place.

**The Career Development Program (CDP)** provides programming geared toward careers in academia. In the "Conversations with Faculty" series, the postdocs were given career insights in an informal and interpersonal format. Sessions included Gaining Teaching Experience with guest panelists Cristina Fernandez Marco, Jaclyn Novatt, and Lisa Prazak; Securing a Faculty Position at a Teaching-Centered Institution with Allison Abbott (Associate Professor, Marquette University); and The Two-Body Problem: The Difficulty Academic Couples Find Obtaining Jobs at the Same University or within a Reasonable Commuting Distance from Each Other with Jennifer Morgan (Associate Scientist, Eugene Bell Center for Regenerative Biology and Tissue Engineering at the Marine Biological Laboratory).

The CDP also hosted a special discussion with Dr. Richard Nakamura, Director of the National Institutes of Health's (NIH) Center for Scientific Review. Dr. Nakamura talked about his career trajectory, his experiences within and outside of the NIH, and his opinions on the future of academic scientific research.

We have also been working on building connections with local colleges and universities to provide teaching experience opportunities for interested postdocs. Informal partnerships are now in place with Molloy College and Long Island University. In addition, selected postdoctoral fellows continue to participate in the Watson School as tutors—either one-on-one or in the classroom setting—thus providing the postdocs with valuable experience in teaching and mentoring.

**The Bioscience Enterprise Club (BEC)** provides information for nonacademic careers and has also been organizing a series of seminars and workshops. The areas of interest include biotechnology, intellectual property, law, regulatory affairs, and venture capitalism. This year's guests included Victoria Aranda, Senior Editor, *Nature Medicine*; Jim Hayward, CEO, Applied DNA; Yvette Seger, Director of Science Policy, Federation of American Societies for Experimental Biology; and Rebecca Bish, Senior Scientific Editor at D.E. Shaw Research.

A discussion of Careers in Medical Genetics was paneled by Dr. David Tegay (New York Institute of Technology), Dr. Bhuma Krishnamachari (New York Institute of Technology), and Dr. Ann-Leslie Zaslav (Stony Brook University), and Jeff Nagel hosted a workshop on management and leadership skills needed to run a lab. BEC also organized a trip to the local biopharmaceutical company Regeneron. In collaboration with the CSHL iPlant initiative and the Software Carpentry, they organized a 2-day bioinformatics boot camp for postdocs at CSHL. BEC and CDP also co-hosted a visit from Dr. Latasha Wright, Staff Scientist and Director of Development at BioBus, a "high-tech laboratory on wheels."

To celebrate National Postdoc Appreciation week, an ice cream social was held on the Delbrück patio. It was a great opportunity for the community to get together for some fun, as well as to network and find out about the ongoing programs we have been conducting. The PDLC also organized an inaugural CSHL Music Festival, featuring some of our very musically talented postdocs.

Through the CSHL Human Resources department, the postdocs have had access to free job-seeker webinars hosted by the Higher Education Research Consortium. The 2014 webinars included "Making Your Case: Writing Cover Letters for Non-Academic Jobs" and "Jobs in the Federal Government."

"Demystifying Science," which began in 2011, enlists postdoctoral fellows to bring their expertise and teaching interests to the educationally diverse staff of the Laboratory. In the 2014 sessions, geared toward a general audience, postdocs have talked about Fighting Cancer with our Own Army, Decoding the Noncoding Genome, and Ebola: Fact and Fiction.

On September 18, 2014, the Rally for Medical Research, a broad coalition of groups from the medical research advocacy community, met with House and Senate offices in Washington, D.C., to urge Congress to make funding for the NIH a national priority. Cristina Aguirre-Chen, a postdoc in Christopher Hammell's lab, represented the CSHL postdoctoral community at this important event.

Since 2003, all postdoctoral fellows and graduate students at the Laboratory have been enrolled 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 that have formed a partnership with NYAS. The aim of the Alliance is to provide career and professional development monitoring for postdoctoral fellows and graduate students in science and engineering, through a series of live events and a dedicated web portal. In addition, the Science Alliance gives graduate students and postdoctoral fellows the opportunity to network with their peers across institutions and with key leaders in academia and industry. This year, the Science Alliance meetings and workshops included Perspectives in STEM, an Evening with Dr. Laurie Leshin; Making the Leap: A Non-Academic Career Planning and Job Search Boot Camp; Supporting Dynamic STEM Careers for Underrepresented Minorities; Building a Collaborative International Network; Scientists Teaching Science; Graduate Student and Postdoc Career Symposium at the

NewYorkBio Annual Conference; Survival Skills for Young Investigators; Journey Through Science Day; Networking and IDPs Workshop; Risky Business: A Pharmaceutical Industry Strategy Workshop; How to Turn Your Science Project into a Product: A Creative Thinking Workshop by The Solution Lab; From Scientist to CSO: Experiencing the Scientific Method as Your Guide to Career Success; Grantsmanship for Graduate Students and Postdocs; But I Have No Skills: Debunking Myths and Exploring Career Options for Ph.D.s; The Craft of Scientific Presentations; The Non-Academic Job Search: Target the Job, Tailor the Approach; Interpersonal Skills: The Keys to Effective Communication and Managing Difficult Conversations; Management Consulting: Using Your Scientific Mind to Solve Business Problems; and An Introduction to Teaching Science Online.

Additionally, Dr. Thomas Magaldi, then the Director of Science Alliance at the NYAS, visited CSHL to give a talk on Preparing for Careers in Science Policy. During the talk, he discussed a variety of science policy careers, what types of fellowships help in exploring those careers, and what can be done during graduate school and/or postdoc to increase the chances of transitioning into the world of policy.

On November 6, the Office of Sponsored Programs, the Development Department and the Postdoc Program Office held a very popular course on grant writing. The sessions included Introduction to Grants and External Funding; The Grant Application; The Grant Abstract/Public Summary; Peer Review Panel Discussion; Research Compliance; Fellowships, Career Development Awards and Resources for Postdocs; and Effective Writing Skills.

Finally, an important measure of our postdoctoral program's success is the ability of postdoctoral fellows to secure positions at the end of their training. In 2014, our fellows took positions at Aarhus University, Adelphi University, Donald Danforth Plant Science Center, Dow AgroSciences, Farmingdale State College, Pfizer, Roche, the University of Basel, the University of Bayreuth, and the University of Michigan.



# UNDERGRADUATE RESEARCH PROGRAM

---

PROGRAM DIRECTORS

Anne Churchland  
Michael Schatz

PROGRAM ADMINISTRATOR

Kimberly Creteur

Established more than 50 years ago, the CSHL Undergraduate Research Program (URP) plays a major role in providing and setting the standard for meaningful hands-on undergraduate research training in biology. The 10-week program begins the first week of June. During the first week, the students attend various orientations and receive a guided historical tour of campus and all the facilities and resources available to them. This ensures a smooth transition into the Laboratory community and research. The URPs work, live, eat, and play among CSHL scientists, and they have a very busy academic and social calendar throughout the remaining nine weeks of the summer. 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 2014 activities included a pizza party with Dr. and Mrs. Watson, dinner with Dr. and Mrs. Stillman, volleyball games, designing the URP T-shirt, competing in the annual Scavenger Hunt, and the ever-famous URP versus PI volleyball match.

The students' scientific development is the most important component of the program. Therefore, in the beginning of the summer, the URPs write an abstract and present a talk on their proposed research. The URP participants work alongside scientists and become increasingly independent throughout the summer. Concluding the program in August, the URP students prepare a final report and present 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.



2014 Undergraduate Research Program Participants

The following 27 students, selected from 819 applicants, took part in the 2014 program:

**Henry Ashworth**, Eckerd College

Advisor: Mikala Egeblad

Funding: Robert H.P. Olney Fellowship

The mystery of lysyl oxidase in pancreatic cancer.

**Patricia Aubel**, San Jose State University

Advisor: W. Richard McCombie

Funding: National Science Foundation

Variant detection with PacBio SMRT sequencing system.

**Syndi Barish**, The College of New Jersey

Advisor: Gholson Lyon

Funding: National Science Foundation

Creation and characterization of an isogenic knockout in Naa50, a catalytic component of amino-terminal acetyltransferase (NAT) A and E in *Saccharomyces cerevisiae*.

**Nikaela Bryan**, University of Maryland, Baltimore County

Advisor: Anne Churchland

Funding: HHMIEXROP

Optogenetic interrogation of mouse posterior parietal cortex during perceptual decision-making.

**Cassandra Burdziak**, Rutgers University, New Brunswick

Advisor: Thomas Gingeras

Funding: National Science Foundation

Characterization of cell-specific fragmenting patterns among exosomal small RNAs.

**Daniel Burkhardt**, University of Massachusetts, Amherst

Advisor: Doreen Ware

Funding: National Science Foundation

Searching for SNPs in stay-green sorghum.

**John Cannon**, Carleton College

Advisor: Anne Churchland

Funding: Burroughs Wellcome Fellowship

Optogenetic approaches to studying perceptual decision-making in the posterior parietal cortex.

**John Simon Chow**, Georgia Institute of Technology

Advisor: Alexander Krasnitz

Funding: National Science Foundation

Convex optimization algorithms for population structure analysis in tumors.

**Michael Dinh**, University of Notre Dame

Advisor: Stephen Shea

Funding: University of Notre Dame URP Scholarship

Olfactory modulation of the auditory cortex by medial amygdala.

**Luz Brielle Dojer**, Boston University

Advisor: Alea Mills

Funding: National Science Foundation

The role of chromodomain helicase DNA-binding protein 5 in neural stem cells.

**Leila Elabbady**, Wellesley College

Advisor: Josh Dubnau

Funding: Alfred L. Goldberg Fellowship

The transposon storm hypothesis of neurodegeneration.

**Carolina Falcon-Campos**, National Polytechnic

Institute, Mexico

Advisor: David Jackson

Funding: Von Stade Fellowship

Identification of novel regulators of cell-to-cell trafficking via plasmodesmata in *Arabidopsis thaliana*.

**Ariel Gewirtz**, Swarthmore College

Advisor: Mickey Atwal

Funding: H. Bentley Glass Fellowship

Ectopic germline gene expression in glioblastoma multiforme and breast cancer.

**Michael Gross**, Cornell University

Advisor: Florin Albeanu

Funding: William Shakespeare Fellowship

Behavioral effects of corticobulbar feedback manipulation in mice.

**Melina-Theoni Gyparaki**, The University of Edinburgh

Advisor: Marja Timmermans

Funding: William Townsend Porter Foundation Scholarship

Functional analysis of small RNA-ARGONAUTE associations and their roles in plant development.

**Margaret Henderson**, Cornell University

Advisor: Partha Mitra

Funding: Joan Redmond Read Fellowship

Improving the precision of stereotactic injections for mapping the mouse brain.

**Samuel Johnson**, Brown University

Advisor: Alexei Koulakov

Funding: Emanuel Ax Fellowship

PCR primer design for mouse olfactory receptors.

**Danxun Li**, University of California, Berkeley

Advisor: Bo Li

Funding: 30th Anniversary URP Scholarship

Decoding reward learning and valuation behavior in cell populations in the globus pallidus.

**Vicki Mercado**, Whittier College

Advisor: David Tuveson

Funding: Garfield Fellowship

Determining the sensitivity of pancreatic cancer cells to endoplasmic reticulum stress.

**Beverly Mok**, University of Cambridge

Advisor: Christopher Vakoc

Funding: **James D. Watson Undergraduate Scholarship**  
Role of mediator complex in AML maintenance.

**Mira Nancheva**, Stanford University

Advisor: **Adam Kepecs**

Funding: **Dorcas Cummings Scholarship**

Optogenetic manipulation of orbitofrontal-ventrostriatal pathway during decision making in rats.

**Lucy Rummler**, Clemson University

Advisor: **Zachary Lippman**

Funding: **National Science Foundation**

Meristem regulation and the *fin* and *fan* mutants in tomato.

**Michael Sayegh**, Harvard College

Advisor: **Bruce Stillman**

Funding: **Former URP Fund Scholarship**

Gene regulation via RB and ORC1 interaction.

**Selin Schamiloglu**, Columbia University

Advisor: **Z. Josh Huang**

Funding: **James D. Watson Fellowship**

Investigating the role of chandelier cells in fear circuitry.

**Rachel Sherman**, Harvey Mudd College

Advisor: **Michael Schatz**

Funding: **National Science Foundation**

Whole-genome assembly and alignment pipeline for unique gene discovery

**Toby Turney**,

University of Notre Dame

Advisor: **Darryl Pappin**

Funding: **Howard Hughes Medical Institute Scholarship**

Improving the yield and purity in a large-scale expression and purification of velocin-N.

**Victoria Wang**, University of Cambridge

Advisor: **Lloyd Trotman**

Funding: **Libby Fellowship**

CRISPR/Cas9 as a genome-editing tool to investigate metastatic prostate cancer.

# PARTNERS FOR THE FUTURE

---

PROGRAM DIRECTOR                      David Jackson

PROGRAM ADMINISTRATOR            Michelle Krull

The Partners for the Future Program, established by Dr. James Watson in 1990, provides an opportunity for talented Long Island high school students to have hands-on experience in biomedical research at Cold Spring Harbor Laboratory. This highly competitive program is open to Long Island high school students in their junior year. Each high school science chairperson may nominate up to three 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 lab. 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 2014–2015 Partners for the Future were chosen from among 54 nominations. They are:

Partner	High School	Laboratory	CSHL Mentor
Luke Antolin	Cold Spring Harbor High School	Adam Kepecs	Hyun Jae Pi
Michael Callahan	Farmingdale High School	Robert Martienssen	Rowan Herridge
Henry Cheung	Seaford High School	Mikala Egeblad	Mario Shields
Melissa Cipolla	Friends Academy	Molly Hammell	Oliver Tam
Noah Davis	Earl L. Vandermeulen High School	Gholson Lyon	Gholson Lyon
Harshil Garg	Bethpage High School	Michael Wigler	Swagatam Mukhopadhyay
Chi Ying “Jillian” Ho	The Stony Brook School	Gholson Lyon	Gholson Lyon
Hannah Johnston	Harborfields High School	Adam Kepecs	Sachin Ranade
Annet Kuruville	East Meadow High School	Adrian Krainer	Tomoki Nomakuchi
Shreeya Panigrahi	Sayville High School	Stephen Shea	Brittany Cazakoff and Dennis Eckmeier
Michelle Shen	Syosset High School	David Jackson	Huyen Bui
Rachel Wesley	Oyster Bay High School	Doreen Ware	Lifang Zhang
Seungmin Yi	The Wheatley School	Mikala Egeblad	Ana Margarida Santos De Almeida



Partners for the Future participants





MEETINGS & COURSES  
PROGRAM



## ACADEMIC AFFAIRS

---

The Cold Spring Harbor Laboratory meetings and courses program attracts scientists from all over the world to communicate, learn, and critique the latest ideas and approaches in the biological sciences. The program consists of advanced laboratory and lecture courses, as well as large meetings and biotechnology conferences that are held almost year-round. The meetings and courses program at the Laboratory attracted strong attendance with almost 7100 meeting participants and almost 1400 course participants (trainees, teaching, and support faculty). The Cold Spring Harbor Asia program including 21 conferences and one summer school has to date attracted more than 3500 participants, bringing the anticipated year-end total for both the United States- and China-based programs to more than 12,000.

The Laboratory held 27 academic meetings this year, which brought together scientists from around the world to discuss their latest research. The spring meeting season culminated in the annual Cold Spring Harbor Symposium, which focused on cognition. The timing of this topic coincided with President Obama's BRAIN initiative to map the workings of the human brain. The Symposium spanned a broad range of approaches toward increased understanding of cognition in humans and other animals including molecular, physiological imaging, and behavioral approaches. Fundamental biological discoveries were balanced with applications relevant to societal well-being including exploration of brain disorders. The 79th Symposium in this historic series attracted around 300 participants, including many notable neuroscientists such as David Anderson, Richard Axel, Cori Bargmann (organizer), Daphne Bavelier (organizer), Patricia Churchland, Karl Deisseroth, Robert Desimone, Allison Doupe, Ann Graybiel, Thomas Jessell, Nancy Kanwisher, David Kleinfeld, Patricia Kuhl, Joseph LeDoux, Eve Marder, Markus Meister, J. Anthony Movshon, Eric Nestler, John Reynolds, Terry Sejnowski (organizer), Michael Stryker, Susumu Tonegawa, and Daniel Wolpert. Dissemination includes the proceedings of the Symposium





published each year by the CSHL Press, and videotaped interviews with leading speakers by editors and journalists attending the Symposium, now available on our Leading Strand YouTube. The Symposium therefore reaches a much wider audience nationally and internationally than can possibly attend.

Cold Spring Harbor Laboratory meetings are unique in assembling a program with very few invited speakers. Instead, meeting organizers are encouraged to select talks from abstracts that are submitted on the basis of scientific merit, ensuring that the latest findings will be presented and that young scientists will have the chance to describe their work. 2014 saw the introduction of several successful new meetings on Avian Model Systems, The PARP Family, and Biological Data Science. The annual meeting on The Biology of Genomes was oversubscribed as usual, attracting almost 100 online subscriptions and more than on Twitter, and many others attracted strong attendance despite the funding climate. Many of these meetings have become essential for those in the field and are held on a biannual basis. Two special meetings convened under the auspices of the CSHL Genentech Center. Conferences on the History of Molecular Biology and Biotechnology addressed the history of messenger RNA and plasmids, respectively, and each attracted rather different audiences than our regular meetings. Partial support for individual meetings is provided by grants from the National Institutes of Health, National Science Foundation, foundations, and companies. Core support for the meetings program is provided by the Laboratory's Corporate Sponsor Program.

The Courses program covers a diverse array of topics in molecular biology, neurobiology, structural studies, and bioinformatics. Instructors update their course curricula annually, invite new speakers who bring a fresh perspective, and introduce new techniques and experimental approaches based on student feedback and progress in the field. Recently introduced courses on Single-Cell Analysis and Synthetic Biology run in parallel with well-established courses in genetics, neuroscience, plant biology, and bioinformatics. New techniques, for example, genome editing using tools such as CRISPR, are introduced as methodologies develop and evolve. We strongly encourage each course to include the latest technical and conceptual developments in their respective fields.

Instructors, lecturers, and assistants come from universities, medical schools, research institutes, and companies around the world to teach at Cold Spring Harbor. Students include advanced graduate students, postdoctoral trainees, principal investigators, and senior scientists from around the world.

A new 1-week lecture course on Genetics and Neurobiology of Language addressed core questions about the bases and origins of speech and language, through talks, interactive sessions, keynotes, and debates, involving leading experts from a range of disciplines. The course integrated the state-of-the-art from complementary perspectives, including development, cognitive models, neural basis, gene identification, functional genomics, model systems, and comparative/evolutionary studies.

Instructors, course assistants, and course lecturers come from universities, medical schools, research institutes, and companies around the world to teach at Cold Spring Harbor. Their excellence and dedication make the course program work so well. We would especially like to thank Drs. Randy Johnson, Gregory Macleod, Cynthia McMurray, Kate O'Connor-Giles, Adrian Rothenfluh, Amy Sater, Xin Sun, Gerald Thomsen, and Stephan Treue, who all retired after many years of service. The sad news that Carlos Barbas, distinguished instructor in the Phage Display course for more than two decades, passed away in early summer was a shock to all of us who knew and respected him and so enjoyed his excellent teaching and wry humor.

Grants from a variety of sources support the courses. The core support provided through the recently renewed grant from the Howard Hughes Medical Institute remains critical to our course program. The courses are further supported by multiple awards from the National Institutes of Health and the National Science Foundation, and additional support for individual courses is

provided from various foundations and other sources. The courses also depend on equipment and reagents that are loaned or donated by a large number of companies, partnerships that are invaluable in ensuring that the courses offer training in the latest technologies.

Now in its fifth year of operation, the Cold Spring Harbor Asia (CSHA) program, under the direction of Dr. Maoyen Chi, is headquartered at the Suzhou Dushu Lake Conference Center, a purpose-built academic conference center on the outskirts of old Suzhou, within a high-technology suburb (SIP). CSHA's scientific program includes large symposia and meetings, training workshops, and Banbury-style discussion meetings. CSHA is a wholly owned subsidiary of CSHL and is not beholden to outside partners in terms of our scientific programming. In 2014, 21 scientific conferences, one summer school, and two special CSHA sessions at Japanese meetings were held, predominantly attracting scientists from the Asia/Pacific region. The program is now making headway in attracting junior Asian scientists and the >15% growth in attendance between 2013 and 2014 bodes well for the future. This program is described in more detail in a separate Annual Report.

Alumni Notes: Winners of the 2014 Nobel Prize in Physiology or Medicine, Edvard and May-Britt Moser, have presented at many CSHL neuroscience meetings (the latter gave the keynote at the 2014 axon guidance meeting only weeks before the prize) and both spoke at the inaugural Crick Symposium in Suzhou in 2010. John O'Keefe gave the keynote at a CSHL/Wellcome Trust neuroscience conference in 2009. Winners of the 2014 Nobel Prize in Chemistry, Eric Betzig and Stefan Hell, have both participated in imaging-related meetings and courses at CSHL over the past decade, and the third laureate in chemistry, William Moerner, presented at the CSHA imaging meeting in Suzhou in 2013. We now count nine Nobel laureates among the student alumni of the course program.

The meetings and courses program staff is composed of a lean 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 services, and other activities. We wish Andrea Stephenson, who ably served as Course Registrar and Office and Operations Manager to the program for many years, the best of luck as she transitions into her new life in North Carolina.

**Terri Grodzicker**

*Dean of Academic Affairs*

**David Stewart**

*Executive Director,*

*Meetings & Courses Program/*

*President, Cold Spring Harbor Asia*

# 79TH COLD SPRING HARBOR LABORATORY SYMPOSIUM ON QUANTITATIVE BIOLOGY

---

## Cognition

---

May 28–June 2      290 participants

ARRANGED BY      **Cori Bargmann**, The Rockefeller University  
**Daphne Bavelier**, University of Geneva and University of Rochester  
**Terrence Sejnowski**, The Salk Institute for Biological Studies  
**David Stewart**, Cold Spring Harbor Laboratory  
**Bruce Stillman**, Cold Spring Harbor Laboratory

The Laboratory selected the theme of Cognition for the 79th Symposium in this historic series largely because of the tremendous advances being made by neuroscientists and psychologists working on cognitive processes at scales varying from the molecular to whole-brain and theoretical studies. Previous Symposia that addressed the brain included the fourth Symposium on Excitation Phenomena (1936), The Neuron (1952), Sensory Receptors (1965), The Synapse (1975), Molecular Neurobiology (1983), The Brain (1990), and Function and Dysfunction in the Nervous System (1996), and thus a return to the general theme of neuroscience was long overdue. Furthermore, a number of national projects into brain function have been initiated during the past few years, including President Barack Obama's announcement of the BRAIN (Brain Research through Advancing Innovative Neurotechnologies) Initiative in April 2013, a collaborative research project akin to the Human Genome Project, with the ultimate goal of mapping the activity of every neuron in the brain by 2025. These projects and initiatives reflect the fact that many neuroscientists are increasingly optimistic that major obstacles in understanding brain function may be overcome through the use of new technologies and approaches and that these advances may shed light on novel approaches to treat psychiatric, neurological, and neurodegenerative diseases.

The 2014 Symposium was designed to span a broad range of approaches toward increased understanding of cognitive processes in the brain including cell/molecular biology; developmental neurobiology; genetics and genomics; electrophysiological approaches; functional neuroimaging at cellular and whole-brain resolutions; computational neuroscience; behavioral, ethological, and psychophysical



G. Buzsak, J. Kubie, S. Tonegawa



A. Pouget, D. Bavelier, P. Churchland, J. DiCarlo

studies; and evolutionary/comparative neuroscience. The program of invited speakers was arranged to ensure that fundamental discoveries were balanced with approaches relevant to societal well-being, including a variety of stratagems for harnessing our increased understanding of brain function to improve treatment of mental illness and brain disorders. Speakers responded to the challenge of putting their own research, whether based on single-neuron studies, circuits, or whole-brain studies, into a broader context, allowing for a tremendous amount of cross-talk across disciplines with much insight gained from this unusually broad meeting. This is one of the strengths of the Symposium, and the 2014 meeting was particularly successful in highlighting emerging connections between molecular/cellular networks and higher brain functions such as language and decision-making.

Opening night speakers included Richard Axel (Columbia University) on innate and learned responses to odors, Allison Doupe (University of California, San Francisco) on vocal motor plasticity, Matthew Rushworth (University of Oxford) on decision-making, and Patricia Kuhl (University of Washington), who addressed language acquisition. One of the three Churchland neuroscientists who spoke at the Symposium, Patricia Churchland (University California, San Diego) addressed “The brains behind morality” in an excellent Dorcas Cummings Lecture for Laboratory friends, neighbors, and Symposium participants in advance of the annual dinner parties. Terry Sejnowski delivered a compelling and illuminating summary at the conclusion of the symposium.

This Symposium was attended by almost 300 scientists from universities around the country and internationally, and the program included 60 invited presentations, six short talks selected from the openly submitted abstracts on the basis of scientific merit, and 130 poster presentations. To disseminate the latest results and discussion of the Symposium to a wider audience, attendees were able to share many of the Symposium talks with their colleagues who were unable to attend using the Leading Strand video archive. A collection of interviews by Karen Carniol (*Cell*), Ann Goldstein (*Neuron*), Charvy Narain (*Nature Neuroscience*), Gary Stix (*Scientific American*), and Jan Witkowski (CSHL Banbury Center), with leading experts in the field, were arranged during the Symposium and distributed as free video from the Cold Spring Harbor Symposium interviews Web site.

This meeting was funded in part by Forest Laboratories and the Swartz Foundation. Contributions from the following companies and foundations provided support for the Symposium: Agilent Technologies, Bristol-Myers Squibb Company, DuPont/Pioneer Hi-Bred International, Genentech, Hudson-Alpha Institute for Biotechnology, Life Technologies (part of Thermo Fisher Scientific), Monsanto Company, and New England BioLabs. Financial support from the corporate sponsors of our meetings program is essential for these Symposia to remain a success, and we are most grateful for their continued support.

## PROGRAM

### Introduction

B. Stillman, *Cold Spring Harbor Laboratory*

### Multisensory Integration and Behavioral Outcomes

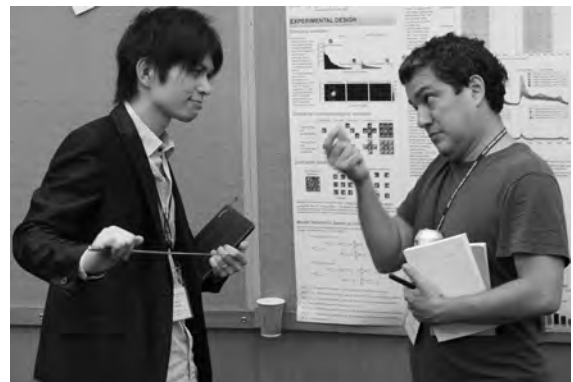
*Chairperson:* K. Svoboda, *HHMI/Janelia Farm Research Campus, Ashburn, Virginia*

### Physiology of Cognition

*Chairperson:* D. Angelaki, *Baylor College of Medicine, Houston, Texas*



J. Watson, H. Hoekstra, A. Gann



A.F. Ebihara, O.A. Gonzalo

**Object Recognition**

*Chairperson: D. Angelaki, Baylor College of Medicine, Houston, Texas*

**Communication, Language, and Higher Cognition**

*Chairperson: A. Doupe, University of California, San Francisco*

**Neural Substrates of Memory**

*Chairperson: A. Doupe, University of California, San Francisco*

**Movement and Motor Control**

*Chairperson: R. Desimone, MIT McGovern Institute for Brain Research, Cambridge, Massachusetts*

**Genes and Circuits**

*Chairperson: C. Dulac, HHMI/Harvard University, Cambridge, Massachusetts*

**Cognitive Disorders and Enhancement I**

*Chairperson: C. Dulac, HHMI/Harvard University, Cambridge, Massachusetts*

**Attention and Value**

*Chairperson: D. Wolpert, University of Cambridge, United Kingdom*

**Dorcas Cummings Lecture: The Brains Behind Morality**

*P. Smith Churchland, University of California, San Diego*

**Visual Perception and Neural Computation**

*Chairperson: A. Graybiel, Massachusetts Institute of Technology, Cambridge*

**Short Talks from Abstracts**

*Chairperson: T. Zador, Cold Spring Harbor Laboratory*

**Cognitive Disorders and Enhancement II**

*Chairperson: E. Nestler, Mount Sinai University, New York*

**Perceptual Judgment and Decision-Making**

*Chairperson: M. Carrasco, New York University, New York*

**Social Cognition and Complex Behavior**

*Chairperson: D. Anderson, HHMI/California Institute of Technology, Pasadena*

**Summary**

*T. Sejnowski, The Salk Institute for Biological Studies*



D. Gardner, D. Kleinfeld, R. Axel



C. Duan, A. Akrami



K. Svoboda, B. Stillman



B. Marlin



Y. Dan, J.A. Movshon

# MEETINGS

---

## Avian Model Systems

---

March 5–8            109 participants

ARRANGED BY        David Clayton, Queen Mary University of London, United Kingdom  
Rusty Lansford, Children's Hospital Los Angeles/University of Southern California  
Andrea Streit, Kings College London, United Kingdom

This meeting brought together internationally leading scientists who reflect the multidisciplinary approaches of the avian field. The avian research community primarily uses chicken, quail, and zebra finch to investigate seminal issues in embryology, behavior, neurobiology, immunology, stem cell biology, and endocrinology to name a few. During the past decade, this has become the key meeting for the research community who use avian model systems to exchange ideas, techniques, and reagents and to form new collaborations. Importantly, it provides a platform for junior scientists to present their research and to interact with scientists working at the frontiers of their respective fields.

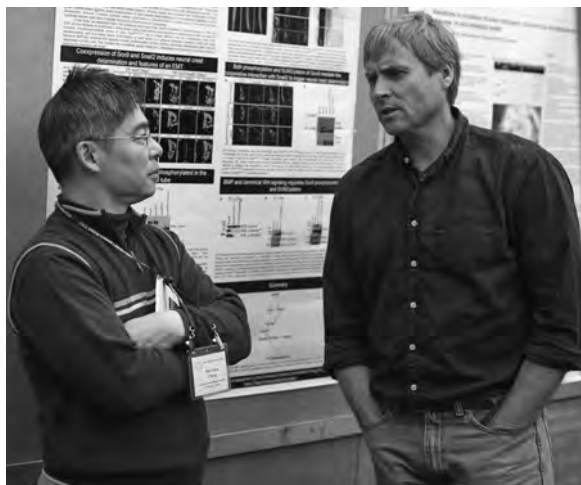
The meeting began with a keynote presentation by Dave Burt detailing the evolution and current state of avian genomics, celebrating a decade since the chicken genome sequence was published. There were approximately 50 poster presentations that mixed well with the wine and cheese.

A wrap-up session included a lively discussion about how large and diverse the avian community has become, integrated with explorations of how best to unite the avian community. Agreed upon Websites were being designed and installed prior to our final CSHL lunch. Plans were made to hold a meeting every 2 years alternating among Asia (Taiwan 2016), Europe (France 2018), and the United States (TBD 2020) to allow more people to attend this meeting and build a strong “avian community.”

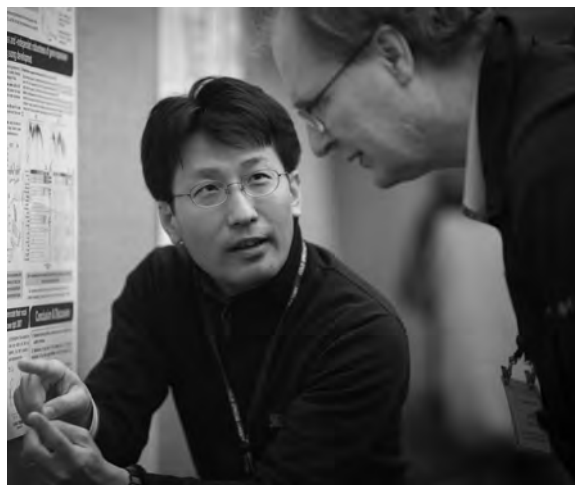
Our Avian Model Systems meeting was preceded by a successful “hands-on” Avian Genomics Workshop (March 4–5, 2014) run in collaboration with Ensembl/European Bioinformatics Institute and Wellcome Trust Sanger Institute. The workshop combined lecture and hands-on problems that were worked through by the students to better understand how to generate and use genomics, transcriptomics, and ChIP epigenomics information and analysis for their own studies.

This meeting was funded in part by the University of Arizona; Developmental Dynamics, Elsevier/Mechanisms of Development and International Society of Developmental Biologists; and the Society for Developmental Biology.





H.C. Cheng, R. Lansford



K. Wada, D. Clayton

## PROGRAM

**Keynote Address: Avian Genomics a Decade after the Publication of the Chicken Genome Sequence**  
D.W. Burt, *University of Edinburgh, United Kingdom*

**Avians as Models for Disease**  
*Chairpersons: A. Nieto, Universidad Miguel Hernandez-CSIC, Alicante, Spain; S. White, University of California, Los Angeles*

**The Shape of Things to Come**  
*Chairpersons: C.M. Chuong, University of Southern California, Los Angeles; P. Kulesa, Stowers Institute for Medical Research, Kansas City, Missouri*

**Evolution: A Bird's Eye View**  
*Chairpersons: R. Schneider, University of California, San Francisco; J. Gros, Institut Pasteur, Paris, France*

**Bird Brains and Behavior**  
*Chairpersons: C. Scharff, Freie Universitat Berlin, Germany; S. London, University of Chicago, Illinois*

**Where Is the Value in Genomics?**  
*Chairpersons: W. Warren, Washington University School of Medicine, St. Louis, Missouri; B. Faircloth, University of California, Los Angeles*

**Birds for the Future: Models for Development**  
*Chairpersons: T. Sauka-Spengler, University of Oxford, United Kingdom; C. Stern, University College London, United Kingdom*

**Stem Cells from an Avian Perspective**  
*Chairpersons: M. McGrew, The Roslin Institute, Edinburgh, United Kingdom; B. Pain, INSERM U846, Bron, France*



M. Skinner, J. George



N. Day

# Systems Biology: Global Regulation of Gene Expression

---

March 18–22      229 participants

ARRANGED BY      **Timothy Hughes**, University of Toronto, Canada  
**John Stamatoyannopoulos**, University of Washington  
**Sarah Teichmann**, European Bioinformatics Institute, United Kingdom

Systems Biology aims to understand the emergent properties of molecular networks in cells by using systematic and global approaches. One of the most actively researched areas of systems biology in recent years has been global regulation of gene expression, which coordinates complex metabolic and developmental programs in organisms. This ninth conference, like those in previous years, captured the continuing rapid progress and many new discoveries in this young and exciting field. The 4-day meeting featured 48 talks and 128 poster presentations covering a broad range of topics. The speakers, poster presenters, and other conference attendees were composed of a mix of students, postdocs, and principal investigators at all levels. Exciting highlights of the meeting were two keynote speeches: Dr. Naama Barkai, who has been a leader in the area of systems biology of gene expression for many years, described the first quantitative dissection of the contributions of initiation and elongation of transcription versus translation to yeast fitness under various growth conditions (phosphate starvation, nitrogen starvation, etc.). Dr. Orkin presented his work on understanding hematopoiesis and unraveling the correct lineage tree, as well as a careful dissection of the switch in expression from fetal to adult globin chains, with relevance to sickle cell anemia.

Many new techniques were featured this year that used next-generation sequencing, arguably the largest advance being in single-cell genomics technologies. This exciting new area takes the regulation of gene expression from averages of hundreds or millions of cells right down to the resolution of a single cell. Many of the presentations also described new approaches to the analysis of data, and these computational advances are driving the field forward at a tremendous pace. One example is the new chromatin immunoprecipitation high-throughput sequencing (ChIP-Seq) normalization technique that allows diverse data sets to be compared collectively (presented by the Ideker lab). Further advances relative to previous



Panel discussion





T. Hughes, S. Teichmann, J. Stamatoyannopoulos



S. Churchman, A. Sidow

years are the ubiquitous inclusion of the genome-wide association study (GWAS) of disease-related mutations, so that many studies of gene expression are now directly linked to human phenotypes.

The integration of genome biology with bioinformatics, genetics, and single-cell approaches at this conference is testament to this very exciting era in research on gene expression regulation.

This meeting was funded in part by the National Cancer Institute, a branch of the National Institutes of Health.

## PROGRAM

### Keynote Address

N. Barkai, *Weizmann Institute of Science, Rehovot, Israel*

### Transcriptional Regulatory Networks

*Chairperson: E. Rothenberg, California Institute of Technology, Pasadena*

### Chromatin/Epigenomics

*Chairperson: H. Chang, HHMI/Stanford University School of Medicine, California*

### Analysis of Rare and Single Cells

*Chairperson: E. Segal, Weizmann Institute of Science, Rehovot, Israel*

### Posttranscriptional Regulation/Technology

*Chairperson: K. Neugebauer, Yale University, New Haven, Connecticut*

### Keynote Address: Blood as a Platform for Systems Biology

S. Orkin, *Dana-Farber Cancer Institute, Boston, Massachusetts*

### Nuclear Architecture

*Chairperson: J. Lieb, Princeton University, Princeton, New Jersey*

### Transcription Factors

*Chairperson: T. Hughes, University of Toronto, Canada*

### *cis*-Regulatory Logic

*Chairperson: M. Bulyk, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts*

### Cell and Tissue Regulation

*Chairperson: J. Stamatoyannopoulos, University of Washington, Seattle*

## PTEN Pathways and Targets

---

March 25–28

93 participants

ARRANGED BY

**Suzanne Baker**, St. Jude Children's Research Hospital

**Anne Brunet**, Stanford University

**Lewis Cantley**, Weill Cornell Medical College

**Pier Paolo Pandolfi**, Beth Israel Deaconess Medical Center and Harvard Medical School of Medicine

This fifth CSHL conference was a dynamic discussion forum, bringing together top scientists studying the PTEN/PI-3 kinase pathway using biochemical and molecular approaches and in vivo model systems. Attending this 2014 meeting was a diverse group of scientists who study various molecular, genetic, and biochemical approaches to the analysis of the PTEN pathway at the cellular level and in development and diseases including cancer, diabetes, and neurological disorders such as autism. Junior and senior researchers discussed their latest research findings and technical approaches toward the analysis of the PTEN pathway. Particular emphasis was given to model systems, disease pathogenesis and therapy.

A total of 46 investigators presented in the seven scientific sessions, with 46 platform and 19 poster presentations and 93 registered attendees. The six platforms and one poster session were marked by cutting-edge dynamic presentations and enthusiastic exchanges of unpublished new results. During the scientific sessions, participants engaged in a vibrant discussion in the question and answer sessions. All of the poster sessions were also very well attended.

Essential funding for the meeting was provided by the National Cancer Institute and the National Institute on Aging, branches of the National Institutes of Health, and by Novartis.



P. Paolo Pandolfi, L. Cantley, A. Brunet, S. Baker



S. Morrison, M. Hansen



D. Sabatini, A. Brunet

**PROGRAM**

**Upstream Signaling in the PTEN Pathways**

*Chairperson: D. Alessi, University of Dundee, United Kingdom*

**Keynote Address**

R. Parsons, *Icahn School of Medicine at Mount Sinai, New York*

**Downstream Signaling in the PTEN Pathway**

*Chairperson: D. Sabatini, Massachusetts Institute of Technology, Cambridge*

**PTEN in Stem Cell Function, Development, and Disease**

*Chairperson: S. Morrison, University of Texas Southwestern Medical Center, Dallas*

**Cancer and PTEN Signaling**

*Chairperson: P.P. Pandolfi, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts*

**Metabolism and Aging**

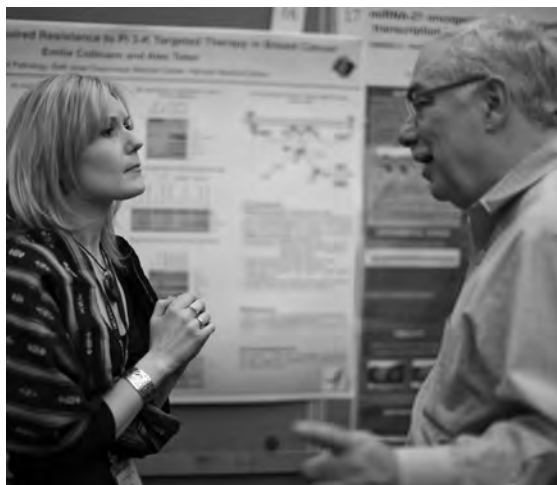
*Chairperson: A. Brunet, Stanford University, California*

**Complex Regulation of PTEN Signaling**

*Chairperson: L. Trotman, Cold Spring Harbor Laboratory*



G. De Sabbata, A. Ortega



E. Collmann, N. Rosen

# Neuronal Circuits

---

April 2–5

167 participants

ARRANGED BY **Dmitri Chklovskii**, Janelia Farm Research Campus  
**Carl Petersen**, Ecole Polytechnique Federale de Lausanne  
**Kristin Scott**, University of California, Berkeley

The anatomical and functional connectivity of neurons underlies the simplest of behavioral decisions to the most complex cognitive tasks. Unprecedented advances in the ability to monitor the activity of large neural populations, to rapidly and precisely manipulate the activity of neurons with optogenetic approaches, and to trace connectivity with synaptic resolution are providing a new systems-level view of neural circuit processing. Coupling recent technological advances with computational modeling and quantitative behavioral approaches has the potential to reveal how the dynamic computations of the brain underlie thought and behavior. This meeting highlights recent progress in elucidating neural circuit function across nervous systems.



D. Chklovskii, C. Petersen, K. Scott

In keeping with the original goals of this meeting, researchers working in a variety of different organisms and systems were brought together to share their scientific and technological advances, focusing on recent advances in neuronal circuit processing. This meeting included six slide sessions, covering a range of topics in the visual system, chemosensory systems, behavior, cortical processing, methods, and plasticity, and a very interactive poster session.

For the fourth meeting of this kind, the response of the field was very enthusiastic. The meeting brought together 167 participants from all over the world, most of whom made either oral or poster presentations. Particularly impressive were the large numbers of students (30%) and postdoctoral fellows (36%) participating in the meeting. Invited talks were given by Ed Boyden, MIT; Marla Feller, University of California, Berkeley; Michael Hausser, University College London; Fritjof Helmchen, University of Zurich; Elly Nedivi, MIT; Michael Reiser, HHMI/Janelia Farm Research Campus; Bernardo Sabatini, HHMI/Harvard Medical School; Scott Sternson, HHMI/Janelia Farm Research



D. Van Essen, J. Vogelstein



E. Aksay, K. Svoboda, F. Helmchen

Campus; Karl Svoboda, HHMI/Janelia Farm Research Campus; David Van Essen, Washington University; Leslie Vosshall, HHMI/Rockefeller University; and Tony Zador, Cold Spring Harbor Laboratory. The meeting provided an important forum for the exchange of ideas and approaches and helped scientists in the field get the most up-to-date information, as well as enabling them to meet, to network, and to establish collaborations. Based on the uniformly enthusiastic comments of the participants, the intensity of the oral and poster sessions, the productive question and answer sessions, and the large crowds that stayed up late to discuss science, the meeting was a great success.

A highlight of the meeting program was the Larry Katz Memorial Lecture, thanks to the generous donation of many colleagues in the field, which recognizes an outstanding research contribution by a graduate student. Although the number of nominations was smaller than in previous years, the overall quality of the 11 nominees was truly spectacular, representing an international cohort of exceptional graduate students. The selection committee was composed of the three current organizers and three past organizers, Cori Bargmann, HHMI/Rockefeller University; Liqun Luo, HHMI/Stanford; and Ed Calloway, Salk Institute. The committee unanimously selected Hidehiko Inagaki, California Institute of Technology, as the 2014 Lecturer. He gave a wonderful presentation about his graduate research. There was a consensus among the organizers and the committee that a history of the Larry Katz prize should be included in the program booklet.

This meeting was funded in part by the National Institute of Neurological Disorders and Stroke, a branch of the National Institutes of Health.

## PROGRAM

### Technology

*Chairperson:* E. Nedivi, *Massachusetts Institute of Technology, Cambridge*

### Vision

*Chairperson:* M. Hausser, *University College, London*

### Dendrites and Plasticity

*Chairperson:* M. Reiser, *HHMI/Janelia Farm Research Campus, Ashburn, Virginia*

### Neocortex

*Chairperson:* C. Petersen, *EPF Lausanne, Switzerland*

### Behavior

*Chairperson:* K. Scott, *University of California, Berkeley*

### Neuromodulation

*Chairperson:* M. Chklovskii, *HHMI/Janelia Farm Research Institute, Ashburn, Virginia*

### Larry Katz Lecture: Neuronal Mechanism of State Control in *Drosophila melanogaster*

H. Inagaki, *HHMI/Janelia Farm Research Campus, Ashburn, Virginia*



K. Podgorski, M. Stephenson Jones



A. Hussain, L. Vosshall

# The PARP Family and Friends: Gene Regulation and Beyond

April 9–12

100 participants

ARRANGED BY

W. Lee Kraus, University of Texas Southwestern Medical Center

Michael Hottiger, University of Zurich, Switzerland

Karolin Luger, Colorado State University

This first CSHL meeting on the poly(ADP-ribose) polymerase (PARP) family members and associated proteins marked 50 years since the discovery of poly(ADP-ribose) (PAR) by Chambon and Mandel in 1963. This was an historic meeting for the PARP field. Although there have been 20 previous meetings during the past four decades, this was the first time that a major scientific conference organization like CSHL sponsored the meeting. As this meeting clearly demonstrated, the field has matured and come into focus. Although the field has historically focused on the roles of PARP proteins in DNA-damage response, studies over the past decade have linked PARP-1 and other family members to the regulation of chromatin structure and transcription. This meeting integrated these two areas of research, highlighting key areas of overlap as well as distinct areas of difference. Other areas of focus at the meeting included gene regulation and chromatin, signaling and cancer, PARPs and RNA, and the biology of ADP-ribosylation (targets, binding proteins, and metabolism). Although the meeting had a decided mechanistic and molecular focus, emerging information about the physiological and pathophysiological roles of PARPs was also covered in detail, including the clinical applications of PARP inhibitors. All of these areas were covered in the poster session as well, providing a forum for continued discussion. Finally, the meeting highlighted new approaches for studying PARP function, including genomics and proteomics.

A goal of this meeting was to bring together researchers who have traditionally been part of the PARP field with those who are new to the field. The meeting



W.L. Kraus, K. Luger, M.O. Hottiger



H. Schuler, S. Krauss, N.W. Chi



A. Ladurner, R. Santoro

was a great success in this regard. In addition to the scientific program, the meeting included a faculty forum over lunch to discuss the state of the field and where it is headed, as well as to plan future activities and initiatives to strengthen and broaden the field. There was a unanimous view and strong support for future CSHL meetings on the PARP family.

This meeting was funded in part by the National Institute of Allergy and Infectious Diseases, a branch of the National Institutes of Health.

## PROGRAM

### Molecular Functions and Mechanisms

*Chairperson: W.L. Kraus, University of Texas Southwestern Medical Center, Dallas*

### Keynote Address: Local and Genome-Wide Dissection of the Mechanisms of Regulated Transcription and the Accompanying Interplay with Chromatin

*J.T. Lis, Cornell University, Ithaca, New York*

### DNA Repair, Gene Regulation, and Chromatin

*Chairperson: P. Chang, Massachusetts Institute of Technology, Cambridge*

### Signaling and Cancer

*Chairperson: S. Smith, New York University School of Medicine, New York*

### Physiology and Pathophysiology

*Chairperson: A. Ladurner, University of Munich, Germany*

### The Biology of ADP-Ribosylation: Targets, Binding Proteins, and Metabolism

*Chairperson: K. Luger, HHMI/Colorado State University, Fort Collins*

### Signaling, Chromatin, and RNA Biology

*Chairperson: M.O. Hottiger, University of Zurich, Switzerland*



J. Pascal, J. Lis



R. Conaway, J. Conaway, J. Pascal

# Gene Expression and Signaling in the Immune System

---

April 22–26

276 participants

ARRANGED BY

**Sankar Ghosh**, Columbia University College of Physicians & Surgeons  
**Dan Littman**, HHMI/New York University School of Medicine  
**Diane Mathis**, Harvard Medical School  
**Stephen Nutt**, The Walter & Eliza Hall Institute, Australia

This meeting was held for the seventh time in 2014, in unseasonable rain and cold temperatures for spring-time in New York. More than 275 registrants, including ~30% from outside the United States and many first-time attendees, participated in a highly engaging and interactive 4-day meeting. Although there are many immunology meetings, the unique aspect of this CSHL meeting is its focus on molecular and biochemical aspects of the development and function of the immune system. In addition, this meeting attracts a broad range of scientists who use the immune system as a model to study basic principles of biological regulation.

Talks were presented by a mix of invited speakers and investigators selected from a group of more than 200 submitted abstracts. A significant number of junior investigators were asked to give talks. Most speakers focused almost exclusively on their unpublished work. There were a number of exciting talks; for example, Amy Wagers (Harvard) presented exciting data on a blood-borne mediator that “rejuvenates” the muscle stem cells of old mice and discussed the cross-talk between these cells and lymphocytes. Ido Amit (Weissman Institute) developed and exploited novel, very sensitive ChIP-Seq (chromatin immunoprecipitation–high-throughput sequencing) technology to test, and ultimately challenge, the classical Waddington model of lineage differentiation. Steve Smale (University of California, Los Angeles) presented exciting studies which showed that by using the technology of RNA-Seq, the composition of the inflammatory gene expression program in macrophages can be determined not just by



D. Mathis, S. Ghosh, S. Nutt, D. Littman



R. Reiner, M. Busslinger



L. CorCovan, G. Belz



transcription, but by the process of splicing of transcripts and their export to the cytosol. Such talks highlighted the relevance of the data discussed at this meeting not only to immunology, but also to the broader scientific community. Oral presentations were supplemented by afternoon poster sessions that were extremely well attended and provided a forum for all meeting participants (including many graduate students and postdocs) to share their most recent data. In addition, attendees took advantage of an engaging piano recital and an evening performance by the ever-evolving Yale-based music group “The Cellmates” led by Richard Flavell.

This meeting was funded in part by the National Institute of Allergy and Infectious Diseases, a branch of the National Institutes of Health.

## PROGRAM

### Stem Cells and Developmental Decisions

*Chairperson: I. Weissman, Stanford University School of Medicine, California*

### Innate Immune Cell Development

*Chairperson: S. Nutt, The Walter and Eliza Hall Institute, Parkville, Australia*

### Adaptive (Lymphoid) Cell Development

*Chairperson: M. Busslinger, Research Institute of Molecular Pathology, Vienna, Austria*

### Chromatin Structure and Epigenetic Regulation

*Chairperson: R. Flavell, Yale University School of Medicine, New Haven, Connecticut*

### Antigen Receptor Gene Assembly and Modification

*Chairperson: D. Schatz, HHMI/Yale University Medical School, New Haven, Connecticut*

### Signaling

*Chairperson: D. Littman, HHMI/New York University School of Medicine, New York*

### Regulation of Immune Cell Function

*Chairperson: D. Mathis, Harvard Medical School, Boston, Massachusetts*

### Regulation of Innate Cell Function

*Chairperson: S. Ghosh, Columbia University, New York*



L. Thompson, H. Roy



C. Luo

# Molecular Chaperones and Stress Responses

---

April 29–May 3      168 participants

ARRANGED BY      **Andrew Dillin**, HHMI/University of California, Berkeley  
**Judith Frydman**, Stanford University  
**F. Ulrich Hartl**, Max-Planck Institute for Biochemistry, Germany

This meeting continues the rich tradition of a forum where cutting-edge analyses of molecular mechanisms of protein folding are presented in the context of its biological importance and health relevance. The high attendance, large number of posters, and lively audience participation in the sessions and in the informal activities that followed are testament to the success of the meeting and to the successful balance of the sessions. The field of chaperone research is evolving and this CSHL meeting is evolving with it and is proving to be the cornerstone meeting for this rich and exciting field. The meeting reflected the many branches of biology and medicine that are now understood to be regulated by molecular chaperones. In addition to well-established biochemical and biophysical methods for studying chaperone function, this meeting featured presentations using single-molecule methods and advanced computational approaches to the study of signaling networks involved in control of chaperone gene expression.

This year's meeting featured three new sessions: one on Spatial Quality Control chaired by Dr. Thomas Nystrom and the other on Misfolding Diseases chaired by Nobel laureate Dr. Stan Prusiner. The former highlights the emerging understanding that recognition and targeting of misfolded proteins to degradation is spatially regulated within the cell and the latter highlighted the close links between protein misfolding and neurodegenerative diseases and cancer. An entire section was devoted to Proteostasis, highlighting the connections between ribosomes and chaperones and novel signaling pathways involved in protein folding homeostasis. Among the cutting-edge themes explored were presentations on systems-level studies focused on understanding the circuitry that maintains protein homeostasis. Talks on the effect of disruption of these networks by cancer or aging emphasized the delicate balance between folding and misfolding required to maintain a healthy proteome.

This meeting was funded in part by the National Institute on Aging, a branch of the National Institutes of Health.



R. Morimoto, P. Genevoux



S. Prusiner



M. Ladjimi, W. Houry

## PROGRAM

### Proteostasis

*Chairpersons:* H. Kampinga, *University Medical Center Groningen, University of Groningen, The Netherlands;* R. Morimoto, *Northwestern University, Evanston, Illinois*

### Spatial Quality Control

*Chairpersons:* T. Nystrom, *Goteborg University, Sweden;* U. Jakob, *University of Michigan, Ann Arbor*

### Diseases

*Chairpersons:* S. Prusiner, *University of California, San Francisco;* S. Lindquist, *Whitehead Institute for Biomedical Research, Cambridge, Massachusetts*

### Machines and Mechanisms

*Chairpersons:* R. Klevit, *University of Washington, Seattle;* M. Hayer-Hartl, *Max-Planck Institute of Biochemistry, Martinsried, Germany*

### Systems and Networks

*Chairpersons:* A. Amon, *Massachusetts Institute of Technology, Cambridge;* J.S. Weissman, *HHMI/University of California, San Francisco*

### Degradation

*Chairpersons:* R. Hampton, *University of California, San Diego;* B.A. Schulman, *HHMI/St. Jude Children's Research Hospital, Memphis, Tennessee*

### ER

*Chairpersons:* T. Sommer, *Max-Delbruck Center for Molecular Medicine, Berlin, Germany;* L. Hendershot, *St. Jude Children's Research Hospital, Memphis, Tennessee*



D. Wolf, C. Enenkel



B. Schulman, S. Misra



M. Kampmann, A. Bertolotti

# The Biology of Genomes

---

May 6–10

562 participants

ARRANGED BY

**Carlos Bustamante**, Stanford University  
**Michel Georges**, University of Liege, Belgium  
**Elaine Mardis**, Washington University School of Medicine  
**Molly Przeworski**, Columbia University

This meeting marked the 27th annual gathering of genome scientists at CSHL. Scientists from around the world attended the meeting, with more than 381 abstracts describing a broad array of topics relating to the production, analysis, and interpretation of genomes from diverse organisms. The meeting built upon the remarkable progress in the sequencing, functional annotation, and analysis of genomes from many human individuals, “model organisms,” other animals including new genomic studies of rabbit, cichlid fish, and an exploration of differential exon usage in primate species. A new session topic focused on translational genomics and genetics, featuring lectures on undiagnosed pediatric disease, sequencing the Utah population, and sequencing of Parkinsons’ disease patients, among others. Because “translation” also applies to agriculture, a lecture about how maize diversity can be used for accelerated breeding was included in this session. The predominance of computational efforts in genomics was outlined by speakers covering a broad range of topics from oncogene discovery, posttranscriptional gene regulation from integrated data analysis, heritability of regulatory variants compared to coding variants in common disease, and single-cell RNA-sequencing studies that examine cell cycle variation. Next-generation sequencing continues to translate into the clinic as outlined by several speakers who described the role of genomics in identifying new cancer drivers, functional screening by CRISPR-Cas9, acquisition of chemoresistance in ovarian cancers, and an overview of mutational processes that play a role in human cancer.

All sessions were well attended, stretching the capacity of the CSHL facilities. Peter Donnelly and Nancy Moran gave the keynote presentations. Lawrence Brody moderated The ELSI (Ethical, Legal, and Social Implications) panel and the panelists included Arti Rai, Claire Laporte, and Richard Gold.

This meeting was funded in part by the National Human Genome Research Institute, a branch of the National Institutes of Health; Fluidigm; Illumina; and Ingenuity Systems.



Z. Chen, S. Shenoy, D. Chang, E. Roberts, M. Ilan



X. Estivill, S. Ossowsk

## PROGRAM

### Functional Genomics

*Chairpersons:* Y. Gilad, *University of Chicago, Illinois*;  
J. Rinn, *Harvard University, Cambridge, Massachusetts*

### Genetics of Complex Traits

*Chairpersons:* M. Ritchie, *Pennsylvania State University, University Park*; N. Wray, *University of Queensland, Brisbane, Australia*

### Cancer and Medical Genomics

*Chairpersons:* R. Karchin, *Johns Hopkins University, Baltimore, Maryland*; M. Stratton, *Wellcome Trust Sanger Institute, Hinxton, United Kingdom*

### Population Genomics

*Chairpersons:* P. Andolfatto, *Princeton University, New Jersey*;  
E. Kenny, *Icahn School of Medicine at Mount Sinai, New York*

### ELSI Panel and Discussion: The Court Has Spoken— What Did It Say? Implications of Recent U.S. Supreme Court Decisions on Diagnostic Testing and Gene Patenting

*Moderator:* L. Brody, *National Human Genome Research  
Institute, National Institutes of Health*

*Panelists:* A. Rai, *Duke University*; C. Laporte, *Foley Hoag,  
LLP*; R. Gold, *McGill University*

### Computational Genomics

*Chairpersons:* D. Pe'er, *Columbia University, New York*;  
M. Stephens, *University of Chicago, Illinois*

### Evolutionary and Nonhuman Genomics

*Chairpersons:* L. Andersson, *Uppsala University, Sweden*;  
J. Tung, *Duke University, Durham, North Carolina*

### Guest Speakers

P. Donnelly, *Wellcome Trust Centre for Human Genetics*;  
N. Moran, *University of Texas, Austin*

### Translational Genomics and Genetics

*Chairpersons:* E. Buckler, *Cornell University and USDA-  
ARS, Ithaca, New York*; A. Butte, *Stanford University,  
California*



M. Michalovova, C. Miller



O. Bahcall, B. Stranger



A.A. Buil, D. Parts

# The Cell Cycle

May 13–17

212 participants

ARRANGED BY

Sue Biggins, Fred Hutchinson Cancer Research Center  
Jacqueline Lees, MIT Center for Cancer Research  
David Toczyski, University of California, San Francisco

This 12th biannual meeting was held this year at Cold Spring Harbor. This conference is internationally recognized for its ability to bring together scientists who study cell cycle regulation in eukaryotes ranging from yeast to humans. As usual, the meeting began with a keynote talk from a leader in the field. We were fortunate to have Angelika Amon, the Kathleen and Curtis Marble Professor of Cancer Biology at MIT and professor of the HHMI, tell us about how aneuploidy causes the accumulation of unfolded proteins, leading to lysosomal degradation defects. After the keynote talk, the evening session continued with six talks focused on cytokinesis and chromosome dynamics. The remainder of the meeting was organized around seven lecture sessions and two poster sessions, which focused on the major stages of the cell division cycle, with an emphasis in each session on the molecular mechanisms that govern cell cycle progression. Many of these sessions emphasized the long-standing problems in cell cycle control, including the transcriptional mechanisms controlling entry in the cell cycle, the mechanisms that initiate and control chromosome duplication, the events that trigger the complex events of chromosome segregation in mitosis, and the numerous checkpoint systems that ensure that cell cycle events occur in the correct order. The meeting represented many approaches, including chemical genetics, kinetics, single-cell assays, biochemistry, microscopy, and computational approaches. There were also a number of talks and posters about development and other specialized aspects of the cell cycle. The meeting connected cell cycle regulation with cancer biology with reports about the functions of oncogenes and tumor suppressors in cell



J. Lees, D. Toczyski, S. Biggins



C. Pantelidou



P. Sandra, C. Schwennicke, G. Müller, K. Engeland, L. Binder, D. Varma



C. Wurzenberger, G. Kops

cycle control. As always, major model systems for cell cycle analysis were represented, and the striking phylogenetic conservation of cell cycle regulatory mechanisms was readily evident. It was another landmark meeting for the cell cycle field, and the participants continue to look forward to equally exciting meetings in future years.

Cold Spring Harbor Laboratory would like to thank Bethyl Laboratories for their generous support of this meeting.

**PROGRAM**

**Keynote Address: Effects of Chromosome Segregation Errors on Cell Physiology**

A. Amon, *Massachusetts Institute of Technology*

**Cytokinesis and Chromosome Dynamics**

*Chairperson:* D. Pellman, *Dana-Farber Cancer Institute, Boston, Massachusetts*

**S Phase**

*Chairpersons:* J. Cook, *University of North Carolina, Chapel Hill*; A. Dutta, *University of Virginia, Charlottesville*

**Checkpoints**

*Chairpersons:* L. Zou, *Massachusetts General Hospital, Harvard Medical School, Boston*; A. Musacchio, *Max-Planck Institute for Molecular Physiology, Dortmund, Germany*

**G<sub>0</sub>-G<sub>1</sub> Control**

*Chairpersons:* P. Sicinski, *Dana-Farber Cancer Institute, Harvard Medical School, Boston, Massachusetts*; N. Dyson, *Massachusetts General Hospital, Boston*

**Mitosis**

*Chairpersons:* I. Cheeseman, *Whitehead Institute, Cambridge, Massachusetts*; J. DeLuca, *Colorado State University, Fort Collins*

**Cell Cycle Decisions**

*Chairpersons:* O. Cohen-Fix, *NIDDK, National Institutes of Health, Bethesda, Maryland*; J. Campisi, *Buck Institute for Research on Aging, Novato, California*

**Metaphase to Anaphase**

*Chairpersons:* M. Rape, *University of California, Berkeley*; J. Ferrell, *Stanford University, California*

**Regulation of CDK**

*Chairpersons:* K. Oegema, *Ludwig Institute for Cancer Research, La Jolla, California*; D. Morgan, *University of California, San Francisco*



S. Biggins, A. Musacchio



B. Mierzwa



J. DeLuca, A. Amon

# Retroviruses

---

May 19–24

409 participants

ARRANGED BY

**John Guatelli**, University of California, San Diego  
**Welkin Johnson**, Boston College

Organizers Drs. John Guatelli and Welkin Johnson are regular attendees who like many of the participants have “grown up” at this meeting and within its distinct community of scientists. Indeed, a characteristic feature of this meeting remains the high number of repeat attendees who come each year, where friendships and collaborations are both initiated and maintained. In keeping with this theme, the organizers chose to recognize with a keynote talk the contributions of one of the strongest supporters of the meeting, Dr. Paul Bieniasz. Paul presented an overview of his many contributions to basic retrovirology with a remarkable blend of the personal and scientific. His talk was very well received by the attendees, many of whom found it inspirational.



W. Johnson, J. Guatelli

The other keynote speaker was Dr. Mary Carrington, who also presented a very well-received and humorous talk that highlighted her work on how immunogenetic variation affects the outcome of disease caused by HIV. This talk, like that of Dr. Bieniasz, fit a continuing trend in the meeting toward an emphasis on the host–pathogen relationship during retroviral infections.

Three annual prizes, originally instituted and endowed by the community of regular attendees with assistance from the CSHL Meetings staff, were also awarded. These prizes memorialize three of our former colleagues who passed away unexpectedly in recent years (The Daniel Wolf Prize, The Ute von Schwedler Prize, and the Andy Kaplan Prize) and serve to encourage up-and-coming scientists at the grad and postdoc/junior levels. The prizes are also a testament to the community nature of the CSH Retroviruses meeting and the continuity of the attendees over the years.

As previously, the meeting was organized into 13 sessions (10 oral and three posters); the presentations featured unpublished, primarily basic, research. The sessions focused on traditional



Z. Williams, I. Muñoz-Arias



H. Levin, S. Hughes





E. Olson, I. Rouzina



C. Münk, K. Sato



R. Sutton, W. Walker

areas, such as specific aspects of the retroviral life cycle, as well as on classes of host defenses, especially specific restriction factors and the viral proteins that counteract them. A general session on evolution and endogenous retroviruses opened the meeting, and a general session on innate and intrinsic immunity to retroviruses closed the meeting. The organizers attempted to distribute talks that could fit in alternative sessions in a manner that led to each session having “something for everyone.” Scientific highlights of the meeting included new insights into the mechanism of action of previously known restriction factors and the viral proteins that counteract them, such as the structure of the ternary complex the restriction factor SAMHD1, the lentiviral protein Vpx, and the cellular factor DCAF; and various aspects of newly recognized antiretroviral host proteins such as the interferon-induced protein Mx2 and the IFITMs. A range of basic information regarding the retroviral life cycle, such as the role of nuclear import in integration site selection and the role of the integrase protein in virion maturation, was also presented. As previously, the meeting also exemplified new techniques of scientific investigation that could be applied both to the discipline of retrovirology and to other work.

This meeting was funded in part by the National Institute of Allergy and Infectious Diseases, a branch of the National Institutes of Health; and Gilead Sciences, Inc.

## PROGRAM

### Evolution and Endogenous Retroviruses

*Chairpersons:* S. Goff, *Columbia University, New York*;  
H. Levin, *NICHD, National Institutes of Health, Bethesda,*  
*Maryland*

### Third Annual Ute von Schwedler Prize for Retrovirology

*Awarded to:* S.J. Rinh, *The Rockefeller University*

### Entry

*Chairpersons:* R. Swanstrom, *University of North Carolina,*  
*Chapel Hill*; G. Melikian, *Emory University, Atlanta,*  
*Georgia*

### Postentry I

*Chairpersons:* Z. Ambrose, *University of Pittsburgh*  
*School of Medicine, Pennsylvania*; S. Ross, *University of*  
*Pennsylvania, Philadelphia*

### Keynote Address: The Influence of Immunogenetic Variation on HIV Disease

M. Carrington, *National Cancer Institute-Frederick*

### APOBEC, VIF, Mutagenesis

*Chairpersons:* V. Simon, *Icahn School of Medicine,*  
*New York*; L. Mansky, *University of Minnesota,*  
*Minneapolis*

### PostEntry II

*Chairpersons:* E. Campbell, *Loyola University-Chicago,*  
*Maywood, Illinois*; S. Wilson, *University of Glasgow,*  
*United Kingdom*

### Eighth Annual Andy Kaplan Prize

*Awarded to:* C. Goujon, *King's College London*

**VPX, SAMHD1, and VPR**

*Chairpersons:* R. Gummuluru, *Boston University School of Medicine, Massachusetts*; F. Diaz-Griffero, *Albert Einstein College of Medicine, Bronx, New York*

**Integration and Integrase**

*Chairpersons:* M. Kvaratskhelia, *The Ohio State University, Columbus*; A. Engelman, *Dana-Farber Cancer Institute, Boston, Massachusetts*

**Keynote Address**

P. Bieniasz, *Aaron Diamond AIDS Research Center, New York*

**Assembly**

*Chairpersons:* K. Bishop, *MRC National Institute for Medical Research, London, United Kingdom*; J. Lingappa, *University of Washington, Seattle*

**Late Replication, VPU, BST2/Tetherin**

*Chairpersons:* S. Neil, *King's College London, United Kingdom*; E. Cohen, *Institut de Recherches Cliniques de Montréal, Canada*

**Innate/Intrinsic Immunity**

*Chairpersons:* D. Evans, *University of Wisconsin, Madison*; M. Li, *University of California, San Diego*

# Glia in Health and Disease

---

July 17–21

247 participants

ARRANGED BY

**Dwight Bergles**, Johns Hopkins University  
**Marc Freeman**, University of Massachusetts

Glial cells are as abundant as neurons in the nervous system, but the diverse roles played by these cells under physiological and pathological conditions are not well understood. In this fifth summer meeting, students and scientists from across the world gathered to discuss recent progress in this rapidly expanding field. Speakers presented exciting new data on a variety of important topics, such as the beneficial and detrimental roles of microglia, glial control of brain metabolism, development and plasticity of myelin, and regulation of the blood brain barrier. The atmosphere was supportive, and there were lively discussions in oral sessions, the poster sessions, and in informal gatherings. Research involving vertebrate and invertebrate model systems was presented along with new developments in state-of-the-art methodologies that are providing new insight into glial cell functions. One of the highlights of the meeting was the keynote address from Dr. Loren Looger, who described new molecular tools that he has developed to monitor and manipulate the behavior of glia *in vivo*. By the conclusion of the meeting, it was clear that information about properties and functions of glia is increasing exponentially, enabled by technological advances in the areas of imaging and molecular genetics. These studies are beginning to yield new insight into the roles of distinct groups of glia in central nervous system (CNS) development, information processing in neural circuits, disease progression, and CNS repair.

This meeting was funded in part by the National Institute of Neurological Disorders and Stroke, a branch of the National Institutes of Health.



M. Freeman, D. Bergles



P. Honsa, S. Kang



N. Gunhanlar, E. Romanelli

**PROGRAM**

**Beneficial and Detrimental Roles of Microglia**

*Chairpersons:* R.M. Ransohoff, *Massachusetts General Hospital, Boston*; W. Gan, *New York University, New York*

**Neuronal Structural Plasticity and Synaptic Pruning**

*Chairpersons:* B. Stevens, *Children's Hospital Boston, Massachusetts*; B. Barres, *Stanford University, California*

**Glial Control of Brain Metabolism**

*Chairpersons:* D. Attwell, *University College London, United Kingdom*; J. Rothstein, *Johns Hopkins University, Baltimore, Maryland*

**Cell-Fate Specification**

*Chairpersons:* R. Rowitch, *University of California, San Francisco*; M. Götz, *Ludwig Maximilians University, Munich, Germany*

**Roles of Astroglial Ca<sup>2+</sup> Signaling**

*Chairpersons:* B. Khakh, *University of California, Los Angeles*; D. Bergles, *Johns Hopkins University, Baltimore, Maryland*

**Keynote Address**

L. Looger, *HHMI/Janelia Farm Research Campus*

**Development and Plasticity of Myelin**

*Chairpersons:* W. Macklin, *University of Colorado, Denver*; J. Chan, *University of California, San Francisco*

**Blood-Brain Barrier Regulation**

*Chairpersons:* R. Daneman, *University of California, San Francisco*; J. Filosa, *Georgia Health Sciences University, Augusta*

**Reactive Gliosis and Glial Scars**

*Chairpersons:* L. Dimou, *Ludwig Maximilians University, Munich, Germany*; R. Giger, *University of Michigan, Ann Arbor*



B. Stevens



M. Iruarizaga, N. Wu



R. Garcia, B. Khakh, J. Petracicz, D. Bergles

# Messenger RNA: From Discovery to Synthesis and Regulation in Bacteria and Eukaryotes

---

August 9–11

128 participants

ARRANGED BY

**James Darnell**, The Rockefeller University  
**Adrian Krainer**, Cold Spring Harbor Laboratory  
**Mila Pollock**, Cold Spring Harbor Laboratory

The approximately 140 attendees at this conference included 40–45 graduate students and post-doctoral fellows, satisfying one important aim of the meeting: to give younger scientists the flavor of important research from the late 1950s onward. In organizing the meeting, Jim Darnell and Adrian Krainer, both scientists who have worked in the field for decades, and Mila Pollock, Executive Director of the Cold Spring Harbor Laboratory Library and Archives, who is dedicated to preserving the history of molecular biology, encouraged all speakers to address how they first arrived at studying an aspect of mRNA formation and function. The talks began with the 1958–1960 history leading to the idea of “messenger RNA” and its discovery as a newly identified molecule in bacteria. Immediately following came the history of the confirmation that the mRNA concept carried over to mammalian cells and the beginnings of establishing the biochemical requirements in eukaryotes.

The first discussions about bacterial mRNA were presented by original participants in the revolution: Jim Watson, Sydney Brenner, Art Pardee, and Matt Meselson. Central to this progress was bacterial genetics, especially that which involved inducible enzymes, notably  $\beta$ -galactosidase, of course, and also T-even bacteriophage molecular biology.

Bacteria, lacking nuclei, offer their mRNA to ribosomes for immediate use, whereas eukaryotic cells have their DNA sequestered in a nucleus, and were soon found to exhibit unanticipated complications to mRNA formation. The first extractions of total mammalian cell RNA revealed previously undiscovered RNA processing. First, giant nucleolar RNA molecules (primary transcripts)



L. Maquat, A. Krainer, M. Pollock, R. Roeder, J. Warner, M. Meselson, J. Darnell, W. Gilbert, B. Stillman



A. Krainer, R. Roberts, O. Anezakow, R. Karni, M. Gierszewska



R. Singer, J. Manley, C. Manley, D. Labbe, N. Sonnenberg

were found to be precursors to shorter cytoplasmic ribosomal RNAs. Large nuclear DNA-like RNAs not related to ribosomes were also identified. Cytoplasmic polyribosomes were found and they contained shorter DNA-like mRNA. Both the large DNA-like RNA and the mRNA had 5' caps as well as 3' poly(A) addition, so RNA processing seemed likely in mRNA formation also. These experiments formed the history of eukaryotic mRNA from ~1963 to 1974.

In 1977, RNA splicing, which uses only disparate parts (exons) of large primary transcripts, while destroying the unused intervening parts (introns), was discovered. This discovery came through studies with adenovirus, a DNA virus that uses the host's nuclear machinery to produce its mRNAs.

This exciting discovery was described for the first time at the 1977 CSHL symposium in June by two research groups. Richard Roberts and Louise Chow, former CSHL workers, and Arnold Berk, representing the MIT lab of Phil Sharp, described the experiments showing that late in adenovirus infection, multiple different segments of a long primary transcript somehow were "spliced" together to make the various mRNAs, exhibiting at once basic splicing and also demonstrating alternative splicing.

Several successive sessions provided details of how these processes—enzymatic synthesis of mRNA precursors, enzymatic 5' and 3' modifications of the transcripts, and the newly recognized RNA chemistry of splicing—were found to be carried out by large protein complexes or ribonucleoprotein plus protein components of the spliceosome.

The complex fates of mRNA in the eukaryotic cytoplasm were the topic in several consecutive sessions. In some cases, there is directed movement of particular mRNAs to specific cellular locales and, most important, the final concentration of eukaryotic mRNA depends not only on the rate of synthesis, but also on the regulated relative stability of different mRNA molecules. Finally, the translation into protein of different mRNAs requires a host of translation factors. The major insights that have been obtained into the mechanisms and regulation of all these events were discussed in the concluding sessions of the meeting.

As noted above, a major purpose of the meeting was to provide for beginning scientists (as well as those currently in mid-career) some first-hand insights into very important discoveries that now range from 30 to 55 years old. From the many positive comments from attendees, it seems that the mission was at least partly accomplished.

We would like to gratefully acknowledge support from the following organizations, which made this conference possible: New England Biolabs, Looger, Rockefeller University, and Cold Spring Harbor Laboratory. Additional support was provided by Lexogen and Life Sciences Foundation.

**PROGRAM**

**Bacterial mRNA, The Idea and Regulation**

*Chairperson: Tom Maniatis, Columbia University Medical Center, New York*

**The Successful Hunt for Eukaryotic mRNA and End Decorations**

*Chairperson: James Darnell, The Rockefeller University, New York*

**Origin of mRNA, Adenovirus to the Rescue**

*Chairperson: Joan Steitz, Yale University/HHMI, New Haven, Connecticut*

**Enzymes and Eukaryotic mRNA**

*Chairperson: Nahum Sonenberg, McGill University, Montreal, Quebec, Canada*

**Mechanisms of Splicing**

*Chairperson: Richard Roberts, New England BioLabs, Ipswich, Massachusetts*

**Nascent RNA: Associated Proteins and Early Events**

*Chairperson: Robert Roeder, The Rockefeller University, New York*

**Transport, Turnover, and Translation of Eukaryotic mRNA**

*Chairperson: Walter Gilbert, Harvard University, Cambridge, Massachusetts*



L. Maquat, R. Overell



A. Pardee, J. Watson



M. Meselson, G. Enikolopov



J. Watson, J. Stietz, W. Gilbert, J. Witkowski, S. Brenner, J. Darnell, K. Maxson, M. Pollack, A. Gann

# Mechanisms and Models of Cancer

---

August 12–16

409 participants

ARRANGED BY

**Benjamin Neel**, Princess Margaret Cancer Center, Canada  
**Karen Vousden**, Cancer Research United Kingdom Beatson Institute  
**William Weiss**, University of California, San Francisco  
**Eileen White**, The Rutgers Cancer Institute of New Jersey

Mechanisms that drive human cancers, along with cell- and animal-based models to test these mechanisms, remain central to our understanding of human malignancies. The current explosion in molecular cataloging of mutation, expression analyses, and epigenetic regulation that illuminate distinct human cancer types has carried with it an improved understanding of cancer metabolism and epigenetics, offering new therapeutic opportunities. Our ability to understand cross-talk between cancer cells and stromal cells continues to improve, leading to immunotherapies that promise cures for diseases that had been rapidly fatal. This fifth meeting held at the Laboratory again convened an international group of investigators whose collective work focused on these themes. Oral and poster presentations provided beautiful examples of how new insights have been gained from application of interdisciplinary approaches that utilize genetics, developmental biology, genomics and proteomics, and model organisms (principally sophisticated mouse models, with significant advances in methods for rapid genomic editing of normal and cancer genomes since our 2012 meeting) to advance the development of rational therapeutics. A keynote address by Gerard Evan clarified roles for MYC not just as a driver for cancer, but also as a cancer engine, as blockade of MYC showed efficacy even in tumors that did not amplify or overexpress this oncoprotein. A keynote by Arnold Levine provided an introduction to p53 signaling, which illuminated new roles for this important tumor suppressor.

This meeting was funded in part by the National Cancer Institute, a branch of the National Institutes of Health.



K. Vousden, E. White, W. Weiss, B. Neel



L. Trotman, P.P. Scaglioni



M. Wilhelm, G. Selivanova, A. Henriksson



## PROGRAM

### Keynote Address: Targeting Cancer's Engines, Not Its Drivers

G.I. Evan, *University of Cambridge, United Kingdom*

### Experimental Therapeutics

*Chairpersons:* T. Letai, *Dana-Farber Cancer Institute, Boston, Massachusetts*; D. Tuveson, *Cold Spring Harbor Laboratory*

### Mouse Models and Stem Cells

*Chairpersons:* C. Abate-Shen, *Columbia University, New York*; L. Parada, *University of Texas Southwestern Medical Center, Dallas*

### Genomics and Epigenomics

*Chairpersons:* N. Jabado, *McGill University Health Center, Montreal, Canada*; J. Bradner, *Dana-Farber Cancer Institute, Boston, Massachusetts*

### Keynote Address: The Evolution of Thymic Lymphomas in p53 Knock-Out Mice

A.J. Levine, *Institute for Advanced Studies and Rutgers Cancer Institute of New Jersey*

### Cancer Drivers and Metabolism

*Chairpersons:* C. Prives, *Columbia University, New York*; L. Cantley, *Weill Cornell Medical College, New York*

### Keynote Address: Oncogenic Ras Signaling: From Biology to Therapeutics

D. Bar-Sagi, *New York University Langone Medical Center*

### RAS

*Chairpersons:* K. Cichowski, *Brigham and Women's Hospital, Boston, Massachusetts*; M. McMahon, *University of California, San Francisco*

### Signaling and Metastases

*Chairpersons:* G. Lozano, *University of Texas MD Anderson Cancer Center, Houston*; Y. Kang, *Princeton University, New Jersey*

### Cancer Targets and Therapeutics

*Chairpersons:* M. Park, *McGill University, Montreal, Canada*; R. Levine, *Memorial Sloan-Kettering Cancer Center, New York*

### Metabolism and Transformation

*Chairpersons:* W. Kaelin, *HHMI/Dana Farber Cancer Institute, Boston, Massachusetts*; E. Gottlieb, *Cancer Research UK, Glasgow, United Kingdom*



S. Lowe, R. Sordella



R. Ferretti, J. Lees, K. Masoumi



M. Hemann, J. Zuber

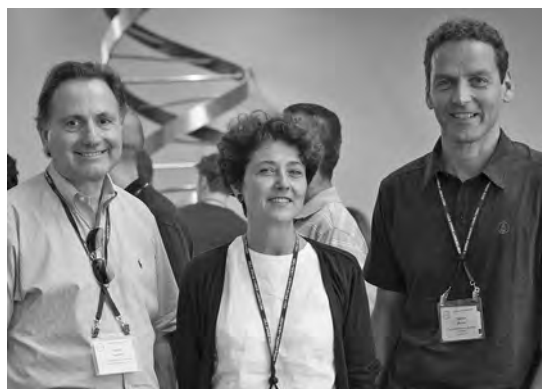
# Nuclear Organization and Function

---

August 19–23      216 participants

ARRANGED BY      **Edith Heard**, Curie Institute, France  
**Martin Hetzer**, Salk Institute for Biological Studies  
**David L. Spector**, Cold Spring Harbor Laboratory

The opening session on Tuesday evening highlighted current studies of cell cycle and DNA repair, with Kim Naysmyth giving a stimulating discussion of the role of cohesin in DNA entrapment in yeast and gene regulation in flies. Jacqueline Barlow (Nussenzweig lab) showed that heavily transcribed and early replicated regions form fragile sites, presumably due to collisions between the transcriptional and replication machineries, and that these sites can be a source of aberrant DNA rearrangement in B-cell lymphomas. Vishnu Dileep (Gilbert lab) presented work on the reorganization of chromatin during early G<sub>1</sub> that reflects the temporal and developmental control of DNA replication. Xiang Deng (Belmont lab) presented insights into the large-scale plasticity of chromatin fibers during DNA replication, using live cell and super-resolution imaging.



D. Spector, E. Heard, M. Hetzer

In the session on the Nuclear Membrane and Nuclear Lamina, U. Kutay discussed the targeting of membrane proteins to the inner nuclear membrane and the role that nuclear localization signal sequences may have in targeting a protein back to the chromatin compartment. Megan King showed that the physical association of chromatin with the inner nuclear membrane provides stiffness necessary to buffer cytoskeletal forces in interphase. Roland Foisner showed that chromatin association sites of A-type lamins and LAP2 $\alpha$  cover long-range genomic regions that unexpectedly are located outside lamin-associated domains (LADs). Using structured illumination microscopy, Takeshi Shimi (Goldman lab) showed that individual lamin filaments have a diameter of 10 nm.



Wine and Cheese Party

Robert Singer led off the session on Biology of Nuclear RNA describing a single-molecule approach to track individual mRNA molecules in living cells. Andy Belmont focused on an innovative technique to show that speckle-proximal domains correlate with gene-rich chromosome “R” bands, and speckle-distal domains correspond to LADs. Gayatri Arun (Spector lab) presented exciting new insights into the role of the Malat1 lncRNA in breast cancer progression. Several other noncoding RNAs were also discussed, including the X-chromosomal FIRRE RNA proposed to be involved in enabling long-range chromosomal interactions, nucleolar Alu-containing RNAs, and centromeric RNAs with a role in telomerase.

In the Chromosome Conformation and Epigenetics session, Genevieve Almouzni described recent work exploring how CenH3/CENP-A in human cells depends on the chaperone DAXX. Elena Ezhkova discussed the impact of polycomb repressive complex 2 (PRC2) proteins in skin stem-cell differentiation control. Filipe Jacinto (Hetzer lab) presented exciting work on the role of Nup153 in the recruitment of the PRC1 complex to developmental genes in embryonic stem cells (ESCs). New insights into the relationship between nuclear position, chromatin state, transcription, and replication timing were presented by Wendy Bickmore. Kyle Eagan (Kornberg lab) and Kerstine Wendt discussed the structure and function of topologically associated domains (TADs).

In the Transcription session, Victor Corces discussed the role of architectural proteins in the three-dimensional organization of the genome and their impact on modulating TAD organization and gene expression during heat shock. Several talks discussed the roles of transcription factor binding in enabling 3D organization and variability in gene expression. Several single-cell or single-molecule studies were presented, for example, by Bas van Steensel, in the context of LADs. A talk from Gerd Blobel’s lab discussed the role of chromosome structure in the mitotic memory of gene expression states.

Jason Brickner led off the session on Chromosome Structure and showed that gene localization and clustering within the yeast nucleus are regulated by transcription factor activation, abundance, and steric interference by transcription repression and modification of chromatin structure at the nuclear pore complex (NPC). Hanhui Ma (Pederson lab) used the CRISPR/Cas9 method to visualize DNA sequences in chromosomes in live cells. Susan Gasser discussed the role of heterochromatin and nuclear organization in *Caenorhabditis elegans* development.

In the session on Nucleocytoplasmic Transport, Broder Schmid (Gorlich lab) showed that Nup98 FG domains spontaneously phase separate into particles with exquisite NPC-like permeability. Christopher Lord (Wente lab) made the interesting observation that yeast mutants lacking GLFG domains of Nup116 display decreased replicative life span and linked nuclear transport to



B. van Steensel, K. Nasmyth



C. Sjögren, S. Buonomo



R. Goldman, R. Singer

mitochondrial function. Greg Matera gave an exciting talk about the engineering of histone genes using bacterial artificial chromosome (BAC)-based transgenic arrays in *Drosophila*.

In the final session on Nuclear Function in Development and Disease, Megan Bodnar-Hogan (Spector lab) showed that the alleles of the *Oct4* gene transiently pair during ESC differentiation, which might be critical for proper gene regulation. Aaron Straight showed that distinct domains of CENP-A act cooperatively to direct CENP-A assembly. Ronan Chaligne (Heard lab) showed that the aberrant expression of genes from the inactive X chromosome in breast cancer is associated with genome-wide depletion of H3K27me<sub>3</sub> and enrichment of H4Ac and thus can be used as a prognostic or diagnostic tool in the clinic. Barbara Meyer discussed condensin-driven remodeling of X-chromosome topology during dosage compensation, which involves higher-order chromatin structure and association with NPCs.

## PROGRAM

### Cell Cycle and DNA Repair

*Chairpersons:* D. Gilbert, *Florida State University, Tallahassee*; K. Nasmyth, *University of Oxford, United Kingdom*

### The Nuclear Membrane and Nuclear Lamina

*Chairpersons:* R. Foisner, *Medical University Vienna, Austria*; U. Kutay, *ETH Zurich, Switzerland*

### Biology of Nuclear RNAs

*Chairpersons:* A. Belmont, *University of Illinois, Urbana*; R. Singer, *Albert Einstein College of Medicine, Bronx, New York*

### Chromosome Conformation and Epigenetics

*Chairpersons:* G. Almouzni, *Institut Curie, Paris, France*; W. Bickmore, *University of Edinburgh, United Kingdom*

### Transcription

*Chairpersons:* V. Corces, *Emory University, Atlanta, Georgia*; B. Van Steensel, *Netherlands Cancer Institute, Amsterdam*

### Chromatin Organization

*Chairpersons:* J. Brickner, *Northwestern University, Evanston, Illinois*; S. Gasser, *Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland*

### Nucleocytoplasmic Transport

*Chairperson:* M. Hetzer, *Salk Institute for Biological Studies, La Jolla, California*

### Nuclear Function in Development and Disease

*Chairpersons:* B. Meyer, *University of California, Berkeley*; A. Straight, *Stanford University, California*



D. Gilbert, S. Buonomo

# Regulatory and Noncoding RNAs

---

August 26–30

286 participants

ARRANGED BY

Gregory Hannon, Cold Spring Harbor Laboratory

Elisa Izaurralde, Max-Planck Institute for Developmental Biology, Germany

Michael Terns, University of Georgia

In the 50 years since RNA was identified as a central component in the flow of genetic information, it has become increasingly clear that RNA is more than a mere messenger and has central roles across numerous biological processes required for cell viability and function. In addition to the well-characterized prokaryotic and eukaryotic small RNAs such as CRISPRs, microRNAs, and Piwi-activating RNAs (piRNAs), eukaryotic genomes are pervasively transcribed to generate thousands of long noncoding RNAs (lncRNAs), which are emerging as key regulators of diverse cellular processes, but their precise physiological function is just beginning to be elucidated. Students and scientists engaged in multidisciplinary discussions regarding the diverse and multifaceted roles of noncoding RNA across species and in a broad range of cellular processes. Sessions focused on recent, unpublished discoveries in the fields of prokaryotic and eukaryotic long and short noncoding RNAs. There were ample opportunities for discussions in the oral sessions, the poster sessions, and in informal settings. There was a strong sense of community, and the conviction that this meeting provides an excellent framework to bolstering discussions of the roles of noncoding RNAs, enhancing our understanding of gene regulation and function.



E. Izaurralde, M. Terns, G. Hannon



V. Ambros, N. Kim



H. Persson, R. Søkilde, I. Newie



N. Proudfoot, A. Molla

**PROGRAM****Keynote Address**

D. Bartel, *HHMI/Whitehead Institute/Massachusetts Institute of Technology*

**Keynote Address**

J. Rinn, *Harvard University*

**Noncoding RNAs in Prokaryotes**

*Chairpersons:* J. Doudna, *University of California, Berkeley;*  
J. Vogel, *University of Würzburg, Germany*

**Transcription and Gene Silencing**

*Chairpersons:* M. Buehler, *Friedrich Miescher Institute, Basel, Switzerland;* N. Proudfoot, *University of Oxford, United Kingdom*

**CircRNA and piRNA Biogenesis**

*Chairpersons:* A. Aravin, *California Institute of Technology, Pasadena;* M. Siomi, *University of Tokyo, Japan*

**Biological Function of piRNAs**

*Chairpersons:* R. Ketting, *Hubrecht Institute, Utrecht, The Netherlands;* D. O'Carroll, *EMBL, Monterotondo, Italy*

**miRNAs: Biogenesis and Mechanisms**

*Chairperson:* N. Kim, *Seoul National University, Korea*

**Silencing Pathways in Plants and Other Species**

*Chairpersons:* X. Chen, *HHMI/University of California, Riverside;* S. Jacobsen, *HHMI/University of California, Los Angeles*

**Long Noncoding RNAs and Editing**

*Chairperson:* A. Lambowitz, *University of Texas, Austin*

# Translational Control

---

September 2–6      355 participants

ARRANGED BY      **Thomas Dever**, National Institutes of Health  
**Rachel Green**, Johns Hopkins University  
**Robert Schneider**, New York University School of Medicine

This meeting attracted participants from around the world and included Ruth Lehmann, Jon R. Lorsch, and Robert H. Singer as keynote speakers, eight platform sessions, and three poster sessions that covered 279 abstracts.

Having celebrated more than a quarter century of Translational Control meetings at CSHL, this meeting demonstrated the continued interest and importance of research on protein synthesis and its control in guiding gene expression in both healthy and diseased cells. Research in protein synthesis ranges from studies examining the molecular intricacies of the ribosome to profiling the genome-wide changes in the distribution of translated mRNAs in different tissues. This research also now includes its many manifestations in human

disease, ranging from viral infection to cancer to neurological diseases when regulation goes awry. By bringing together a wide variety of researchers with different areas of expertise, this meeting serves a foundational role for a growing community of established and young investigators.

This year, the program was organized into seven main topics. The first session on Initiation opened on Tuesday. A highlight was the keynote address by Jon Lorsch, an established member of the field and the new director of the National Institute of General Medical Sciences at the National Institutes of Health. In his talk, Jon described proposals to investigate alternate funding programs for researchers, as well as studies from his and collaborator laboratories revealing a high-resolution structure of the ribosome in the act of selecting the start codon for translation. The following sessions included two additional keynote addresses by Rob Singer and Ruth Lehmann, as well as 82 short talks selected from among the submitted abstracts. In addition, 194 abstracts presented as posters in three sessions were well attended in both Bush Hall and the Hershey Building and prompted many discussions and exchange of ideas.



R. Schneider, R. Green, T. Dever



R. Schneider, I. Mohr



W. Gilbert, A. Jacobson



S. Blanchard, A. Ferguson



B. Bukau



R. Luna, A. Marintchev

New technologies had a prominent role in the meeting. Cryo-electron microscopy (EM) structures of initiating and elongating ribosome complexes were presented at near-atomic resolution, reflecting the recent advances in direct detection technologies that have revolutionized the EM field. Ribosomal profiling studies examining the genome-wide distribution of ribosomes on mRNAs in a cell provided new insights into translational control, mRNA turnover, ribosome recycling following termination, and the impact of differentiation or activating signals on the repertoire of translated mRNAs. Intriguingly, two talks revealed linkages between promoter usage for mRNA transcription, localization of the mRNA in a cell, and mRNA translation.

Additional highlights of the meeting included new insights into the mechanism and physiological importance of translational control and an emerging role of translational regulation in disease including viral infection, neurological diseases such as fragile-X syndrome, and many forms of human cancer. Interestingly, translation factors were identified as nexus points for anticancer therapy and novel drug development. Unanticipated functions were reported for components of the translational machinery including a role for ribosomal proteins in directing the translation of key developmental control mRNAs. In a twist of function, microRNAs were shown to activate translation of select mRNAs, and translational changes at egg activation and during development were linked to control of mRNA polyadenylation and to inactivation of translational repressors, respectively. Molecular and biophysical studies provided new insights into ribosomal frameshifting including an intriguing link to microRNAs and to mRNA turnover; and finally, ribosome ubiquitylation was linked to protein quality control, and mRNA decay was reported to have important roles in autoregulatory pathways, selenium bioavailability, stem-cell fate, and T-cell activation.

In summary, the 2014 meeting continues the key role of the Translational Control meeting as the major focal point for the ongoing excitement and growth of this field of research.

## PROGRAM

### Initiation

*Chairperson:* G. Pavitt, *University of Manchester, United Kingdom*

### Turnover

*Chairperson:* T. Inada, *Tohoku University, Sendai, Japan*

### Ribosomes

*Chairperson:* D. Wilson, *University of Munich, Germany*

### Elongation and Termination

*Chairperson:* S. Blanchard, *Weill Cornell Medical School, New York*

### Development

*Chairperson:* E. Gavis, *Princeton University, New Jersey*

### Disease

*Chairperson:* P. Todd, *University of Michigan, Ann Arbor*

### Regulation I

*Chairperson:* C. Calkhoven, *University Medical Center Groningen, The Netherlands*

### Regulation II

*Chairperson:* N. Ingolia, *University of California, Berkeley*



# Epigenetics and Chromatin

September 9–13 366 participants

ARRANGED BY **Shelley Berger**, University of Pennsylvania  
**Robert Kingston**, Massachusetts General Hospital and Harvard Medical School  
**Juerg Mueller**, Max-Planck Institute of Biochemistry, Germany

This second meeting following the successful start of this conference in 2012 brought together researchers working on different aspects of chromatin biology and epigenetics. The focus was on molecular mechanisms by which chromatin and chromatin-modifying protein complexes regulate gene transcription. The spectrum of the oral and poster presentations illustrated that the field has been making progress through studies that range from atomic resolution structure determination of chromatin proteins to the functional analyses of regulatory networks at the organism level.

The high attendance, large numbers of posters, and the sustained involvement of the participants throughout the meeting were testament to its success. Scientists from 31 countries attended, and for many of these, it was their first visit to Cold Spring Harbor. Several of the oral presentations selected from abstracts and given by junior principal investigators, postdocs, or students turned out to be scientific highlights, showing the success of this meeting format as a presentation platform for junior researchers. The emergence of novel mechanistic themes in the field, the progress in methodologies, and the presentations by junior scientists allowed for sustained interest by the large and diverse audience throughout the different sessions.

These sessions covered the function and mechanisms of chromatin-modifying enzymes and the role of chromatin modifications in gene activation and repression (three sessions). The functions of chromatin proteins and noncoding RNAs in genome- and chromosome-wide regulation of gene expression were discussed in two different sessions and the processes that mediate long-range interactions and three-dimensional genome organization were discussed in a further session. The remaining two sessions focused on chromatin dynamics and remodeling and on the role of chromatin in controlling developmental processes at the organism level.



R. Kingston



C. Muchardt, S. Oberdoerffer, P. Hosseini



B. Lesch, J. Cocquet



R. Young

Among the highlights of the meeting were presentations of the first crystal structure of a histone-modifying enzyme bound to a nucleosome, of novel approaches to investigate the dynamics of the genesis of chromatin states following DNA replication, about the function of DNA-binding proteins with “boundary” function in delimiting chromatin domains, or about the role of chromatin-modifying proteins in aging. These talks, together with presentations reporting unexpected findings about the function of particular histone modifications, or snapshots providing insight into the three-dimensional organization of chromosomes, illustrated that chromatin still offers vast uncharted territory that waits to be explored.

## PROGRAM

### ncRNA

*Chairperson:* S. Grewal, *NCI, National Institutes of Health, Bethesda, Maryland*

### Keynote Address: X-Chromosome Inactivation in Development and Cancer

*E. Heard, Institut Curie, Paris, France*

### Enzymes on Chromatin

*Chairpersons:* A. Akhtar, *Max-Planck Institute of Immunology and Epigenetics, Freiburg, Germany*; C. Peterson, *University of Massachusetts Medical Center, Worcester*

### Keynote Address: Analysis of the Role of Spt6 in Controlling Transcriptional Accuracy in Yeast

*F. Winston, Harvard Medical School, Boston, Massachusetts*

### Chromatin Genome-Wide

*Chairpersons:* R. Young, *Whitehead Institute/Massachusetts Institute of Technology, Cambridge*; L. Boyer, *Massachusetts Institute of Technology, Cambridge*

### Covalent Modification/Activation

*Chairpersons:* F. Winston, *Harvard Medical School, Boston, Massachusetts*; J. Tyler, *University of Texas MD Anderson Cancer Center, Houston*

### Long-Range Interactions

*Chairperson:* E. Heard, *Curie Institute, Paris, France*

### Repression

*Chairpersons:* K. Helin, *Biotech Research & Innovation Centre (BRIC), Copenhagen, Denmark*; K. Zaret, *University of Pennsylvania, Perelman School of Medicine, Philadelphia*

### Chromatin Gymnastics

*Chairpersons:* A. Shilatifard, *Stowers Institute for Medical Research, Kansas City, Missouri*; G. Narlikar, *University of California, San Francisco*

### Development and Inheritance

*Chairpersons:* W. Bickmore, *MRC Human Genetics Unit, Edinburgh, United Kingdom*; B. Cairns, *University of Utah Huntsman Cancer Institute, Salt Lake City*



N. Blackledge



I. Gonzalez, C. Alabert



M. Timmers, R. Baas

# Axon Guidance, Synapse Formation, and Regeneration

---

September 16–20 285 participants

ARRANGED BY **Greg Bashaw**, University of Pennsylvania  
**Alain Chedotal**, Vision Institute, INSERM, Paris  
**Yishi Jin**, University of California, San Diego

The human brain has billions of nerve cells (neurons), and each neuron is typically connected to hundreds of other neurons via synapses in a highly precise fashion. This complex neural wiring underlies the ability of humans, and other animals, to interact with the outside world, to learn, and to perform complex behaviors. Defects in the development of neural connections are being increasingly linked to the cause of neurological disease. One of the major challenges for the field of neuroscience is to understand how nerve connections are made accurately and reliably. In the past decade, our understanding of the mechanisms that control axon growth and guidance, synaptogenesis, and the remodeling of neural circuits during development has progressed rapidly from phenomenology to the identification of specific molecular control mechanisms.

This ninth meeting in the series focused on key issues in axon guidance, circuit formation, synaptogenesis and axon regeneration, and included sessions devoted to particular problems in the assembly, plasticity, and repair of the nervous system, with speakers chosen from among the participants submitting abstracts by session chairs who are leaders in the field. The response of the field to this conference was one of overwhelming enthusiasm with more than 285 participants—240 of whom submitted abstracts despite several Gordon Conferences held late in the summer on similar topics and the concurrent CSHL–Asia Neurobiology meeting. Forty-nine abstracts were selected for talks, in seven sessions, the remaining abstracts being presented as posters. Advanced and starting assistant professors, postdoctoral fellows, and graduate students were well represented as speakers and participants. Session chairs at the meeting were well balanced between men and women, and the meeting had a clear international presence with participants from Europe, Australia, and Asia. Many of the outstanding talks this year were by graduate students. All of the major areas of research in the field were covered, as were all of the major approaches (cellular, physiological, anatomical, molecular, dynamic imaging, biochemical, and genetic). In addition, there were three keynote addresses. The first was given by Professor Cori Bargmann, a prominent figure in the field of neuronal development, as well as in circuit contribution to complex behaviors in *Caenorhabditis elegans*. She reviewed her work on peptide modulation of behavior and also described her work as co-chair of the President’s Brain Initiative Committee.

The second and third keynote lectures were a pairing of speakers who have made profound and transformative contributions to dissecting neural circuits underlying complex behaviors. Karl Deisseroth focused on recent advances in optical techniques to dissect circuit function (optogenetics, a field that he founded), as well as the application of “Clarity” to the anatomical analysis and three-dimensional reconstruction of complex tissues. May-Britt Moser spoke on her transformative work on “grid cells,” special cells in the entorhinal cortex that endow mammalian species with a sense of place.



G. Bashaw, Y. Jin, A. Chedotal



E. Stoeckli, Y. Zou, A. Sitko, T. Sakural, J. Frei



B. Grill, A. Sagasti

Also for the second time, this meeting held an informal event on professional development. Topics included how to search for mentors (postdoc or faculty), how to prepare for job interviews (academic or industry), how to write grants, managing teaching and research, how to balance life and work, and how to handle paper reviews. Two special discussion sessions were dedicated to publishing by the attending journal editors and to careers as a physician-scientist by investigators with an M.D. or M.D.-Ph.D. degree and clinical experience. About half of the senior investigators participated in the event. The students and postdocs expressed high appreciation for the close interaction and valuable advice they gained through this event. The business meeting was conducted with open discussion.

Overall, this meeting provided an important forum for ideas and approaches and helped scientists in the field get the most up-to-date information, as well as enabling them to meet, to network, and to establish collaborations. As at the previous meeting, a need was identified for an additional poster session because the number and quality of abstracts submitted for posters was so high. The three poster sessions were extremely well attended. In addition, fewer talks were scheduled in the evening, and a talk session after the evening keynote was eliminated to allow time for discussion and attendee interactions. Based on the enthusiastic comments of the participants, the intensity of the oral and poster sessions, and the large crowds that stayed up late every night at the bar to discuss science, the meeting was a great success.

This meeting was funded in part by National Institute for Neurological Disorders and Stroke, a branch of the National Institutes of Health.

## PROGRAM

### Axon to Synapse I

*Chairpersons:* W. Grueber, *Columbia University, New York*;  
A. Sagasti, *University of California, Los Angeles*

### Synapse to Circuit I

*Chairpersons:* E. Nedivi, *Massachusetts Institute of Technology, Cambridge*; P. Scheiffele, *University of Basel, Switzerland*

### Special Lecture

C. Bargmann, *The Rockefeller University*

### Regeneration and Disease I

*Chairpersons:* A. DiAntonio, *Washington University School of Medicine, St. Louis, Missouri*; L.-H. Tsai, *HHMI/Massachusetts Institute of Technology, Cambridge*

### Axon to Synapse II

*Chairpersons:* F. Charron, *IRCM/McGill University, Montréal, Canada*; G. Lopez-Bendito, *UMH-CSIC, Alicante, Spain*

**Synapse to Circuit II**

*Chairpersons: D. Colon-Ramos, Yale University, New Haven, Connecticut; K. McAllister, University of California, Davis*

**Regeneration and Disease II**

*Chairpersons: C. Bagni, Catholic University of Leuven, Belgium; Z. He, Children's Hospital, Boston, Massachusetts*

**Special Lectures**

*K. Deisseroth, Stanford University, Stanford, California; M.-B. Moser, Kavli Institute for Systems Neuroscience, Norway*

**Axon to Synapse III**

*Chairpersons: S. Guthrie, King's College London, United Kingdom; L. Richards, University of Queensland, St. Lucia, Australia*



L. Boubakar



L. Richards, K. Brose

# The History of Plasmids

---

September 21–23 128 participants

ARRANGED BY **Stanley N. Cohen**, Stanford University School of Medicine  
**Stanley Falkow**, Stanford University School of Medicine  
**Richard Novick**, New York University  
**Dhruba Chattoraj**, National Cancer Institute  
**Chris Thomas**, University of Birmingham, United Kingdom  
**Jan Witkowski**, Cold Spring Harbor Laboratory

This meeting was the third meeting held under the auspices of the Genentech Center for the History of Molecular Biology and Biotechnology at Cold Spring Harbor Laboratory. It was attended by scientists and historians of science drawn from the United States and abroad. Keynote presentations were given on the opening evening by Stanley N. Cohen, Stanley Falkow, and Richard Novick, covering the basic biology of plasmids and their importance in public health and in biotechnology. These broad categories were covered in more detail during the rest of the meeting. Among the many notable participants were three whose work contributed to the biotechnology industry: Stanley Cohen, who, with Herb Boyer, developed plasmids as vectors for cloning; Ananda Chakrabarty, whose Supreme Court case opened the door to the patenting of life forms; and Marc van Montague, who, with Jeff Schell, developed cloning techniques for plants. A highlight of the meeting was a roundtable discussion of the historians' and scientists' views on the history of plasmids and history of science in general.

The meeting was made possible by generous support from the Society for Microbiology, Life Sciences Foundation, MedImmune Inc., Monsanto, and anonymous donors.

## PROGRAM

### Antibiotic Resistance

*Chairperson:* C. Thomas, *University of Birmingham, United Kingdom*

### Bacterial Toxins and Microcins: Plasmids and Pathogenicity

*Chairperson:* G. Jacoby, *Lahey Clinic, Burlington, Massachusetts*

### Plasmid Host Range and Evolution

*Chairperson:* R. Curtiss III, *Arizona State University, Tempe*

### The Nature of Plasmids

*Chairperson:* B. Funnell, *University of Toronto, Ontario, Canada*

### Transposition and Genetic Recombination

*Chairperson:* J. Mekalanos, *Harvard Medical School, Boston, Massachusetts*

### Plasmid-Borne Genes and Metabolic Functions

*Chairperson:* G. Wagner, *Uppsala University, Sweden*

### Roundtable: Historians' and Scientists' Views and General Discussion on the Role of Plasmids in the Overall History of the Biological Sciences

*Chairperson:* J. Witkowski, *Banbury Center, Cold Spring Harbor Laboratory*

### Plasmids and Plants

*Chairperson:* J. Davies, *University of British Columbia, Vancouver, Canada*

### Plasmid DNA Replication

*Chairperson:* S. Brom, *Universidad Nacional Autónoma de México, Cuernavaca, Morelos Mexico*

### Interbacterial Transferability

*Chairperson:* D. Chattoraj, *National Cancer Institute, Bethesda, Maryland*

### Stability, Partitioning, and Post-Segregation Lethality

*Chairperson:* D. Sherratt, *University of Oxford, United Kingdom*

### Plasmids as Whole Systems

*Chairperson:* E. Top, *University of Idaho, Moscow*

### Session 13: Future Directions in Plasmid Biology



Final Group



S. Cohen



J. Watson, M. Yarmolinsky



R. Novick, G. Jacoby



M. Jones, J. Davies



D. Clewell, R. Curtiss

# Molecular Genetics of Aging

September 29–October 3     277 participants

ARRANGED BY

Vera Gorbunova, University of Rochester  
Malene Hansen, Sanford-Burnham Medical Research Institute, La Jolla  
David Sinclair, Harvard Medical School

Aging is the most prominent risk factor for developing a number of serious diseases, ranging from neurodegeneration to cancer. Research from a number of laboratories has shown that the genetic makeup, along with the environment, plays a prominent role in the aging process. Among many significant findings is the identification and characterization of several molecular pathways that influence the rate of aging and susceptibility to age-related disorders in diverse species, from yeast to humans.

This conference provided a stimulating and informative forum for the latest results and emerging ideas in aging research at the molecular, cellular, and organismal levels. A session on longevity genes and human biology provided an update on the latest results on specific gene loci and their age-related effects. Dr. Cynthia Kenyon, who spoke about the new aging-focused company Calico, closed the session. Another session focused on DNA repair and senescence and highlighted new mechanisms that drive cells to stop proliferating with time. A new initiative, the next session centered on comparative biology of aging and nonstandard animal models. This was an interesting session highlighting research using a variety of species, including killifish and naked mole rats. Following the first poster session (featuring more than 100 posters), the next oral session was devoted to how cells maintain homeostasis, for example, by ensuring functional mitochondria. The conference subsequently featured a session on epigenetics, discussing age-related changes in DNA/chromatin organization. An emerging theme that came up in several talks of this session was activation of retrotransposons with age. Two subsequent sessions were devoted to the popular topic of metabolism, and how metabolic alterations have



V. Gorbunova, M. Hansen



L. Guarente, D. Lombard



J. Weise, H. Feldser



H. Aquilaniu, R. Schmookler



strong effects on aging. These sessions were split between a second poster session and a session on immunity and stem cells, emphasizing the importance of these cellular defense mechanisms and compartments for the aging process. The final session focused on discussing nutrient sensing and interventions and highlighted the latest findings on noninvasive approaches to interfering with the aging process.

The conference was attended by researchers from many different fields who collectively contributed to a very vibrant and exciting discussion forum focused on obtaining a better understanding of the molecular mechanisms of aging so that better treatments can be developed against age-related diseases.

This meeting was funded in part by the National Institute on Aging, a branch of the National Institutes of Health; the Glenn Foundation for Medical Research; and Calico.

## PROGRAM

### Longevity Genes/Human Biology

*Chairpersons:* D. Chen, *University of California, Berkeley*;  
R. Shmookler-Reis, *University of Arkansas for Medical Sciences, Little Rock*

### DNA Repair/Senescence

*Chairpersons:* L. Niedernhofer, *Scripps Florida, Jupiter*;  
S. Ahmed, *University of North Carolina, Chapel Hill*

### Comparative Biology/Nonstandard Models

*Chairpersons:* A. Seluanov, *University of Rochester, New York*;  
V. Gladyshev, *Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts*

### Homeostasis

*Chairpersons:* T. Finkel, *NHLBI, National Institutes of Health, Bethesda, Maryland*; A. Dillin, *HHMI/University of California, Berkeley*

### Epigenetics

*Chairpersons:* J. Sedivy, *Brown University, Providence, Rhode Island*; V. Lunyak, *Buck Institute for Research on Aging, Novato, California*

### Metabolism I

*Chairpersons:* D. Cai, *Albert Einstein College of Medicine, Bronx, New York*; J. Auwerx, *Ecole Polytechnique Fédérale de Lausanne, Switzerland*

### Immunity/Stem Cells

*Chairpersons:* A. Wagers, *Harvard University, Cambridge, Massachusetts*; S. Morrison, *HHMI/University of Texas Southwestern Medical Center, Dallas*

### Metabolism II

*Chairpersons:* L. Guarente, *Massachusetts Institute of Technology, Cambridge*; M. Donath, *University Hospital Basel, Switzerland*

### Nutrient Sensing and Interventions

*Chairpersons:* J. Mitchell, *Harvard School of Public Health, Boston, Massachusetts*; S. Simpson, *University of Sydney, Australia*



M. Gill, T. Hoppe



S. Libert, V. Gladyshev

# Germ Cells

October 7–11 198 participants

ARRANGED BY **Stephen DiNardo**, Perelman School of Medicine at the University of Pennsylvania  
**Mary Ann Handel**, The Jackson Laboratory  
**E. Jane Hubbard**, New York University School of Medicine

*Sessions and Science:* The significance of this meeting derives from its emphasis on high-resolution analysis from in vivo studies using cellular, molecular, and/or genomic approaches. Importantly, this is coupled to an examination of germ cell biology comparatively among diverse organisms. The meeting met these historically high standards. Attendance was high, which hit the average of the last six meetings. Such participation in the currently dismal funding climate demonstrates how strongly the community supports this valued meeting and its format.

The co-organizers fashioned several new session themes. For example, exciting progress unifying diverse biological systems led to a novel session on “Evolution and Conservation,” bookended by outstanding invited speaker talks on evolutionary origins of meiosis and germ cell specification (Logsdon, Extavour). Another new session focused on “Energetics and External Influences on the Germ Line,” which featured surprising roles of mitochondria and hypoxia (Capel, Hurd/Lehmann, Oatley), along with additional metabolic influences on germ cells (Arkov, Drummond-Barbosa). These new sessions were complemented by outstanding work across traditionally strong areas including progress on fertilization and sperm receptor candidates (Dean); surprising parallels between vertebrate and invertebrate oogenesis (Greenstein, Lei/Spradling), as well as the nature of conserved germ granules (Seydoux); the plasticity of sexual fate (Matunis, Van Doren); meiotic crossover selection and licensing (Hawley, Campos-Ortera, Yanowitz); and syndromes associated with assisted reproduction (Bartolomei). Of special note was a talk involving high school science outreach with implications for public policy on dietary supplements (O’Reilly).



M.A. Handel, S. DiNardo, E.J. Hubbard



M. Amoyel, L. Wingert



G. Fanourgakis, M. Stanzone

*Diversity:* The meeting was balanced for experimental systems with 22 talks on mouse or human systems, 13 on *Caenorhabditis elegans*, 15 on the fruit fly, and six on other models. This allowed participants to appreciate deeply conserved mechanisms and underscored a great strength of the meeting: significant inclusion of nonmammalian systems and a clear focus on basic biology. Additionally, the meeting provided ample opportunity for early-career scientists: The 56 talks included 10 student, 15 postdoctoral, eight junior, and five mid-career faculty (gender balance, 31 male, 25 female), with nine international speakers. New perspectives were also fostered, as seven of the invited slots had not given invited speaker talks previously.

Finally, two successful changes in format were implemented to encourage discussion. First, the average length of speaking time for each session was reduced, increasing time for questions and discussion. Second, organizers gave CSHL T-shirts to the first two students posing questions at each session in order to encourage questions from this underrepresented group. As a consequence, talks averaged 10–12 questions, with significant time for discussion from various perspectives.

This meeting thus continues to be a great success as judged by the criteria of exemplary science, the breadth and depth of its discussion, and attendance. The co-organizers have always been drawn from the broader germline community, each with keen appreciation of the strengths of this meeting. Thus, the co-organizers selected Bob Braun, who eminently meets these criteria, as one of the 2016 co-organizers. While in the past the current co-organizers would select the three future co-organizers, this year's mandate was to select only one. We are confident that Bob will fill out the co-organizer slate with colleagues who appreciate the necessity of maintaining organism and system diversity.

This meeting was funded in part by National Institute of Child Health and Human Development, a branch of the National Institutes of Health; the Lalor Foundation; and the March of Dimes Birth Defects Foundation.

## PROGRAM

### Germ Cell Specification and Identity

*Chairpersons:* B. Capel, *Duke University Medical Center, Durham, North Carolina*; C. Extavour, *Harvard University, Cambridge, Massachusetts*

### Gametes, Fertilization and Egg Embryo Transition

*Chairpersons:* M. Van Doren, *Johns Hopkins University, Baltimore, Maryland*; J. Yanowitz, *Magee-Women's Research Institute, Pittsburgh, Pennsylvania*



R.S. Hawley, A. Spradling



Q. En, Y.J. Wu

**Protecting the Germ Line and Offspring**

*Chairpersons:* S. Arur, *MD Anderson Cancer Center, Houston, Texas*; X. Chen, *Johns Hopkins University, Baltimore, Maryland*

**Germ Granules and Posttranscriptional Regulation**

*Chairpersons:* D. Katz, *Emory University, Atlanta, Georgia*; M. Wilkinson, *University of California, San Diego*

**Energetics and External Influences on Germ Cells**

*Chairpersons:* J. Dean, *National Institutes of Health, Bethesda, Maryland*; M. Colaiacovo, *Harvard Medical School, Boston, Massachusetts*

**Germline Stem Cells and Niches**

*Chairpersons:* E. Matunis, *Johns Hopkins University School of Medicine, Baltimore, Maryland*; P. Cohen, *Cornell University, Ithaca, New York*

**Meiotic Initiation and Meiosis**

*Chairpersons:* J. Logsdon, *University of Iowa, Iowa City*; M. Lilly, *NICHD, National Institutes of Health, Bethesda, Maryland*

**Evolution and Conservation**

*Chairpersons:* M.L. King, *University of Miami School of Medicine, Florida*; D. Greenstein, *University of Minnesota, Minneapolis*

# Nuclear Receptors and Disease

---

October 28–November 1

141 participants

ARRANGED BY

**Ron Evans**, HHMI/Salk Institute for Biological Studies

**Karen Knudsen**, Thomas Jefferson University

**Anastasia Kralli**, The Scripps Research Institute

**Susanne Mandrup**, University of Southern Denmark

This fifth meeting has been held biennially since 2006. Nuclear receptors are ligand-regulated transcription factors that have key roles in cell growth and differentiation, development, and physiology. Their deregulated activity is implicated in the development or progression of different types of cancer, developmental disorders, and metabolic diseases. Accordingly, nuclear receptors can also serve as targets for therapeutic intervention in many of these diseases. The meeting provided a unique forum for basic and clinical researchers in the field to discuss new developments in areas ranging from basic molecular mechanisms and cellular functions to physiological functions, pathological roles in cancer and metabolic disease, and potential as therapeutic targets.

The meeting started with two outstanding keynote talks that highlighted the molecular mechanisms versus clinical implications overarching theme. Peter Fraser presented new approaches that reveal large-scale chromosome structure, map long-range chromatin interactions, and link three-dimensional organization to control of gene expression. Peter Nelson discussed how changes in androgen receptor signaling can enable prostate cancer to progress to recurrent “castration-resistant” disease. The opening session was followed by eight oral sessions and two poster sessions that continued on the interplay of basic science with roles of nuclear receptors in diseases. Sessions 2 and 4 focused on epigenetic regulation and the roles of chromatin, cofactors, and posttranslational modifications in determining nuclear receptor activity at different cellular, physiological, and disease models. Session 8 furthered the basic mechanism side by reports on genomic binding of receptors and structural insights of receptors and receptor cofactors. Session 7 focused on



R. Evans, S. Mandrup



C. Helsen, D. Boyanova



C. Wall, H. Uhlenhaut



J. Francois, A.E. Swinstead



I. Rogatsky, M. Lazar

cancer, and Sessions 5 and 9 displayed the latest developments in the field of metabolism. Session 10 on the final day closed with talks on inflammation and age-related diseases. Overall, there were 47 talks: 28 long talks from invited speakers and the four organizers and 19 short talks chosen from abstracts and aiming at enhancing exposure of younger investigators. The two poster sessions (73 posters presented in Sessions 3 and 6) further enhanced the program by highlighting research in all areas and creating an active setting for interactions and further discussions.

All sessions were characterized by the presentation of unpublished research and open and lively discussions. The format of this meeting, centered on a class of molecules that function in diverse pathways and are implicated in a wide range of diseases, enabled the cross-fertilization of ideas developed in one biological system but widely applicable to others, the emergence of common themes and concepts across diseases, and a cross-education of experts in cancer, metabolic disease, and age-related diseases that is essential in our efforts to harness the regulatory capacities of these receptors in therapeutic approaches.

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.

## PROGRAM

### Keynote Address: Splicing, Repression, and Co-Option:

**How the AR Retains Command of Prostate Cancer Mutiny After Mutiny**

P. Nelson, *University of Washington School of Medicine, Seattle*

### Keynote Address: Long-Range Interactions, Chromosome Structure, and Genome Organization Links to Gene Expression Control

P. Fraser, *Babraham Institute, Cambridge, United Kingdom*

### Chromatin and Transcriptional Networks

*Chairperson:* R. Evans, *The Salk Institute for Biological Studies, La Jolla, California*

### Cofactors and Mechanisms

*Chairperson:* D. McDonnell, *Duke University, Durham, North Carolina*

### Metabolism

*Chairperson:* A. Kralli, *The Scripps Research Institute, La Jolla, California*

### Cancer

*Chairperson:* K. Yamamoto, *University of California, San Francisco*

### Structure and Ligands

*Chairperson:* C. Rochette-Egly, *IGBMC/INSERM, Illkirch, France*

### Metabolic Disorders

*Chairperson:* M. Lazar, *University of Pennsylvania, Philadelphia*

### Inflammation and Aging

*Chairperson:* B. Desvergne, *University of Lausanne, Switzerland*

# Biological Data Science

---

November 5–8      200 participants

ARRANGED BY      **Anne Carpenter**, Broad Institute of Harvard/Massachusetts Institute of Technology  
**Michael Schatz**, Simons Center for Quantitative Biology, Cold Spring Harbor  
**Matt Wood**, Amazon Web Services

This inaugural conference was the first in a new biennial series focusing on the infrastructure, software, and algorithms needed to analyze large data sets in biological research. It brought together researchers from diverse areas of biology and diverse computational domains, with interest in large-scale data analysis in common. The goal of the meeting was for attendees to step a bit outside their usual domain and glean lessons from other areas of research. In this, it was a success. With a heavy emphasis on genome sequence and transcriptional expression data, attendees nonetheless spanned numerous data-intensive fields such as imaging, proteomics, and text mining. The talks and poster presentations were similarly diverse.

Two hundred researchers participated in the inaugural meeting, with 107 abstracts submitted from academic, government, and industry laboratories by researchers who regularly mine large data sets for genomics, imaging, translational, and clinical projects. There were 10 invited talks and two keynote lectures, as well as 33 additional talks selected from submitted abstracts and 74 poster presentations. The first keynote lecture was given by David Haussler, who described the timeliness, need for, and overall design of the Global Alliance for Genomics and Health. The second keynote was a master lecture given by Kristin Lauter, who described new methods for securely but privately sharing data sets. The meeting participants had a strong presence on Twitter, with all of the talks opting to be discussed on Twitter and live-tweeted by a number of audience members (hashtag #biodata14).



J. Williams, D. Haussler, A. Carpenter, M. Schatz, D. Church

**PROGRAM**

**Keynote Address: Global Exchange of Human Genetic Data for Medicine and Research**  
 D.H. Haussler, *HHMI/University of California, Santa Cruz*

**Data and Data Mining I**  
*Chairpersons:* D. Church, *Personalis, Inc. Menlo Park, California*; M. Massie, *University of California, Berkeley*

**Data and Data Mining II**  
*Chairperson:* O. Troyanskaya, *Princeton University, New Jersey*

**Compute Infrastructure**  
*Chairpersons:* R. Stevens, *Argonne National Lab/University of Chicago, Illinois*; J. Taylor, *Johns Hopkins University, Baltimore, Maryland*

**Algorithmics**  
*Chairpersons:* B. Langmead, *Johns Hopkins University, Baltimore, Maryland*; M. Carneiro, *Broad Institute, Cambridge, Massachusetts*

**Biological Software**  
*Chairpersons:* J. Williams, *Cold Spring Harbor Laboratory*; R. Margolis, *NIDDK, National Institutes of Health, Bethesda, Maryland*

**Master Lecture: Homomorphic Encryption as a Tool to Preserve Privacy in Genomic Computation**  
 K.E. Lauter, *Microsoft Research, Redmond, Washington*

**Human Biology**  
*Chairpersons:* E. Perakslis, *Harvard Medical School, Boston, Massachusetts*; M. Gerstein, *Yale University, New Haven, Connecticut*



S. Payne, S. Tsuji



H. Sussman, A. Mortazavi



A. Cox, Y. Huang



M. Burnett, S. Sadedin



H.T. Wu, D. Ray



# Personal Genomes: Discovery, Treatment, and Outcomes

---

November 12–15 156 participants

ARRANGED BY **Ann Daly**, Newcastle University, United Kingdom  
**Nicholas Katsanis**, Duke University  
**Deanna Kroetz**, University of California, San Francisco  
**James R. Lupski**, Baylor College of Medicine

This seventh meeting again combined both the content of the Personal Genomes Meeting and that of the Pharmacogenomics Meeting to bring together scientific communities interested in gleaning medically actionable information from individual personal genomes. The meeting was entitled “Personal Genomes: Discovery, Treatment, and Outcomes” and was a further step toward Precision Medicine that utilizes genomic information for guiding differential diagnoses and potentially optimizing and directing therapeutic interventions based on individual genetic/genomic variation. Keynote speaker Richard Lifton described a series of clinical examples showing the value of exome sequencing in diagnosis and treatment of rare diseases. He provided a range of fascinating examples and showed clearly that exome sequencing gives valuable insights into clinical issues. For example, sequencing of pseudohyperaldosteronism type II cases revealed that this phenotype is associated with dominant mutations in one of four different genes. Further study of these mutations has led to identification of a novel pathway involved in angiotensin II signaling, and improved understanding of this pathway explains why high dietary potassium lowers blood pressure. Rick showed data that gave insights into how human genetic approaches and the discovery of loss-of-function “protective variants” could provide the gene entry into therapeutic targets. The keynote was followed by two further talks. Misha Angrist provided a personal, thought-provoking and mainly positive perspective on the Personal Genomes project, which he described as “mostly probabilistic but occasionally deterministic.” We then switched to pharmacogenomics with Alan Shuldiner, who described the importance of the *CYP2C19* genotype to the outcome of clopidogrel treatment and the ongoing challenges for clinical implementation. This varied opening session encompassing the two main strands of the meeting was a great start.

On the second day, we covered clinical sequencing and the interpretation of genome variation, with emphasis on moving from populations to individuals, followed by a session on somatic cell genomics with emphasis on cancer. In both of these sessions, the challenges of data interpretation and clinical translation were discussed. In the morning session, one highlight included a study showing how the combination of rare null variants and a common risk allele at a locus could result in the common complex trait of scoliosis. A poster session with 41 posters also took place in the early evening. We then moved to population sequencing, which



J. Lupski, D. Kroetz, A. Daly



G. Oliver, R. Weinshilboum, L. Wang



B. Ozenberger, J. Hu



D. Deacon, S. Waszak, W. Mueller

emphasized multiethnic groups and minority populations, and to pharmacogenomics on day 3. The utility of exome sequencing in pharmacogenomics in addition to Mendelian diseases was demonstrated by studies on warfarin dosing and adverse drug reactions that detected several novel genetic risk factors. The ethics panel session, which is always a feature of the meeting, involved an Oxford-style debate with the motion “Implementation of DNA sequencing in clinical medicine: The bar has been set too high,” with Munir Pirmohamed in favor and Bob Nussbaum against. At the start of the debate, the audience voted in favor of the motion, but by the end of session, they had been persuaded by Bob to reject it. Both Munir and Bob did a terrific job arguing their position!

On the final day, a second keynote lecture was given by Nancy Cox, who spoke on moving beyond coding sequences to investigate regulatory variants using “multiomics” approaches. Most of the examples she showed were concerned with complex diseases, but these new approaches can also be applied to pharmacogenomics and Mendelian disorders. Further talks on moving beyond coding sequences were also included in this exciting final session.

Once again, the meeting was a celebration of the advances in genomic technologies, their application to personal genomes, and how the variation discovered can transition to clinical application. It is clear that excellent progress is being made on many different fronts using genome sequencing, but considerable challenges in both moving into noncoding regions and clinical implementation remain.

This meeting was funded in part by Regeneron Pharmaceuticals, Inc., and Quest Diagnostics/Athena Diagnostics.

## PROGRAM

### Patient Genomes: Discovery and Personal Perspectives

*Chairperson: J. Lupski, Baylor College of Medicine, Houston, Texas*

### Keynote Address: Genes, Genomes, and the Future of Medicine

*R. Lifton, Yale University School of Medicine*

### Clinical Sequencing

*Chairperson: N. Katsanis, Duke University, Durham, North Carolina*

### Interpreting Genome Variation I: From Populations to Individuals

*Chairperson: S. Leal, Baylor College of Medicine, Houston, Texas*

### Somatic Cell Genomics

*Chairperson: D. Kroetz, University of California, San Francisco*

### Population-Based Sequencing

*Chairperson: A. Shuldiner, University of Maryland School of Medicine, Baltimore*

### Pharmacogenetics

*Chairperson: A. Daly, Newcastle University Medical School, United Kingdom*

### Ethics Panel

*Chairperson: N. Katsanis, Duke University, Durham, North Carolina*

**Interpreting Genomic Variation II: Beyond Coding Sequence**

*Chairperson: N. Cox, University of Chicago, Illinois*

**Keynote Address: New Ways of Integrating Genome Variation and Genome Function to Inform Pharmacogenomics Discovery and Translation**

*J. Cox, University of Chicago, Illinois*



E. Woodahl, M. Angrist



F. Ouellette, M. Cherry

# Neurodegenerative Diseases, Biology, and Therapeutics

December 3–6      179 participants

ARRANGED BY      **Karen Duff**, Columbia University Medical Center  
**Jeffrey Rothstein**, Johns Hopkins University School of Medicine  
**John Trojanowski**, University of Pennsylvania School of Medicine

Up to one-half of those aged 85 years or older will develop one or multiple debilitating neurodegenerative diseases (NDDs) of the central nervous system. These diverse NDDs include Alzheimer's disease (AD), Parkinson's disease (PD), frontotemporal lobe degeneration (FTLD), Huntington's disease (HD), and amyotrophic lateral sclerosis (ALS). To reflect some of the newest and most relevant research while providing a diverse base of current advances in understanding these neurodegenerative diseases, the topics in this meeting included prion-like spread of diseases including AD, PD, and tauopathies, new biology of the ALS/FTLD C9ORF72 mutation, new thoughts about biomarkers and comorbidities, the role of RNA in the pathogenesis of neurodegenerative diseases, the role of glia (especially their potential as targets for therapy), and a number of therapeutic targets and approaches including immune therapy, small-molecule drugs, and gene inactivation strategies.

Although most forms of neurodegenerative disease occur in the absence of obvious heritability or identifiable genetic mutations, it has been possible during the past 25 years to discover uncommon genetic mutations as well as risk-modifying DNA changes in some neurodegenerative diseases, and predictable causative changes in others. For some of the neurodegenerative diseases, common genetic abnormalities lead to a spectrum of disease phenotypes, which is clearly seen for FTLD and ALS. New insight into disease mechanisms and possible overlapping therapeutic approaches are implicated. From these findings, transgenic technology has rapidly led to the development of mouse, fruit fly, and nematode model systems that partly recapitulate the clinical



J. Rothstein, K. Duff, J. Trojanowski



K. Spiller, S. Porta, K. Ring



J. Buxbaum, M. Hoffman

abnormalities of the human diseases as well as some of the hallmark molecular and morphological pathology of the conditions. Of particular relevance in this respect is the emergence of new ideas about the spread of disease entities within the brain and from the periphery, similar to that seen in the prion diseases, which has become a dominant topic in AD, FTLN, and PD research. New tools using human-induced pluripotent cells derived from these degenerative diseases were also discussed. This new and exciting field of research was well represented.

This biannual meeting series started in 2000, and as in previous years, the explicit goal of the meeting focuses on identifying disease pathways and facilitating the translation of “break-through” science into effective medicines. At this year’s meeting, numerous academic labs and commercial drug discovery organizations presented data on novel compounds, clinical trial results, or new druggable pathways for AD, PD, FTLN, HD, ALS, spinal muscular atrophy, spinocerebellar ataxia, and prion diseases. Indeed, of the 179 participants, 19 came from industry. In recent years, various molecular, biochemical, and cell-based screens have led to the development of small-molecule, peptide, and oligonucleotide compounds that show promise in neurodegenerative disease models. Importantly, the emergence of gene inactivation strategies using stabilized antisense oligonucleotides has shown particular promise both as a tool for studying pathogenesis and, more importantly, as a clinically relevant therapeutic agent that could be used to knock down pathological proteins or pathways affected in several diseases.

For this meeting, platform sessions were organized around common technological themes. Chairpersons, invited speakers, and speakers selected from submitted abstracts were drawn from the academic and pharma sectors, representing 19 countries. The discussion of new unpublished data was emphasized and adequate time was left for discussion of each presentation. Poster presentations also covered a wide range of neurodegenerative disease pathways, new animal and insect models, and novel therapeutic insights. Posters were displayed for an extended period during the meeting, and poster viewing was especially encouraged during the cocktail hours before the evening meals. In general, the meeting was most notable for its breadth of coverage of different neurodegenerative diseases, insightful and novel presentations, as well as in-depth discussions of each presentation—and the insights gleaned from lively interaction among the diverse participants.

This meeting was funded in part by National Institute of Neurological Disorders and Stroke, a branch of the National Institutes of Health; and Merck Research Laboratories.



I. Ahmed, Y. Wairkar



J. Jankowsky, T. Maniatis

**PROGRAM**

**Prion-Like Neurodegeneration**

*Chairperson: V. Lee, University of Pennsylvania School of Medicine, Philadelphia*

**AD/PD Pathology**

*Chairperson: K. Duff, Columbia University, New York*

**TDP43/C9orf72: RNA Metabolism in Neurodegenerative Disease**

*Chairperson: J. Rothstein, Johns Hopkins University School of Medicine, Baltimore, Maryland*

**Submitted Abstracts**

*Chairperson: J. Trojanowski, University of Pennsylvania School of Medicine, Philadelphia*

**Neurons and Glia: Therapeutic Opportunities**

*Chairperson: S. Przedborski, Columbia University, New York*

**Neurons and Glia: Models**

*Chairperson: R. Ransohoff, Cleveland Clinic, Ohio*

**Submitted Abstracts**

*Chairperson: J. Trojanowski, University of Pennsylvania School of Medicine, Philadelphia*

**Gene Inactivation and Protein-Lowering Strategies**

*Chairperson: T. Miller, Washington University, St. Louis, Missouri*

# Blood Brain Barrier

---

December 10–13 130 participants

ARRANGED BY **Britta Engelhardt**, University of Bern, Switzerland  
**Stefan Liebner**, Goethe University Clinic, Germany  
**Ryan Watts**, Genentech, Inc.

Highlighted at this meeting were advances in the cellular and molecular aspects of blood brain barrier (BBB) development, function, and disease. The conference was well attended, particularly by young scientists (graduate students and postdoctoral fellows), who also had a prominent role in giving oral presentations. This made for a lively discussion of unpublished data, giving foresight into the future of BBB research.

As in years past, the conference had a substantial representation of genetic model organisms, imaging, in vitro models, and powerful molecular genetics that highlighted invertebrate and vertebrate BBB physiology and development. A major focus of the conference was the emerging field of BBB maintenance, with a focus on molecular mechanisms of BBB modulation for drug transport, and a growing focus on the BBB in disease.

Two keynote addresses were given by Richard Daneman on BBB development and function and Maiken Nedergaard on the newly described glymphatic system in regulation of fluid flow in the central nervous system. As is traditional at CSHL meetings, selection of material for oral and poster presentation was made by the organizers and individual session chairs on the basis of scientific merit.

This meeting was funded in part by National Institute of Neurological Disorders and Stroke, a branch of the National Institutes of Health; and Forest Laboratories, Inc.



S. Liebner, R. Watts, B. Engelhardt



R. Dieguez, H.A. Andrews



J. Smith, J. Abbot

## PROGRAM

### BBB Development

*Chairpersons: J. Nathans, Johns Hopkins School of Medicine, Baltimore, Maryland; B. Vanhollebeke, Université Libre de Bruxelles, Gosselies, Belgium*

### The Functional BBB

*Chairpersons: D. Agalliu, University of California, Irvine; C. Gu, Harvard Medical School, Boston, Massachusetts*

### Keynote Address

*R. Daneman, University of California, San Francisco*

### BBB in a Dish

*Chairpersons: E. Shusta, University of Wisconsin, Madison; J. Wong, Houston Methodist Research Institute, Texas*

### Molecular BBB Transport

*Chairpersons: P.-O. Freskgard, F. Hoffmann-La Roche AG, Basel, Switzerland; J. Yu Zuchero, Genentech, South San Francisco, California*

### The Diseased BBB

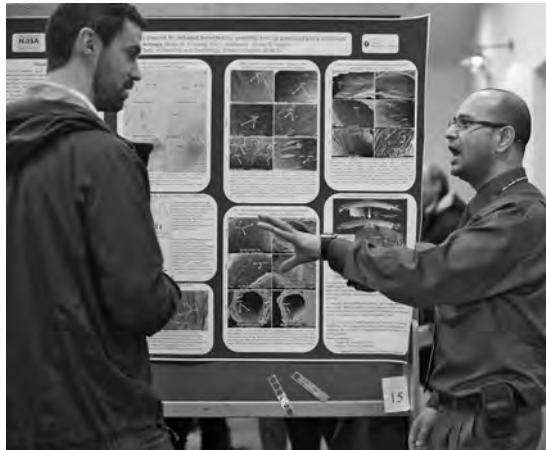
*Chairpersons: R. Klein, Washington University School of Medicine, St. Louis, Missouri; R. Bell, Pfizer, Inc., Cambridge, Massachusetts*

### Keynote Address

*M. Nedergaard, University of Rochester Medical Center*

### Pericytes, Astrocytes, and the BBB

*Chairpersons: J. Abbott, King's College London, United Kingdom; B. Appel, University of Colorado School of Medicine, Aurora*



M. Allingham, N. Acharya



F. Elwood, M. Allingham



# POSTGRADUATE COURSES

---

## Workshop on Leadership in Bioscience

---

March 14–16

**INSTRUCTORS**      C. Cohen, Science Management Associates, Newton, Massachusetts  
                              D. Kennedy, Workshop Co-Facilitator, Cambridge, Massachusetts

This highly interactive 3.5-day workshop develops the skills students need to lead and interact effectively with others, in both one-on-one and group settings. It focused on techniques, situations, and challenges that relate specifically to leading and managing in the scientific workplace. It also emphasized learning by doing and involved role-playing, giving and receiving feedback, and group problem solving. Much of the learning was peer-to-peer. Participants were expected to discuss their own experiences and listen to others as they discuss theirs.

The workshop helped participants identify areas where they needed guidance and growth, as well as how to capitalize on areas of strength. Participants had the opportunity to share their experiences and challenges with others and to receive feedback and guidance from others with experience in leading scientists in a variety of settings. At the end of the course, participants were linked through a unique on-line community in which they could continue learning from one another and from the course instructors.

This workshop was supported with funds provided by American Express Philanthropy.



PARTICIPANTS

Anczukow-Camarda, O., Ph.D., Cold Spring Harbor  
Laboratory

Balu, D., Ph.D., McLean Hospital, Belmont, Massachusetts

Cash, K., Ph.D., Northeastern University, Boston,  
Massachusetts

Crenshaw, S., Ph.D., University of North Carolina, Chapel  
Hill

De Las Penas, A., Ph.D., IPICYT, Mexico

De Marco Garcia, N., Ph.D., Weill Cornell Medical College,  
New York

Dixon, S., Ph.D., Stanford University, California

Dumond, J., Pharm., University of North Carolina Eshelman  
School of Pharmacy, Chapel Hill

Dunn, F., Ph.D., University of Washington, Seattle

Eisenhoffer, G., Ph.D., University of Utah, Salt Lake City

Fredriksson, L., Ph.D., Karolinska Institutet, Sweden

Hon, G., Ph.D., Ludwig Institute for Cancer Research,  
San Diego, California

Knodler, L.L., Ph.D., Washington State University,  
Pullman

Koles, K., Ph.D., Brandeis University, Waltham, Massachusetts

Limonta, D., Ph.D., Pedro Kouri Institute of Tropical  
Medicine, Cuba

Malone, J., Ph.D., University of Connecticut, Storrs

Payen, C., Ph.D., University of Washington, Seattle

Pi, H.J., Ph.D., Cold Spring Harbor Laboratory

Sansom, S., Ph.D., University of Oxford, United Kingdom

Short, S., Ph.D., University of North Carolina, Chapel Hill

Valdes-Ferrer, S., Ph.D., Feinstein Institute for Medical  
Research, Manhasset, New York

Vogel, C., Ph.D., New York University, New York

Weidner, J., Ph.D., Karolinska Institutet, Sweden

Will, B., Ph.D., Albert Einstein College of Medicine, Bronx,  
New York

Wong, A., Ph.D., Janelia Farm, Ashburn, Virginia

Zheng, X., Ph.D., University of North Carolina, Chapel Hill

# Protein Purification and Characterization

---

April 1–15

**INSTRUCTORS**

- A. Courey, University of California, Los Angeles
- S.-H. Lin, MD Anderson Cancer Center/University of Texas, Houston
- M. Marr, Brandeis University, Waltham, Massachusetts
- S. Nechaev, University of North Dakota School of Medicine, Grand Forks

**ASSISTANTS**

- A. Abbas, University of North Dakota School of Medicine, Grand Forks
- J. Cao, University of California, Los Angeles
- M. Chambers, University of California, Los Angeles
- M. Donovan, Brandeis University, Waltham, Massachusetts
- Y.-C. Lee, MD Anderson Cancer Center, Houston, Texas
- H.-C. Liu, MD Anderson Cancer Center, Houston, Texas
- A. Samarakkody, University of North Dakota School of Medicine, Grand Forks
- M. Spellberg, Brandeis University, Waltham, Massachusetts

This course was for scientists who were not familiar with techniques of protein isolation and characterization. The course was extremely rigorous and included laboratory work during the day, as well as lectures with discussions and student talks in the evenings. Students are typically graduate students, postdoctoral scholars, staff scientists, and professors with specialized scientific expertise who now need to learn about protein purification and characterization. The course emphasized laboratory strategies and current best practices in the field.

Each student became familiar with major techniques in protein purification by performing four separate isolations: (1) a regulatory protein (calmodulin) from mammalian tissue (cow brain);



(2) a sequence-specific DNA-binding protein (transcription factor AP1) from HeLa cell nuclei; (3) a recombinant protein overexpressed as inclusion bodies in *Escherichia coli*; and (4) a membrane-bound protein (insulin receptor) from rat liver.

Students were divided into four groups of four and spent 3 days in each of the above four modules. In addition to the primary purification in a given module, a number of relevant characterizations were performed on each protein, giving students experience with immunological and biochemical assays, peptide mapping, protein sequencing, and mass spectrometry. Students learned bulk fractionation and electrophoretic and chromatographic techniques, including precipitation by salts, pH, and ionic polymers; ion exchange, gel filtration, hydrophobic interaction, and reverse phase chromatography; lectin affinity, ligand affinity, oligonucleotide affinity, and immunoaffinity chromatography; polyacrylamide gel electrophoresis and electroblotting; and high-performance liquid chromatography. Students also learned procedures for solubilizing proteins from inclusion bodies and refolding them into active monomeric forms and for solubilizing membrane proteins by detergent screens. Evenings in the course involved lectures about topics that complement the laboratory curriculum: protein structure; modification of proteins; methodologies for protein purification, stabilization, and characterization; high-throughput protein purification; and applications of protein biochemistry to molecular cell biology and cancer research.

This course was supported with funds provided by the National Cancer Institute.

## PARTICIPANTS

Bakman, I., B.Sc., Brooklyn College, The City University of New York, Brooklyn

Blaby, C., Ph.D., University of California, Los Angeles

Bowling, J., B.A., St. Jude Children's Research Hospital, Memphis, Tennessee

Cirak, S., Ph.D., Children's National Medical Centre, Washington D.C.

Criado-Marrero, M., B.S., Ponce School of Medicine, Ponce, Puerto Rico

Dahmen, J., Ph.D., Matrix Genetics, Seattle, Washington

Doerfel, M., Ph.D., Cold Spring Harbor Laboratory

Jones, J., B.S., Idaho State University Meridian Health Science Center, Meridian

Lansky, S., B.S., The Hebrew University of Jerusalem, Israel

Masocha, W., Ph.D., Kuwait University, Safat

Nielsen, N., M.S., University of Oslo, Norway

Rafique, S., Ph.D., Centre for Applied Molecular Biology, Thokar niaz baig, Lahore, Pakistan

Roda, R., Ph.D., New York University, New York

Turner, B., M.S., University of Kansas, Lawrence

Uddin, Y., B.S., Georgia Institute of Technology, Atlanta

Wu, Q., Ph.D., Cold Spring Harbor Laboratory

## SEMINARS

Courey, A., University of California, Los Angeles: Welcome and Introduction to Course. System-wide analyses of Groucho and SUMO in *Drosophila*.

Pappin, D., Cold Spring Harbor Laboratory: Introduction to mass spectrometry of proteins. Quantitative approaches to mass spectrometry of proteins.

Love, J., Albert Einstein School of Medicine, New York: High-throughput production of eukaryotic soluble and membrane proteins.

Marr, M., Brandeis University, Boston, Massachusetts: Introduction to protein purification. Multilevel control of gene expression response to cellular stress.

Nechaev, S., University of North Dakota, Grand Forks: Analysis of EMT genes reveals a role of polII pausing in gene activation.

Lin, S.-H., MD Anderson Cancer Center, Houston, Texas: Secretome analysis of prostate cancer bone metastasis.

# Quantitative Imaging: Cells to Molecules

---

April 2–15

**INSTRUCTORS**     **H. Elliott**, Harvard Medical School, Cambridge, Massachusetts  
**M. Krummel**, University of California, San Francisco  
**H. Shroff**, National Institutes of Health, Bethesda, Maryland  
**J. Waters**, Harvard Medical School, Boston, Massachusetts

**ASSISTANTS**     **K. Corbin**, University of California, San Francisco  
**A. Ettinger**, University of California, San Francisco  
**A. Kumar**, National Institutes of Health, Bethesda, Maryland  
**T. Lambert**, Harvard Medical School, Boston, Massachusetts  
**H. Pinkard**, University of California, San Francisco

This course focused on advanced quantitative fluorescence microscopy techniques used for imaging a range of biological specimens, from cells to single molecules. The course was designed for cell and molecular biologists with little or no microscopy experience who wish to begin using microscopy in their own research. Students gained a theoretical understanding of, and hands-on experience with, state-of-the-art equipment used in quantitative fluorescence microscopy, including laser-scanning and spinning-disk confocal microscopy, deconvolution methods, total internal fluorescence microscopy (TIRF), superresolution methods (structured illumination, STORM, and PALM), multiphoton microscopy, light sheet microscopy, and digital image processing and analysis. Students learned how to design and implement a wide range of imaging experiments using these techniques. They also used the techniques to address specific quantitative questions and then discussed the results daily as a group, learning to troubleshoot the common problems that occur in the course of a quantitative imaging experiment. Among the lectures presented



were microscopy basics, image analysis basics, CCD and sCMOS, cameras, confocal microscopy, multiphoton microscopy, deconvolution, TIRF, single-molecule imaging, imaging ratiometric “biosensors” (including FRET), light sheet microscopy, and superresolution techniques. Students also learned guidelines for choosing fluorescent proteins and worked with live samples requiring environmental control. Additional lecturers for 2014 included George Patterson (National Institutes of Health), Paul Goodwin (GE Healthcare), and Adam Cohen (Harvard University).

This course was supported with funds provided by the National Cancer Institute.

## PARTICIPANTS

Banala, S., Ph.D., Janelia Farm Research Campus, Ashburn, Virginia	Law, K., B.S., Icahn School of Medicine at Mount Sinai, New York
Duquette, P., B.Sc., McGill University, Montreal, Canada	Nelson, S., B.S., University of Colorado School of Medicine, Aurora
Eun, Y., Ph.D., Harvard University, Cambridge, Massachusetts	Ondracka, A., Diplomat., The Rockefeller University, New York
Finn, C., Ph.D., Rocky Mountain Laboratories, Hamilton, Montana	Renvoise, B., Ph.D., NINDS, National Institutes of Health, Bethesda, Maryland
Heinritz, J., B.S., Yale University/HHMI, New Haven, Connecticut	Ross, K., Ph.D., University of Strathclyde, Glasgow, United Kingdom
Hempel, A., Ph.D., University of California, San Diego	Ruckh, T., Ph.D., Northeastern University, Boston, Massachusetts
Johnson, R., B.S., SUNY Albany and Wadsworth Center NYSDOH, Albany, New York	Slaidina, M., Ph.D., New York University School of Medicine, New York
Jordi, J., Ph.D., University of Zurich, Switzerland	
Komorowski, M., Ph.D., Polish Academy of Sciences, Warszawa, Poland	

## SEMINARS

Cohen, A., Harvard University, Cambridge, Massachusetts: Bringing bioelectric phenomena to light.	Transmitted light microscopy. Fluorescence microscopy. Quantitative digital imaging. Live-cell imaging.
Patterson, G., National Institutes of Health, Bethesda, Maryland: Fluorescent proteins.	Krummel, M., University of California, San Francisco: Confocal microscopy. Live confocal microscopy, FRAP, and photoactivation. TIRF. The immune response in four dimensions. Multiphoton.
Shroff, H., National Institutes of Health, Bethesda, Maryland: Superresolution imaging. Light sheet microscopy. Faster and sharper: New technologies for 4D imaging.	Goodwin, P., Applied Precision, Inc., Issaquah, Washington: Deconvolution.
Elliott, H., Harvard Medical School, Cambridge, Massachusetts: Image segmentation. Ratio imaging and FRET. Image analysis.	Lambert, T., Harvard Medical School, Boston, Massachusetts: Digital Cameras.
Waters, J., Harvard Medical School, Boston, Massachusetts: Quantitative microscopy basics. Objective lenses.	

# Cell and Developmental Biology of *Xenopus*

---

April 4–15

**INSTRUCTORS**     K. Liu, Kings College, London, United Kingdom  
A. Sater, University of Houston, Texas  
G. Thomsen, Stony Brook University, Stony Brook, New York

**ASSISTANTS**     R. Miller, The University of Texas Medical School, Houston  
J. Oomen-Hajagos, Stony Brook University, Stony Brook, New York  
B. Steventon, Institut Pasteur, Paris, France  
T. Nakayama, University of Virginia, Charlottesville

*Xenopus* is the leading vertebrate model for the analysis of gene function in development. The combination of lineage analysis, gene-knockout strategies, experimental manipulation of the embryo, and genomic/bioinformatic techniques makes it ideal for studies on the molecular control of embryo patterning, morphogenesis, and organogenesis. Moreover, recent advances in *Xenopus* genomics offer new opportunities to integrate computational strategies with experimental approaches, including genome editing-based strategies for gene knockout. The course combined intensive laboratory training with daily lectures from recognized experts in the field. Students learned both emerging technologies and classical techniques to study gene function in *Xenopus* development. An important element was the informal interaction between students and course faculty.

Technologies covered included oocyte and embryo culture, lineage analysis and experimental manipulation of embryos, time-lapse imaging of morphogenesis, gain- and loss-of-function analysis using mRNAs, antisense oligonucleotides, and the CRISPR/Cas9 system, whole-mount in situ



hybridization, immunocytochemistry, genomics and bioinformatics, chromatin immunoprecipitation, preparation of transgenic embryos, and use of *Xenopus tropicalis* for genetic analyses. This course was designed for those new to the *Xenopus* field, as well as for those wanting a refresher course in the emerging technologies. The course was open to investigators from all countries.

This course was supported with funds provided by the National Institute of Child Health and Human Development.

#### PARTICIPANTS

Benlamara, S., B.S., Institute of Systems and Synthetic Biology, Evry, France

Buescher, J., Ph.D., Nationwide Children's Hospital–Research Institute, Columbus, Ohio

Chang, L.-S., M.S., German Cancer Research Center, Heidelberg, Germany

Easterling, M., B.S., Washington State University, Pullman

Guay, J., B.S., Tufts University, Medford, Massachusetts

Hayes, M., B.S., University of Iowa, Iowa City

Kudra, R., B.A., University of Connecticut, Storrs

Mazzeo, I., M.S., Universita di Trento, Trento, Italy

Michal, J., M.S., Washington State University, Pullman

Pasternak, M., M.S., MRC Laboratory of Molecular Biology, Cambridge, United Kingdom

Talhouarne, G., B.S., Carnegie Institution for Science, Baltimore, Maryland

Van Itallie, E., B.A., Harvard Medical School, Boston, Massachusetts

Warren, C., B.A., Albert Einstein College of Medicine, Bronx, New York

Yaklichkin, S., Ph.D., Baylor College of Medicine, Houston, Texas

#### SEMINARS

Sater, A., University of Houston, Texas: microRNAs and the specification of the *Xenopus* eye field. *Xenopus*: History and prospects.

El-Hodiri, H., Ohio State University, Columbus: The retinal homeobox (Rx) gene plays essential roles in the regenerating retina.

Cousin, H., University of Massachusetts, Amherst: The *Xenopus* cranial neural crest: The perfect model system to study collective cell migration and craniofacial evo-devo.

Thomsen, J., Stony Brook University, Stony Brook, New York: The organizer.

Wallingford, J., University of Texas, Austin: The awesome power of living imaging in *Xenopus*.

Liu, K., Kings College, London, United Kingdom: Craniofacial anomalies: Modeling gene function in the neural crest.

Gilchrist, M., National Institute for Medical Research, London, United Kingdom: Using high-resolution RNA-Seq data to untangle mechanisms of gene activation in early development.

Khokha, M., Yale University, New Haven, Connecticut: Gene discovery in humans with birth defects, functional analysis in *Xenopus*: A powerful combination.

Klein, P., University of Pennsylvania School of Medicine, Philadelphia: Mechanisms of Wnt signaling in early development.

Miller, R., The University of Texas Medical School, Houston: Shaping kidney tubules through Wnt signaling and cilia.

Keller, R., University of Virginia, Charlottesville: Cell motility and mechanobiology of early *Xenopus* morphogenesis.

Grainger, R., University of Virginia, Charlottesville: Analysis of eye formation in *Xenopus*: Model system for genetic and genomic studies of determination, induction, and organogenesis.

Blythe, S., Princeton University, New Jersey: Imminent catastrophe: Coordination of DNA replication and zygotic genome activation.



# Workshop on Schizophrenia

---

June 4–10

**INSTRUCTORS**     A. **Abi-Dargham**, Columbus University College of Physicians & Surgeons, New York  
                              J. **Hall**, Cardiff University School of Medicine, Cathays, United Kingdom  
                              A. **Sawa**, Johns Hopkins University, Baltimore, Maryland

**ASSISTANTS**        K. **Delevich**, Cold Spring Harbor Laboratory  
                              T. **Lancaster**, Cardiff University, United Kingdom

This workshop provided students with the most current understanding of the molecular, cellular, and neural systems underlying the disturbances in brain function in these devastating illnesses. During the 7-day workshop, students learned about the clinical aspects of schizophrenia, schizoaffective disorder, and bipolar disorder. They also explored in detail the genetic and neurobiological underpinnings of these complex psychiatric disorders. The Workshop included sessions focused on the clinical syndrome, basic neurobiology, cognitive neuroscience, neuroimaging, genetics and genomics, endophenotypes, and gene expression and gene modulation. In addition to hearing about the most recent research in these areas, controversial topics and challenges to basic assumptions in the field were explored and discussed. A diverse faculty 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 did it help them build the foundation for their future research, it also introduced them to many potential collaborators working to understand schizophrenia and similar disorders from different perspectives. Although these 7 days featured intense lecture sessions throughout, 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 a tennis court.

This Workshop was supported with funds provided by Cold Spring Harbor Laboratory and the Stanley Foundation.

#### PARTICIPANTS

Andrade, G., Ph.D., Albert Einstein College of Medicine, Bronx, New York  
 Argyelan, M., Ph.D., The Zucker Hillside Hospital, Glen Oaks, New York  
 Birkenow, M., M.S., University of Copenhagen, Denmark, Valby, Denmark  
 Bohlken, M., Ph.D., University Medical Center, Utrecht, The Netherlands  
 Boyd, F., B.S., University of Arizona, Phoenix  
 Choueiry, J., Ph.D., University of Ottawa, Canada  
 Clifton, N., Ph.D., Cardiff University, United Kingdom  
 Deakin, J., M.B.Phil., University of Cambridge, United Kingdom  
 Ferraro, L., M.S., Section of Psychiatry, Palermo, Italy  
 Gibson, L., Ph.D., Temple University, Philadelphia, Pennsylvania

Lam, M., Ph.D., Institute of Mental Health, Singapore  
 Mihaljevic, M., M.D., Clinic of Psychiatry, Clinical Centre of Serbia, Belgrade, Serbia  
 Odonez, A., M.D., NIMH, National Institutes of Health, Bethesda, Maryland  
 Parker, K., Ph.D., University of Iowa, Iowa City  
 Passeri, E., Ph.D., Johns Hopkins, Baltimore, Maryland  
 Primerano, A., M.D., Medical Genetics University Hospital Mater Domini, Catanzaro, Italy  
 Sarpal, D., M.D., Hofstra North Shore–LIJ School of Medicine, Glen Oaks, New York  
 Scharf, S., F. Hoffmann-La Roche Ltd., Basel, Switzerland  
 Schnakenberg, A., Ph.D., Indiana University, Bloomington  
 Solis, M., Ph.D., Schizophrenia Research Forum, Seattle, Washington  
 Sykes, L., Ph.D., Cardiff University, United Kingdom

#### SEMINARS

Hall, J., Cardiff University School of Medicine, Cathays, United Kingdom: Course overview.  
 Murray, R., Kings College, London, United Kingdom: Clinical syndrome and risk factors.  
 Yolken, B., Johns Hopkins Children's Center, Baltimore, Maryland: Epidemiology of schizophrenia and psychiatric disorders.  
 Cannon, T., Yale University, New Haven, Connecticut: Biomarkers of vulnerability and progression in the psychosis prodrome.  
 Seidman, L., Massachusetts General Hospital, Boston: Lifespan developmental evolution of neurocognition in schizophrenia.  
 Fletcher, P., University of Cambridge, United Kingdom: A cognitive understanding of psychosis.  
 Weinberger, D., Lieber Institute for Brain Development, Baltimore, Maryland: Epistasis in schizophrenia genetics: What's missing is not heritability.  
 Corvin, A., Trinity College Dublin, Dublin, Ireland.  
 Malhotra, A., The Zucker Hillside Hospital, Glen Oaks, New York: Pharmacogenetic approaches to schizophrenia.

Law, A., University of Colorado, Denver, Aurora: Neuregulin/ErbB signaling in neurodevelopment, cognition, and schizophrenia: From human to rodent.  
 Lewis, D., University of Pittsburgh, Pennsylvania.  
 Akbarian, S., Icahn School of Medicine at Mount Sinai, New York: Epigenetic dysregulation in schizophrenia.  
 Sawa, A., Johns Hopkins University School of Medicine, Baltimore, Maryland.  
 Brennand, K., Mount Sinai School of Medicine, New York: Using induced pluripotent stem cells to study schizophrenia.  
 Moore, H., Columbia University, New York.  
 Meyer-Lindenberg, A., University of Heidelberg, Germany.  
 Abi-Dargham, A., Columbia University, New York: Imaging dopamine transmission in schizophrenia with PET. Careers in Academia: Highlights and challenges.  
 Desmond, N., National Institutes of Health, Bethesda, Maryland.  
 O'Donnell, P., Pfizer, Cambridge, Massachusetts: Careers in the pharmaceutical industry.

# Advanced Bacterial Genetics

---

June 4–24

**INSTRUCTORS**     **D. Hughes**, Uppsala University, Sweden  
                              **B. Lazazzera**, University of California, Los Angeles  
                              **F. Yildiz**, University of California, Santa Cruz

**ASSISTANTS**        **S. Elgamal**, The Ohio State University, Columbus  
                              **E. Garmendia**, Uppsala University, Sweden  
                              **D. Zamorano Sanchez**, University of California, Santa Cruz

This course presented logic and methods used in the genetic dissection of complex biological processes in diverse bacteria. Laboratory methods included classical mutagenesis using transposons, mutator strains, and chemical mutagens; recombineering with single- and double-stranded DNA; detection of gene expression changes using various reporter genes; the mapping of mutations using genetic and physical techniques; modern approaches to the generation and analysis of targeted gene disruptions and fusions using polymerase chain reaction (PCR) and cloning methods; epitope insertion mutagenesis; site-directed mutagenesis; and fluorescence microscopy. Key components of the course used sophisticated genetic methods in the analysis of model bacteria (including *Escherichia coli*, *Salmonella*, *Bacillus subtilis*, and *Vibrio cholerae*), and the use of the wealth of new genomic sequence information to motivate these methods. Invited lecturers presented various genetic approaches to study bacterial mechanisms of metabolism, development, and pathogenesis.

The course admitted 16 students, both foreign and United States residents, from diverse backgrounds and career levels for intensive (but fun) instruction in microbial genetics.

This course was supported with funds provided by the National Science Foundation.



## PARTICIPANTS

Acar, H., B.S., Institute of Science and Technology,  
Klosterneuburg, Austria

Aubee, J., B.S., Howard University College of Medicine,  
Washington, D.C.

Coates, J., B.S., Emory University, Atlanta, Georgia

de Gier, J.-W., Ph.D., Stockholm University, Sweden

Harty, C., B.S., Geisel School of Medicine at Dartmouth,  
Hanover, New Hampshire

Jensen, L., B.Sc., Technical University of Denmark, Denmark

Kingsbury, D., DACVII, University of California, Davis

Kumar, V., Ph.D., Stony Brook University, Stony Brook,  
New York

Teschler, J., B.S., University of California, Santa Cruz

Turnbull, K., B.Sc., Newcastle University, Newcastle upon  
Tyne, United Kingdom

van Vliet, S., M.Sc., Eawag, Dübendorf, Switzerland

Wei, Z., B.S., Harvard University, Cambridge,  
Massachusetts

Williams, T., B.S., University of North Carolina, Charlotte

Wolff, K., Ph.D., Merck Research Laboratories, Kenilworth,  
New Jersey

Zemke, A., Ph.D., University of Pittsburgh, Pennsylvania

Zhou, J., Ph.D., Dartmouth College, West Lebanon,  
New Hampshire

## SEMINARS

Camilli, A., Tufts University, Boston, Massachusetts: Genetic  
analysis of the *Vibrio cholera* life cycle.

Gross, C., University of California, San Francisco: Reaching  
for functional genomics in bacteria.

Kearns, D., Indiana University, Bloomington: A bifunctional  
flagellar clutch disables motility in *Bacillus subtilis* biofilm.

Zhu, J., University of Pennsylvania, Philadelphia: Genetic  
tools to study gene expression during bacterial pathogen  
infection.

Blokesch, M., Swiss Federal Institute of Technology,  
Lausanne, Switzerland: Visualizing the DNA uptake  
process in naturally competent *Vibrio cholera*.

Laub, M., Massachusetts Institute of Technology,  
Cambridge: Mapping, rewiring, and evolving protein–  
protein interfaces in bacteria.

Gitai, Z., Princeton University, New Jersey: Mechano-  
microbiology: How mechanical forces shape bacterial cells  
and communities.

# Mouse Development, Stem Cells, and Cancer

---

June 4–24

**INSTRUCTORS** X. Sun, University of Wisconsin, Madison  
R. Johnson, MD Anderson Cancer Center, Houston, Texas

**CO-INSTRUCTORS** D. Wellik, University of Michigan Medical Center, Ann Arbor  
M. Lewandoski, National Cancer Institute, Frederick, Maryland

**ASSISTANTS** H. Adams, MD Anderson Cancer Center, Houston, Texas  
Y. Furuta, RIKEN, Center for Development Biology, Kobe, Japan  
R. Kadzik, University of Pennsylvania, Perleman School of Medicine, Philadelphia  
B. Larsen, University of Michigan, Ann Arbor  
D. Rux, University of Michigan, Ann Arbor  
Y. (Tina) Zhang, University of Wisconsin, Madison

This intensive lecture and laboratory course was designed for biologists 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, embryonic, adult, and induced pluripotent stem cells, and cancer biology. The laboratory portion of the course provided an extensive hands-on introduction to engineering of mouse models, stem cell technologies, and molecular analysis of mouse embryos and tissues.

The experimental techniques included pronuclear microinjection, isolation, culture, and manipulation of pre- and postimplantation embryos; embryo transfer; embryo electroporation and roller culture; embryo and sperm freezing; genetic manipulation of embryonic stem cells including genome by CRISPR/Cas9, production of chimeras, the generation and differentiation of



embryonic stem cells as well as induced pluripotent stem cells; isolation of mouse embryonic fibroblasts; teratoma formation assay; vibratome and cryosectioning; in situ hybridization; immunostaining; vascular injections; skeletal preparation; organ explant culture; confocal imaging; and live time-lapse microscopy of early embryos, explanted tissues, and tumors in vivo.

This course was supported with funds provided by the National Cancer Institute.

## PARTICIPANTS

Balmer, S., Ph.D., Icahn School of Medicine, Mount Sinai, New York  
 Bishop, K., Ph.D., Maine Medical Center, Scarborough  
 Carmona-Fontaine, C., Ph.D., Memorial Sloan-Kettering Cancer Center, New York  
 Cheng, L., M.S., Mayo Clinic, Rochester, North Carolina  
 Chung, Y.R., M.S., Memorial Sloan-Kettering Cancer Center, New York  
 Dingwall, H., B.S., Harvard University, Cambridge, Massachusetts  
 Gordon, P., B.B.M.E., University of Utah, Salt Lake City

He, M., Ph.D., Cold Spring Harbor Laboratory  
 Korde, A., M.S., Ph.D., Louisiana State University, Baton Rouge  
 Lim, X., Ph.D., Institute of Medical Biology, Singapore  
 Osorio Da Silva, L., Ph.D., The University of Hong Kong, Hong Kong  
 Plaza Reyes, A., M.S., Karolinska Institutet, Stockholm, Sweden  
 Prunskaitė-Hyyryläinen, R., M.Sc., Ph.D., University of Oulu, Finland  
 Ramjee, V., M.D., University of Pennsylvania, Philadelphia

## SEMINARS

Ralston, A., University of California, Santa Cruz; Early mouse embryo-derived stem cells and chimeras.  
 Tabin, C., Harvard Medical School, Boston, Massachusetts: Using mouse genetics to explore the development and evolution of vertebrate morphology.  
 Largaespada, D., University of Wisconsin, Madison: *Sleeping Beauty* and cancer.  
 Threadgill, D., North Carolina State University, Raleigh: Quantitative trait analysis in mice.  
 Solter, D., Institute of Molecular Biology, Singapore: Historical overview of mouse as a model system. Mouse preimplantation.  
 Wellik, D., University of Michigan Medical Center, Ann Arbor: Hox genes: Musculoskeletal development and repair.  
 Yancopoulos, G., Regeneron Pharmaceuticals, Inc., Tarrytown, New York: Growth factors and cancer.  
 Hang, H., The Jackson Laboratory, Bar Harbor, Maine: CRISPR/Cas9 and pluripotency.  
 Maillard, I., University of Michigan, Ann Arbor: Mouse hematopoiesis.  
 Rivera, J., University of Massachusetts Medical School, Worcester: Postimplantation mouse development.  
 Huang, J., Cold Spring Harbor Laboratory: Mapping neural circuits.  
 Sanes, J., Harvard University, Cambridge, Massachusetts: Neural genetics.

Capecchi, M., University of Utah School of Medicine, Salt Lake City: Homologous recombination and knockout mice.  
 Lewandoski, M., NCI, National Institutes of Health, Frederick, Maryland: Mouse genetics Technologies I. Somitogenesis.  
 Egeblad, M., Cold Spring Harbor Laboratory: Cancer imaging in mice.  
 Klein, O., University of California, San Francisco: Dental and intestinal stem cells.  
 Tam, P., Children's Medical Research Institute, Australia: Mouse gastrulation.  
 Trainor, P., Stowers Institute for Medical Research, Kansas City, Missouri: Mouse craniofacial development and disease.  
 Soriano, P., Mt. Sinai School of Medicine, New York: Signaling. Mouse genetics Technologies II.  
 Johnson, R., MD Anderson Cancer Center, Houston, Texas: Liver cancer.  
 Lovell-Badge, R., MRC National Institute for Medical Research, United Kingdom: Mouse sex determination.  
 Lowe, S., Memorial Sloan-Kettering Cancer Center, New York: Probing cancer drivers and dependencies using mouse models and RNAi.  
 Huppert, S., Cincinnati Children's Hospital Medical Center, Ohio: Liver development.  
 Sun, X., University of Wisconsin, Madison: Lung.

# Ion Channels and Synaptic Transmission

---

June 4–24

**INSTRUCTORS**      **S. Brenowitz**, Janelia Farm Research Campus/HHMI, Auburn, Virginia  
**I. Duguid**, University of Edinburgh, United Kingdom  
**P. Kammermeier**, University of Rochester Medical Center, New York

**CO-INSTRUCTOR**      **T. Branco**, University College London, United Kingdom

**ASSISTANTS**      **A. Che**, University of Connecticut, Storrs  
**K. Kuchibhotla**, New York University School of Medicine, New York  
**J. Lueck**, University of Iowa Carver College of Medicine, Iowa City  
**J. Schiemann**, University of Edinburgh, United Kingdom  
**V. Stempel**, Charite Berlin, Germany  
**N. Wanaverbecq**, Air-Marseille University, Marseille, France

The primary goal of this course was to investigate, through lectures and laboratory work, the properties of ion channels that allow neurons to carry out their unique physiological functions in a variety of neural systems. Areas of particular interest included channels that (1) are activated by the neurotransmitter at central and peripheral synapses, (2) are activated by voltage changes in axons and dendrites, (3) respond to neuromodulators with changes in functional properties, (4) are developmentally required and regulated, or (5) are light-gated and engineered to express in specific neural subtypes. The research interests of guest lecturers reflect these areas of emphasis.

The laboratory component of the course introduced students to state-of-the-art electrophysiological approaches for the study of ion channels in their native environments. Hands-on exercises



included patch-clamp recording of ion channel activity in acutely isolated or cultured cells or neurons in brain slice preparations. Different recording configurations were used (e.g., whole-cell, cell-attached, and dendritic patches, and using voltage- and current-clamp configurations) to examine macroscopic or single-channel activity. Similarly, various methods of ligand and drug application were demonstrated. The advantages and disadvantages of each method, preparation, and recording technique were considered with respect to the specific scientific questions being asked. Admissions priority is given to students and postdocs with a demonstrated interest, specific plans, and a supportive environment to apply these techniques to a defined problem.

This course was supported by the Howard Hughes Medical Institute.

## PARTICIPANTS

Anatharam, A., Ph.D., Wayne State University, Detroit, Michigan

Bishop, H., B.S., University of California, Davis

Costa, K., M.S., Goethe University Frankfurt, Frankfurt am Main, Germany

Daie, K., Ph.D., Weill Cornell Medical College, New York

Evans, D., B.Sc., MRC Laboratory of Molecular Biology, Cambridge, United Kingdom

Fok, A., Ph.D., New York University School of Medicine, New York

Mersch, D., Ph.D., University of Lausanne, Switzerland

Mironova, Y., B.S., University of Michigan, Ann Arbor

Morrie, R., B.S., University of California, Berkeley

Wang, L., B.S., University of Michigan, Ann Arbor

Wulf, M.-A., M.Med., University Hospital Zurich, Switzerland

Yan, L., B.S., The Scripps Research Institute, San Diego, California

## SEMINARS

Lee, A., Janelia Research Campus, Ashburn, Virginia:

Whole-cell recordings in freely moving rodents.

Lee, A., University of Iowa, Iowa City: Cav1 Ca<sup>2+</sup> channel at ribbons synapses.

Plested, A., Leibniz Institute for Molecular Pharmacology, Berlin, Germany: Superactivation of AMPA receptors by auxiliary proteins.

Montell, C., University of California, Santa Barbara: Control of animal behavior and decision making with TRP channels.

Nimigeon, C., Weill Cornell Medical College, New York: Insights into gating and selectivity mechanisms in potassium channels from prokaryotic channels.

Shepherd, G., Northwestern University, Chicago, Illinois: Synaptic circuit organization of mouse motor cortex.

Colecraft, H., Columbia University, New York: Ion channel modulation. Calcium channel engineering.

Sjostrom, J., McGill University, Montreal, Quebec, Canada: Spike-timing-dependent plasticity in neocortex.

Trussell, L., Oregon Health and Science University, Portland: Electrical synaptic transmission redefines the principal cell–interneuron relationship in an auditory microcircuit.

Isom, L., University of Michigan, Ann Arbor: Role of sodium channel SCN1A and CN1B in inherited epilepsy.

Shah, M., University College London, United Kingdom: Presynaptic HCN channels in the entorhinal cortex: Function, plasticity, and trafficking.

Xu-Friedman, M., University of Buffalo, SUNY, New York: Regulation of release probability at an auditory synapse.

Hausser, M., University College London, United Kingdom: Single-neuron computation.

Kammermeier, P., University of Rochester Medical Center, Rochester, New York: Intro/testing potential.

Duguid, I., University of Edinburgh, Germany: Intro/resting potential.

Nicoll, R., University of California, San Francisco: Long-term potentiation: 40 years on.

Brenowitz, S., Janelia Farm Research Campus/HHMI, Auburn, Virginia: V-clamp-Rs/analysis-Igor.

Oertner, T., Center for Molecular Neurobiology Hamburg (ZMNH), Germany: Synapses: Killed by their lazy lifestyle?





Grigorev, K., Bioinformatics Institute of Saint Petersburg, Russia  
Jarvis, J., Ph.D., Coriell Institute for Medical Research,  
Camden, New Jersey  
Lezhnina, K., M.S., M.V. Lomonosov Moscow State  
University, Russia  
Makeeva, O., Ph.D., Head of Research Planning Group,  
Tomsk, Russia  
McEachin, Z., Ph.D., Emory University/Georgia Institute of  
Technology, Atlanta  
Miranda-Angulo, A., Ph.D., University of Antioquia,  
Medellin, Colombia  
Rotunno, M., Ph.D., National Institutes of Health, Bethesda,  
Maryland

Saha Mandal, A., Ph.D., University of Calgary, Calgary,  
Canada  
Terry-Lorenzo, R., Ph.D., Sunovion Pharmaceuticals,  
Marlborough, Massachusetts  
Torres, J., Ph.D., University of Chicago, Illinois  
Umoh, M., B.A.S., Emory University, Atlanta, Georgia  
Verbel, D., M.P.H., Eisai Inc., Woodcliff Lake, New Jersey  
Wang, W., Ph.D., University of Minnesota, Twin Cities,  
Minneapolis  
Xu, K., Ph.D., Bristol Myers-Squibb Company, Pennington,  
New Jersey

#### SEMINARS

Sinsheimer, J., David Geffen School of Medicine at the  
University of California, Los Angeles: Statistics 101.  
Borecki, I., Washington University School of Medicine,  
St. Louis, Missouri: Introduction to genetic epidemiology.  
Loos, R., Mount Sinai School of Medicine, New York:  
Gene–environment interaction and common disease.  
Breen, G., Kings College London, United Kingdom: miRNA  
and microsatellites.  
Al-Chalabi, A., Kings College, and Laura Almasy, Texas  
Biomedical Research Institute: Informal session. Basic  
association strategies. Any additional topics.  
Neale, B., Massachusetts General Hospital, Boston: Exome  
arrays and imputation.

McCombie, W., Cold Spring Harbor Laboratory: Next-  
generation sequencing.  
Bestor, T., Columbia University, New York: Epigenetic basis  
of genotype-independent phenotypic variation.  
Fromer, M., Icahn School of Medicine at Mount Sinai, New  
York: Fully leveraging exome sequencing data for disease  
association.  
Sham, P., The University of Hong Kong, Hong Kong, China:  
Power considerations.  
Knight, J., Centre of Addiction and Mental Health, Toronto,  
Canada: Complex association strategies: Using additional  
data.

# Statistical Methods for Functional Genomics

---

June 18–July 1

**INSTRUCTORS**     **H. Bussemaker**, Columbia University, New York  
**S. Davis**, National Institutes of Health, Bethesda, Maryland  
**O. Elemento**, Weill Cornell Medical College, New York

**ASSISTANTS**     **V. FitzPatrick**, Columbia University, New York  
**E. Giannopoulou**, Weill Cornell Medical College, New York  
**C. Rastogi**, Columbia University, New York

During 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 understood 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.

Detailed lectures and presentations by instructors and guest speakers were combined with hands-on computer tutorials. The methods covered in the lectures were applied to example high-throughput data sets.

This course was supported with funds provided by the National Institute of General Medical Sciences.

## PARTICIPANTS

Alexander Rascon, C., M.Phil., The Rockefeller University,  
New York

Almada, A., Ph.D., Harvard University, Cambridge,  
Massachusetts



Blythe, S., Ph.D., Princeton University, New Jersey  
 Cedernaes, J., M.D., Uppsala University, Sweden  
 De Kumar, B., Ph.D., Stowers Institute for Medical Research,  
 Kansas City, Missouri  
 Deshotel, M., B.S., University of Utah, Salt Lake City  
 Fitzgerald, D., B.A., SUNY at Albany, New York  
 Hassan, S., Ph.D., Oregon Health and Science University,  
 Portland  
 Hu, Y., Ph.D., Bristol-Myers Squibb, Pennington, New Jersey  
 Huisman, S., B.S., Delft University of Technology, The  
 Netherlands  
 Jacob, L., Ph.D., Memorial Sloan-Kettering Cancer Center,  
 New York  
 James, T., M.Sc., Karolinska University Hospital, Stockholm,  
 Sweden  
 Konopacki, C., B.S., Gerstner Sloan-Kettering Graduate  
 School, New York

McGeachy, A., B.A., Berkeley University, California  
 Medway, C., Ph.D., Mayo Clinic, Jacksonville, Florida  
 Murtha, M., Ph.D., New York University School of  
 Medicine, New York  
 Nettles, K., Ph.D., The Scripps Research Institute, Jupiter,  
 Florida  
 Nurmi, E., Ph.D., University of California, Los Angeles  
 Pashos, E., Ph.D., University of Pennsylvania, Philadelphia  
 Schmidt, J., M.Sc., Friedrich Miescher Institute for  
 Biomedical Research, Basel, Switzerland  
 Storvall, H., M.S., Karolinska Institutet, Stockholm, Sweden  
 Walrad, P., Ph.D., University of York, Heslington, United  
 Kingdom  
 Wells, C., Ph.D., Clemson University, Clemson, South  
 Carolina  
 Will, B., Ph.D., Albert Einstein College of Medicine, Bronx,  
 New York

## SEMINARS

Culhane, A., Harvard University School of Public Health,  
 Boston, Massachusetts: Exploratory data analysis.  
 Kundaje, A., Stanford University, California: Lecture and Lab:  
 Working with ENCODE and Roadmap epigenomics data.  
 Leslie, C., Memorial Sloan-Kettering Cancer Center, New  
 York: Posttranscriptional regulation.  
 Hodges, E., Cold Spring Harbor Laboratory: DNA  
 methylation, noncoding transcription, and a basis for  
 identifying active enhancers.  
 Stamatoyannopoulos, J., University of Washington School  
 of Medicine, Seattle: Decoding the human genome—  
 Part I: Mapping and analysis of regulatory DNA.  
 Decoding the human genome—Part II: Encoding  
 complex regulatory programs.  
 Pickrell, J., New York Genome Center, New York: Joint  
 analysis of functional genomic data and genome-wide  
 association studies of 18 human traits.

Kinney, J., Cold Spring Harbor Laboratory: Using deep  
 sequencing to characterize the biophysical mechanisms of a  
 transcriptional regulatory sequence.  
 Carlson, M., Fred Hutchinson Cancer Research Center,  
 Seattle, Washington: Bioconductor annotation and data  
 resources.  
 Reimers, M., Virginia Commonwealth University,  
 Richmond: Epigenomics. Integrative analysis.  
 Irizarry, R., Johns Hopkins Bloomberg School of Public  
 Health, Baltimore, Maryland: Batch effects, experimental  
 design, and methylation data analysis.  
 Lappalainen, T., New York Genome Center/Columbia  
 University, New York: Understanding functional variation  
 in the human genome with transcriptome sequencing.

# Frontiers of Techniques in Plant Science

---

June 27–July 17

**INSTRUCTORS**     M. Johnson, Brown University, Providence, Rhode Island  
S.-H. Shiu, Michigan State University, E. Lansing  
M. Timmermans, Cold Spring Harbor Laboratory

**ASSISTANTS**     P. Atkins, University of Minnesota, St. Paul  
Y. Chen, Michigan State University, E. Lansing  
V. Jawahir, University of Missouri, St. Louis  
N. Panchy, Michigan State University, E. Lansing

This course provided an intensive overview of topics in plant genetics, physiology, biochemistry, development, and evolution as well as hands-on experiences in molecular, analytical, computational, and high-throughput approaches to understanding plant biology. It emphasized recent results from model organisms including *Arabidopsis*, maize, tomato, and a variety of other plants and provided an introduction to current methods used in basic and applied plant biology. The course was designed for scientists with some experience in molecular techniques or in plant biology who wished to work with plants using the latest technologies. The course consisted of a vigorous lecture series, hands-on laboratories, informal discussions, and social activities designed to foster an intense but enjoyable learning environment. The instructors and a stellar group of speakers who are acknowledged leaders in their fields presented up-to-the-moment research and led laboratory sessions on a wide range of topics in plant research.



The seminar series included plant morphology and anatomy, development, evolution, light and circadian 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, transient gene expression, applications of fluorescent proteins, genome editing, and chromatin immunoprecipitation. Students gained hands-on experience on computational tools and environments for genome assembly, transcriptomics, construction of gene regulatory networks, identification of quantitative trait loci, mapping by sequencing, and mathematical modeling of development and hormone action. The course included several short workshops on important themes in plant research. Throughout the course, students interacted individually and informally with the speakers to further enrich the learning experience.

This course was supported with funds provided by the National Science Foundation.

#### PARTICIPANTS

- Akoz, G., M.Sc., University of Veterinary Medicine and GMI, Vienna, Austria  
 Campbell, J., B.A., University of Minnesota, St. Paul  
 Castro Guerrero, N., Ph.D., University of Missouri, Columbia  
 Choudury, S., B.S., The Ohio State University, Columbus  
 Conrad, J., B.A., Dartmouth College, Hanover, New Hampshire  
 Duarte, G., Ph.D., State University of Campinas (Unicamp), Campinas, Brazil  
 Gray, S., Ph.D., University of California, Davis  
 Keshishian, E., Ph.D., Auburn University, Auburn, Alabama  
 Lee, S., B.S., Rutgers, The University of New Jersey, New Brunswick  
 Lee, U.-S., M.S., University of Pennsylvania, Philadelphia  
 Luginbuehi, L., M.S., John Innes Centre, Norwich, United Kingdom  
 Monroe, J., B.S., Colorado State University, Fort Collins  
 Rains, M., B.Sc., Algoma University and Queen's University, Sault Ste. Marie, Canada  
 Swift, J., M.S., New York University, New York  
 Volkova, P., M.A., Russian Institute of Agricultural Radiology and Agroecology, Obninsk, Russia  
 Warczak, T., Ph.D., Dartmouth College, Hanover, New Hampshire.

#### SEMINARS

- Kramer, E., Harvard University, Cambridge, Massachusetts: How-to-EvoDevo: Pitfalls and considerations in comparative development. *Aquilegia* as a test case for developing new model systems.  
 Grotewold, E., Ohio State University, Columbus: Control of plant gene expression: From the parts to the network.  
 Howe, G., Michigan State University, E. Lansing: Plant–insect interactions.  
 Doebley, J., University of Wisconsin, Madison: Evolutionary genetics of maize.  
 Maloof, J., University of California, Davis: Light signaling.  
 Sinha, N., University of California, Davis: Introduction to plant biology. Using gene coexpression networks to identify developmental regulators.  
 Peck, S., University of Missouri, Columbia: New insights into non-self recognition between hosts and pathogens.  
 Brady, S., University of California, Davis: Regulation of root development in tomato.  
 Harmer, S., University of California, Davis: Circadian rhythms.  
 Brutnell, T., Donald Danforth Plant Science Center, St. Louis, Missouri: A systems approach to understanding and engineering C4 photosynthesis.  
 Grossniklaus, U., University of Zurich, Switzerland: Plant reproduction from a female gametophyte's perspective.

# *Drosophila* Neurobiology: Genes, Circuits, and Behavior

---

June 29–July 17

**INSTRUCTORS**     K. Kaun, Brown University, Providence, Rhode Island  
G. Macleod, Florida Atlantic University, Jupiter  
K. O'Connor-Giles, University of Wisconsin, Madison  
A. Rothenfluh, University of Texas Southwestern Medical Center, Dallas

**ASSISTANTS**     J. Bruckner, University of Wisconsin, Madison  
E. Imler, University of Arizona, Tucson  
S. Gratz, University of Wisconsin, Madison  
A. Narayanan, University of Texas Southwestern Medical Center, Dallas

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. The two and a half-week course was designed to introduce students to a wide variety of topics and techniques, including the latest approaches to study 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 function or focus on specific techniques and approaches to study fly neurobiology. 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 around student-led projects, where gene mutants are discovered and analyzed throughout the course utilizing the different morphological and physiological measurements and behavioral paradigms at hand. This included electrophysiological and *in vivo* calcium recordings, anatomical examination, and circuit mapping, as well as numerous quantitative behavioral



measures. Collectively, the course provided a comprehensive and practical introduction to modern experimental methods for studying the *Drosophila* nervous system.

The course was supported by funds provided by the National Institute on Drug Abuse and the Howard Hughes Medical Institute.

#### PARTICIPANTS

Arzan Zarin, A., Ph.D., University of Oregon/HHMI,  
Eugene

Braco, J., B.S., Wake Forest University, Winston-Salem,  
North Carolina

Chin, J., B.Sc., Temasek Life Sciences Laboratory,  
Singapore

Davis, S., B.A., Baylor College of Medicine, Houston, Texas

Fischer, C., B.S., Yale University, New Haven, Connecticut

Giachello, C., Ph.D., University of Manchester, United  
Kingdom

Gonzalez, A., B.S., Universidad Central del Caribe,  
Bayamon, Puerto Rico

Louis, T., Ph.D., The Scripps Research Institute, Jupiter,  
Florida

Privman, E., B.S., University of Virginia, Charlottesville  
Soto Padilla, A., M.Sc., University of Groningen, The  
Netherlands

Vaccaro, A., B.S., ESPCI Paris Tech, Paris, France

Wan, Y., B.S., Janelia Farm Research Campus/HHMI,  
Ashburn, Virginia

#### SEMINARS

Rothenfluh, A., University of Texas Southwestern, Dallas:  
Using genetics to study behavior in *Drosophila*: Alcohol as  
a case study.

Keene, A., University of Reno, Nevada: Circadian rhythms  
and sleep in *Drosophila*.

Collins, C., University of Michigan, Ann Arbor: Axon injury  
and repair.

Lee, C.-h., NICHD, National Institutes of Health, Bethesda,  
Maryland: Visual system development.

McKellar, C., and Albin, S., Janelia Farm Research Campus,  
Ashburn, Virginia: Brain anatomy and genetic tools for  
mapping circuits underlying behavior.

Kravitz, E., Harvard University, Cambridge, Massachusetts:  
Aggression in *Drosophila*.

Levitan, E., University of Pittsburgh, Pennsylvania: Imaging  
dense core vesicles and neurosecretion.

Turner, G., Cold Spring Harbor Laboratory: Studying neural  
activity in the central nervous system.

Macleod, G., Florida Atlantic University, Jupiter: Calcium  
and pHlorin imaging to measure neuronal activity.

Darnell, B., The Rockefeller University, New York: Applying  
advanced techniques in molecular neurosciences to the  
human brain.

Dubnau, J., Cold Spring Harbor Laboratory: *Drosophila*  
learning and memory.

O'Connor-Giles, K., University of Wisconsin, Madison:  
Genetic approaches to understanding synaptogenesis.

Rolls, M., Penn State University, University Park,  
Pennsylvania: Neuronal polarity.



# Advanced Techniques in Molecular Neuroscience

---

July 1–17

**INSTRUCTORS**      C. Lai, Indiana University, Bloomington  
J. Loturco, University of Connecticut, Storrs

**ASSISTANTS**      S. Cerceo-Page, Lieber Institute at Johns Hopkins, Baltimore, Maryland  
F. Chen, University of Connecticut, Storrs  
C. Fricano, Dartmouth College, Lebanon, New Hampshire  
M. Girgenti, University of Connecticut, Storrs  
K. Griffin, Indiana University, Indianapolis  
J. Lubner, University of Connecticut, Storrs  
S. Maloney, Washington University School of Medicine, St. Louis, Missouri  
A. Mitchell, Icahn School of Medicine at Mount Sinai, New York  
M. Rannals, Lieber Institute for Brain Development, Baltimore, Maryland  
D. Roosien, University of Michigan, Ann Arbor  
M. Sakai, Indiana University, Bloomington  
K. Sakers, Washington University School of Medicine, St. Louis, Missouri  
T. Tekieli, University of Michigan, Ann Arbor  
M. Williams, Dartmouth College, Lebanon, New Hampshire  
A. Zyburu, Indiana University, Bloomington

This laboratory and lecture course was designed to provide neuroscientists at all levels with a conceptual and practical understanding of several of the most advanced techniques in molecular neuroscience. 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 topics such as an introduction to the design and use of animal virus vectors in neurobiology; the use of CRISPR genome editing and RNA interference (RNAi) approaches for regulating the expression of specific genes in neurons; practical exercises in gene delivery systems including mammalian cell infection and transfection and electroporation techniques for targeted gene transfer *in vivo*; an introduction to overall strategies, use, and design of bacterial artificial chromosome (BAC) transgenic vectors; real-time reverse transcription–polymerase chain reaction (RT-PCR) analyses; assays of chromatin and chromatin structure in neurons; and mRNA isolation from specified neural subtypes by TRAP.

This course was supported with funding from the Howard Hughes Medical Institute.

#### PARTICIPANTS

Asiminas, A., B.Sc., The University of Edinburgh, United Kingdom  
 Behm, M., M.S., Stockholm University, Sweden  
 Cancer, M., M.Sc., Uppsala University, Sweden  
 Carter, C., Ph.D., University of Iowa/HHMI, Iowa City  
 de Lartigue, G., Ph.D., University of California, Davis  
 Harkness, J., B.A., Oregon Health & Science University, Portland  
 Jurkin, J., Ph.D., Institute of Molecular Biotechnology, Vienna, Austria  
 Kim, C., Ph.D., Janelia Farm Research Institute/HHMI, Ashburn, Virginia

Lin, T., B.S., Georgia Regents University, Augusta  
 McMichael, G., M.Phil., The University of Adelaide, North Adelaide, Australia  
 Nie, E., B.S., University of California, Los Angeles  
 Osorio, J., M.D., University of Rochester Medical Center, New York  
 Punal, V., Ph.D., Duke University, Durham, North Carolina  
 Thomas, E., B.S., Yale University, New Haven, Connecticut  
 Vayndorf, E., Ph.D., University of Alaska, Fairbanks  
 Yu, X.-W., Ph.D., National Institutes of Health, Bethesda, Maryland

#### SEMINARS

Shaeffer, A., Icahn School of Medicine at Mount Sinai, New York: Epigenetic control of neuronal functional diversity.  
 Barres, B., Stanford University School of Medicine, California: What do astrocytes do?  
 Stevens, B., Children's Hospital Boston, Harvard Medical School, Boston, Massachusetts: Wiring and unwiring the brain: Role of glia and the classical complement cascade.

Mayford, M., The Scripps Research Institute, La Jolla, California: Genetic control of memory representation.  
 Darnell, R., The Rockefeller University/HHMI, New York: Applying advanced techniques in molecular neuroscience to the human brain.  
 Akbarian, S., Icahn School of Medicine at Mount Sinai, New York: Exploring higher-order chromatin and regulatory noncoding DNA in the brain.

# Single-Cell Analysis

---

July 2–15

**INSTRUCTORS** J. Eberwine, University of Pennsylvania Perelman School of Medicine, Philadelphia  
C. McMurray, University of California, Berkeley

**ASSISTANTS** P. Buckley, Janssen Research & Development, LLC, Spring House, Pennsylvania  
J. Hyman, Lawrence Berkeley National Laboratory, Berkeley, California  
K. Frankel, Lawrence Berkeley Laboratory, Berkeley, California  
D.-Y. Lee, Lawrence Berkeley Laboratory, Berkeley, California  
J. Morris, University of Pennsylvania, Philadelphia  
A. Polyzos, Lawrence Berkeley National Laboratory, Berkeley, California

The goal of this 2-week course was to familiarize students with the most recent cutting-edge technologies for characterization of single cells. Important in this process was highlighting the advantages of analyzing single cells in isolation and in their natural microenvironment. Sections of the course were taught by scientists who are experts in particular areas of single-cell analysis. Topics covered included quantitative single-cell analysis by RNA-Seq, genomic DNA analysis, epigenetics, RNA localization analysis, proteomics, protein translation, and metabolomics. The techniques taught included real-time live-cell quantification where appropriate. In addition, multiple techniques to isolate specific cell populations and individual cells were taught. The course included the use of model systems such as mouse, *Drosophila*, *Caenorhabditis elegans*, *Aplysia*, and *Planaria*.

This course was supported with funding from the Howard Hughes Medical Institute.



## PARTICIPANTS

- Chatzi, A., Ph.D., Weill Cornell Medical College, New York  
 Dergalev, A., Sp.S., A.N. Bach Institute of Biochemistry,  
 Moscow, Russia  
 Grabocka, E., Ph.D., New York University Langone Medical  
 Center, New York  
 Hard, J., B.S., Karolinska Institute, Solna, Sweden  
 Kaminskaya, A., Ph.D., Federal Medical Biological Agency,  
 Saint Petersburg, Russia  
 Lobanova, E., Ph.D., Albert Eye Research Institute, Durham,  
 North Carolina  
 Ma, K., B.S., The University of Texas, Austin  
 Naguro, I., Ph.D., The University of Tokyo, Japan  
 Peschansky, V., Ph.D., University of Miami, Miller School of  
 Medicine, Florida  
 Salmen, F., M.S., Royal Institute of Technology, Solna, Sweden
- Scurry, A., B.S., University of Nevada School of Medicine,  
 Reno  
 Shoykhet, M., Ph.D., Washington University School of  
 Medicine, St. Louis, Missouri  
 Soderberg, L., M.S., Royal Institute of Technology, Solna,  
 Sweden  
 Teng, M., B.S., Sanger Institute, Cambridge, United  
 Kingdom  
 Vanaerschot, M., M.S., Institute of Tropical Medicine,  
 Antwerp, Belgium  
 Yamauchi, K., B.S., University of California, Berkeley  
 Zheng, F., B.A., Purdue University, West Lafayette,  
 Indiana  
 Zigoneanu, I., Ph.D., University of North Carolina,  
 Chapel Hill

## SEMINARS

- Larabell, C., Lawrence Berkeley Laboratory, Berkeley,  
 California: X-ray high-resolution imaging.  
 McMurray, C., Lawrence Berkeley Laboratory, Berkeley,  
 California: Metabolomics.  
 Eberwine, J., University of Pennsylvania, Philadelphia:  
 Transcriptomics.  
 Sweedler, J., University of Illinois, Chicago: Proteomics.  
 Kim, J., University of Pennsylvania, Philadelphia:  
 Biocomputation.
- McConnell, M., University of Virginia, Charlottesville:  
 Genomics.  
 Albritton, N., University of North Carolina, Chapel Hill:  
 Cell manipulation.  
 Sims, P., Columbia University, New York: MicroFluidics.  
 Finkbeiner, S., University of California, San Francisco: Cell  
 dynamics.

# Computational Neuroscience: Vision

---

July 11–24

**INSTRUCTORS**     **G. Boynton**, University of Washington, Seattle  
**G. Horwitz**, University of Washington, Seattle  
**J. Pillow**, The University of Texas, Austin  
**S. Treue**, German Primate Center, Goettingen, Germany

Computational approaches to neuroscience will produce important advances in our understanding of neural processing. Prominent success will come in areas where strong inputs from neurobiological, behavioral, and computational investigation can interact. The theme of the course was that an understanding of the computational problems, the constraints on solutions to these problems, and the range of possible solutions can help guide research in neuroscience. Through a combination of lectures and hands-on experience with MATLAB-based computer tutorials and projects, this intensive course examined visual information processing from the retina to higher cortical areas, spatial pattern analysis, motion analysis, neuronal coding and decoding, attention, and decision-making.

This course was supported with funding from the Howard Hughes Medical Institute.

## PARTICIPANTS

Bakst, L., Ph.D., University of Washington, Seattle  
Bondy, A., Ph.D., NEI, National Institutes of Health,  
Bethesda, Maryland

Hür, J., M.S., German Primate Center, Goettingen,  
Germany  
Itthipuripat, S., Ph.D., University of California, San Diego



- Kar, K., Ph.D., Rutgers University, Newark, New Jersey  
 Katz, L., Ph.D., University of Texas, Austin  
 Kim, M., Ph.D., New York University, New York  
 Morimoto, M., M.S., Janelia Farm/HHMI, Ashburn, Virginia  
 Nitzany, E., M.S., Weill Cornell Medical College/Cornell University, Ithaca, New York  
 Olsson, C., Ph.D., New York University, New York  
 Ponce, C., Ph.D., Harvard Medical School, Boston, Massachusetts  
 Popovkina, D., Ph.D., University of Washington, Seattle  
 Ravi, S., Ph.D., University of Southern California, Los Angeles  
 Roth, N., B.A., University of Pennsylvania, Philadelphia  
 Sabastian, S., B.S., University of Texas, Austin
- Sedigh-Sarvestani, M., Ph.D., University of Pennsylvania, Philadelphia  
 Snow, M., M.D./Ph.D., Albert Einstein College of Medicine, Bronx, New York  
 Tikidji-Hamburyan, A., Ph.D., Stanford University, California  
 Trott, A., Ph.D., Harvard Medical School, Boston, Massachusetts  
 Vincent, J., Ph.D., University of Washington, Seattle  
 Vincent, J., Ph.D., Harvard Medical School, Boston, Massachusetts  
 Wei, X.-X., Ph.D., University of Pennsylvania, Philadelphia  
 Yao, X., Ph.D., University of Southern California, Monterey Park  
 Yates, J., Ph.D., The University of Texas, Austin

## SEMINARS

- Movshon, J.A., New York University, New York: The elements of early vision.  
 Pillow, J., The University of Texas, Austin: Intro to neural encoding models and generalized linear models (GLMs).  
 Chichilnisky, E.J., The Salk Institute, La Jolla, California: Retina.  
 Simoncelli, E., New York University, New York: Encoding/decoding of visual information.  
 Horwitz, G., University of Washington, Seattle: White noise analysis of V1 neurons.  
 Boynton, G., University of Washington, Seattle: fMRI in the LGN and V1. Psychophysics/signal detection theory.  
 Palmer, S., University of Chicago, Hyde Park, Illinois: Information in the early visual system.  
 Heeger, D., New York University, New York: Normalization as a canonical neural computation.  
 Cohen, M., University of Pittsburgh, Pennsylvania.  
 Li, Z., Cornell University, Ithaca, New York.  
 Priebe, P., Northwestern University, Chicago, Illinois: Spike threshold, inhibition, and stimulus selectivity in primary visual cortex.
- Brainard, D., University of Pennsylvania, Philadelphia: Computational color vision.  
 Fairhall, A., University of Washington, Seattle.  
 Olshausen, B., University of California, Berkeley: Computational models of perception and scene analysis.  
 Rust, N., University of Pennsylvania, Philadelphia: The neural mechanisms responsible for identifying and finding objects.  
 DeAngelis, G., University of Rochester, New York: Linking neurons to behavior: Motion and depth.  
 Treue, S., German Primate Center, Goettingen, Germany  
 Serences, J., University of California, San Diego: Attentional modulations in human cortex.  
 Ma, W.J., New York University, New York: Psychophysics and Bayesian modeling.  
 Churchland, A., Cold Spring Harbor Laboratory: Integrating vision with other modalities: Insights into optimal behavior and the structure of neural populations.  
 Kiani, R., New York University, New York.

# Proteomics

---

July 15–28

**INSTRUCTORS**     M. Bereman, North Carolina State University, Raleigh  
M. Cilia, Cornell University, Ithaca, New York  
I. Cristea, Princeton University, New Jersey  
D. Pappin, Cold Spring Harbor Laboratory

**ASSISTANTS**     M. Alexander, Cornell University, Ithaca, New York  
H. Budayeva, Princeton University, New Jersey  
B. Diner, Princeton University, New Jersey  
T. Greco, Princeton University, New Jersey  
T. Hassan, Cold Spring Harbor Laboratory  
P. Pinheiro, Cornell University, Ithaca, New York  
K. Rivera, Cold Spring Harbor Laboratory  
J. Wilson, Cold Spring Harbor Laboratory

This intensive laboratory and lecture course focused on cutting-edge proteomic approaches and technologies. Students gained practical experience purifying and identifying protein complexes and posttranslational modifications. In a section focused on quantitative whole-proteome analyses or top-down proteomics, students gained hands-on experience using two-dimensional gel electrophoresis and mass spectrometry analysis. Students used differential in-gel electrophoresis (DIGE) for gel-based protein quantification. Differentially expressed proteins were recognized by statistical methods using advanced gel analysis software and identified using MALDI mass spectrometry. For shotgun proteomic analysis sections or bottom-up proteomics, students used label-free and covalent isotopic-labeling quantitative approaches to differentially profile changes in protein complexes and whole proteomes. Students were trained in high-sensitivity microcapillary



liquid chromatography coupled with nanospray-ESI and tandem mass spectrometry analysis and learned both single-dimension and multidimensional separation methods. In a section focused on targeted proteomics, they learned to analyze and process shotgun proteomic data for the development of SRM/MRM assays that accurately identify and quantify targeted proteins. Students designed transitions for selected peptides and performed SRM/MRM mass spectrometry assays. They learned to process and interpret the acquired data to measure changing quantities of targeted proteins in a variety of biological samples. For all sections of the course, a strong emphasis was placed on data analysis. A series of outside lecturers discussed various proteomics topics including imaging by mass spectrometry, de novo sequence analysis, advanced mass spectrometry methods, protein arrays, and functional proteomics. 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 with funds provided by the National Institute of Child Health and Human Development.

#### PARTICIPANTS

- |   |   |
|---|---|
| Cendrowski, S., Ph.D., National Biodefense Analysis and Countermeasures Center, Frederick, Maryland | McDonald, C., D.V.M., Cornell University, Ithaca, New York              |
| Chamoun, R., Ph.D., McGill University, Ste. Anne de Bellevue, Canada                                | Medzihradzsky, M., Ph.D., University of Heidelberg, Heidelberg, Germany |
| Cheng, Z., B.S., New York University, New York  | Reynolds, N., Ph.D., Yale University, New Haven, Connecticut            |
| Engle, D., Ph.D., Cold Spring Harbor Laboratory   | Shi, Z., B.S., Stanford University, California                          |
| Garrett, A., Ph.D., Center of Disease Control and Prevention, Atlanta, Georgia                      | Simsek, D., Ph.D., Stanford University, California                      |
| Harold, A., B.S., Duke University, Durham, North Carolina   | Suryawanshi, A., Ph.D., Institute of Life Sciences, Bhubaneswar, India  |
| Hong, M., M.S., Icahn School of Medicine at Mount Sinai, New York                                   | Tomlinson, R., M.S., AstraZeneca, Waltham, Massachusetts                |
| Lee, E.B., Ph.D., Seoul National University College of Medicine, South Korea                        | Uehling, J., M.S., Duke University, Durham, North Carolina              |

#### SEMINARS

- |   |  |
|---|--|
| Nesvizhskii, A., University of Michigan, Ann Arbor:<br>Analysis of affinity purification mass spectrometry data for reconstruction of protein–protein interactions.                                 | Medzihradzsky, K., University of San Francisco, California:<br>How to interpret/evaluate MS/MS spectra: Focus on PTM assignments.  |
| Pappin, D., Cold Spring Harbor Laboratory: Identification of proteins using tandem mass spectrometry.   | Bereman, M., North Carolina State University, Raleigh:<br>Introduction to mass spectrometry. Systems suitability for LC-MS/MS.   |
| Muddiman, D., North Carolina State University, Raleigh:<br>Thing 1 and Thing 2 (and Thing 3): ESI, MALDI, and hybrids of the two. Controlling chaos: Adventures in absolute protein quantification. | MacCoss, M., University of Washington, Seattle:<br>Introduction to peptide quantification and targeted proteomics. Beyond data-dependent acquisition: Sometimes we need to go backward before we can go forward. |
| Conlon, F., University of North Carolina, Chapel Hill:<br>Functional investigation of the cardiac interactomes.   |  |
| Clauser, K., Broad Institute of MIT and Harvard, Cambridge, Massachusetts: De novo interpretation of tandem mass spectra.   |  |



# Eukaryotic Gene Expression

---

July 22–August 11

**INSTRUCTORS**     K. Adelman, NIH/NIEHS, Research Triangle Park, North Carolina  
G. Narlikar, University of California, San Francisco  
A. Shilatifard, Stowers Institute for Medical Research, Kansas City, Missouri  
D. Taatjes, University of Colorado, Boulder

**ASSISTANTS**     B. Allen, University of Colorado, Boulder  
E. Bowman, NIH/NIEHS, Durham, North Carolina  
D. Gilchrist, NIEHS, Research Triangle Park, North Carolina  
R. Rickels, Stowers Institute for Medical Research, Kansas City, Missouri  
C. Zhou, University of California, San Francisco

This course was designed for students, postdocs, and principal investigators who had recently ventured into the exciting area of gene regulation. The course focused on state-of-the-art strategies and techniques used in the field. Emphasis was placed both on *in vitro* and *in vivo* protein–DNA interactions and on novel methodologies to study gene regulation. Students made nuclear extracts, performed *in vitro* transcription reactions, and measured RNA levels using primer extension. Students isolated transcription factor complexes and assessed their activity in functional assays. In addition, students learned techniques for the assembly and analysis of chromatin *in vitro*. These included transcription assays, chromatin footprinting, and chromatin remodeling assays.

During the past few years, the gene regulation field has developed *in vivo* approaches to study gene regulation. Students learned widely used techniques such as quantitative reverse transcription–polymerase chain reaction (qRT-PCR) and chromatin immunoprecipitation (ChIP). They also used RNA interference (RNAi) for specific knockdown experiments. Determining cellular



gene expression profiles has been accelerated tremendously by microarray and sequencing technology. Students received hands-on training in performing and interpreting results from microarrays, ChIP-Seq, and RNA-Seq data sets.

This course was supported with funds provided by the National Cancer Institute.

#### PARTICIPANTS

- |   |   |
|---|---|
| <p>Benegiamo, G., M.Sc., Salk Institute for Biological Studies, La Jolla, California</p> <p>Bockstaller, M., M.Sc., Ludwig-Maximilians-Universitaet Muenchen, Munich, Germany</p> <p>Chen, H., Ph.D., National Institutes of Health, Bethesda, Maryland</p> <p>Coulon, A., Ph.D., National Institutes of Health, Bethesda, Maryland</p> <p>Ganguly, S., Ph.D., National Cancer Institute, Bethesda, Maryland</p> <p>Handley, A., B.Sc., Ludwig-Maximilians-Universitaet Muenchen, Munich, Germany</p> <p>Guss, K., Ph.D., Dickinson College, Carlisle, Pennsylvania</p> | <p>Labbe, D., Ph.D., Harvard Medical School/DFCI, Boston, Massachusetts</p> <p>Lauridsen, F., B.A., University of Copenhagen, Denmark</p> <p>Muthurajan, U., Ph.D., Colorado State University, Fort Collins</p> <p>Novatt, J., Ph.D., Cold Spring Harbor Laboratory</p> <p>Palchaudhuri, R., Ph.D., Harvard University, Cambridge, Massachusetts</p> <p>Park, J., Ph.D., Harvard Medical School, Boston, Massachusetts</p> <p>Pera-Gresely, B., Ph.D., Weill Cornell Medical College, New York</p> <p>Wolter, J., B.S., Arizona State University, Tempe</p> <p>Yu, S., B.S., Stockholm University, Sweden</p> |
|---|---|

#### SEMINARS

- |  |  |
|--|--|
| <p>Meyer, B., University of California/HHMI, Berkeley: Dosage compensation in <i>Caenorhabditis elegans</i>.</p> <p>Wu, C., Janelia Farm Research Campus/HHMI, Ashburn, Virginia: Interplay between chromatin structure and transcription.</p> <p>Burge, C., Massachusetts Institute of Technology, Cambridge: Genomic analysis of splicing patterns using RNA-Seq reveals interesting evolutionary and regulatory strategies.</p> <p>Bentley, D., University of Colorado School of Medicine, Aurora: Cotranscriptional mRNA processing.</p> <p>Furlong, E., European Molecular Biology Laboratory, Heidelberg, Germany: Development and activity of enhancers in <i>Drosophila</i>.</p> <p>Kugel, J., University of Colorado, Boulder: Fundamental gene regulatory strategies and their regulation by noncoding RNAs.</p> <p>Tyler, J., University of Texas, MD Anderson Cancer Center, Houston: Regulation of chromatin assembly during DNA replication and transcription.</p> | <p>Kadonaga, J., University of California, San Diego: Novel core promoter motifs work within specialized transcription systems.</p> <p>Lis, J., Cornell University, Ithaca, New York: Pausing of RNA polymerase II.</p> <p>Whetstine, J., Massachusetts General Hospital and Harvard Medical School, Charlestown: Histone demethylases in health disease.</p> <p>Neugebauer, K., Yale University, New Haven, Connecticut: Functional coupling between transcription and splicing.</p> <p>Di Croce, L., Center for Genomic Regulation, Barcelona, Spain: Development control of chromatin modifying enzymes.</p> <p>Shiekhattar, R., University of Miami, Coral Gables, Florida: Integrator complex as a general regulator of polII transcription.</p> <p>Kingston, R., Massachusetts General Hospital and Harvard Medical School, Boston: Noncoding RNAs and Polycomb.</p> <p>Berger, S., University of Pennsylvania, Philadelphia: Changes in chromatin architecture and transcription that occur during aging.</p> |
|--|--|

# Yeast Genetics and Genomics

---

July 22–August 11

**INSTRUCTORS**      **G. Brown**, University of Toronto, Ontario, Canada  
**M. Dunham**, University of Washington, Seattle  
**M. Gartenberg**, Robert Wood Johnson Medical School, Piscataway, New Jersey

**ASSISTANTS**      **M. Borrie**, Robert Wood Johnson Medical School, Piscataway, New Jersey  
**M. Sanchez**, University of Washington, Seattle  
**T. Sing**, University of Toronto, Ontario, Canada

This course is a modern, state-of-the-art laboratory course designed to teach students the full repertoire of genetic approaches needed to dissect complex problems in the yeast *Saccharomyces cerevisiae*. Combinations of classical and modern genetic approaches were emphasized, including the isolation and characterization of mutants, tetrad analysis, complementation, and mitotic recombination. Students learned array-based methods, next-generation sequencing, and genome-based methods of analysis facilitated by the yeast genome sequence, the 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 used classical approaches and modern whole-genome sequencing to gain experience in identifying and interpreting various kinds of genetic interactions including suppression and synthetic lethality (including SGA). Students immersed in yeast genomics and performed and interpreted experiments using DNA arrays, whole-genome sequencing, and multiplexed DNA



barcode sequencing. Students gained first-hand experience in modern cytological approaches such as epitope tagging and imaging yeast cells using indirect immunofluorescence, green fluorescent protein (GFP) fusions, and a variety of fluorescent indicators for different subcellular organelles. Lectures on fundamental aspects of yeast genetics were presented along with seminars given by prominent experts in the field on topics of current interest.

This course was supported with funds provided by National Science Foundation and the Howard Hughes Medical Institute.

#### PARTICIPANTS

Barber, J., Ph.D., RIKEN Institute, Wako, Japan  
 Ewald, F., M.S., University of Gothenburg, Sweden  
 Fine, R., B.S., University of Virginia, Charlottesville  
 Hanson, S., Ph.D., University College Dublin, Ireland  
 Jayashankar, V., M.S., University of California, Irvine  
 Kumar, A., Ph.D., University of Miami, Florida  
 Ledbetter, R., M.S., Utah State University, Logan  
 Lothrop, A., Ph.D., Tufts University, Medford, Massachusetts

McMillen, D., Ph.D., University of Toronto Mississauga, Canada  
 Kyle Mohler, K., B.S., The Ohio State University, Columbus  
 Sampaio, N.M., B.S., Colorado State University, Ft. Collins  
 Shen, D., B.S., Lehigh University, Bethlehem, Pennsylvania  
 Silva, G., Ph.D., New York University, New York  
 Stenberg, S., B.Sc., Norwegian University of Life Sciences, As, Norway  
 Thomas, V., Northwestern University, Evanston, Illinois  
 Wu, Y., M.S., University of Toronto, Canada

#### SEMINARS

Gladfelter, A., Dartmouth College, Hanover, New Hampshire: Protein aggregation organizes cytosol for the cell cycle and symmetry breaking.  
 Stillman, B., Cold Spring Harbor Laboratory: Biochemistry and genetics of DNA replication in *Saccharomyces cerevisiae*.  
 Myers, C., University of Minnesota, Minneapolis: Lessons from large-scale analysis of genetic interactions in yeast.  
 Drubin, D., University of California, Berkeley: Harnessing actin dynamics for trafficking events in budding yeast.  
 Zakian, G., Princeton University, Princeton, New Jersey: Getting from there to here, the history of a scientific project.  
 Keasling, J., University of Berkeley, Lawrence Berkeley National Laboratory, and Joint BioEnergy Institute, Emeryville, California: Metabolic engineering of yeast for production of fuels and chemicals.  
 McCusker, J., Duke University Medical Center, Durham, North Carolina: The 100-genomes strains population genomics, genotypic/phenotypic variation, and emergence of an opportunistic pathogen.

Costanzo, M., Stanford University School of Medicine, California: Harnessing the awesome power of yeast with the *Saccharomyces* Genome Database.  
 Rando, O., University of Massachusetts Medical School, Worcester: Structural biology of the yeast genome.  
 Philippsen, P., University of Basel, Switzerland: Cell growth but no divisions: The lifestyle of a group of multinucleated filamentous fungi controlled by a budding-yeast-like genome.  
 Rothstein, R., Columbia University Medical Center, New York: Transformation, recombination, and choreography of the DNA damage response.  
 Emr, S., Cornell University/Weill Institute, Ithaca, New York: Sorting out protein traffic: A geneticist's journey from genes to proteins to functions.  
 Haase, S., Duke University, Durham, North Carolina: Network control of cell-cycle transcription program.  
 Jaspersen, S., Stowers Institute for Medical Research, Kansas City, Missouri: Divide and conquer: The nuclear envelope and mitosis.

# Imaging Structure and Function in the Nervous System

---

July 23–August 11

**INSTRUCTORS** J. Waters, Allen Institute for Brain Science, Seattle, Washington  
K. Zito, University of California, Davis  
F. Albeanu, Cold Spring Harbor Laboratory  
P. Tsai, University of California, San Diego, La Jolla

**ASSISTANTS** C. Johnson, University of California, Berkeley  
J. Lambert, University of California, Davis  
R. Larsen, Allen Institute for Brain Science, Seattle, Washington  
F. Munoz-Cuevas, University of Maryland, Baltimore  
J. Zhuang, Allen Institute for Brain Science, Seattle, Washington

Advances in light microscopy and digital image processing and the development of a variety of powerful fluorescent probes 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 the theoretical and practical knowledge to use emerging imaging technologies. The primary emphasis of the course was on vital light microscopy. Students learned the principles of light microscopy, as well as use of different types of cameras, laser-scanning systems, functional fluorophores, delivery techniques, and digital image-processing software. In addition to transmitted light microscopy for viewing cellular structure, the course examined a variety of molecular probes of cell function, including calcium-sensitive dyes and optogenetic and photoactivatable molecules. Particular weight was given to multiphoton laser-scanning



microscopy and to biological fluorophores, especially green fluorescent protein (GFP) and its variants. We used a spectrum of neural and cell biological systems, including living animals, brain slices, and cultured cells.

Applicants had a strong background in the neurosciences or in cell biology. In their personal statements, applicants specified (1) their experience with optical techniques, (2) how they would apply optical methods in their current projects, (3) the microscope systems available to them, and (4) their long-term goals in learning more about optical methods.

This course was supported in part by the Howard Hughes Medical Institute and the Burroughs Wellcome Fund.

## PARTICIPANTS

- |  |  |
|--|--|
| Bicker, S., Ph.D., University of Marburg, Marburg an der Lahn, Germany               | Moore, A., B.S., Brandeis University, Waltham, Massachusetts             |
| Bouvier, G., M.S., Institut de Biologie de l'Ecole Normale Supérieure, Paris, France | Nelson, A., B.A., Duke University Medical Center, Durham, North Carolina |
| Crisp, S., Ph.D., University College London, United Kingdom                          | O'Farrell, F., B.Sc., University College London, United Kingdom          |
| Deutsch, D., Ph.D., Princeton University, New Jersey                                 | Ohayon, S., Ph.D., Massachusetts Institute of Technology, Cambridge      |
| Dosumu-Johnson, R., B.S., Harvard Medical School, Boston, Massachusetts              | Sato, T., M.D., University College London, United Kingdom                |
| Gabitto, M., B.S., Columbia University, New York                                     | Schroeder, S., Ph.D., University College London, United Kingdom          |
| Issa, E., Ph.D., Massachusetts Institute of Technology, Cambridge                    | Wijetunge, L., Ph.D., University of Edinburgh, United Kingdom            |

## SEMINARS

- |   |   |
|---|---|
| Dombeck, D., Northwestern University, Evanston, Illinois: Awake imaging in the hippocampus.   | Zito, K., University of California, Davis: Photolysis.  |
| O'Connor, D., Johns Hopkins University School of Medicine, Baltimore, Maryland: Extended Ca imaging: Kinetics and spike detection.  | Deisseroth, K., Stanford University, California: Optogenetics: Recent advances/CLARITY.   |
| Engert, F., Harvard University, Cambridge, Massachusetts: 2P microscopy; lasers.  | Loew, L., University of Connecticut Health Center, Farmington: Imaging voltage.   |
| Albeanu, F., Cold Spring Harbor Laboratory: Intrinsic imaging.  | Ji, N., Janelia Farm Research Campus/HHMI, Ashburn, Virginia: Deep imaging/adaptive optics.   |
| Lanni, F., Carnegie-Mellon University, Pittsburgh, Pennsylvania: Nature of light, coherence, wave description, polarization. Incoherent emission pt source, PSF/Airy pattern, Fourier/OTF, aberrations. | Griesbeck, O., Max-Planck Institute of Neurobiology, Martinsried, Germany: XFP calcium indicators.  |
| Waters, J., Northwestern University, Chicago, Illinois: Noise and detectors.  | Tsai, P., University of California, San Diego, La Jolla: Ray tracing, refraction, lenses. Fluorescence/Fred's rocks, Jablonski diagrams, filters. |
| Lichtman, J., Harvard University, Cambridge, Massachusetts: Confocal microscopy.  | Yasuda, R., Max-Planck Florida Institute for Neuroscience, Jupiter: FRET and FLIM.  |
| Waters, J., Harvard Medical School, Boston, Massachusetts: Cameras and PMTs; noise (sources and troubleshooting) and saturation.  | Kuhlman, S., Carnegie Mellon University, Pittsburgh, Pennsylvania: In vivo 2p guided patch.   |
| Mertz, J., Boston University, Massachusetts: Contrast: Phase, DIC, De Senarmont DIC, Dodt tube.   | Emiliani, V., CNRS, INSERM, University Paris Descartes Paris, France: SLM.  |
| Biteen, J., University of Michigan, Ann Arbor: STORM/PALM, STED, structured illumination.   | Murthy, V., Harvard University, Cambridge, Massachusetts: Holographic photoactivation.  |
|   | Denk, W., Max-Planck Institute for Medical Research, Heidelberg, Germany: Block-face EM/connectome.   |

# Synthetic Biology

---

July 21–August 11

**INSTRUCTORS**     K. Haynes, Arizona State University, Tempe  
J. Lucks, Cornell University, Ithaca, New York  
P. Peralta-Yahya, Georgia Institute of Technology, Atlanta  
D. Savage, University of California, Berkeley  
J. Tabor, Rice University, Houston, Texas

**ASSISTANTS**     R. Davis, Arizona State University, Tempe  
E. Olson, Rice University, Houston, Texas  
V. Peterson, Georgia Institute of Technology, Atlanta  
M. Takahashi, Cornell University, Ithaca, New York  
R. Yokoo, University of California, Berkeley

Synthetic biology is a discipline wherein living organisms are genetically programmed to carry out desired functions in a reliable manner. This field takes inspiration from our ever-expanding ability to measure and manipulate biological systems, as well as the philosophical reflections of Schrödinger and Feynman that physical laws can be used to describe and rationally engineer biology to accomplish useful goals. After all, cells are the world's most sophisticated chemists, and their ability to learn to adapt to changing environments offers enormous potential to 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 was supported by the Howard Hughes Medical Institute and the U.S. Office of Naval Research.



## PARTICIPANTS

- Bhatnagar, P., Ph.D., Baylor College of Medicine, Houston, Texas
- Brechun, K., B.S., University of Toronto, Canada
- Canady, T., B.A., Carnegie Mellon University, Pittsburgh, Pennsylvania
- Cavaleiro, A., M.S., Novo Nordisk Foundation Centre for Biosustainability, Horsholm, Denmark
- Chessher, A., Ph.D., Manchester Institute of Biotechnology, Manchester, United Kingdom
- Coffman, C., Ph.D., University of Missouri, Columbia
- Ghosh, S., M.S., CSIR-Indian Institute of Chemical Biology, Kolkata, India
- Hecht, A., Ph.D., National Institute of Standards and Technology, Stanford, California
- Inda, M., Ph.D., IBR, Rosario, Argentina
- Kennedy, L., Ph.D., United States Naval Academy, Annapolis, Maryland
- Kumaran, R.I., Ph.D., Cold Spring Harbor Laboratory
- Lopatkin, A., B.S., Duke University, Durham, North Carolina
- Munson, M., Ph.D., National Institute of Standards and Technology, Stanford, California
- Raghavan, S., Ph.D., Biomedical Sciences Institute, Singapore
- Rasaputra, K., Ph.D., University of Houston, Texas
- Woodsworth, D., B.Sc., British Columbia Cancer Research Center, Vancouver, Canada

## SEMINARS

- Ellington, A., University of Texas, Austin: Directed evolution as synthetic biology.
- Klavins, E., University of Washington, Seattle: Computation with multicellular systems.
- Wang, H., Columbia University, New York: Multiplex genome engineering in cells and communities.
- Keasling, J., University of California, Berkeley/Lawrence Berkeley National Laboratory, Emeryville: Synthetic (yeast) biology for synthetic chemistry.
- Hoff, K., Genomatica, Inc., San Diego, California: An industrial synthetic biology.
- Dunlop, M., University of Vermont, Burlington: Engineering feedback control systems in microbes.
- Palmer, M., Stanford University, California: Responsible innovation.
- Jewett, M., Northwestern University, Evanston, Illinois: Engineering biology by design thinking outside the cell.
- Gonzalez, R., Rice University, Houston, Texas: Engineering biology for fuel and chemical production.
- Murray, R., California Institute of Technology, Pasadena: Rapid prototyping and debugging of biomolecular circuits using cell-free breadboards.
- Weiss, R., Massachusetts Institute of Technology, Cambridge: Mammalian cell engineering.
- Qi, S., University of California, San Francisco: Editing and regulation of the mammalian genome using CRISPR.



# Genetics and Neurobiology of Language

---

July 28–August 3

**INSTRUCTORS**     S. Fisher, Max-Planck Institute for Psycholinguistics, Nijmegen, The Netherlands  
D. Poeppel, University Maryland College Park, College Park  
K. Watkins, Oxford University, United Kingdom

Why are children able to acquire highly sophisticated language abilities without needing to be taught? What are the neurobiological and neurophysiological processes that underpin human speech and language and how do they go awry in developmental and acquired disorders? Which genetic factors contribute to this remarkable suite of human skills and are there evolutionary precursors that we can study in other species? This unique new CSHL course addressed core questions about the bases and origins of speech and language through talks, interactive sessions, keynotes, and debates involving leading experts from a range of disciplines. It integrated the state-of-the-art from complementary perspectives, including development, cognitive models, neural basis, gene identification, functional genomics, model systems, and comparative/evolutionary studies.

This course was supported with funding from the Nancy Lurie Marks Family Foundation.

## PARTICIPANTS

Akimoto, Y., Ph.D., RIKEN Brain Science Institute,  
Wako-shi, Saitama, Japan  
Bergelson, E., Ph.D., University of Rochester, New York  
Chesters, J., D.Phil., University of Oxford, United Kingdom  
Frale, E., Ph.D., University of California, Los Angeles  
Getz, H., Ph.D., Georgetown University, Washington, D.C.  
Guidi, L., D.Phil., University of Oxford, United Kingdom

Hayes, R., M.S., University of Pittsburgh, Pennsylvania  
Hobson, H., D.Phil., University of Oxford, United Kingdom  
Kavaklioglu, T., M.S., Max-Planck Institute for  
Psycholinguistics, Nijmegen, The Netherlands  
Krishnan, S., Ph.D., Birkbeck, University of London, United  
Kingdom  
Martinez Alvarez, A., M.A., University of Barcelona, Spain



Mody, M., Ph.D., Lurie Center for Autism, Lexington, Massachusetts  
 Nora, A., M.Psych., Brain Research Unit, Finland  
 Oseki, Y., M.A., New York University, New York  
 Pagliarini, E., M.A., University of Milano-Bicocca, Milan, Italy  
 Payne, H., Ph.D., University of College, London, United Kingdom

Rendall, A., Ph.D., University of Connecticut, Storrs  
 Riva, V., Ph.D., Scientific Institute "Eugenio Medea," Bosiso Parini (Lecco), Italy  
 Swanson, M., Ph.D., University of North Carolina, Chapel Hill  
 Williams, A., M.A., New York University, New York  
 Yablonski, M., M.Sc., Bar-Ilan University, Ramat Gan, Israel  
 Zhang, L., Ph.D., New York University, New York

## SEMINARS

- Aslin, R., University of Rochester, New York: Speech perception and word recognition in infants and adults.  
 Werker, J., University of British Columbia, Vancouver, Canada: Mechanisms of development and plasticity in speech perception.  
 McAllister Byun, T., New York University, New York: Perception-production relations in speech acquisition and disorders: A view from American English.  
 Watkins, K., Oxford University, United Kingdom: Similarities and differences in the neural correlates of developmental disorders of speech and language.  
 Rice, M., The University of Kansas, Kansas City: Growth of language in children with and without SLI: Moving toward epigenetic models of growth strengths and weaknesses.  
 Emmorey, K., San Diego State University, California: The impact of distinct sensory-motor systems on the neurobiology of language: Signed versus spoken language.  
 MacSweeney, M., University College London, United Kingdom: The impact of deafness on the neurobiology of spoken language.  
 Pylkkanen, L., New York University, New York: Combinatory semantics in the cognitive neuroscience of language.  
 Fedorenko, E., Massachusetts Institute of Technology, Cambridge: The language network and its place within the broader architecture of the human mind and brain.  
 Dronkers, N., University of California, Davis: What has the study of adult aphasia taught us about language?  
 Davis, M., MRC Cognition and Brain Sciences Unit, Cambridge, United Kingdom: The neurobiology of spoken word recognition.
- Bishop, D., University of Oxford, United Kingdom: Using genetics as a tool to explore relations between language and other domains.  
 Fisher, S., Max-Planck Institute for Psycholinguistics, Nijmegen, The Netherlands: Genes, speech, and language: Translating the genome in human neuroscience.  
 Vernes, S., Max-Planck Institute for Psycholinguistics, Nijmegen, The Netherlands: The building blocks of language: From molecules to neuronal circuits.  
 Poeppel, D., New York University, New York: How can we develop linking hypotheses between genetics, computational neuroscience, and language research?  
 Hagoort, P., Max-Planck Institute for Psycholinguistics, Nijmegen, The Netherlands: The neurobiology of language beyond the information given.  
 Paracchini, S., University of St Andrews, Fife, United Kingdom: Dyslexia and laterality: Is there a link?  
 Enard, W., Ludwig-Maximilians University, Munich, Germany: Mice, chimpanzees, and the molecular basis of speech.  
 Scharff, C., Freie Universität Berlin, Germany: FoxPs in birds and insects.  
 Kayser, C., University of Glasgow, Scotland: The neural mechanisms of speech encoding: From oscillations to neural populations.  
 Fitch, T., University of Vienna, Austria: A comparative approach to the neurobiology of language.  
 Arnold, K., University of St Andrews, Fife, United Kingdom: The search for language in nonhuman primates.

# Brain Tumors

---

August 6–12

**INSTRUCTORS**      C. Gladson, Cleveland Clinic, Ohio  
A. Lasorella, Columbia University, New York  
S. Majumder, University of Texas MD Anderson Cancer Center, Houston

**ASSISTANT**      A. Marisetty, University of Texas MD Anderson Cancer Center, Houston

This 1-week discussion course provided a clinical overview of brain tumors and emphasized molecular mechanisms involved in their growth. Special emphasis was placed on neural differentiation, signaling mechanisms, DNA replication, chromatin modulation, stem cells, mouse models, genomics, imaging techniques, genetically modified mouse techniques, nanotechnology, mechanism-based therapeutic strategies, and biobanks and ethical concerns. Attendees were able to interact with senior investigators one-to-one in an informal environment. All nonfaculty students were considered for a generous stipend depending on stated need. Applications were invited from medical and graduate students, postdoctoral fellows, faculty, and clinicians and scientists.

This course was supported with funding from the American Brain Tumor Association.

## PARTICIPANTS

Badeaux, M., Ph.D., Massachusetts General Hospital, Boston      Dhruv, H., Ph.D., Translational Genomics Research  
Braun, S., Ph.D., Lund University, Sweden      Institute, Phoenix, Arizona



- Ferraro, G., Ph.D., Harvard Medical School, Charlestown, Massachusetts
- Garrós-Regúlez, L., Ph.D., Biodonostia Institute, San Sebastián, Spain
- Kaur, A., M.Sc., University of Turku, Finland
- Klopper, J., Ph.D., Massachusetts General Hospital, Boston
- Lu, L., M.D., MD Anderson Cancer Center Genetics, Houston, Texas
- Raklli, V., Ph.D., Karolinska Institutet, Stockholm, Sweden
- Seano, G., Ph.D., Harvard Medical School/Massachusetts General Hospital, Boston
- Selvadurai, H., Ph.D., The Hospital for Sick Children, Toronto, Canada
- Shahzad, U., Ph.D., The Hospital for Sick Children, Toronto, Canada
- Wang, C.-L., Ph.D., Cold Spring Harbor Laboratory
- Weissmann, S., Ph.D., University of Copenhagen, Denmark

## SEMINARS

- Sawaya, R., MD Anderson Cancer Center, Houston, Texas: Extent of resection and survival in glioblastoma: Why and how?
- Fuller, G., MD Anderson Cancer Center, Houston, Texas: Clinical aspects of brain tumors: Oncologic surgical neuropathology.
- Bondy, M., Baylor College of Medicine, Houston, Texas: Epidemiology of glioma: Current state of the science.
- Reilly, K., National Cancer Institute, Bethesda, Maryland: Mouse models of brain cancer to study genetic risk factors.
- Mietz, J., National Cancer Institute, Bethesda, Maryland.
- Sawaya, R., MD Anderson Cancer Center, Houston, Texas: The comprehensive neuro-oncology center: Challenges and opportunities.
- Wechsler-Reya, R., Duke University Medical Center, Durham, North Carolina.
- Baker, S., St. Jude Children's Research Hospital, Memphis, Tennessee: Molecular and clinical subgroups of pediatric high-grade glioma.
- Eberhart, C., Johns Hopkins University School of Medicine, Baltimore, Maryland: The role of notch signaling in brain tumors.
- Jabado, N., McGill University Health Center, Montreal, Quebec.
- Roussel, M., St. Jude Children's Research Hospital, Memphis, Tennessee: Pediatric medulloblastoma.
- Rakesh Jain, R., Massachusetts General Hospital, Boston.
- Gladson, C., Cleveland Clinic, Ohio.
- Pieper, R., University of California, San Francisco: The deceptively complicated story of mutant IDH-mediated drug resistance.
- James, C.D., Northwestern University, Evanston, Illinois: Therapeutic hypothesis testing using brain tumor engraftment models.
- Holland, E., Fred Hutchinson Cancer Research Center, Seattle, Washington: The evolution of non-CIMP GBM subtypes.
- Gutmann, D., Washington University School of Medicine, St. Louis, Missouri: Leveraging neurofibromatosis-1 to understand pediatric glioma pathogenesis.
- Enikolopov, G., Cold Spring Harbor Laboratory: Modes of division and differentiation of neural stem cells.
- Dirks, P., Hospital for Sick Children/University of Toronto, Canada.
- Kornblum, H., University of California Molecular and Medical Pharmacology, Los Angeles: Neurospheres: Making a mess of the data.
- Rich, J., Cleveland Clinic, Ohio: Brain tumor stem cells.
- Snyder, E., Sanford/Burnham Institute, La Jolla, California.
- Lasorella, A., Columbia University, New York.
- Majumder, S., University of Texas MD Anderson Cancer Center, Houston: Cellular flexibility regulates both normal development and cancer.
- Furnari, F., University of California, San Diego: EGF receptors in glioma biology: Therapeutic targets and multifaceted drivers of pathogenicity.
- Berger, M., University of California, San Francisco.
- Aldape, K., The University of Texas, Houston.
- Brennan, C., Memorial Sloan-Kettering Cancer Center, New York.
- Iavarone, A., Columbia University Medical Center, New York: The master regulators and drivers of oncogenesis of glioblastoma.
- Gilbert, M., MD Anderson Cancer Center, Houston, Texas.

# X-Ray Methods in Structural Biology

---

October 13–28

**INSTRUCTORS**     **W. Furey**, V.A. Medical Center and University of Pittsburgh, Pennsylvania  
**G. Gilliland**, Janssen R&D, LLC, Johnson & Johnson, Radnor, Pennsylvania  
**A. McPherson**, University of California, Irvine  
**J. Pflugrath**, The Woodlands, Texas

**ASSISTANTS**     **A. Armstrong**, Janssen R&D, LLC, Spring House, Pennsylvania  
**B. Hintze**, Duke University Medical Center, Durham, North Carolina

Crystallography and X-ray diffraction yield a wealth of structural information unobtainable through other methods. This intense laboratory/computational course focused on the major techniques used to determine the three-dimensional structures of macromolecules. It was designed for scientists with a working knowledge of protein structure and function, but who are new to macromolecular crystallography. Topics covered included basic diffraction theory, crystallization (proteins, nucleic acids, and complexes), crystal characterization, X-ray sources and optics, synchrotrons, cryocrystallography, data collection, data reduction, heavy atom substructure determination, multiple isomorphous replacement, single and multiwavelength anomalous diffraction, molecular replacement, solvent flattening, noncrystallographic symmetry averaging, electron density interpretation, molecular graphics, structure refinement, structure validation, coordinate deposition, and structure presentation. Participants learned through extensive hands-on experiments in fully equipped labs. They crystallized multiple proteins and determined their crystal structures by several methods while learning through extensive lectures on crystallographic theory and informal discussions behind the techniques.

Applicants were familiar with the creation and editing of simple text files on Linux workstations using a screen-based editor (either vi or emacs).



This course was supported with funds provided by the National Institute of General Medical Sciences.

#### PARTICIPANTS

- Arturo, E., B.S., Drexel University College of Medicine, Philadelphia, Pennsylvania  
 Chandrasekarar, V., Ph.D., University of Utah, Salt Lake City  
 Coey, A., Ph.D., Stanford University, California  
 Doyle, L., B.S., Fred Hutchinson Cancer Research Center, Seattle, Washington  
 Feng, J., Ph.D., Pfizer, Inc., San Diego, California  
 Henderson-Stull, M.N., Ph.D., Augsburg College, Minneapolis, Minnesota  
 Hermida Aponte, D., M.S., University of Copenhagen, Denmark  
 Jones, J., B.S., Idaho State University, Meridian  
 Lizak, C., Ph.D., University of British Columbia, Vancouver, Canada  
 Lofgren, M., Ph.D., Harvard Medical School, Boston, Massachusetts  
 Mansoor, S., Ph.D., Oregon Health & Science University, Portland  
 Nag, S., Ph.D., Yale University, New Haven, Connecticut  
 Mayclin, S., Ph.D., National Institutes of Health/New York University, Bethesda, Maryland  
 Parthasarathy, G., M.S., Merck and Company, West Point, Pennsylvania  
 Sorensen, A., M.S., Aalborg University and Novo Nordisk A/S, Aalborg, Denmark  
 Wasmuth, E., B.S., Gerstner Sloan-Kettering Graduate School, New York

#### SEMINARS

- McPherson, A., University of California, Irvine: Crystallization of macromolecules I. Crystallization of macromolecules II. Symmetry, periodicity, unit cells, space groups, Miller planes, and lattices. Waves, vectors, and complex numbers. Fundamental diffraction relationships and Bragg's Law. Diffraction patterns, reciprocal space, and Ewald's Sphere. Fourier transforms and the electron density equation. Patterson methods. The art of capillary crystal mounting. Heavy atoms and anomalous scatterers.
- Perrakis, A., Netherlands Cancer Institute, Amsterdam: Automated model building and rebuilding: From ARP/WARP to PDB\_REDO.
- Thorn, A., MRC Laboratory of Molecular Biology, Cambridge, United Kingdom: An introduction to SHELXC/D/E. The basics of twinning in crystals of macromolecules.
- Smith, C., Stanford University, California: Synchrotron data collection and femtosecond crystallography.
- Tronrud, D., Oregon State University, Corvallis: Macromolecular refinement I. Macromolecular refinement II.
- Gilliland, G., Janssen R&D, LLC, Johnson & Johnson, Radnor, Pennsylvania: Course overview: The structure determination of biological macromolecules.
- Kleywegt, G., European Bioinformatics Institute, Hinxton, Cambridge, United Kingdom: Just because it's in Nature, doesn't mean it's true... (macromolecular structure validation).
- Holton, J., University of California, San Francisco: Tips and tricks for improving diffraction.
- Pflugrath, J., The Woodlands, Texas, and Borek, D., University of Texas Southwestern Medical Center at Dallas: X-ray data collection and processing. Cryocrystallography. Cryocrystallography and beyond. Scaling and merging synchrotron data.
- Richardson, J., and Hintze, B., Duke University Medical Center, Durham, North Carolina: Detection and repair of model errors using Molprobity. Structure presentation.
- Westbrook, J., Rutgers University, Piscataway, New Jersey: Automating PDB deposition.
- Caffrey, M., Trinity College Dublin, Ireland: Crystallizing membrane proteins for structure-function studies using lipidic mesophases.
- Adams, P., University of California, Berkeley: Structure refinement. PHENIX overview.
- Emsley, P., MRC Laboratory of Molecular Biology, Cambridge, United Kingdom: Model-building tools in Coot.
- Read, R., University of Cambridge, United Kingdom: Using SAD data in Phaser. Molecular replacement: New structures from old.
- Sweet, R., Brookhaven National Laboratory, Upton, New York: A slightly different view of fundamental crystallography. X-ray sources and optics.
- Jackson, R., Montana State University, Bozeman: Crystal structure of the CRISPR RNA-guided surveillance complex from *Escherichia coli*.
- Terwilliger, T., Los Alamos National Laboratory, New Mexico: Automated structure solution and model building.

Hendrickson, W., Columbia University, New York: MAD and SAD phasing.

Furey, W., V.A. Medical Center and University of Pittsburgh, Pennsylvania: Anomalous data collection considerations. Patterson group therapy. Isomorphous replacement and anomalous scattering. Solvent flattening/phase

combination. MAD phasing: A classical approach. The role of direct methods in macromolecular crystallography. Noncrystallographic symmetry averaging.

Ji, X., National Cancer Institute, Frederick, Maryland: Structural view of double-stranded RNA processing, mechanism, and evolution of RNase III.

# Programming for Biology

---

October 13–28

INSTRUCTOR      S. Prochnik, DOE–Joint Genome Institute, Walnut Creek, California

CO-INSTRUCTOR   S. Robb, Stowers Institute for Medical Research, Kansas City, Missouri

ASSISTANTS      S. Ahrendt, University of California, Riverside  
D. Messina, Cofactor Genomics, St. Louis, Missouri  
S. Rynearson, University of Utah, Salt Lake City  
D. Triant, University of Virginia, Charlottesville  
K. Youens-Clark, University of Arizona, Tucson

Web-based tools are no longer enough for today's biologist who needs to access and analyze large data sets from myriad sources in disparate formats. The need to design and program custom analysis pipelines is becoming ever more important as new technologies increase the already exponential rate at which biological data are generated. Because this 2-week course was designed for lab biologists with little or no programming experience, students were able to complete it with the bioinformatics and scripting skills necessary to exploit the abundance of biological data. The prerequisite for the course was basic knowledge of UNIX; some scripting experience was helpful, but not necessary. Lectures and problem sets from previous years that covered this material were available online for students to study before starting the course.

The course taught Perl, a scripting language that is easy to learn and efficient to use. Perl also has a vast array of ready-built modules (such as Bioperl) that are designed to solve common





biological problems. We started with a week of introductory coding and continued with a survey of available biological libraries and practical topics in bioinformatics, taught by invited experts, and ended with a final group project. Formal instruction was provided on every topic. Students also worked to solve problem sets covering common scenarios in the acquisition, validation, integration, analysis, and visualization of biological data. For their final projects, which ran during the second week of the course, students proposed problems using their own data and worked with each other and the faculty to solve them. They learned how to design, construct, and run powerful and extensible analysis pipelines in a straightforward manner. Final projects have formed the basis of publications as well as public biological websites (see, for example, <http://bio.perl.org/wiki/Deobfuscator>). The students were provided with a library of reference books that they took 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.

This course was supported by the National Human Genome Research Institute.

#### PARTICIPANTS

- |  |  |
|--|--|
| <p>Alaba, O., M.S., International Institute of Tropical Agriculture, Ibadan, Nigeria</p> <p>Beaudoin, J.-D., Ph.D., Yale University, New Haven, Connecticut</p> <p>del Campo, J., Ph.D., Institut de Ciències del Mar–CSIC, Barcelona, Spain</p> <p>Gayden, T., Ph.D., McGill University, Montreal, Canada</p> <p>Gogakos, T., B.S./M.S., The Rockefeller University, New York</p> <p>Gryganskyi, A., Ph.D., Lambert Spawn Co., Coatesville, Pennsylvania</p> <p>Hafner, M., Ph.D., NIAMS/National Institutes of Health, Bethesda, Maryland</p> <p>Hehenberger, E., Ph.D., University of British Columbia, Vancouver, Canada</p> <p>Hernandez, M., Icahn School of Medicine at Mount Sinai, New York</p> | <p>Irisarri, I., Ph.D., University of Konstanz, Germany</p> <p>Jackson, S., Ph.D., National Institute of Standards and Technology, Gaithersburg, Maryland</p> <p>Kepler, R., Ph.D., United States Department of Agriculture–ARS, Beltsville, Maryland</p> <p>Kosova, G., Ph.D., Massachusetts General Hospital, Boston</p> <p>McLaughlin, P., Ph.D., Drexel University, Philadelphia, Pennsylvania</p> <p>Monahan, K., Ph.D., Columbia University, New York</p> <p>Orkin, J., Ph.D., Washington University, St. Louis, Missouri</p> <p>Schwessinger, B., Ph.D., University of California, Davis</p> <p>Ann Sutton, L., Ph.D., Uppsala University, Sweden</p> <p>Tavera-Mendoza, L., Ph.D., Dana-Farber Cancer Institute/Harvard Medical School, Boston, Massachusetts</p> <p>Yue, M., Ph.D., University of Pennsylvania, Philadelphia</p> <p>Zhao, B., B.A., University of Cambridge, United Kingdom</p> |
|--|--|

#### SEMINARS

- |   |   |
|---|---|
| <p>Moore, B., University of Utah, Salt Lake City: VAAST.</p> <p>Haas, B., Broad Institute, Cambridge, Massachusetts: RNA transcript reconstruction.</p> <p>Robinson, J., Broad Institute, Cambridge, Massachusetts: IGW.</p> <p>Stajich, J., University of California, Riverside: Introduction to NGS.</p> <p>Schatz, M., Cold Spring Harbor Laboratory: Genome assembly.</p> | <p>Thomas, P., University of Southern California, Los Angeles: Protein function annotation.</p> <p>Cain, S., Ontario Institute for Cancer Research, Toronto, Canada: Jbrowse.</p> <p>Marques-Bonet, T., Universitat Pompeu Fabra, Barcelona, Spain: Structural variation.</p> <p>Pearson, W., University of Virginia, Charlottesville: Sequence similarity searching.</p> |
|---|---|

# Computational and Comparative Genomics

---

October 29–November 4

**INSTRUCTORS**      **W. Pearson**, University of Virginia, Charlottesville  
                             **L. Stubbs**, University of Illinois, Urbana

**ASSISTANTS**        **D. Miller**, Stowers Institute for Medical Research, Kansas City, Missouri  
                             **L. Mills**, University of Virginia, Charlottesville

This course presented a comprehensive overview of the theory and practice of computational methods for the identification and characterization of functional elements from DNA sequence data. The course focused on approaches for extracting the maximum amount of information from protein and DNA sequence similarity through sequence database searches, statistical analysis, and multiple sequence alignment. Additional topics included alignment and analysis of “next-gen” sequencing data, with applications from metagenomic, RNA-Seq, and ChIP-Seq experiments; the Galaxy environment for high-throughput analysis; 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 genome features. The course combined lectures with hands-on exercises; students were encouraged to pose challenging sequence analysis problems using their own data. The course was designed for biologists seeking advanced training in biological sequence and genome analysis, computational biology core resource directors and staff, and individuals in other disciplines (e.g., computer science) who wish to survey current research problems in biological sequence analysis. Advanced programming skills were not required. The primary focus of this course was the theory and practice of algorithms in



computational biology, with the goals of using current methods more effectively for biological discovery and developing new algorithms

This course was supported by the National Human Genome Research Institute.

#### PARTICIPANTS

Abbas, H., Ph.D., University of Texas MD Anderson Cancer Center, Houston  
 Ai-Yin Liu, I., Ph.D., Duke University, Durham, North Carolina  
 Alvarez, C., M.D., Universidad Nacional Autonoma de Mexico, Mexico City  
 Ball, R., Ph.D., The Jackson Laboratory, Bar Harbor, Maine  
 Brady, L., Ph.D., University of Pennsylvania, Philadelphia  
 Engin, H.B., Ph.D., University of California, San Diego  
 Garrett, E., Ph.D., Washington University, St. Louis, Missouri  
 Guerra Amorim, C.E., Ph.D., Columbia University, New York  
 Jain, R., M.D., University of Pennsylvania School of Medicine, Philadelphia  
 Jain, R., Ph.D., Yale University, New Haven, Connecticut  
 Josefowicz, S., Ph.D., The Rockefeller University, New York  
 Lee, A., Ph.D., University of British Columbia, Vancouver, Canada  
 Lee, J., Ph.D., University of Minnesota Medical School, Duluth  
 Livnat, A., Ph.D., Virginia Tech, Blacksburg  
 Mazutis, L., Ph.D., Harvard University, Cambridge, Massachusetts  
 McGee, W., B.A., Northern University, Chicago, Illinois  
 McKeithen-Mead, S., M.S., Yale University, New Haven, Connecticut  
 Megquier, K., B.A., Broad Institute, Cambridge, Massachusetts  
 Memi, F., Ph.D., University of Connecticut Health Center, Farmington  
 Moir, R., Ph.D., Albert Einstein College of Medicine, Bronx, New York  
 Savitskaya, E., Ph.D., Skolkovo Institute of Science and Technologies, Moscow, Russia  
 Taffner, S., M.S., Washington University, St. Louis, Missouri  
 Weishaupt, H., Ph.D., Uppsala University, Uppsala, Sweden  
 Williams, J., B.S., Cold Spring Harbor Laboratory

#### SEMINARS

Mackey, A., University of Virginia, Charlottesville: Genome annotation (HMM Basics). SNP discovery and variation. RNA-Seq. Gene lists to pathways. Assembling genomes and transcriptomes.  
 Miller, D., Stowers Institute for Medical Research, Kansas City, Missouri: Unix command line.  
 Hawkins, D., Ludwig Institute for Cancer Research, La Jolla, California: Chromatin states 1: Analysis of histone modifications.  
 Stormo, G., Washington University School of Medicine, St. Louis, Missouri: Modeling motifs: Collecting data. From motifs to regulatory networks.  
 Taylor, J., Emory University, Atlanta, Georgia: Galaxy for high-throughput analysis. Galaxy visualization. Probing higher-dimension chromatin structure.  
 Mills, L., University of Virginia, Charlottesville: Introduction to galaxy/genomics file formats. Genome browsers.  
 Stubbs, L., University of Illinois, Urbana: Introduction to genome biology. Integrating genomics data sets for biological inference.  
 Hardison, R., Pennsylvania State University, State College: Genomics of gene regulation 1: Epigenomic features and assays. Genomics of gene regulation 2: Epigenomic features and assays.  
 Pearson, W., University of Virginia, Charlottesville: Introduction and Overview. Protein evolution and sequence similarity searching. Practical sequence similarity searching. Multiple sequence alignment. PSSMs, HMMs, and Pfam. PSSMs to phenotype.

# Antibody Engineering and Phage Display

---

November 5–18

**INSTRUCTORS**     **D. Siegel**, University of Pennsylvania School of Medicine, Philadelphia  
                              **G. Silverman**, New York University School of Medicine, New York

**SENIOR ASSISTANT** **S. Kacir**, University of Pennsylvania School of Medicine, Philadelphia

**ASSISTANTS**        **S. Gilgunn**, Dublin City University, Dublin, Ireland  
                              **K. Noren**, New England Biolabs, Ipswich, Massachusetts  
                              **M.A. Pohl**, Oncobiologics, Cranbury, New Jersey

Recent advances in the generation and selection of antibodies from combinatorial libraries allow 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 phage and 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. Antibodies were selected from the library by panning. Production, purification, and characterization of Fab fragments expressed in *Escherichia coli* were also covered. Epitopes were selected from peptide libraries and characterized.

The lecture series, presented by a number of invited speakers, emphasized polymerase chain reaction (PCR) of immunoglobulin genes, the biology of filamentous phage and the utility of surface expression libraries, expression of antibodies in *E. coli* and mammalian cells, antibody



structure and function, the immunobiology of the antibody response, and recent results on the use of antibodies in therapy. The theory and practical implications for selection from phage-displayed libraries of random peptides, cDNA products, and semisynthetic proteins were also explored.

This course was supported with funds provided by the Howard Hughes Medical Institute.

#### PARTICIPANTS

Adams, R., Ph.D., Cold Spring Harbor Laboratory  
 Åstrand, M., M.S., KTH Royal Institute of Technology,  
 Stockholm, Sweden  
 Elias, M., Ph.D., Inventox, Inc., Wynnewood, Pennsylvania  
 Gordon, S., B.S., Stanford University, California  
 Günaydin, G., Ph.D., Karolinska Institutet, Stockholm,  
 Sweden  
 Hernandez, D., M.S., Sackler Institute of Biomedical Science  
 at NYU, New York  
 Koti, M., Ph.D., Queen's University, Kingston, Canada  
 Maier, N., Dipl., University of Potsdam, Germany  
 Moran, K., M.Sc., Dublin City University, Ireland

Pelzek, A., M.S., New York University School of Medicine-  
 Sackler, New York  
 Robinson, M.P., B.S., Cornell University, Ithaca, New York  
 Santich, B., B.A., Gerstner Sloan-Kettering Graduate School,  
 New York  
 Stephenson, H., B.S., Sutro Biopharma, S. San Francisco,  
 California  
 Suzuki, M., M.D., Memorial Sloan-Kettering Cancer Center,  
 New York  
 Theunissen, J.-W., Ph.D., IGENICA, Burlingame, California  
 Vierira-Pires, R., Ph.D., Centro de Neurociências e Biologia  
 Celular, Coimbra, Portugal

#### SEMINARS

Payne, A., University of Pennsylvania, Merion Station:  
 Shared autoantibody VH gene usage in pemphigus  
 vulgaris: Implications for disease development.  
 Rader, C., The Scripps Research Institute-Florida, Jupiter:  
 Cancer and approaches to therapeutic antibody development.  
 Noren, C., New England Biolabs, Beverly, Massachusetts:  
 Peptide display libraries.  
 Siegel, D., University of Pennsylvania Medical Center,  
 Philadelphia: Human phage display libraries, cell-surface  
 panning, translational medicine applications, and llamas.

Phage display, metastatic cancer, and personalized  
 medicine.  
 Silverman, G.J., New York University Medical Center,  
 New York: Human repertoire evolution and microbial  
 influences.  
 Wilson, I., Scripps Research Institute, La Jolla, California:  
 Broad neutralization of viral pathogens: Implications for  
 vaccine design.  
 Boyd, S., Stanford University, California: B-cell repertoire  
 sequencing in human vaccine responses.

# Advanced Sequencing Technologies and Applications

---

November 11–23

**INSTRUCTORS**     E. **Mardis**, Washington University School of Medicine, St. Louis, Missouri  
G. **Marth**, University of Utah School of Medicine, Salt Lake City  
W. **McCombie**, Cold Spring Harbor Laboratory  
A. **Quinlan**, University of Virginia, Charlottesville  
M. **Schatz**, Cold Spring Harbor Laboratory

**ASSISTANTS**        B. **Ainscough**, Washington University, St. Louis, Missouri  
A. **Farrell**, University of Utah, Salt Lake City  
T. **Garvin**, Cold Spring Harbor Laboratory  
M. **Griffith**, Washington University School of Medicine, St. Louis, Missouri  
O. **Griffith**, Washington University School of Medicine, St. Louis, Missouri  
J. **Gurtowski**, Cold Spring Harbor Laboratory  
V. **Magrini**, Washington University School of Medicine, St. Louis, Missouri  
S. **McGrath**, Washington University School of Medicine, St. Louis, Missouri  
J. **Walker**, Washington University School of Medicine, St. Louis, Missouri  
A. **Ward**, University of Utah School of Medicine, Salt Lake City

During the last decade, large-scale DNA sequencing has markedly impacted the practice of modern biology and is beginning to affect the practice of medicine. With the introduction of several advanced sequencing technologies, costs and timelines have been reduced by orders of magnitude, enabling investigators to conceptualize and perform sequencing-based projects that heretofore were prohibitive. Furthermore, the application of these technologies to answer questions previously not experimentally approachable is broadening their impact and application.



This intensive 2-week course explored applications of next-generation sequencing technologies, with a focus on commercially available methods. Students were instructed in the detailed operation of several revolutionary sequencing platforms, including sample preparation procedures, general data handling through pipelines, and in-depth data analysis. A diverse range of the types of biological questions enabled by next-generation sequencing technologies was explored including DNA resequencing of known cancer genes to compare tumor and normal DNA, de novo DNA sequencing and assembly of bacterial genomes, RNA sequencing, and others that were tailored to the student's research areas of interest. Guest lecturers highlighted their own applications of these revolutionary technologies.

We encouraged applicants from a diversity of scientific backgrounds including molecular evolution, development, neuroscience, cancer, plant biology, and microbiology.

This course was supported by the National Human Genome Research Institute and the Howard Hughes Medical Institute, with major equipment provided by Illumina and Life Technologies.

#### PARTICIPANTS

Broner, S., B.Sc., University of Southern Denmark, Sonderborg

Contreras, R., Ph.D., University of Michigan, Ann Arbor

Cottrell, C., Ph.D., Washington University School of Medicine, St. Louis, Missouri

Fox, P.K., B.A., University of Washington, Seattle

Gibbs, L., B.S., University of North Texas Health Science Center, Fort Worth

Hlouskova, P., B.C., Centre European Institute of Technology MU, Brno, Czech Republic

Lindblom, K., B.S., Duke University, Durham, North Carolina

Lum, C., M.D., University of Hawaii, Honolulu

Ngun, T., Ph.D., University of California, Los Angeles

Oliver, D., B.S., University of South Carolina, Columbia

Przybyl, J., M.Sc., Stanford University, California

Shih, A., Ph.D., Memorial Sloan-Kettering Cancer Center, New York

Teich, A., Ph.D., Columbia University, New York

Tnah, L.H., Ph.D., Forest Research Institute Malaysia, Selangor, Malaysia

Trimmer, C., Ph.D., Monell Chemical Senses Center, Philadelphia, Pennsylvania

Vega Melendez, C., Ph.D., University of North Carolina, Greensboro

Waanders, E., Ph.D., Radboud University Medical Center, Nijmegen, The Netherlands

Watkins, D., Ph.D., SAHMNRI, Adelaide, Australia

Wray, B., B.S., Field Museum of Natural History, Chicago, Illinois

Zolnik, C., M.S., Fordham University, Bronx, New York

#### SEMINARS

Moore, B., University of Utah, Salt Lake City: VAAST and pVAAST algorithms for genetic analysis.

Mardis, E., Washington University School of Medicine, St. Louis, Missouri: NGS-based cancer genomics.

Dewar, K., McGill University and Genome QC Innovation Centre, Montreal, Canada: Genome assembly of PacBio + Illumina data.

Elemento, O., Weill Cornell Medical College, New York: Introduction to epigenomics analysis.

# The Genome Access Course

---

**INSTRUCTORS**     **E. Hodges**, Cold Spring Harbor Laboratory  
                          **G. Howell**, The Jackson Laboratory  
                          **B. King**, Mount Desert Island Biological Laboratory  
                          **J. Ward**, Middlebury College

**LECTURERS**        **A. Gordon**, Whitehead Institute for Biomedical Research  
                          **M. Schatz**, Cold Spring Harbor Laboratory  
                          **M. Zody**, New York Genome Center

The Genome Access Course (TGAC) is an intensive 2-day introduction to bioinformatics that was held three times in 2014 and trained more than 110 participants in total. Registration for the course is open to all on a first-come, first-served basis subject to basic eligibility requirements. The core of TGAC covers the manipulation and analysis of sequence data using publicly available tools. The course is broken into modules designed to give a broad overview of a given topic, with ample time for hands-on exercises chosen by the instructors. Each module includes three parts: (1) a discussion of theory and methods, (2) coverage of software and Web resources, and (3) use of selected tools in practice. The modular design allows the instructors to tailor the curriculum to the interests of the students. Modules in 2014 included Genome Sequencing and Assembly, Gene Prediction, the UCSC Genome Browser, Ensembl, Comparative Genome Analysis, Gene Set Enrichment and Pathway Analysis, High-Throughput Sequence Analysis, The Galaxy Project, and RNA-Seq Analysis Using R. Students were encouraged to bring questions and data from their own research projects to the course, and they were also encouraged to contact instructors with additional questions once they returned to their home institutions.





As in 2013, one of the 2014 TGACs was held in Manhattan at the New York Genome Center (NYGC). It marked the second collaborative training project with the NYGC, for which CSHL is an institutional founding member. The Manhattan course had a very high enrollment of 42 students, over half of whom were from local New York institutions. CSHL plans to offer the Manhattan course again in August/September 2015 in collaboration with the New York Genome Center, and annually thereafter as long as demand for the course continues.

### March 31–April 2 (Cold Spring Harbor Laboratory)

**Students: 35**

#### PARTICIPANTS

Cloud, V., Stowers Institute for Medical Research  
Dus, M., Skirball Institute, New York University School of Medicine  
Ghazi, I., University of Hyderabad, India  
Grobler, Y., New York University School of Medicine  
Haines, A., Norfolk State University  
Hrncir, T., Institute of Microbiology, Academy of Sciences CR  
Huang, L., ISIS Pharmaceuticals  
Kelly, T., Massachusetts Institute of Technology  
Khamit-Kush, K., Human Genome Sequencing Center Pre-Graduate Education Training, Baylor College of Medicine  
Kim, J., Mayo Clinic College of Medicine  
Liang, Y., The Rockefeller University  
Liu, P., Booz Allen Hamilton  
Lum, C., University of Hawaii, John A. Burns School of Medicine  
Morris, L., Memorial Sloan-Kettering Cancer Center

Murray, D., Baylor College of Medicine  
Okonkwo, L., Winthrop University Hospital  
Orshinsky, A., University of Minnesota  
Pascuzzi, P., Purdue University  
Perron, G., University of Ottawa  
Rebecca, V.H., Lee Moffitt Cancer Center  
Renwick, N., The Rockefeller University  
Richter, S., University of Toronto  
Rupp, L., University of Pennsylvania  
Santiago-Vazquez, L., University of Houston, Clear Lake  
Sengupta, S., North Shore University Health System  
Solano, V., Montclair State University  
Spinler, J., Baylor College of Medicine  
Yoo, S., STCube Pharmaceuticals Inc.  
Yoon, H., Mayo Graduate School  
Zhu, Y., North Shore University Health System/University of Chicago



## September 2–4 (New York Genome Center)

Students: 42

## PARTICIPANTS

Anike, F., North Carolina Agricultural and Technical State University  
 Cañas Villamar, I., Instituto Potosino de Investigacion Científica y Tecnológica  
 Dao, F., Memorial Sloan-Kettering Cancer Center  
 de Planell Saguer, M., Columbia University  
 Farris, H., The MITRE Corporation  
 Frank, M., New York Genome Center  
 Gnatenko, D., State University of New York, Stony Brook  
 Guerra Amorim, C., Columbia University  
 Hernandez, E., Memorial Sloan-Kettering Institute  
 Isikhuemhen, O., North Carolina Agricultural and Technical State University  
 Kaur, P., Georgia State University  
 Kim, I., Seoul National University  
 Kim, J., The Rockefeller University  
 Kolander, T., Newport Laboratories  
 Kwan, J., University of Illinois, Chicago  
 Lee, K., Pfizer  
 LeFrancois, M., New York University Langone Medical Center  
 Li, Q., Memorial Sloan-Kettering Cancer Center  
 Liu, Y., Stony Brook University  
 Mallipattu, S., Stony Brook University  
 Moore, K., New York University Langone Medical Center  
 Nerbonne, J., Washington University School of Medicine  
 Niescierowicz, K., International Institute of Molecular and Cell Biology  
 Orange, D., The Rockefeller University  
 Peled, M., New York University Medical Center  
 Rahman, K., New York University School of Medicine  
 Rinehold, K., New York University Langone Medical Center  
 Rodrigues, A., Federal University of São Paulo  
 Sabando Eguizabal, A., Stony Brook University  
 Scott, D., Icahn School of Medicine, Mount Sinai  
 Siebert, A., University of Rochester  
 Singh, R., Weill Cornell Graduate School of Medical Sciences  
 Solomon, W., Maimonides Cancer Center  
 Sovegni, G., Federal University of São Paulo  
 Still, D., California State Polytechnic University, Pomona  
 Taub, D., Southwestern University  
 Taylor, R., Northwestern University Feinberg School of Medicine  
 VanSolingen, C., New York University Langone Medical Center  
 Varma, H., Columbia University  
 Vogel, T., Washington University in St. Louis  
 Yimlamai, D., Boston Children's Hospital



**November 12–14 (Cold Spring Harbor Laboratory)****Students: 31**

## PARTICIPANTS

Ang, R., Icahn School of Medicine, Mount Sinai  
 Barrett, T., New York University School of Medicine  
 Boland, J., Leidos Biomedical Research, Inc.  
 Byun, M., Washington University, St Louis  
 Calvert, M., Temasek Lifesciences Laboratories  
 Chaudhuri, D., Boston Children's Hospital  
 Christensen, C., Dana-Farber Cancer Institute  
 Cullen, M., Leidos Biomedical Research, Inc.  
 De Maio, A., Baylor College of Medicine, Neurological  
 Research Institute  
 Deck, K., University of Wisconsin, Madison  
 Elfekih, S., Commonwealth Scientific and Industrial  
 Research Organization  
 Fisher, E., New York University School of Medicine  
 Gomes, A., Icahn School of Medicine, Mount Sinai  
 Gonzalez, C., Columbia University  
 Gorres, K., Yale University  
 Jeong, Y., University of Illinois at Chicago  
 Kataoka, A., Grocery Manufacturers Association  
 Khong, S., Stanford University  
 Lee, D.W., Yonsei University  
 Lim, H., Hunter College of the City University of New York  
 Lipatov, M., Stony Brook University  
 Liu, J., Sanofi Pharmaceuticals  
 Meng, X., Baylor College of Medicine  
 OHara, J., University of Maryland School of Public Health  
 Paczkowska, M., Medical University of Bialystok  
 Picoraro, J., Columbia University Medical Center  
 Ramkhelawon, B., New York University Langone Medical  
 Center  
 Spatafora, G., Middlebury College  
 Venkatasubramanian, M., University of Cincinnati  
 Viereg, J., University of Chicago  
 Wiener, B., Newport Laboratories, Inc.

**The Laboratory acknowledges the generosity of the following companies that loaned equipment and reagents to the various courses:**

AB SCIEX LLC	Diagenode	Luigs & Neumann	Roche Applied Science
Agilent Technologies Inc.	Drummond Scientific Company	Mightex	Roche Nimblegen
ALA Scientific Instruments	eBioscience	Millipore Corporation	Sage Science Inc.
A-Systems	Electron Microscopy Science	Miltenyi Biotec Inc.	Sakura Finetek USA Inc.
Andor Technology	Eppendorf North America	Molecular Devices	Santa Cruz Biotechnology Inc.
Applied Precision	Fotodyne Inc.	Nanodrop Technologies	Scientifica Ltd.
Applied Scientific Instrumentation Inc.	GE Healthcare	NanoTemper Technologies	SciGene
Berthold Technologies USA, LLC	Hamamatsu Photonic Systems	Narishige International USA Inc.	Singer Instruments
BioLegend	Hampton Research	Nasco	SONY Biotechnology Inc.
Bioline USA	Harlan	Neta Scientific	Spectra-Physics Lasers Inc.
Bio-Rad Laboratories	Harvard Apparatus Inc.	New England Biolabs Inc.	Spectral Applied Research
BioTeck	Heka Instruments	Newport Corporation	Stemgent
BioVision	Hellma USA Inc.	Nikon Inc.	Sutter Instruments
Bitplane	Illumina Inc.	Nonlinear Dynamics	Taconic Farms
Carl Zeiss Inc.	Integrated DNA Technologies	NSK America Corporation	Tecan US, Inc.
Charles River Laboratories Inc.	Invitrogen	Olympus America Inc.	The Jackson Laboratories
Chroma Technology Corporation	Kapa Biosystems	Parkell	The Mathworks Inc.
Clontech	Lafayette Instrument	PerkinElmer Life and Analytical Sciences	Thermo Fisher Scientific
Coherent Laser	Leica Biosystems Inc.	Photometrics	Thor Labs
ConOptics	Leica Microsystems Inc.	Pressure Biosciences	Tokai Hit
Covaris, Inc.	Life Technologies	Prior Scientific	Tonbo Biosciences
CrystalLaser	LI-COR	Promega Corporation	TotalLab Ltd.
Crystalgen Inc.	Lonza Pharma-Bioscience Solutions	Qiagen Inc.	Vector Laboratories Inc.
Dage-MTI		Rapp Optoelectronic GmbH	Warner Instruments
		RedShirt Imaging LLC	Waters Corporation
			World Precision Instruments
			Zymo Research Corporation

# 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 an opportunity for the exchange of ideas in an informal setting.

Speaker	Title	Host
<b>January</b>		
Kenneth Zaret, Penn Institute for Regenerative Medicine, University of Pennsylvania	Nucleosomes, enhancers, and a basis for mitotic memory.	Thomas Gingeras
Stephen Harrison, Harvard Medical School	Antibody affinity maturation: From molecular genealogy to molecular structure.	Leemor Joshua-Tor
Richard Tsien, New York University Langone Medical Center	Autism: Related genes and signal transduction pathways.	Bo Li
<b>February</b>		
David Liu, Ph.D., Harvard University	Integrating chemistry and evolution to illuminate and program biology.	Gholson Lyon
<b>March</b>		
Jay Keasling, Ph.D., University of California, Berkeley, Lawrence Berkeley National Laboratory	Synthetic biology for synthetic chemistry.	Ian Peikon
Gregory Petsko, Weill Cornell Medical College	How Parkinson's disease starts and how it might be stopped.	Leemor Joshua-Tor
Xiaowei Zhuang, Ph.D., Harvard University	Bioimaging at the nanoscale: Single-molecule and superresolution fluorescence microscopy.	Leemor Joshua-Tor
<b>April</b>		
Robert Malenka, M.D., Ph.D., Stanford University	Neural mechanisms underlying reward and aversion.	Pavel Osten
Daniel Haber, M.D., Ph.D., Massachusetts General Hospital Cancer Center, Harvard Medical School	Molecular characterization of single circulating tumor cells.	Raffaella Sordella
<b>October</b>		
Michael Bevan, John Innes Centre	Understanding and exploiting plant genomes for crop improvement.	Richard McCombie
Roberto Malinow, Ph.D., University of California, San Diego	Synapses in health and disease.	Bo Li
Fang Xu, University of British Columbia	Genetic and molecular analysis of resistance protein-mediated plant immunity.	Dave Jackson
Navdeep Chandel, Northwestern University	Mitochondria as signaling organelles.	Raffaella Sordella
<b>November</b>		
Bart De Strooper, KU Leuven, The Netherlands	$\beta$ and $\gamma$ secretases: Understanding their biology and how this defines the therapeutic options in Alzheimer disease.	Linda Van Aelst
Paul Sternberg, California Institute of Technology	Molecular circuitry for sleep, sex, and predation.	Chris Hammell
Molly Przeworski, Columbia University	An evolutionary perspective on human germline mutation.	Mickey Atwal
Scott Keeney, Weill Cornell Graduate School of Medical Sciences	A dangerous game: Controlling DNA breaks in meiosis.	Bruce Stillman

**IN-HOUSE SEMINAR PROGRAM**

Cold Spring Harbor Laboratory'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.

<b>Speaker</b>	<b>Title</b>
<b>January</b>	
Stephen Shea	Inhibitory circuits for social communication in mice.
Chris Hammell	Oscillatory gene expression and cell fate specifications.
Gholson Lyon	Deep brain stimulation, psychiatric genetics, and iPS cell models of disease.
Chris Vakoc	Chromatin regulators as cancer dependencies.
<b>February</b>	
Florin Albeanu	Feedforward and feedback signals in mammalian olfaction.
Lloyd Trotman	What signals the end of signals?
<b>March</b>	
Tom Gingeras	Conservation and novel functions of noncoding RNAs.
Pavel Osten	Mapping behavior and counting cells.
Camila Dos Santos	Epigenetic control of normal and malignant mammary gland development.
<b>April</b>	
Senthil Muthuswamy	Cell polarity and 3D organoids: A trend becomes mainstay in cancer biology.
<b>October</b>	
Erkan Karakas (Furukawa Lab)	Structural characterization of NMDA receptor ion channels.
Mario Penzo (Li Lab)	Neuronal circuits controlling the formation and recollection of fear memories.
Olga Anczukow-Camarda (Krainer Lab)	Alternative-splicing misregulation in breast cancer.
<b>November</b>	
Cora MacAlister (Lippman Lab)	Life at the tip: Protein glycosylation in pollen tubes and protonema.
Keerthi Krishnan (Huang Lab)	Elucidating the pathogenesis of Rett syndrome, a neurological disorder.
Wee Siong Goh (Hannon Lab)	Just add human: Finding the real function of mouse piRNAs.
<b>December</b>	
Bruce Stillman	ATP-driven machines control the cell division cycle.
Gayatri Arun (Spector Lab)	Probing the therapeutic potential of Malat1 in mammary tumor progression.



BANBURY CENTER



# BANBURY CENTER

## EXECUTIVE DIRECTOR'S REPORT

---

We were relieved after the overwhelming program of 2013 that 2014 was a little less hectic, but we were nevertheless still busy dealing with 15 events and 500 participants.

Participants in the meetings were drawn from 37 states, and as usual, four states—California, Maryland, Massachusetts, and New York—accounted for 60% of participants. Banbury meetings continue to have strong international participation, with 19% of participants coming from 21 countries. Thirty-five percent of participants were female, an unusually high percentage. The number of female participants has doubled over the years since 1988.

### Education

The majority of the Banbury Center program is devoted to meetings dealing with research topics in biomedical research—meetings that might be considered as high-level continuing education for scientists. But each year, the Center also provides a venue for the education of young scientists who are in the early stages of their careers.

One such meeting was the Workshop on Leadership in BioScience, taught by Carl M. Cohen (Science Management Associates) and Danielle Kennedy (Worklab Consulting LLC). David Stewart was awarded a 4-year grant from American Express to support the program. The course is attended primarily by senior postdocs who are going to run their own laboratory and by young faculty. They receive training in, for example, the characteristics of a good manager, how to interact with people, and how to control the dynamics of meetings.

The Boehringer Ingelheim Fonds holds two retreats each year for its fellows, one in Europe and one in the United States. Rather like the Workshop on Leadership in BioScience course, the Boehringer fellows receive intense instruction in writing, making presentations, and the skills



Cocktails on Robertson House patio



needed to carry out research. The Foundation first came to Banbury in 2005 and in alternate years until 2011, and now comes annually.

Finally, there was the NIMH Brain Camp. Tom Insel, director of the National Institute of Mental Health, has long been passionate about the need to base psychiatry and the treatment of psychiatric disorders on the findings of neuroscience. To that end, beginning in 2009, Insel has brought 24 of his brightest clinical psychiatric fellows to Banbury to expose them to the highest-quality research and to persuade them that they should consider research careers. Speakers have included Eric Nestler, Karl Deisseroth, Eric Kandel, Steven Paul, and Huda Zoghbi.



Robertson House

### Banbury Meetings on Health

A second meeting demonstrating the power of personal experience was *Rhabdomyosarcoma: A Critical Review of Research and What It Means for Developing Therapies*. Rhabdomyosarcoma is the most common soft tissue sarcoma of childhood, but despite four decades of advances in therapy, the outcome for metastatic or relapsed disease is particularly poor. This meeting, funded by families determined that others should not suffer, examined all aspects of rhabdomyosarcoma biology with three questions in mind: What are the key areas for future research? What can be done to ensure funding for research? How can promising laboratory findings be turned to developing drugs for clinical trials?

The meeting with the rather prosaic title *Interpreting Personal Genomes: How Are We to Set Appropriate Statistical Standards for Identifying Pathogenic Genetic Variants?* dealt with a critical issue in personal medicine: How can we determine whether a genetic variant may be responsible for an observed disease phenotype? Current tools for identifying variants as potentially disease-causing are far from optimal, raising the risk of false clinical diagnoses. Participants in this meeting examined current methods and discussed how standards could be set for ensuring that appropriate analyses are performed.

Compared with other disorders, research on and care of the mentally ill is notoriously underfunded relative to the impact the illnesses have on society. Other disorders, such as AIDS, have strong advocates, but the stigma associated with mental illness seems to inhibit advocacy. Not so for Glenn Close, whose sister and nephew are afflicted with bipolar disorder. Close's Bring Change 2 Mind Foundation is dedicated to reducing stigma and has embarked on developing a "college toolbox" to eliminate stigma that can be implemented in a 4-year university course. This meeting brought together individuals whose knowledge and expertise can help plan a pilot project. It was a fascinating meeting in which Close played a very active role.

Another meeting dealing with mental health was *Lewy Body Dementia: Current Status, Future Directions*.



Sammis, Spring 2014



Reception at Robertson House



Coffee break

This is the second most common cause of cognitive impairment after Alzheimer's disease. Although a great deal is known about the molecular processes of Alzheimer's, much less is known about Lewy body dementia (LBD) despite the fact that it makes up 30% of all dementia cases. Participants in the meeting critically reviewed the current state of knowledge of the genetics of LBD and the usefulness or otherwise of current clinical, imaging, and biological markers. There was much discussion of how global research efforts on LBD could be delivered in conjunction with the goals of NINDS, NAPA, the G8 Summit Objectives, and related initiatives.

Finally, *The Genetics of Pain and Pain Inhibition: Where to from Here?* considered what is the most prevalent human health problem, with a lifetime prevalence of almost one in two. Association studies and exome sequencing studies of chronic pain disorders are now being published, and rare genetic variants responsible for pain disorders have been identified. Participants reviewed the relative merits of the association studies and single-gene approaches for the study of chronic pain, along with various current therapies in the field.

### Other Notable Meetings

Epigenetic regulation plays a pivotal role in plant development and offers a largely untapped resource for crop improvement strategies aimed at enhancing productivity. The *Epigenetics and Agriculture* meeting focused on epigenetic mechanisms of gene regulation and their roles in heterosis, epigenetic programming of plant reproduction, transgenerational inheritance, and adaptation to abiotic and biotic stresses. Participants also explored the needs of agricultural biotechnology, and how epigenetic research can help efforts to manipulate gene expression in crops toward increasing sustainable food and feed production, to meet the needs of a growing population. The meeting brought leading researchers in plant epigenetics together with scientists representing agricultural biotechnology companies.

One of the more unusual meetings was *Interdisciplinary Symposium on Creativity*, organized by Suzanne Nalbantian, C.W. Post College, Westbury, New York. This was the fourth, and the third held at Banbury, in a series of meetings bringing together neuroscientists and scholars from the humanities. The neuroscientists described the functioning of the brain in creative acts of scientific discovery or aesthetic production. The comparatists described instances of creativity in the composition of major literary works, of musical compositions, or of works of visual art.

## Banbury Center History

As Banbury approaches the 40th anniversary of its founding, we are trying to produce as true a record as possible of the activities of the Banbury Center since the first meeting in 1978. The current database goes back only to 1987 when I arrived, so we have made a second database covering the years 1978–1986 using the information recorded in the annual reports. However, as these record only those who gave talks at Banbury and not all attendees, we are going back through the files using other information such as housing lists. The historical database currently lists 7000 participants, and the contemporary database lists more than 11,000.

A second project concerns the black-and-white photographs of Banbury meetings taken from 1978 to 2003 when digital photography began. These images have been inaccessible because they exist only as negatives, and there is no catalog of them and no way of knowing whether a particular scientist is in an image. Now the negatives have been scanned and entered into a database. The difficult job remains—to identify every individual in more than 20,000 negatives!

The default position regarding publication of reports of Banbury Center meetings is that we do not do so. This is in part so as not to deter people from attending by making submission of a manuscript a prerequisite for participating in a Banbury meeting; we also want to encourage presentation of unpublished material. However, if the participants choose to develop a publication, they are strongly encouraged to do so. With the help of Richard Sever, we have created a category called “Banbury White Paper” in the CSHL Press’s *Perspectives* series edited by Richard. Publications in 2014 included:

Erlich Y, et al. 2014. Redefining genomic privacy: Trust and empowerment. *PLOS Biology* **12**: 1–5.

Hettmer S, et al. 2014. Rhabdomyosarcoma: Current challenges and their implications for developing therapies. *Cold Spring Harb Perspect Med* **4**: a025650.

Lee FS, et al. 2014. Adolescent mental health—Opportunity and obligation: Emerging neuroscience offers hope for treatments. *Science* **346**: 547–549.

## Acknowledgments

For more than 6 years, Hakon Heimer has been helping develop meetings in neuroscience and mental health. Hakon’s background is in neurobiology, but for many years, he has been a writer and advocate for mental health research. He established the Schizophrenia Research Forum, the top online source for information for both the public and professionals. Hakon is a member of



the National Institute of Mental Health Council and of the National Institutes of Health Council of Councils. Hakon's help has been invaluable in thinking of topics for meetings and in finding funding for them. The current funding situation is so difficult that for every meeting we hold, we have to be working on at least six others concurrently—an impossible task for one person.

And, of course, Banbury could not function at the level it does without the hard work and help of others: Janice Tozzo and Pat Iannotti in the Banbury office and Basia Polakowski at Robertson House. Culinary Services, Facilities, and the Meetings Office play key roles in the operation of the Center, and Jose Covera, Joe McCoy, and Fredy Vasquez keep the Banbury estate looking beautiful. The meetings could not take place without the enthusiasm and hard work of the organizers, the contributions made by all participants, the generosity of the Laboratory's Corporate Sponsors and the other donors who fund our meetings, and the Laboratory's scientists who continue to support the Center.

**Jan A. Witkowski**  
*Executive Director*

**BANBURY CENTER MEETINGS**

<i>Date</i>	<i>Title</i>	<i>Organizer(s)</i>
February 23–26	Banbury Summit III: Genetics Training in the Genomic Era	Bruce Korf, ACMG Work Group
March 28–April 2	Communicating Science	Claudia Walther, Sandra Schedler
April 6–9	Connections and Communications in the Brain	Bijan Pesaran, Nicolas Brunel
April 16–18	The College Toolbox Project: Eliminating Stigma	Bernice Pescosolido
April 21–23	Defeating Ovarian Cancer	Ronald Buckanovich, Gordon Mills, David Tuveson
May 10–12	NIMH Brain Camp VI	Joyce Chung, Thomas Insel
May 13–16	Rhabdomyosarcoma: A Critical Review of Research and What It Means for Developing Therapies	Charles Keller, Amy Wagers
June 22–25	The Genetics of Pain and Pain Inhibition: Where to from Here?	Jeffrey Mogil, Clifford Woolf
September 3–5	High-Performance Computing in Undergraduate Biology Education: Scanning the Landscape	David Micklos, Dan Stanzione
September 7–10	The Immune System and Cancer	Glenn Dranoff, Douglas Fearon
September 15–18	Interpreting Personal Genomes: How Are We to Set Appropriate Statistical Standards for Identifying Pathogenic Genetic Variants?	David Goldstein, Daniel MacArthur
October 21–24	Interdisciplinary Symposium on Creativity	Suzanne Nalbantian
October 26–29	ROS in Biology and Cancer	Arne Holmgren, Navdeep Chandel, Nicholas Tonks, David Tuveson
November 9–12	Epigenetics and Agriculture	Brian Hauge, Rob Martienssen
November 16–19	Lewy Body Dementia: Current Status, Future Directions	James Galvin, Ian McKeith

# BANBURY CENTER MEETINGS

## Banbury Summit III: Genetics Training in the Genomic Era

February 23–26

FUNDED BY American College of Medical Genetics

ARRANGED BY B. Korf, University of Alabama, Birmingham and the ACMG Work Group

The field of medical genetics is in a rapid state of flux as genomic approaches revolutionize the diagnosis of both rare and common genetic conditions, and insights into pathogenesis open possibilities for treatment of an increasing number of disorders. These changes not only require new approaches to training medical geneticists, but also raise questions about the scope of practice of medical geneticists versus other medical specialists, as genomic tests become increasingly available and accessible. These issues may warrant changes in the system of training of medical geneticists. This conference brought together major stakeholders in the medical genetics community to review current approaches to training and consider development of new approaches.



G. Herman, B. Korf, J. Witkowski

**Introductory Remarks:** J.A. Witkowski, Banbury Center, Cold Spring Harbor Laboratory

**Welcome and Overview:** B. Korf, University of Alabama, Birmingham



**SESSION 1: Medical Genetics Education and Training**

**Current Status of Clinical Genetics Training and Board Certification**

ABMG	M. Blitzer, University of Maryland School of Medicine, Baltimore
ACMG	B. Korf, University of Alabama, Birmingham
RRC	R. Sutton, Baylor College of Medicine, Houston, Texas
AMGF	M. Watson, American College of Medical Genetics, Bethesda, Maryland
APHMG	L. Demmer, Carolinas Medical Center, Charlotte, North Carolina

**Current Status of Laboratory Genetics Training and Board Certification**

ABMG	M. Blitzer, University of Maryland School of Medicine, Baltimore
Cytogenetics	K. Rao, University of North Carolina, Chapel Hill J. Gastier-Foster, Nationwide Children's Hospital, Columbus, Ohio
Molecular Genetics	J. Feldman, Wayne State University School of Medicine, Detroit, Michigan J. Gastier-Foster, Nationwide Children's Hospital, Columbus, Ohio
Biochemical Genetics	M. Blitzer, University of Maryland School of Medicine, Baltimore
ACMG	M. Watson, American College of Medical Genetics, Bethesda, Maryland
APHMG	L. Demmer, Carolinas Medical Center, Charlotte, North Carolina
ASHG	C. Morton, Brigham & Women's Hospital, Boston, Massachusetts

**SESSION 2: New Opportunities and Challenges in a Genomics World**

**Integration of Genetics into Medical Specialties**

**Moderator:** L. Demmer, Carolinas Medical Center, Charlotte, North Carolina  
Cancer Genetics: W. Chung, Columbia University, New York  
Cardiovascular Genetics: A. Roberts, Boston Children's Hospital, Boston, Massachusetts  
Neurogenetics: B. Korf, University of Alabama, Birmingham

**Genomics and Genome Sequencing**

Introduction: J. Feldman, Wayne State University School of Medicine, Detroit, Michigan  
Genetic Counseling for Genome Sequencing: Genetic Counselors  
Training Medical Geneticists for Genome Sequencing: R. Sutton, Baylor College of Medicine, Houston, Texas



J. Hoskovec, A. Matthews, S. Hahn, R. Bennett

Should There be a Medical Genomics Training Pathway?:

J. Feldman, Wayne State University School of Medicine,  
Detroit, Michigan

Discussion

#### **Training in Therapeutics and Clinical Trials**

**Moderator:** B. Korf, University of Alabama, Birmingham

Introduction: B. Korf, University of Alabama, Birmingham

Biochemical Genetics: G. Herman, Nationwide Children's  
Hospital Research Institute, Columbus, Ohio

Small-Molecule and Other Treatments for Genetic Disorders,  
e.g., Rasopathies: A. Roberts, Boston Children's Hospital,  
Boston, Massachusetts

Discussion

#### **Future of Laboratory Training**

**Moderators:** J. Gastier-Foster, Nationwide Children's Hospital,  
Columbus, Ohio, and M. Blitzler, University of Maryland,  
Baltimore

Future of Biochemical Genetics Training: R. Sutton, Baylor  
College of Medicine, Houston, Texas

Training in Laboratory Genomics: K. Rao, University of  
North Carolina, Chapel Hill

Discussion

#### **Review of Meeting Statement**

Continued discussion and action items.



# Communicating Science

---

March 28–April 2

FUNDED BY **Boehringer Ingelheim Fonds Foundation for Basic Research in Medicine**

ARRANGED BY **C. Walther**, *Boehringer Ingelheim Fonds, Mainz, Germany*  
**S. Schedler**, *Boehringer Ingelheim Fonds, Mainz, Germany*

The *Boehringer Ingelheim Fonds* has an international program of support for Ph.D. fellowships, and it first brought its fellows to the Banbury Center for their annual North American retreat in 2005. It has been a great pleasure to have them return, and their 2014 stay at Banbury was the seventh occasion that they have been here. At Banbury, the fellows receive intensive instruction in matters such as giving presentations and writing papers—topics usually learned by default (and often poorly) during graduate research.

**Introductory Remarks:** J.A. Witkowski, *Banbury Center, Cold Spring Harbor Laboratory*

**Opening Remarks and All About BIF (Part 1):** C. Walther, *Boehringer Ingelheim Foundation, Mainz, Germany*

K. Achenbach, *Boehringer Ingelheim Foundation, Mainz, Germany*: Communication: Why and how?

N. LeBrasseur, *DNA Medical Communications, New York*: Writing techniques and how to structure papers; Writing assignment 1.

B. Tansey, *Vanderbilt University, Nashville, Tennessee*: Preparing and delivering a scientific talk; Group A: 4-min Powerpoint presentations, videotaped with replay and feedback.

N. LeBrasseur, *DNA Medical Communications, New York*: Discussion of writing assignment 1; Writing assignment 2.

B. Tansey, *Vanderbilt University, Nashville, Tennessee*: Group B: 4-min Powerpoint presentations, videotaped with replay and feedback.

Group A: Time for preparing 3-min presentations.

Group B: 4-min presentations, Group A: 3-min preparation.



Group A: 3-min Powerpoint presentations, videotaped with replay and feedback.

Group B: Time for preparing 3-min presentations.

N. LeBrasseur, DNA Medical Communications, New York: Return and discussion of writing assignment 2.

B. Tansey, Vanderbilt University, Nashville, Tennessee: Group B: 3-min PowerPoint presentations, videotaped with replay and feedback.

S. Pfeiffer, Lazard Frères & Co. LLC, New York: Career talk.

K. Ris-Vicari, Katie Ris-Vicari Graphic Design, Bethpage, New York, and M. Hansen, Nature Publishing Group, New York: How to design figures.

C. Walther, Boehringer Ingelheim Foundation, Mainz, Germany: All about BIF: Part 2 and feedback.

**Guided Walking Tour on CSHL Campus**

# Connections and Communications in the Brain

---

April 6–9

FUNDED BY            **The Swartz Foundation, Setauket, New York**

ARRANGED BY        **B. Pesaran, New York University, New York**  
**N. Brunel, University of Chicago, Illinois**

Over many decades, neuroscience has been deeply influenced by evidence that specific behavioral processes are localized to particular brain regions. Anatomical studies have revealed how the brain is organized into different regions. Functional studies have shown how activity in different brain regions is specific for different behavioral processes. Focal damage to the brain has been shown to result in remarkably precise behavioral deficits. However, it is also clear that brain areas interact to guide behavior by communicating with each other, and they do so over neuronal projection systems formed by large populations of neurons that are not localized to specific brain regions. To understand how different areas cooperate to engage in a particular behavior, we must understand both the structure of the connectivity between these areas and the rules that govern the communication between them. The goal of this workshop was to encourage researchers using experimental and theoretical approaches to bring new data, tools, concepts, and ideas to bear on understanding the mechanisms and functional significance of communication in the brain.

Welcoming Remarks: J.A. Witkowski, Banbury Center, Cold Spring Harbor Laboratory

Introduction: N. Brunel, University of Chicago, Illinois  
                  B. Pesaran, New York University, New York



**SESSION 1**

H. Kennedy, INSERM, Bron, France: What inconsistent data fail to tell you about cortical networks.

A. Burkhalter, Washington University School of Medicine, St. Louis, Missouri: Neuronal network of the mouse visual cortex.

H. Monyer, University of Heidelberg, Germany: GABAergic cells in the hippocampal–entorhinal formation and their role in spatial coding and memory.

G. Buzsáki, New York University, New York: Communication by spikes in the hippocampal–entorhinal system.

**SESSION 2**

X.-J. Wang, New York University, New York: Building a large-scale model of the primate cortex: Structure and dynamics.

D. Bassett, University of Pennsylvania, Philadelphia: Probing human brain network dynamics during learning.

O. Jensen, Radboud University, Nijmegen, The Netherlands: Temporal coding organized by coupled  $\alpha$  and  $\gamma$  oscillations prioritize visual processing.

S. Bressler, Florida Atlantic University, Boca Raton, Florida: Beta synchrony and top-down feedforward processing in visual expectation.

**General Discussion****SESSION 3**

M. Siegel, University of Tuebingen, Germany: Spectral fingerprints of large-scale neuronal interactions in the human brain.

P. Fries, Ernst Strungmann Institute for Neuroscience, Frankfurt, Germany: Brain-wide and cell-type-specific synchronization at the service of attention.

J. Maunsell, Harvard University, Boston, Massachusetts: Cerebral cortex as a look-up table.

C. Gray, Montana State University, Bozeman: Distributed cortical processing during visual working memory.

**SESSION 4**

N. Kopell, Boston University, Massachusetts: Cortical rhythms facilitate bottom-up and top-down processing.

T. Akam, Centro Champalimaud, Lisboa, Portugal: Neural codes supporting oscillatory control of effective connectivity among brain regions.

D. Battaglia, Institute for System Neuroscience, Marseille, France: Collective dynamics of multiscale circuits shape functional interactions.

T. Buschman, Princeton University, New Jersey: Neural dynamics of cognitive flexibility.

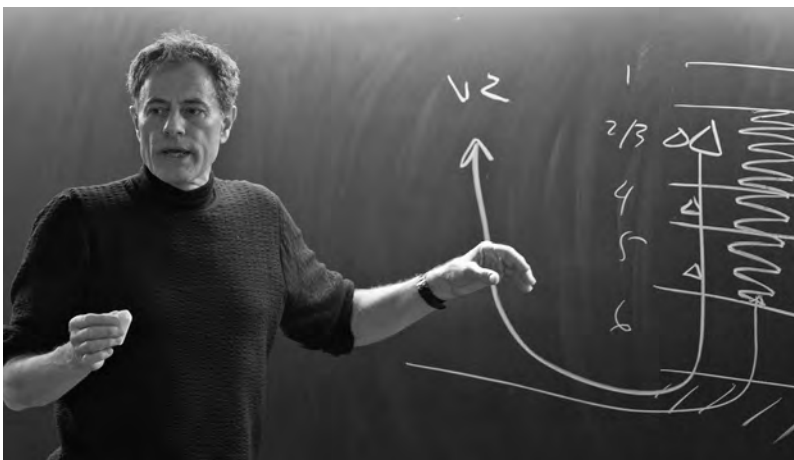
S. Fusi, Columbia University, New York: High dimensional neural representations in prefrontal cortex.

**General Discussion****SESSION 5**

B. Pesaran, New York University, New York: A role for coherent neural activity in coordination and decision making.

T. Akam, Centro Champalimaud, Lisboa, Portugal: Neural codes supporting oscillatory control of effective connectivity among brain regions.

A. Compte, IDIBAPS, Barcelona, Spain: Stimulus fluctuations and top-down feedback can account for the dynamics of choice probability in MT.

**General Discussion****Closing Remarks**

A. Burkhalter



H. Cohen

# The College Toolbox Project: Eliminating Stigma

---

April 16–18

FUNDED BY **Bring Change 2 Mind, Indiana University Foundation, The Margaret Clark Morgan Foundation and individual participants**

ARRANGED BY **B. Pescosolido, Indiana University, Bloomington, Indiana**

Since the mid-1990s, there has been a resurgence of research in the United States on the stigma attached to mental illness. That research has provided clear findings for efforts to reduce prejudice and discrimination. “Contact”—whether face-to-face or video—has been shown to be the most effective means of reducing stigma. Older individuals tend to harbor more out-of-date notions, while younger individuals have become more open to discussing these issues.

As a result, to eliminate stigma, Bring Change 2 Mind (BC2M) has decided to embark on the development of a “college toolbox” that can be designed and implemented as a 4-year project that will be piloted at Indiana University and disseminated through a national effort.

This meeting brought together individuals whose knowledge and expertise can help plan a pilot project. Participants included the best minds in research and intervention on stigma, as well as administrators, student groups, and the community. By the end of the meeting, it was expected that the group would have developed a clear outline of what the next steps should be and a timeline for their implementation.



B. Pescosolido

**Welcoming Remarks:** J.A. Witkowski, Banbury Center, Cold Spring Harbor Laboratory, B. Pescosolido, Indiana University, Bloomington, and G. Close, Bring Change 2 Mind, San Francisco, California



**SESSION 1: BC2M's Vision**

G. Close and P. Harrington, Bring Change 2 Mind, San Francisco, California: What do we know about mental illness and/or stigma among college students?

**Group Discussion**

C. Boyer, J. Lee, and B. Pescosolido, Indiana University, Bloomington

**SESSION 2: The Foundations of the College Toolbox Program**

B. Pescosolido, Indiana University, and G. Close and P. Harrington, Bring Change 2 Mind, San Francisco, California

**SESSION 3: Basic Study Goals and Design**

J. Martin, Y.Y. Ahn, B. Pescosolido, and E. Wright, Indiana University

**SESSION 4: What Do We Know About Stigma Interventions?**

B. Angell, Rutgers University, New Brunswick, New Jersey  
S. Evans-Lacko, University of London, United Kingdom

N. Bonfine, Northeast Ohio Medical University, Rootstown, Ohio

H. Heimer, Schizophrenia Research Forum, Providence, Rhode Island

J. Macphee, The Jed Foundation, New York

R. Kellar, The Margaret Clark Morgan Foundation, Hudson, Ohio

S. Barnett, Indiana University, Bloomington

**Group Discussion**

**SESSION 5: By Students, For Students**

L. Fasone, M. Oppenheim, R. Green, R. Martinez, and A. Parrill, Indiana University, Bloomington

**Group Discussion**

**SESSION 6: Next Steps, Action Steps**

B. Pescosolido, Indiana University, Bloomington, Indiana: Developing partnerships, getting permissions, setting program rollout.



R. Green, H. Heimer, G. Close, R. Martinez, Y.Y. Ahn

# Defeating Ovarian Cancer

---

April 21–23

FUNDED BY **Jonathan Gray; Cold Spring Harbor Laboratory Office of Development**

ARRANGED BY **R. Buckanovich, University of Michigan, Ann Arbor**  
**G. Mills, MD Anderson Cancer Center, University of Texas, Houston**  
**D. Tuveson, Cold Spring Harbor Laboratory**

This was the first Cold Spring Harbor Laboratory meeting on ovarian cancer. The organizers thought that the time was right for a critical review of current research that could illuminate new avenues for basic and translational research. Among the topics discussed were the molecular pathology of human ovarian cancer, models of ovarian cancer, and the best current and investigational approaches to patient management: cytotoxic, targeted, epigenetic, immunological, metabolic, and stem-cell-directed therapies.

Welcoming Remarks: **J.A. Witkowski, Banbury Center, Cold Spring Harbor Laboratory**



R. Buckanovich, J. Carstens

Introduction: **D. Tuveson, Cold Spring Harbor Laboratory**



**SESSION 1: Models: Cell Lines PDXs and Gemms**

- D. Dinulescu, Brigham and Women's Hospital, Boston, Massachusetts: Can animal models of disease improve early detection of tumors and precursor lesions?
- T. Ince, University of Miami Miller School of Medicine, Florida: Novel ovarian cell culture systems.
- K. Cho, University of Michigan Medical School, Ann Arbor: Mouse models in the context of the dualistic pathway of ovarian cancer pathogenesis.

**SESSION 2: TME**

- P. Sabbatini, Memorial Sloan-Kettering Cancer Center, New York: Tumor vaccines in ovarian cancer.
- J. Wolchok, Memorial Sloan-Kettering Cancer Center, New York: Immune modulators for cancer therapy: Assessing antagonists and agonists.

**SESSION 3: Traditional and Targeted Therapy for Ovarian Cancer**

- S. Domchek, Abramson Cancer Center, Philadelphia, Pennsylvania: Germline genetics and implications for ovarian cancer therapeutics.
- U. Matulonis, Dana-Farber Cancer Institute, Boston, Massachusetts: Combinations of biologic agents + PARP inhibitors and immunotherapy approaches to ovarian cancer.
- D. Bowtell, Peter MacCallum Cancer Centre, East Melbourne, Australia: Primary and acquired resistance in high-grade serous cancer.

**SESSION 4: Genetics/Epigenetics**

- D. Huntsman, BC Cancer Agency Research Centre, Vancouver, British Columbia: Endometriosis-associated ovarian cancers.
- D. Levine, Memorial Sloan-Kettering Cancer Center, New York: Clinically relevant genomic signatures and pathways.
- J. Brenton, Li Ka Shing Centre, Cambridge, United Kingdom: Monitoring disease response and genomic change with circulating tumor DNA.
- D. Solit, Memorial Sloan-Kettering Cancer Center, New York: Insights from the study of extraordinary responders.

**SESSION 5: CSC/Heterogeneity**

- S. Shah, University of British Columbia, Vancouver, Canada: Phylogenetic portraits of high-grade serous ovarian cancers.
- B. Neel, Princess Margaret Research/University Health Network, Toronto, Canada: Patient-derived xenografts for evaluating ovarian cancer drug response and tumor initiating cells.
- R. Buckanovich, University of Michigan, Ann Arbor: An ovarian cancer stem-like cell hierarchy and why it matters.

**SESSION 6: Circulating Tumor Cells**

- D. Tuveson, Cold Spring Harbor Laboratory: Discussion on advances.

**General Discussion**



D. Bowtell, C. Bales



S. Shah, J. Watson



# NIMH Brain Camp VI

---

May 10–12

FUNDED BY National Institute of Mental Health, National Institutes of Health

ARRANGED BY J. Chung, National Institute of Mental Health, NIH, Bethesda, Maryland  
T. Insel, National Institute of Mental Health, NIH, Bethesda, Maryland

Cold Spring Harbor Laboratory is renowned worldwide for its education programs, from high school level to the highest professional levels. One of the Banbury Center's contributions is to host the NIMH-sponsored "Brain Camp." The goal of the Brain Camp is to identify areas of neuroscience that are of interest and relevance to psychiatrists and to communicate these to a small group of outstanding psychiatry residents and research fellows. Some of the most distinguished and thoughtful neuroscientists in the country have come to the meetings as guest speakers. The goal of the series of meetings is to develop a neuroscience curriculum that can eventually be shared with psychiatry training programs around the country.

## Introductions and Opening Session

T. Insel, National Institute of Mental Health, NIH, Bethesda, Maryland: Psychiatry 2024.

S. Paul, Weill Cornell Medical College, New York, New York: Alzheimer's disease: A prototype for drug discovery and diagnostics in neuropsychiatry.

S. Lisanby, Duke University School of Medicine, Durham, North Carolina: Device-based neuromodulation: From engineering bench to bedside.

## Roundtable Discussion with All Participants

### SESSION 1: New Diagnostics/New Therapeutics

M. State, University of California, San Francisco: From genes to therapeutics in neurodevelopmental disorders.

S. Vinogradov, University of California, San Francisco: Neuroscience-informed cognitive training for impaired neural systems in schizophrenia.

### SESSION 2: New Diagnostics/New Therapeutics

C. Tamminga, Southwestern Medical Center, Dallas, Texas: From biomarkers to biotypes in psychosis.

### SESSION 3: New Diagnostics/New Therapeutics

E. Leibenluft, National Institute of Mental Health, Bethesda, Maryland: From DSM Bipolar disorder to RDoC irritability: A nosologic journey.

B. Teachman, University of Virginia, Charlottesville, Virginia: Regaining Control: Interventions for automatic and strategic cognitive biases in adult anxiety.

D. Pine, National Institute of Mental Health, Bethesda, Maryland: RDoC and mechanism-based therapeutics: Attention bias modification in pediatric anxiety.

# Rhabdomyosarcoma: A Critical Review of Research and What It Means for Developing Therapies

---

May 13–16

FUNDED BY Friends of the TJ Foundation, Inc., The Michelle Paternoster Foundation for Sarcoma Research

ARRANGED BY C. Keller, Oregon Health & Science University, Portland  
A. Wagers, Harvard Medical School, Boston, Massachusetts

Rhabdomyosarcoma is the most common soft tissue sarcoma of childhood, but despite four decades of advances in chemotherapy, radiation, and surgery, the outcome for metastatic or relapsed disease is particularly poor. Why is this? What are the biological characteristics of these recurring tumors? Can these characteristics be exploited for new therapies? This was an auspicious time to ask these questions. Recent changes to the NCI Cancer Therapy Evaluation Program and the Children’s Oncology Clinical Trial process, in conjunction with FDA incentives, are facilitating the movement of basic science discoveries from bench to clinical trial. The meeting provided an opportunity for researchers to present data, to review critically the state of the field, to highlight areas for future research, and to establish new collaborations.



C. Keller, A. Billin

Welcome and Background: B. Stillman, Cold Spring Harbor Laboratory

Introduction: C. Keller, Oregon Health & Science University, Portland  
A. Wagers, Harvard University, Boston, Massachusetts



### SESSION 1

**Chairperson: A. Wagers**, Harvard University, Boston, Massachusetts

- L. Wexler, Memorial Sloan-Kettering Cancer Center, New York: Sleeping Beauty, The Sequel: The search for Prince Charming.
- L. Helman, National Cancer Institute, Bethesda, Maryland: Development of novel combination targeted therapies for rhabdomyosarcoma.
- J. Khan, National Cancer Institute, Bethesda, Maryland: The application of Omics to identify novel targets and treatments for rhabdomyosarcoma.
- P. Houghton, Nationwide Children's Research Institute, Columbus, Ohio: Exploiting the IGF-mTOR pathway for treatment of rhabdomyosarcoma.
- C. Keller, Oregon Health & Science University, Portland: Three novel target-therapy pairs for potential clinical trials within 18 months.

### SESSION 2

**Chairperson: P. Houghton**, Nationwide Children's Research Institute, Columbus, Ohio

- T.A. Partridge, Children's National Medical Center, Washington DC: A model of reversible context-dependent rhabdomyosarcoma in the dystrophic mouse.
- B. Schaefer, University Children's Hospital, Zurich, Switzerland: Cancer stem cells in RMS: Fact or fiction?
- J. Shipley, Institute of Cancer Research, Sutton, United Kingdom: Histone methylation status and differentiation therapy.
- D. Cornelison, University of Missouri, Columbia: Comparative medicine-Eph/eprin expression profiles in RMS samples from canine and human patients.

### General Discussion



R. Rota

### SESSION 3

**Chairperson: D. Langenau**, Massachusetts General Hospital, Charlestown

- A. Wagers, Harvard University, Boston, Massachusetts: A transplant-based model for rhabdomyosarcoma in mice.
- F. Barr, National Cancer Institute, Bethesda, Maryland: The molecular correlates of fusion status in rhabdomyosarcoma.
- D.C. Guttridge, Ohio State University, Columbus: Regulation and function of NF-kB in rhabdomyosarcoma.
- C.M. Linardic, Duke University, Durham, North Carolina: Hippo pathway signaling in ARMS.

### General Discussion

### SESSION 4

**Chairperson: L. Helman**, National Cancer Institute, Bethesda, Maryland

- R. Pal, Texas Tech University, Lubbock: Predictive modeling of drug sensitivity.
- S. Hettmer, Charite-Universitaets-Medizin Berlin, Berlin, Germany: Asparagine homeostasis and its contributions to rhabdomyosarcoma growth.
- G. Pavlath, Emory University, Atlanta, Georgia: Nuclear transport receptors and myogenesis.
- R. Rota, Ospedale Pediatrico Bambino Gesù, Rome, Italy: Potential cross-talk involving Notch and epigenetic pathways in rhabdomyosarcoma.
- A.R. Ehrlich, Focus on Rhabdo, Plainsboro, New Jersey: An ad hoc framework for community: Scientific collaboration for rhabdomyosarcoma survivability.

### General Discussion



P. Houghton, J. Shipley

## SESSION 5

**Chairperson:** B. Schaefer, University Children's Hospital, Zurich, Switzerland

D. Langenau, Massachusetts General Hospital, Charlestown: Self-renewal mechanisms in embryonal rhabdomyosarcoma.

A. Hayes-Jordan, MD Anderson Cancer Center, Houston, Texas: The potential role of TOX-4 in rhabdomyosarcoma: Implications for metastatic disease.

Z. Li, Genomics Institute of Novartis Research Foundation, San Diego, California: Study and screen rhabdomyosarcoma in NIBR.

## Discussion: Charting Future Research

- What have we learned at this meeting?
- What are the key areas for research that will advance our understanding of rhabdomyosarcoma?
- What can be done to ensure funding for research?
- How can we fund the development of promising laboratory findings into drugs for clinical trials?

# The Genetics of Pain and Pain Inhibition: Where to from Here?

---

June 22–25

FUNDED BY           The Mayday Fund

ARRANGED BY       J. Mogil, McGill University, Montreal, Quebec, Canada  
                          C. Woolf, Harvard University, Boston, Massachusetts

Chronic pain is the most prevalent human health problem, with a lifetime prevalence of almost one in two. It is now 10 years since the first human genetic association studies of pain began appearing in the literature; association studies and exome sequencing studies of chronic pain disorders are also now being published. At the same time, progress has been made in identifying rare genetic variants responsible for monogenic pain disorders, both loss-of-function congenital insensitivity to pain and gain-of-function (e.g., erythromelalgia).

The time was right for a meeting to address the relative merits of the association studies and single-gene approaches for the study of chronic pain, along with various current practices in the field. Some of the questions that may be considered are the following: Will pain researchers ever be able to amass cohort sizes appropriate for replicable genome-wide association studies? What is the optimal way to phenotype patients and controls in pain genetics studies? Are genetic studies in model organisms broadly translatable to human clinical pain?

This discussion meeting reviewed the findings and promise of human genetic studies of pain, with the primary aim to compare the usefulness of rare versus common variant approaches in this field.

Welcoming Remarks: J.A. Witkowski, Banbury Center, Cold Spring Harbor Laboratory



**SESSION 1: Introductory Session**

C. Woolf, Boston Children's Hospital, Boston, Massachusetts: Why pain genes?

J. Mogil, McGill University, Montreal, Canada: History of pain genetics.

**SESSION 2: Lessons from Other Fields**

T. Grosser, University of Pennsylvania, Philadelphia: Pharmacogenetics.

A. Bowcock, National Heart & Lung Institute, London, United Kingdom: Molecular genetics of inflammatory diseases.

**General Discussion**

**SESSION 3: Single-Gene Pain Trait**

G. Woods, Cambridge Institute for Medical Research, United Kingdom: SCN9A loss of function.

S. Waxman, Yale University, New Haven, Connecticut: SCN9A gain of function.

M. Ferrari, Leiden University, The Netherlands: FHM genes.

I. Belfer, University of Pittsburgh, Pennsylvania: COMT: A complex story.

**General Discussion**

**SESSION 4: Complex Pain Genetics Session**

L. Diatchenko, McGill University, Montreal, Canada: Complex pain genetics: Progress so far.



M. Ferrari



C. Spellman, P. Thye, C. Stebbins



I. Belfer, L. Diatchenko, A. Apkarian

W. Maixner, University of North Carolina, Chapel Hill: OPFERA update.

V. Anttila, Harvard University, Boston, Massachusetts: Migraine GWAS.

**SESSION 5:** Translational Pain Genetics Session

A. van den Maagdenberg, Leiden University Medical Centre, The Netherlands: Transgenic mouse models of single-gene disorders: Migraine as the example.

G. Peltz, Stanford University, California: Haplotype mapping in mice.

M. Costigan, Harvard University, Boston, Massachusetts: Expression profiling and target validation.

G. Neely, Garvin Institute of Medical Research, Sydney, Australia: Pain genetics in nonmammalian organisms.

**SESSION 6:** Phenotyping

C. Nielsen, Norwegian Institute of Public Health, Oslo, Norway: What can/should be measured in large cohorts?

V. Apkarian, Northwestern University, Chicago, Illinois: Brain imaging and genetic studies.

**SESSION 7:** Breakout Groups: What Is the Best Way Forward?

Four breakout groups 1.5 hours. Group people and give each a mandate?

**SESSION 8:** New Approaches

D. Bennett, University of Oxford, United Kingdom: Applying RNA-Seq to both animal and human pain models.

R. Ratan, Burke Medical Research Institute, White Plains, New York: Epigenetics.

**Report Back to Other Groups**

**Group Review, Conclusion, and Summary**

# High-Performance Computing in Undergraduate Biology Education: Scanning the Landscape

---

September 3–5

FUNDED BY Alfred P. Sloan Foundation and the National Science Foundation

ARRANGED BY D. Micklos, DNA Learning Center, Cold Spring Harbor Laboratory  
D. Stanzione, University of Texas, Austin

“Big data” are rapidly becoming the currency of modern biology, including data from DNA/RNA sequencing, ecology and range mapping, remote sensing, automated phenotyping, visualization, and modeling. The generation and subsequent analysis of these data via high-performance computing (HPC) have traditionally been confined to research institutions. However, many underutilized data sets are now freely available to anyone with an Internet connection, and the technology for generating local data sets is coming within reach of faculty and students in primarily undergraduate institutions (PUIs).

Generating and analyzing large data sets hold the promise of bringing undergraduate students up-to-the-minute with biological research and of learning firsthand the modern synthesis of biochemistry and bioinformatics. Big data analysis can potentially support course-based research, which scales authentic student research to reach into introductory biology courses. Large data analysis also provides opportunities for distributed projects in which many students look at different aspects of the same problem. The time is right to discuss *IF* and *HOW* access to HPC resources can be extended to undergraduate institutions.



D. Micklos





The meeting format alternated between “perspective briefings” that presented the current state of knowledge, with free discussions that focused on the practicalities of extending HPC to undergraduate education. The meeting drew together 30 leaders and stakeholders from undergraduate biology education, HPC, and funding agencies to set a course for the future.

**Welcoming Remarks:** J.A. Witkowski, Banbury Center, Cold Spring Harbor Laboratory

**Introduction:**

D. Micklos, DNA Learning Center, Cold Spring Harbor Laboratory: Making HPC egalitarian.

T. Woodin, National Science Foundation, Arlington, Virginia: Vision and change.

**SESSION 1: The Biological Landscape**

M. Schatz, Cold Spring Harbor Laboratory: Genome structure and function.

B. Heidorn, University of Arizona, Tucson: Small data sets and biodiversity informatics.

T. Brettin, Argonne National Laboratory, Lemont, Illinois: Genomics and systems analysis.

J. Watkins, University of Arizona, Tucson: Statistical modeling of biological phenomena.

M. Cabrera, University of Puerto Rico, Mayaguez: Mathematical modeling and optimization in biology.

L. Fletcher, Austin Community College, Texas: Workforce development in biotechnology.

V. Byrd and L. Tanner, Clemson University, South Carolina: Broadening participation in next-generation computing.

B. Panoff, Shodor Education Foundation, Durham, North Carolina: Supporting computational thinking.

B. Barnett, National Center for Genome Analysis Support, Bloomington, Indiana: Support for genome analysis.

H. Neeman, Oklahoma University, Norman: HPC in plain english, starting from scratch.

**SESSION 2: The HPC Landscape**

D. Stanzione, Texas Advanced Computing Center, Austin: HPC, XSEDE, and lessons from iPlant.

P. Blood, Pittsburgh Supercomputing Center, Pennsylvania: Engaging genomics researchers, developers, and gateways through XSEDE.

B. Rekepalli, Oak Ridge National Laboratory, Knoxville, Tennessee: Small cluster applications in science gateways.

S. Gordon, Ohio Supercomputer Center, Columbus: XSEDED education program.

K. Gaither, Texas Advanced Computing Center, Austin: Visualization and girls in STEM.

**SESSION 4: The Funding Landscape**

**Recommendations, Funding Priorities, Collaborations, and Distributed Projects**

**SESSION 3: The Undergraduate Research/Education Landscape**

J. Jungck, University of Delaware, Newark: BioQUEST curriculum consortium and HPC.

S. Donovan, University of Pittsburgh, Pennsylvania: QUBES: Supporting quantitative thinking across the biology curriculum.

C. Ghiban, M. Khalfan, and J. Williams, Cold Spring Harbor Laboratory; U. Hilgert, University of Arizona, Tucson: A GUI for biological HPC.

J. Brusslan, California State University, Long Beach; B. Buckner, Truman State University, Kirksville, Missouri; S. Lewis, Prairie View A&M University, Texas; J. Seto, New York City College of Technology, Brooklyn, New York: iPlant tools in course-based research.



J. Williams

# The Immune System and Cancer

---

September 7–10

FUNDED BY **Oliver Grace Cancer Fund**

ARRANGED BY **G. Dranoff**, Dana-Farber Cancer Institute, Boston, Massachusetts  
**D. Fearon**, Cold Spring Harbor Laboratory

The recent reports that immunotherapy may induce durable responses in some cancer patients prompted the organization of a meeting to discuss multiple aspects of the biology of the interaction between the immune system and cancer. The topics that were discussed included imaging the tumor microenvironment, monitoring tumor immune reactions, current immunological therapies in human cancer, and newer approaches to immune interventions.



G. Dranoff, A. Korman

Welcoming Remarks: **J.A. Witkowski**, Banbury Center, Cold Spring Harbor Laboratory

Introduction: **G. Dranoff**, Dana-Farber Cancer Institute, Boston, Massachusetts



### SESSION 1

**Chairperson: S. Topalian**, Johns Hopkins University, Baltimore, Maryland

- A. Korman, Bristol-Myers Squibb, Redwood City, California: Antitumor activity of immunomodulatory antibodies.
- S. Topalian, Johns Hopkins University, Baltimore, Maryland: PD-1 pathway blockade in cancer therapy.
- T. Mempel, Harvard Medical School, Boston, Massachusetts: Local regulation of the T-cell response in the tumor microenvironment.
- S. Riddell, Fred Hutchinson Cancer Research Center, Seattle, Washington: Adoptive cellular therapy.

#### General Discussion

### SESSION 2

**Chairperson: D. Felsher**, Stanford University School of Medicine, California

- D. Bar-Sagi, New York University School of Medicine, New York: Mapping the immune landscape during early pancreatic neoplasia.
- S. Leach, Memorial Sloan-Kettering Cancer Center, New York: IL-17, gut microbiome, and pancreatic cancer.
- E. Sahai, Cancer Research UK, London, United Kingdom: Imaging the tumor microenvironment and localized responses to therapy.
- D. Felsher, Stanford University School of Medicine, California: MYC inactivation elicits an innate and adaptive immune response.
- B. Van den Eynde, Ludwig Institute for Cancer Research, Brussels, Belgium: Tumor-induced apoptosis of tumor-specific T lymphocytes in an autochthonous melanoma model.

#### General Discussion

### SESSION 3

**Chairperson: R. Schreiber**, Washington University School of Medicine, St. Louis, Missouri

- R. Schreiber, Washington University School of Medicine, St. Louis, Missouri: The importance of tumor-specific mutant antigens in endogenous and therapeutically induced immune responses to cancer.
- M. Krummel, University of California, San Francisco: The dynamics of tumor surveillance at primary and metastatic sites.
- D. Irvine, Koch Institute for Integrative Cancer, Cambridge, Massachusetts: Engineering cancer vaccine potency through lymph node targeting.
- G. Dranoff, Dana-Farber Cancer Institute, Boston, Massachusetts: Engineering improved cancer vaccines.
- G. Nolan, Stanford University School of Medicine, California: Single-cell proteomics and genomics at high scale.

#### General Discussion

### SESSION 4

**Chairperson: E. Vivier**, Centre d'Immunologie de Marseille-Luminy, Marseille, France

- E. Vivier, Centre d'Immunologie de Marseille-Luminy, Marseille, France: Natural killer cells: Development, regulation, and treatment.
- L. Lanier, University of California, San Francisco: Immune evasion mediated by tumor-derived lactate dehydrogenase induction of NKG2D ligands on myeloid cells in cancer patients.
- K. Wucherpfenning, Dana-Farber Cancer Institute, Boston, Massachusetts: Isolation of antibodies from patients responding to cancer immunotherapy.



D. Bar-Sagi



G. Trinchieri, E. Vivier

A. Rudensky, Memorial Sloan-Kettering Cancer Center, New York: Regulatory T cells in cancer.

A. Prevost-Blondel, Cochin Institute, INSERM, Paris, France: Immunoregulatory properties of NOS2: Impact on gd T cells and PMN-MDSC.

**General Discussion**

**SESSION 5**

**Chairperson:** G. Trinchieri, National Cancer Institute, Frederick, Maryland

G. Trinchieri, National Cancer Institute, Frederick, Maryland: Microbiota and cancer therapy.

M. Egeblad, Cold Spring Harbor Laboratory: Cancer cells hijack neutrophils' pathogen eradicating function to promote metastasis.

D. Fearon, Cold Spring Harbor Laboratory: T-cell exclusion: A dominant means for tumoral immune suppression.

**Wrap-Up Discussion**

# Interpreting Personal Genomes: How Are We to Set Appropriate Statistical Standards for Identifying Pathogenic Genetic Variants?

---

September 15–18

FUNDED BY Cold Spring Harbor Laboratory Corporate Sponsor Program

ARRANGED BY D. Goldstein, Center for Human Genome Variation, Duke University, Durham, North Carolina  
D. MacArthur, Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston

Genome-scale sequencing—both exome and whole-genome sequencing—are already being used in the clinical setting for both molecular diagnosis and gene discovery in severe disease patients. Current sequencing methods produce vast amounts of data that must be sifted for variants, but appropriate standards are still lacking for the robust identification of pathogenic variants and for assessing the evidence for new disease-linked genes. There is increasing concern that this lack of standards may lead to a proliferation of false positive claims, contaminating the literature and mutation databases and influencing clinical decisions.

The meeting had two aims. First, the participants reviewed the different approaches taken to assess potential for pathogenicity, including statistical and population genetic evidence, functional evaluation, and assessment of clinical “similarity” among patients with similar presentations.

Second, the meeting reviewed how to ensure that the highest standards are maintained both in the published literature and in any evaluations that influence clinical decision making. Although



D. Goldstein, D. Valle



many elements of what was covered have been addressed in other settings, a Banbury meeting affords a unique environment to review the central questions in detail. Moreover, the meeting ensured that the full range of relevant expertise was represented.

**Introduction:** J.A. Witkowski, Banbury Center, Cold Spring Harbor Laboratory

## SESSION 1

**Chairperson:** R. Myers, HudsonAlpha Institute for Biotechnology, Huntsville, Alabama

**Topic 1:** Interpreting Patient Genomes: Real World Experiences

- A. Beaudet, Baylor College of Medicine, Houston, Texas: Challenges in clinical implementation.
- M. Bamshad, University of Washington, Seattle: Mendelian Disease Center.
- D. Valle, The Johns Hopkins School of Medicine, Baltimore, Maryland: Mendelian Disease Center 2.
- W. Chung, Columbia University, New York: Lessons learned from clinical implementation of exome sequencing.

### General Discussion

**Topic 2:** Framing the Discussion

- T. Manolio, National Human Genome Research Institute, Rockville, Maryland: NHGRI programs relevant to variant interpretation.
- H. Rehm, Brigham and Women's Hospital, Cambridge, Massachusetts: The new ACMG guidelines for interpreting sequence variants: Where are we today and what's next?
- D. MacArthur, Massachusetts General Hospital, Boston: Lessons from previous efforts.

### General Discussion

## SESSION 2

**Chairperson:** J. Hirschhorn, Boston Children's Hospital, Massachusetts

**Topic 3:** Evaluating Evidence for Variant Pathogenicity Determination

- D. Conrad, Washington University School of Medicine, St. Louis, Missouri: Statistical approaches for the  $n = 1$  problem.
- S. Petrovski, Duke University, Durham, North Carolina: Bioinformatic approaches for prioritizing variants.
- D. MacArthur, Massachusetts General Hospital, Boston: Resources for assessing variant frequency and likely pathogenicity.
- A. Allen, Duke University, Durham, North Carolina: Statistical approaches for identifying statistical anomalies in patient genomes.
- K. Samocha, Massachusetts General Hospital, Boston: Assessing de novo mutations and gene constraint.

### General Discussion

**Topic 4:** What Is the Best Model for Sharing Genomic and Phenotypic Data?

R. Sever, Cold Spring Harbor Laboratory: Molecular case studies journal.

## SESSION 3

**Chairperson:** H. Rehm, Brigham and Women's Hospital, Cambridge, Massachusetts

**Topic 5:** Evaluating Evidence for Implicating Genes in Disease

- M. Daly, Massachusetts General Hospital, Boston: Classes of evidence for implicating new genes in human disease.
- J.-L. Casanova, The Rockefeller University, New York: The value of single families in identifying new disease genes.

**Topic 6:** Computational Tools

- G. Cooper, HudsonAlpha Institute for Biotechnology, Huntsville, Alabama: Overview of bioinformatics tools for variant classification.
- C. Bustamante, Stanford School of Medicine, California: Machine-learning algorithms and other high-throughput approaches to variant assessment.

**Topic 7:** Use of Phenotype to Inform Pathogenicity

- D. Adams, National Human Genome Research Institute, Bethesda, Maryland: Using deep phenotyping to inform DNA-sequence-variant classification: Practical challenges.
- L. Biesecker, National Human Genome Research Institute, Rockville, Maryland: POST-HOC phenotyping and pathogenicity.

### General Discussion

## SESSION 4

**Chairperson:** D. MacArthur, Massachusetts General Hospital, Boston

**Topic 8:** Functional Assessment of Genetic Variation

- L. Pennacchio, Lawrence Berkeley National Laboratory, Berkeley, California: Beyond exomes: Noncoding in vivo function and human disease.
- R. Xavier, Broad Institute, Boston, Massachusetts: Genes to biology.

**Topic 9:** Systems to Support Variant and Case Data Aggregation

- H. Rehm, Brigham and Women's Hospital, Cambridge, Massachusetts: Advances in variant databases.

M. Brudno, University of Toronto, Ontario, Canada: Phenotyping tools and the matchmaker exchange project.

**Topic 10:** Approaches to Support Variant and Gene Curation

M. Lebo, Brigham and Women's Hospital, Cambridge, Massachusetts: Clinical approaches to variant interpretation.

C. Cassa, Brigham and Women's Hospital, Boston, Massachusetts: Literature mining.

B. Funke, Massachusetts General Hospital, Cambridge: Clinical approaches to gene evidence evaluation.

**General Discussion**

**SESSION 5**

**Chairperson:** L. Biesecker, National Human Genome Research Institute, Bethesda, Maryland

**Topic 11:** Striking the Right Balance in Interpreting Personal Genomes

E. Worthey, Medical College of Wisconsin, Wauwatosa: A review of our 2014 lessons learned and the challenges and opportunities for 2015.

D. Goldstein, Duke University, Durham, North Carolina: Lessons from 150 diagnostic exomes.

**Topic 12:** The Way Forward: Discussion and Summary

**Discussion Leaders:**

D. Goldstein, Duke University, Durham, North Carolina

H. Rehm, Brigham and Women's Hospital, Cambridge, Massachusetts

D. MacArthur, Massachusetts General Hospital, Boston



H. Rehm



A. Allen

# Interdisciplinary Symposium on Creativity

---

October 21–24

FUNDED BY           The Daniel & Joanna S. Rose Foundation; Mr. & Mrs. Howard Phipps

ARRANGED BY       S. Nalbantian, Long Island University, Brookville, New York

This conference brought together a small group of comparatist scholars from the humanities with neuroscientists for an interdisciplinary investigation of creativity. The neuroscientists described the functioning of the brain in creative acts of scientific discovery or aesthetic production. The comparatists described instances of creativity that they analyze in the composition of major literary works, of musical compositions, or of works of visual art.

The conference covered such topics as mechanisms of creativity: How creativity is linked to brain structure and function; MRI tracking creativity in the brain; Inputs of specific brain regions; Components of creativity—memory, emotion, decision making, and intelligence; Pathology and creativity; Creativity and its reception: Empathy with reader, viewer, collaborator.



S. Nalbantian

**Introduction:** J.A. Witkowski, Banbury Center, Cold Spring Harbor Laboratory





**SESSION I: Mechanisms of Creativity in the Brain**

- S. Nalbantian, Long Island University, Brookville, New York: Introduction to the symposium.
- J.-P. Changeux, Kavli Brain-Mind Institute, La Jolla, California: Creativity in art: A neuronal hypothesis.
- R. Stickgold, Harvard Medical School, Boston, Massachusetts: The roles of REM sleep, dreaming, and quiet wake in the creative process.
- R. Jung, University of New Mexico, Albuquerque: Interacting networks of creative cognition: Perspectives from structural neuroimaging.

**SESSION II: Components of Creativity in the Arts**

- S. Nalbantian, Long Island University, Brookville, New York: Creativity in modernist literary writers: Acts of mental transformation.
- E. Fiorentini, Humboldt University, Berlin, Germany: The mind's eye?: Questions about imaging and creativity.
- M. Hussey, Pace University, New York: Significant form and the "aesthetic emotion": Clive Bell, Roger Fry, and Virginia Woolf.
- J. Wirtz, Hunter College, New York: Interdisciplinary convergence on writerly invention.

**SESSION III: Outlier Cases for Studying Creativity: From Pathology to Genius**

- P. Matthews, Imperial College, London, United Kingdom: Insights into creativity from diseases of the brain.
- S. Henke, University of Louisville, Kentucky: Posttraumatic fiction: Twentieth-century pathological writers and their creativity.

- N. Andreasen, University of Iowa, Iowa City: The creative brain: The neuroscience of genius.

**SESSION IV: Moments of Creativity; Methodologies for the Study of Creativity**

- M. Beeman, Northwestern University, Evanston, Illinois: Eureka: The "aha" moments in the creative process.

**Discussion by Participants of Useful Methodologies for Studying Creativity**

**Creativity in Music, Film, and Neuroscience**

- Talks by B. Adolphe and A. Gambis, introduced by S. Nalbantian, Grace Auditorium

**SESSION V: Creativity: Its Reception, Its Reduction**

- P. Schneck, University of Osnabruck, Germany: Henry James and the creative process: Demons, nuggets, and the stewpot of the imagination.
- D. Wehrs, Auburn University, Alabama: Literary innovation and re-imagining memory: Chivalric romance, renaissance genre, and plasticity.
- J. Bickle, Mississippi State University, Starkville: Are creative humans poor recallers? A functional hypothesis suggested by ruthlessly reductive molecular neuroscience.

**Concluding Round Table Discussion of All Participants**

- Includes discussion of plans for a volume on creativity to be edited by S. Nalbantian and P. Matthews.



E. Fiorentini



D. Wehrs



J. Bickle

# ROS in Biology and Cancer

---

October 26–29

FUNDED BY            Oliver Grace Cancer Research Fund

ARRANGED BY        A. Holmgren, Karolinska Institute, Stockholm, Sweden  
                              N. Chandel, Northwestern University, Chicago, Illinois  
                              N. Tonks, Cold Spring Harbor Laboratory  
                              D. Tuveson, Cold Spring Harbor Laboratory

This meeting delved more deeply into the biology of reactive oxygen species (ROS) to explore its role in cancer genesis and relevance for therapeutics than did the 2013 Banbury meeting. In addition to ROS's ability to stimulate cancer initiation, it is now apparent that the cellular anti-oxidant machinery has important roles in protecting cancer cells from oxidative damage. Furthermore, the anti-oxidant machinery can be up-regulated in response to oncogenes and may confer drug resistance and "stemness." Such observations suggest that redox modulation may offer a novel approach for selective targeting of cancer cells. Participants explored the chemical, biochemical, and genetic facets of ROS biology in relation to cancer, with the goal of determining whether ROS can be manipulated *in vivo* to alter cancer pathogenesis and the response of cancer cells to therapy.

Welcoming Remarks: J.A. Witkowski, Banbury Center, Cold Spring Harbor Laboratory

Introduction: D. Tuveson, Cold Spring Harbor Laboratory



## WHAT ARE ROS AND WHERE IS THE ROS COMING FROM?

### SESSION 1: Redox Biology

- M. Ristow, Swiss Federal Institute of Technology, Zurich, Switzerland: Mitohormesis: How mitochondrial ROS production promotes metabolic health and life span.
- M. Haigis, Harvard Medical School, Boston, Massachusetts: Posttranslational control of mitochondria.
- G. Shadel, Yale University School of Medicine, New Haven, Connecticut: Novel roles for mitochondrial ROS in signaling and disease pathology and potential roles for mitochondria in activating the immune system in the context of cancer.

### SESSION 2: ROS Biology and Chemistry

- C. Winterbourn, University of Otago, Christchurch, New Zealand: Identifying cellular targets for reactive oxidants and the influence of peroxiredoxins.
- Y. Janssen-Heininger, University of Vermont, Burlington: Targeting S-glutathionylation chemistry in lung cancer.
- T. Miller, IC MedTech Corporation, El Cajon, California: Can targeted ROS = affordable new anti-neoplastics?

## CAN TARGETED ROS MODULATION = AFFORDABLE NEW ANTINEOPLASTICS?

### SESSION 3: ROS Pathways and Redox

- M. Murphy, MRC Mitochondrial Biology Unit, Cambridge, United Kingdom: Complex I as a source of superoxide during ischaemia-reperfusion injury.
- H. McNeil, Lunenfeld-Tanenbaum Research Institute, Mount Sinai Hospital, Toronto, Ontario, Canada: Fat cadherins regulation of mitochondrial function.
- P. Puigserver, Dana-Farber Cancer Institute, Boston, Massachusetts: The PGC1 pathway programs mitochondrial and ROS detoxification in tumor cells.

### SESSION 4: ROS Methods and Therapeutic Development

- T. Dick, German Cancer Research Center, Heidelberg, Germany: Mechanisms of regulatory protein thiol oxidation.
- E. White, The Cancer Institute of New Jersey, New Brunswick: Role of mitochondria quality control in cancer.

### General Discussion

- Moderator:** M. Espey, National Cancer Institute, NIH, Rockville, Maryland

## WHAT IS ROS DOING?

### SESSION 5: Exogenous and Endogenous Anti-Oxidants and Cancer

- A. Holmgren, Karolinska Institute, Stockholm, Sweden: Thioredoxin and peroxiredoxin in ROS signaling.
- C. Neumann, University of Pittsburgh Cancer Institute, Pennsylvania: Prdx1.
- N. Chandel, Northwestern University, Chicago, Illinois: Targeting redox for cancer therapy.

- P. Schumacker, Northwestern University, Chicago, Illinois: Mitochondrial ROS in cancer: Initiators, facilitators, or innocent bystanders?

### SESSION 6: NRF2 and ROS Regulation

- D. Tuveson, Cold Spring Harbor Laboratory: Nrf2 in pancreatic cancer.
- E. Schmidt, Montana State University, Bozeman: Functional antagonism of Nrf2 activity by the disulfide reductase systems.
- D. Zhang, University of Arizona, Tucson: Nrf2 regulation and its dual role in cancer.

### SESSION 7: ROS Signaling

- D. Pappin, Cold Spring Harbor Laboratory: A double-labeling approach for quantitative cysteine proteomics.



T. Dick, C. Neumann

**SESSION 8: ROS Sensors**

N. Hay, University of Illinois, Chicago: Selective targeting of cancer cells with ROS inducers: The Achilles' heel of Akt.

A. Kimmelman, Dana-Farber Cancer Institute, Boston, Massachusetts: Redox balance in pancreatic cancer.

**General Discussion**

**Moderators:** D. Tuveson, Cold Spring Harbor Laboratory; N. Chandel, Northwestern University, Chicago, Illinois; A. Holmgren, Karolinska Institute, Stockholm, Sweden.

**SESSION 9: ROS and Therapies I**

J. Doroshow, National Cancer Institute, Bethesda, Maryland: NADPH oxidases and cancer.

P. Huang, MD Anderson Cancer Center, Houston, Texas: ROS in cancer therapeutics.

M. Bergo, University of Gothenburg, Sweden: Antioxidants and lung cancer progression.

**SESSION 10: ROS and Therapies II**

L. Trotman, Cold Spring Harbor Laboratory: Redox therapy in prostate cancer.

B. Stockwell, Columbia University, New York: Metabolic control of lipid peroxidation and cell death.

**Discussion**

**Moderator:** S. Biller, Agios Pharmaceuticals, Cambridge, Massachusetts

**Final Meeting Summary**

N. Chandel, Northwestern University, Chicago, Illinois

A. Holmgren, Karolinska Institute, Stockholm, Sweden

D. Tuveson, Cold Spring Harbor Laboratory



C. Winterbourn



H. McNeil, M. Murphy



T. Miller

# Epigenetics and Agriculture

---

November 9–12

FUNDED BY Cold Spring Harbor Laboratory Corporate Sponsor Program

ARRANGED BY B. Hauge, Monsanto Company, Chesterfield, Missouri  
R. Martienssen, Cold Spring Harbor Laboratory

Epigenetic regulation has a pivotal role in plant development, the response to the signals from the environment, and natural variation of gene expression. Epigenetics therefore offers a largely untapped resource for crop improvement strategies aimed at enhancing productivity through the selection of favorable epialleles, plant adaptation to abiotic and biotic stresses, and strategies for durable efficacious expression of transgenes. The goal of this meeting was to explore these opportunities by bringing together leading researchers in the plant epigenetics field with scientists representing agricultural biotechnology companies.

The meeting focused on epigenetic mechanisms of gene regulation and their roles in heterosis, epigenetic programming of plant reproduction, transgenerational inheritance, and adaptation to abiotic and biotic stresses.

Participants also explored the needs of agricultural biotechnology, and how epigenetic research can help efforts to manipulate gene expression in crops toward enhancing sustainable food and feed production, to meet the needs of a growing population.

Welcoming Remarks: J.A. Witkowski, Banbury Center, Cold Spring Harbor Laboratory

Introduction: V. Shukla, The Bill & Melinda Gates Foundation, Seattle, Washington

From Insight to Impact: How Do We Translate Epigenetics Research into Developing World Agriculture?





V. Colot, D. Weigel



E. Dennis

### SESSION 1: Epialleles: Utility and Transgenerational Inheritance

**Chairperson: R. Martienssen**, Cold Spring Harbor Laboratory

D. Weigel, Max-Planck Institute for Developmental Biology, Tübingen, Germany: Computational methods for detection of DNA methylation differences and their application to the study of natural and spontaneous variation.

V. Colot, Institut de Biologie de l'École Normale Supérieure, Paris, France: Stability and potential of transgenerational epigenetics in plant breeding.

S. Jacobsen, University of California, Los Angeles: Epigenetic mechanisms in *Arabidopsis*.

E. Richards, Boyce Thomas Institute, Ithaca, New York: 3D epigenetics and stress responses.

R. Schmitz, University of Georgia, Athens: Exploring genome-wide patterns of DNA methylation throughout the plant kingdom.

J. Hollick, Ohio State University, Columbus: Paramutation in *Zea mays*.

### SESSION 2: Heterosis, Polyploidy, and Chromosome Dosage

**Chairperson: D. Jackson**, Cold Spring Harbor Laboratory

E. Dennis, CSIRO Plant Industry, Canberra, Australia: Heterosis and the role of epigenetics.

D. Baulcombe, University of Cambridge, United Kingdom: Paramutation in hybrid tomato.

Z. Lippman, Cold Spring Harbor Laboratory: Optimization of crop productivity using induced mutations in the florigen flowering pathway.

J. Chen, University of Texas, Austin: Maternal small RNAs and seed development.

N. Springer, University of Minnesota, St. Paul: Variation for DNA methylation patterns in maize populations.

C. Pikaard, Indiana University, Bloomington: Epigenetic inheritance and RNA silencing.

### SESSION 3: GxE Epigenetic Contribution to Stress Adaptation

**Chairperson: D. Ware**, Cold Spring Harbor Laboratory

J. Paszkowski, University of Cambridge, United Kingdom: Genetic determinants of epiallelic switches.

J. Gutierrez-Marcos, University of Warwick, Coventry, United Kingdom: Myths and facts about transgenerational epigenetic inheritance in plants.

U. Grossniklaus, University of Zurich, Switzerland: Selection of epigenetic variation in *Arabidopsis*.

O. Mittelsten Scheid, Gregor Mendel Institute, Vienna, Austria: Epigenetics and genetics: A complex relationship.

J.-K. Zhu, Purdue University, West Lafayette, Indiana: Epigenetic antisilencing mechanisms in plants.

#### General Discussion

### SESSION 4: Gametogenesis, Apomixis, and Imprinting

**Chairperson: P. Schnable**, Iowa State University, Ames

R. Martienssen, Cold Spring Harbor Laboratory: Genome reprogramming and epigenetic inheritance.

M. Gehring, Whitehead Institute for Biomedical Research, Cambridge, Massachusetts: Epigenetics and seed development.

D. Grimanelli, Institut Recherche pour le Développement, Montpellier, France: Epigenome dynamics and reproductive development.

F. Berger, Gregor Mendel Institute of Molecular Plant Biology, Vienna, Austria: Higher-order chromatin structure and its impact on genome activities in plants.

#### General Discussion

**SESSION 5: Agricultural Applications and Translational Epigenetics**

**Chairperson: T. Osborn**, Monsanto, Chesterfield, Missouri

F. Van Ex, Bayer CropScience, Zwijnaarde, Belgium: Exploring the epigenetic potential of crops.

S. Mackenzie, University of Nebraska, Lincoln: Heritable epigenomic reprogramming in crop species by plastid perturbation.

P. Schnable, Iowa State University, Ames: How well does DNA variation predict phenotypes in maize? Can we improve these predictions?

J. Reinders, DuPont Experimental Station, Johnston, Iowa: Assessing the impact of epigenetic variation on phenotypic variation in maize.

**Conclusion and Future Work**



S. Mackenzie



E. Richards, D. Baulcomb

# Lewy Body Dementia: Current Status, Future Directions

---

November 16–19

FUNDED BY           The Dana Foundation; Prothena Biosciences, Inc.; Sandy and Nelson DeMille

ARRANGED BY       J. Galvin, New York University, New York  
I. McKeith, Newcastle University, Newcastle upon Tyne, United Kingdom

Lewy body dementia (LBD) is the second most common cause of cognitive impairment after Alzheimer's disease (AD), affecting more than 1.3 million Americans and perhaps over 4 million people worldwide. However, much less is known about LBD despite the fact that the prevalence rate of LBD approaches 5% in the elderly population and makes up 30% of all dementia cases.

We are only just beginning to understand the genetics of LBD. A family history of dementia may be more common in LBD compared with healthy older adults, suggesting some form of inheritance pattern. Mutations in at least four genes are now associated with LBD— $\alpha$ -synuclein,  $\beta$ -synuclein, glucocerebrosidase, and new candidate loci linked to chromosome 2. There is increasing evidence that although MRI may have a limited role in discriminating LBD from AD, nuclear imaging studies using novel markers with single photon emission computerized tomography (SPECT) and positron emission tomography (PET) scans may be able to improve diagnoses and be used as outcomes for clinical trials.

Participants in the meeting critically reviewed the current state of knowledge of the genetics of LBD and the usefulness or otherwise of current clinical, imaging, and biological markers. There was much discussion of how global research efforts on LBD could be delivered in conjunction with the goals of NINDS, NAPA, the G8 Summit Objectives, and related initiatives.

Welcoming Remarks: J.A. Witkowski, Banbury Center, Cold Spring Harbor Laboratory

## Introductions and Work Plan:

I. McKeith, Newcastle University, Newcastle upon Tyne, United Kingdom

J. Galvin, New York University, New York





### SESSION 1

**Chairperson:** J. Leverenz, Cleveland Center for Brain Health, Ohio

A. Singleton, National Institute on Aging, Bethesda, Maryland: DLB genetics: Progress and priorities.

J. Bras, University College London, United Kingdom: Genetic risk for DLB: A genome-wide assessment of genetic variability in DLB.

J. Leverenz, Cleveland Center for Brain Health, Ohio: Genetics of cognition in the Lewy body disorders.

### SESSION 2

**Chairperson:** D. Dickson, Mayo Clinic, Jacksonville, Florida

D. Dickson, Mayo Clinic, Jacksonville, Florida: Neuropathology of diffuse Lewy body disease: Familial and sporadic forms.

G. Halliday, University of New South Wales and Neuroscience Research, Australia: The dynamics of Lewy body formation in Lewy body diseases.

### SESSION 3

**Chairperson:** O. El-Agnaf, HBK University, Doha, United Arab Emirates

P. Brundin, Van Andel Institute, Grand Rapids, Michigan: What is the possible role of cell-to-cell transfer of  $\alpha$ -synuclein in the pathogenesis of Lewy body dementia?

J. Duda, Philadelphia VA Medical Center, Pennsylvania: Olfactory dysfunction in sporadic and genetic forms of Lewy body disorders.

O. El-Agnaf, HBK University, Doha, United Arab Emirates: Detection of  $\alpha$ -synuclein pathogenic conformations as a strategy for biomarker development for Lewy body diseases.

**General Discussion: Genetics and Pathology**

**SESSION 4:** Coordination of DLB Research Initiatives to Include NAPA/ADRD, ADC-DLB Module, JPND, G8, and Others

**Chairpersons:** T. Montine, University of Washington, Seattle, and D. Aarsland, Karolinska Institute, Huddinge, Sweden

### SESSION 5

**Chairperson:** J. O'Brien, University of Cambridge, United Kingdom

D. Arslan, Karolinska Institute, Huddinge, Sweden: Prognosis of DLB and the potential of predictive biomarkers in DLB.

J.-P. Taylor, Institute of Neuroscience, Newcastle, United Kingdom: Neurophysiological biomarkers in Lewy body dementia.

J. O'Brien, University of Cambridge, United Kingdom: Neuroimaging biomarkers for Lewy body dementia.

E. Masliah, University of California, La Jolla: Novel therapeutics for Lewy body disease in transgenic models: Bridging the gap.

**General Discussion**

### SESSION 6

**Chairperson:** D. Weintraub, University of Pennsylvania, Philadelphia

A. Muhs, AC Immune SA, Lausanne, Switzerland: Identification of small-molecule inhibitors of  $\alpha$ -synuclein aggregate toxicity.

R. Mills, ACADIA Pharmaceuticals, Inc. San Diego, California: Novel antipsychotics in Lewy body disease.

E. Mori, Tohoku University, Sendai, Japan: Cholinesterase inhibitors in Lewy body dementia.

D. Weintraub, University of Pennsylvania, Philadelphia: Design and conduct of clinical trials in Lewy body dementia.

**Discussion: Management Strategies for LBD**

J. O'Brien, University of Cambridge, United Kingdom

J.-P. Taylor, Institute of Neuroscience, United Kingdom

### SESSION 7

**Chairperson:** J. Galvin, New York University, New York

B. Boeve, Mayo Clinic, Rochester, Minnesota: REM sleep behavior disorder as an early manifestation of evolving dementia with Lewy bodies.



E. Mori

- J. Galvin, New York University, New York: Improving the detection and diagnosis of LBD in the office setting.
- I. McKeith, Newcastle University, Newcastle upon Tyne, United Kingdom: Prodromal DLB: Diagnosing, managing, and trialling it.
- A. Taylor, Lewy Body Dementia Association, Inc., Lilburn, Georgia: The role of the nonprofit in LBD education and research.

**SESSION 8**

**Chairperson: B. Boeve**, Mayo Clinic, Rochester, Minnesota  
**Proposal for DLB 6 Consortium Meeting**  
B. Boeve, D. Dickson, J. Leverenz, and I. McKeith

**Wrap-Up Discussion**



J. O'Brien



E. Masliah, O. El-Agnaf





DNA LEARNING CENTER



# DNA LEARNING CENTER

## EXECUTIVE DIRECTOR'S REPORT

---

### ADMINISTRATION

Lauren Corrieri  
Anne Greengard  
Mary Lamont  
Valerie Meszaros  
David Micklos  
Carolyn Reid

### INSTRUCTION

Elna Carrasco  
Heather Cosel-Pieper  
Antonia Florio  
Natalia Hanson  
Emily Harrison  
Ashleigh Jackobel  
Melissa Lee  
Christine Marizzi  
Kathleen McAuley

### BIOMEDIA

Amanda McBrien  
Erin McKechnie  
Bruce Nash  
Brooke Roeper  
Shreya Shah  
Mona Spector  
Xiaoqun Catherine Zhang  
Cornel Ghiban  
Mohammed Khalfan  
Susan Lauter  
Jason Williams  
Chun-hua Yang

A fall meeting of computer scientists and undergraduate biology educators at CSHL's Banbury Center came to a surprising conclusion: There is plenty of compute time available on high-performance computers for undergraduate research projects, but most teaching faculty do not know how to access or use it. Representatives of the National Science Foundation's (NSF) supercomputer system, XSEDE (Extreme Science and Engineering Discovery Environment), along with two major cyberinfrastructure projects—NSF's *iPlant Collaborative* and the Department of Energy's KnowledgeBase (KBase)—agreed that they would make increased efforts to reach faculty at primarily undergraduate institutions.

High-performance computing (HPC) generally refers to supercomputers composed of hundreds of thousands of individual processors (identical to a high-end desktop processor) linked in a parallel system that optimizes communication speed between the individual processors. HPC involves splitting a job into tasks that are performed by different processors (parallelization) and also assigns repeated, or looped, tasks (vectorization). Extremely intensive calculations may engage a large proportion of the available processors; however, smaller jobs are assigned to one or a few nodes, or clusters, of processors. Thus, HPC begins with clusters of parallel processors—often including some “cloud” services—that can perform in minutes or hours complex calculations that would take a desktop computer days or weeks to finish, if at all.

Biology is awash with data from genome sequencing. The cost of DNA sequencing—and, by extension, of producing a sequenced genome—has decreased 10,000-fold since 2007. Beijing Genomics Institute, a genome sequencing “factory” in Shenzhen and Hong Kong, China, can sequence several hundred eukaryotic genomes per day, including a human genome for as little as \$3000. Mike Schatz, a genome informaticist at CSHL, said that within a decade, the “data tsunami” of DNA sequences will grow to an exabyte (1 billion gigabytes) annually—equivalent to the information carried by 200 million DVDs. Indeed, 95% of graduate students, postdoctoral fellows, and research faculty we have surveyed at workshops and professional meetings say that they are currently analyzing large data sets or expect to in the near future.

At least DNA data makes use of a standard file format, called FastQ. Unfortunately, there are few standards for many other types of biological data—from phenotyping, field studies, remote sensing, range mapping, modeling, and visualization. According to Bryan Heidorn of the University of Arizona, 80% of biological data is being created by smaller projects—without computational infrastructure, data standards, or data federation. These data are often “trapped” in lab notebooks or in Word, Excel, or PDF files that cannot readily be accessed. For these reasons, only about 8% of nongenomic ecological data is shared. This creates a massive amount of biological “dark data.” Like the dark matter of the universe, “we know it is there or used to be there, but we can't see it.”

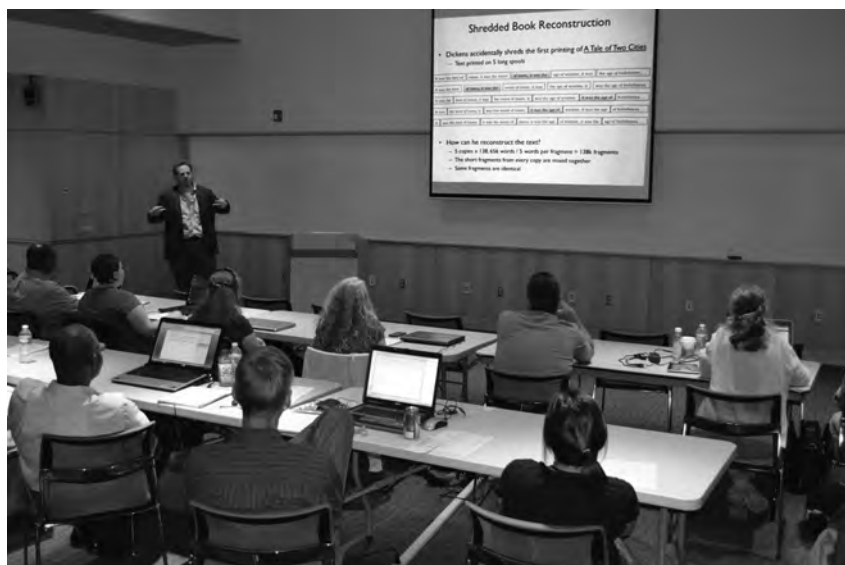
Unlike the physical sciences, which have a long tradition of large-scale, computationally intensive projects, biology has only recently entered the HPC world. Henry Neeman of the University of Oklahoma said, “The undergraduate biology education community is in the same position today as the physical science education community was some years ago—aware of the value and potential of HPC, but still figuring out what resources and capabilities are available, and how to use them.”

Neeman explained that there are “owners and orphans” in the HPC world. The “owners” are high-level research programs with annual allocations of hundreds of millions of core hours. Owner jobs, which may command thousands of compute nodes at once, get priority scheduling. The “orphans” are small-scale users who use nodes that are not scheduled by owners. Orphans can access XSEDE through a gateway—such as *Galaxy*, a popular bioinformatics web-based graphical interface, or *iPlant*—or can formally request an education allocation. A “start-up” education allocation of 150,000 core hours is enough to process at least 200 RNA sequence (RNA-Seq) experiments capable of yielding publishable results. So, although HPC resources are oversubscribed by owners, there is ample orphan bandwidth for undergraduate biology faculty. However, biological projects only account for 10% of educational allocations on XSEDE.

Amid calls to make undergraduate biology education more engaging, data sets from genomics, phenomics, range mapping, modeling, and visualization can provide authentic research opportunities for large numbers of undergraduate courses. Although the nation’s supercomputers have enough capacity to support a major expansion in undergraduate biology research, faculty need to be made aware of computational resources and trained to use them to analyze big data.

### Answering the Call for RNA Sequence Analysis

RNA-Seq analysis is viewed as an “entry level” analysis of a whole eukaryotic genome, providing a snapshot of genes expressed by an organism that can be used to annotate new genes and gene variants or to document differential gene expression. Tools for RNA-Seq analysis are the most frequently requested genome applications at the *iPlant Collaborative* and *Galaxy*. The biochemical expertise for extracting RNA needed for RNA-Seq analysis is available at virtually any undergraduate institution and sequencing costs under \$3000.



Mike Schatz speaks with the RNA-Seq Working Group.

The computational challenge of analyzing even a modest RNA-Seq experiment, however, is daunting. The simplest RNA-Seq experiment—two experimental conditions and biological replicates—generates 30–50 million sequence “reads,” each of 75–150 nucleotides. Aligning these reads to a sequenced genome, assembling the reads into gene models, and calculating differences in gene expression would take 27 days for a laptop computer with a single processor. Thus, we have made a major commitment to develop a simple way for biology faculty and students to access the HPC resources to analyze RNA-Seq data sets.

As part of the *iPlant Collaborative*, a 10-year NSF project to develop computer infrastructure to support biological research, the DNALC developed *DNA Subway* with several “lines” to simplify the analysis of genes and genomes. In 2014, we completed the Green Line of our *DNA Subway* as an educational workflow for RNA-Seq analysis that taps into the XSEDE national high-performance computing infrastructure. Based on the popular Tuxedo Protocol, the Green Line stages faculty and student jobs on the Stampede supercomputer at the Texas Advanced Computing Center, where an average RNA-Seq experiment takes about 40 hours on 16 processors (cores), or 640 core hours. This represents 0.015% of available computational power on this supercomputer!

This *iPlant* infrastructure coordinates with our \$537,000 NSF grant, “Infrastructure and Training to Bring NGS Analysis into Undergraduate Education,” under which we are helping undergraduate faculty to develop novel RNA-Seq data sets for use in course-based and independent student research. In June 2014, faculty members from predominantly undergraduate teaching institutions (PUI) in 10 states convened at CSHL for a Working Group Retreat. With support from Illumina and sequencing services performed by the CSHL Genome Center, faculty obtained 50 RNA-Seq transcriptomes covering ~4.2 billion paired-end DNA sequencing reads. Projects were diverse, included eight different species, and examined gene expression changes in (1) *Apis mellifera* flight muscle during life-stage transitions; (2) leaf development and senescence in *Arabidopsis thaliana*; (3) retina development in *Gallus gallus*; (4) caprine testes during juvenile development to puberty; (5) maize in response to cold stress; (6) retinas of mice with retinitis pigmentosa; (7) differentiation of rat neural precursors during maternal immune activation; (8) seed abscission in *Sorghum bicolor*; (9) floral inflorescence genes in banana/plantains; and (10) *Brassica rapa* gibberellic acid (*gad*) mutants upon gibberellic acid exposure. The beta version of the Green Line was rigorously tested during the retreat; 837 jobs were submitted to TACC, consuming 695 hours’ total CPU time.

Also during the Retreat, Working Group members planned classroom implementation and contributed content to a new website, *RNA-Seq for the Next Generation* ([www.rnaseqforthenextgeneration.org](http://www.rnaseqforthenextgeneration.org)), which was launched in December. The site already features a Profile section, where each Working Group member has a dedicated page that includes a video introduction and abstract, experimental details and data access information, and course descriptions and resources. The RNA-Seq data and profile page content can serve as a foundation and reference for other undergraduate faculty as they incorporate RNA-Seq analysis into their teaching. Post-retreat, we have conducted monthly videoconferences (utilizing Adobe Connect meeting rooms) to update Working Group faculty on data analysis progress and changes to the website and to share evolving instructional strategies. A multifaceted evaluation program will assess



Eric Antoniou gives a tour of the CSHL Genome Center to RNA-Seq faculty.



effects on student learning, interests, and attitudes across different classroom and student research settings. By year's end, Working Group members had already used their RNA-Seq data sets with more than 1000 students in a variety of classes—including introductory biology, anatomy and physiology, genetics, cell and molecular biology, genomics, agricultural breeding, and independent research.

### iPlant Collaborative

The *iPlant Collaborative* cyberinfrastructure includes high-performance computing, data storage, software, and human resources. *iPlant's* mission is to empower researchers and educators to use cyberinfrastructure to solve grand challenges in biology—large questions that integrate large-scale data from genomes, phenotypes, and environments to predict how organisms and ecosystems function. *iPlant's* second 5-year plan was ratified by the National Science Board in fall 2014, bringing total funding to \$100 million. As lead for the Education, Outreach, and Training (EOT) component, the DNALC is involved with a number of activities that are transitioning *iPlant* from infrastructure development to widespread use and sustainability into the future.

In addition to the release of the Green Line for RNA-Seq analysis, we continued to improve the usefulness of *DNA Subway*. Blue Line, used primarily for DNA barcoding, received an upgraded alignment viewer that shows the quality scores of polymorphisms, and phylogenetic tree building was improved with an easy toggle to designate outgroups. We made significant progress in integrating WebApollo, a faster and easier-to-use version of the Apollo gene annotation tools. There were 20,832 site users and 3782 new registrations in 2014, a 37% and 44% increase, respectively. This increase in traffic gave us a grand total of 28,934 projects for 2014, a 58% increase over 2013!

We continued to deliver two-day training workshops at sites nationwide, reaching 124 researchers at five *Tools & Services* workshops and 60 faculty at three *Genomics in Education* workshops. One-hour monthly “Get Started with *iPlant*” webinars reached 168 attendees, and Livestream feeds from in-person workshops reached an additional 96 participants. Realizing that we can only reach a fraction of *iPlant's* 20,000 registered users with in-person training, we initiated a *train-the-trainers* program to extend our reach. Working with 29 faculty at workshops at the University of California, Davis, and the U.S. Department of Agriculture's Big Data Initiative, we devised a model curriculum that will be expanded in 2015. Our goal will be to develop networks of “*iPlant Helpers*” that can act as local campus “experts” to help others get started with *iPlant*.

The EOT team presented papers and posters to a number of different audiences. *DNA Subway: Making Genome Analysis Egalitarian*, presented at XSEDE14, showcased the Green Line's use of the Stampede supercomputer—currently the world's seventh most powerful computer. *Cyberinfrastructure for Life Sciences: iAnimal Resource for Genomics and Other Data Driven Biology* and *DNA Subway: An Educational Bioinformatics Platform for Gene and Genome Analysis* were presented at the World Congress of Genetics Applied to Livestock Production. Participation in the European Conference on Computational Biology and organization of CSHL meetings on *Biological Data Science* and *High-Performance Computing in Undergraduate Education* (described earlier) advanced the DNALC's leadership in link biology and computation.

We also developed online materials for asynchronous learning. The new Online Learning Center ([www.iplantcollaborative.org/learning-center](http://www.iplantcollaborative.org/learning-center)) features more than 50 tutorials and 100 YouTube videos. These help users get started with applications and analysis tools available in the *iPlant* Discovery Environment and *iPlant* cloud computing service, Atmosphere. The new *iPlant*

*Academy* is a place for biology faculty to exchange lesson plans, present case studies of how they use *iPlant* tools in their teaching, and showcase student projects.

Evaluation continues to guide and inform EOT efforts and document our success in reaching our intermediate audience of biological researchers who do not have specialized expertise in computation. Our preworkshop surveys in 2014 indicated that 49% of attendees classify themselves at the beginner's level of bioinformatics skills, whereas surveys across a wide variety of meetings and conferences now indicate a whopping 95% of researchers currently work with or anticipate working with large data sets involving genome sequencing and RNA-Seq. This great interest in analysis of these specialized data sets underscores the need for increased access to training and help.

### New Developments around the World

Following 3 years of collaboration, in May, we established Beijing 166 High School as a licensed DNA Learning Center. The project is funded by the Dongcheng School District, which serves 100,000 students in the ancient core of Beijing, adjacent to the Forbidden City and Tiananmen



Beijing Dave teaching. Dave Micklos leads students in a PCR lab.



Beijing Peking Man. Catherine Zhang leads Beijing students on a trip to see the Peking Man site.



Beijing 166 High School course graduation.

Square. The licensing agreement is part of Principal Wang Lei's master plan in strengthening the leadership of Beijing 166 as the city's only designated high school of biology.

Under the contract, we increased the number of summer training weeks from three to 10—and added student and teacher training workshops in Beijing—with the objective of providing 260 student-weeks of training. Seventy-two middle and high school students attended four 2- or 3-week courses at the DNALC in the summer. Middle school students completed modules on *Fun with DNA*, *World of Enzymes*, and *Forensic Detectives*. High school students completed modules on *DNA Science* plus *DNA Barcoding Research* or *Human Genomics*. Eighteen students designed and conducted independent research projects exploring the biodiversity of Cold Spring Harbor's ecosystem using DNA barcoding, including fieldwork, lab work, computer analyses, and a presentation of research findings. Some summer camp students met Dr. James Watson during a special media event in August, which was covered by journalists from more than ten media agencies.

In October, we taught two 4-day student courses on *Human Genomics* at Beijing 166 School. Sixty-three students used polymerase chain reaction (PCR) to look at DNA variations in their own DNA, highlighting concepts in population genetics, genotype–phenotype analysis, and human evolution. The courses culminated with a field trip to the Peking Man World Heritage site outside of Beijing. We also conducted 2-day teacher training workshops for 340 biology teachers from the Beijing area.

We rigorously evaluated the students who participated in courses conducted at the DNALC and at the Beijing 166 School—and we compared their genetic content knowledge and literacy with New York students taking the same courses. The results of these studies support our contention that the DNALC courses provide Chinese students with an exposure to biology practice and concept development that is largely missing from even the best Chinese schools. The results of this study suggest that Beijing high school students are less “scientifically mature” than their Long Island counterparts—with less general biological knowledge and less exposure to the scientific process, critical thinking, and “soft” skills of scientific collaboration and communication. The results also suggest that science enrichment programs—such as the DNALC courses—can bolster Chinese students' familiarity with science process and enquiry, and prepare them for science study at American universities.

In November, we initiated serious negotiations with Suzhou Industrial Park (SIP) to establish a licensed DNALC in its BioBay biotechnology park. SIP mayor Barry Yang offered his strong support for the development of eight teaching labs and support spaces plus exhibit and cafeteria facilities in three existing buildings in a grassy park in the center of BioBay. The agreement will build upon the existing relationship with SIP, with which the Laboratory currently operates CSH Asia Conferences. Like the CSHL conferences, the DNALC at SIP would be directly operated by CSHL. However, due to the ongoing liberalization in China and strong support from the SIP mayor and Communist Party secretary, the DNALC at SIP will be set up as a nongovernmental organization (NGO). In addition to being tax exempt, the NGO designation should make monetary transfers between SIP and CSHL much easier. The NGO will share core management and accounting services with CSH Asia Conferences. We expect to sign a formal agreement in spring 2015, with the objective of bringing the facility into operation in late 2015.

We also worked with a high-level consortium on plans for two licensed centers in Mexico due to open in late 2015 and 2016. One center will be a major attraction of the “Codes of Life Health Park,” an EPCOT-like exposition in Chapultepec Park—Mexico City's “Central Park”—and the largest urban green space in the western hemisphere. The Mexico City Project is headed by Armando Barriguete, a world expert on nutrition, Jose Cordova, Minister of Health under former President Felipe Calderone, and Hugo Scherer, the campaign manager for current President Enrique Peña Nieto. The other licensed DNALC will be located in the new “Agri-food Research Center” (CIDAM) in Morelia, the capital of Michoacán State. Michoacán

is the most important agricultural region in Mexico—the world's largest producer of avocados—and the Ministry of Agriculture has invested \$10 million to develop CIDAM as a biotechnology research center to buttress the local agricultural sector. The DNALC project has strong support from Antonio Jara, a physicist who was president of the University of Michoacán and now is governor of the state.

## DNA Center NYC

The success of the *Harlem DNA Lab* and DNALC-inspired centers around the world provided the impetus to establish a dedicated center in New York City. This will allow us to further extend the DNALC “brand” to NYC, using a successful formula of academic year field trips, in-school instruction and teacher training, and follow-up support—perfected over 25 years. The development of *DNALC NYC* is a \$25 million goal of CSHL's current \$250 million 125th Anniversary Campaign. The catalyst for the venture came in April 2012 with a \$6 million lead gift from CSHL Trustee Laurie Landeau. This was followed in 2013 by a \$10 million gift from the Thompson Family Foundation and a \$3 million grant from the Alfred P. Sloan Foundation. An additional \$750,000 of support has come from the Hearst Foundation, the Booth Ferris Foundation, and the Achelis Foundation. The strategy is to use \$3 million for renovation, with drawdown from a \$22 million endowment funding about half of annual operating costs and providing scholarships for at least half of the student participants.

During the year, we worked intensively with Cushman & Wakefield Realtors to identify a 7,000–10,000 square foot facility with good street visibility and easy access to public transport. VVA and Gilman Consulting reviewed the complicated New York City Department of Buildings codes. We visited and reviewed approximately 20 properties, ultimately developing preliminary plans or offers on six. Safety zoning, cost, or landlord issues scuttled all of the best prospects. So, we were heartened when we re-opened negotiations with City University of New York (CUNY) Chief Operating Officer Allan Dobrin, Vice Chancellor for Research Gillian Small, and New York City College of Technology President Russell Hotzler for space on their City Tech Campus in Brooklyn.

The *DNA Center NYC* leadership team was significantly strengthened in April with the hiring of Annie Greengard as Manager of Science Education Partnerships. Annie reports to the CSHL Vice President of Development, Charlie Prizzi, and is charged to work closely with the DNALC to help establish, sustain, and operate a Manhattan location. Annie is working to establish a donor base in NYC by identifying, cultivating, and soliciting individuals, foundations, and corporations. A major effort is to build the *DNALC NYC* Council, whose objective is to help integrate the Center into the life of NYC. The Council will be challenged to raise funds to support the Center on an annual basis to complement the Center's endowment. Members are introducing the initiative to opinion leaders in business, government, education, and science who can help advance CSHL's mission to bring New Yorkers into the gene age. Annie is also spearheading efforts to grow our network of *Partner Member* (previously *Charter Member*) schools, whose prepayment for activities eases annual budgeting.

## Genomic Approaches in BioSciences

During the year, we wound down on NSF's Advanced Technological Education project to train college faculty members to implement experiments that integrate four major technologies of the genome era: PCR, DNA sequencing, RNA interference (RNAi), and bioinformatics. DNALC staff taught a workshop on the CSHL campus in August, bringing the workshop total to 13 and exceeding our goal of providing direct training to 288 faculty. Underrepresented minorities composed 21.7% of participants, and 67.2% were females. We maintained a mix of educators from

three educational levels: about one-fourth high school, one-half community college, and one-fourth 4-year college or university.

We concentrated primarily on supporting second-round workshops that were taught by faculty trained by us in the initial term of the grant. Six summer workshops were taught by collaborators at Universidad del Turabo (Gurabo, Puerto Rico), Bluegrass Community and Technical College (Lexington, Kentucky), Madison Area Technical College (Madison, Wisconsin), and Kennedy-King College (Chicago, Illinois). A preliminary analysis of surveys of teachers trained in first- and second-round workshops offers support for the “train-the-trainer” model. Post-workshop knowledge gains and confidence in teaching lab methods and bioinformatics were strikingly similar for participants instructed by DNALC staff at primary workshops and those instructed by faculty trainers at supplementary workshops. Slight differences for certain categories are most likely due to differences in emphasis at the supplementary workshops, which were customized according to local instructional needs. Most importantly, the follow-up survey conducted 9–12 months later showed that similar or higher proportions of supplementary workshop participants reported having incorporated workshop labs, bioinformatics, and biotech careers into their teaching.

As part of a collaboration with Stony Brook University, doctoral student Caren Gough is measuring the impact of the *Genomic Approaches* workshops on faculty confidence over time. The *Genomic Approaches in Biosciences* Workshops are unique in that they, by design, include a mix of educators from high schools, community colleges, and 4-year universities. In a preliminary analysis, Ms. Gough found that there were significant gains (with medium to large effect sizes) in both knowledge and confidence pre- to post-workshop. Statistical comparisons revealed similar gains across the three teacher groups. High school, community college, and university faculty had similar outcomes, indicating that the benefits were comparable for all participants. High school teachers felt as confident as teachers from other levels—even for difficult bioinformatics computer labs. Importantly, 75% of respondents to a reflection survey agreed that having teachers from



*Genomic Approaches in Biotechnology* locations of home institutions.

multiple education levels working together was a positive aspect of the workshop. Focus groups corroborated that participants benefitted from sharing their learning experience with teachers from the other levels.

## Genomic Approaches in Biotechnology Participant Home Institutions

### DNA Barcoding

DNA barcoding programs at the DNALC have now expanded into three separate programs. The newest is *Barcode Long Island: Exploring Biodiversity in a Unique Urban Landscape (BLI)*, which is funded by a \$1.2 million National Institutes of Health (NIH) Science Education Partnership Award. In collaboration with Stony Brook University (SBU), Brookhaven National Laboratory (BNL), and the American Museum of Natural History (AMNH), *BLI* will provide students with real and relevant research experience while they contribute to our knowledge of Long Island biodiversity.

After advertising to our extensive network of school districts and holding several information sessions, we held two *BLI* training workshops for 73 teachers planning to mentor student teams. By the end of the year, 24 teams with 71 students from 10 Long Island schools had submitted research proposals. The two schools from Nassau County and eight schools from Suffolk are diverse, with minority enrollment ranging from 5% to 91% (24% average). The students have proposed DNA barcoding projects on animals (six), plants (six) and fungi (one) and will present their findings at a symposium in spring 2015. A website to support all aspects of the project has been developed that allows for project proposal submission, review, revision, and acceptance—features that are necessary to make distributed barcoding feasible.

*BLI* builds upon our experience with two student research programs in New York City: the *Urban Barcode Project* competition (*UBP*) and the *Urban Barcode Research Program* (*UBRP*). In the *UBP*, students work with teacher mentors on DNA barcoding projects during the school year and compete for a grand prize. In the *UBRP*, students complete summer workshops and then conduct 55+ hours of research directly with scientist mentors.

Thirty-eight *UBP* teams—145 students, 53% African American or Latino—presented research posters at the American Museum of Natural History (AMNH) on May 27. These projects were supported by 14 footlocker kits and 42 weekend or after-school “open labs” at *Harlem DNA Lab* and Genspace. The students’ projects included mapping wildlife found in NYC, detecting food



Antonia Florio demonstrates DNA isolation during a *Barcode Long Island* teacher workshop.



*Barcode Long Island* teacher using a pestle to grind his plant sample for DNA barcoding.

fraud, determining biodiversity of NYC green spaces, identifying traded plant and animal species, and exploring the effects of biodiversity on human health. When asked to compare the experience with other research experiences—mostly eighth grade exit projects or school science fairs—*UBP* students ( $n = 145$ ) feel that the DNA barcoding programs provide more “real world” science (82% of students), more chance for hands-on experience (84%) and to learn science (75%), more opportunity to develop critical thinking (71%) and independent inquiry skills (73%), and more understanding of the scientific process (68%). DNA barcoding increases students’ interest in studying science or pursuing a career in science (72%), while still being more fun than other research experiences (69%).

A jury of 21 experts in biodiversity, conservation biology, DNA barcoding, and education selected the top six finalist teams (13 students) who gave oral presentations at AMNH on June 2. Mark Stoeckle, Senior Research Associate in the Program for the Human Environment at The Rockefeller University, gave the keynote address: “DNA Barcoding: The Thrill of Discovery.” The *UBP* grand prize winner was Eric Bovee, from The High School for Environmental Studies, under the direction of mentors J.T. Boehm and Marissa Bellino. Their project studied the international and domestic seahorse trade to see if vendors abide by trade regulations. After measuring and extracting DNA from hundreds of seahorses, Eric found that several vulnerable species are sold and most domestically traded seahorses are under the 10-cm size limit. Both of these findings suggest more work needs to be done to protect vulnerable seahorse populations.

Twenty *UBRP* teams (38 students) presented their projects as posters and oral presentations on April 29 at the AMNH. These students were selected from the 130 high-achieving NYC high school students who completed *UBRP* preparatory courses in *Conservation Genetics* and *DNA Barcoding*. The students worked with scientist mentors from institutions all over NYC, including the AMNH, CUNY, Fordham University, and Albert Einstein Medical Center. The *UBRP* projects spanned studies of food fraud, biodiversity, conservation genetics, phylogenetics (evolution), behavior, and the first DNALC microbiome studies. The *UBRP* was the first science fair for 32% of students. Of those who had previously participated in a science fair or competition (68%), the vast majority felt that the *UBRP* experience *increased* their interest in a career in science (86%).

The NYC teams collected and processed more than 1000 samples for DNA sequencing—an average of 17 samples per team—with more than 1600 single sequences provided by GENEWIZ, Inc. and 600,000 sequence reads produced by next-generation sequencing methods. Importantly, *UBP* and *UBRP* students produced an additional 12 novel DNA sequences that did not match existing data in the international database, GenBank. These are being published to GenBank with the students as authors, which will make the data freely available for use by other researchers.

During the summer, 48 middle school students attended the *Backyard Barcoding* camp, and 39 high school students stepped up to *DNA Barcoding Research*. The advanced students explored the aquatic biodiversity of the intertidal zone of CSHL’s campus, collecting 158 samples that produced 30 novel sequences to be published to GenBank. The students developed project proposals, collected samples from Cold Spring Harbor, extracted and amplified specimen DNA, and analyzed DNA sequences. We also partnered with the NYC Department of Education to offer two DNA barcoding courses at the Environmental Study Center, where teams from four public high schools assessed the biodiversity of Brooklyn’s Marine Park. The NYC students collected 115 samples, with 15 novel sequences to be published to GenBank. Each barcoding workshop closed with oral presentations of student findings. Most of the barcoding sequences came from small aquatic invertebrates, species that are hard to classify using traditional taxonomy, highlighting the power of DNA barcoding.

## High School Research Programs

In 2014, 19,298 students attended field trips at our three facilities: Dolan DNALC, *DNALC West*, and *Harlem DNA Lab*. We reached an additional 9046 students through in-school instruction by DNALC staff, and more than 1300 students attended week-long DNALC summer camps. Grants from Bank of America and National Grid Foundation supported programs for more than 2000 underserved students from Long Island school districts including Brentwood, Uniondale, William Floyd, Central Islip, and Valley Stream public schools. An additional 15 students from IS 59 in Queens participated in an ongoing collaboration with NSLIJ at *DNALC West*, which culminated at the New York Hall of Science's *STEMtastic* day where students showcased their lab skills.

Of the 2314 students who visited *Harlem DNA Lab*, 85% received scholarships from the William Townsend Porter Foundation, including students from two schools co-located within the John S. Roberts Educational Complex: Coalition School for Social Justice and Harlem Village Academy Leadership School. Supported by the New York City Department of Education (NYC DOE) Office of School Programs and Partnerships, 83 high school and 71 middle school students attended camps at the *Harlem DNA Lab* and the NYC DOE Environmental Study Center during school breaks. The footlocker kit rental program at the *Harlem DNA Lab* continued to provide teachers with equipment and reagents to implement basic genetics, biotechnology, and DNA barcoding labs. Footlocker kits were used by 1032 students, 117 of whom were conducting independent research under the *UBP* and *UBRP*. Sixty-four percent were underrepresented minorities.

The *Partner Membership* (formerly *Charter Membership*) program continues to flourish in NYC. Founding member Chapin School is now in the second year of an advanced elective, *Molecular Genetics*, which included a pilot project using metabarcoding to investigate the diversity of bacteria in their school. The third-prize *UBP* team from Convent of the Sacred Heart School presented *Nice to Meat You: Using DNA Barcoding to Detect Mislabeling in the Meat Industry*. St. David's School finished its first year in the program with 5th grade poster presentations, including a few DNA barcoding projects. Marymount School began their first year with in-school labs for upper and middle school girls, as well as DNALC participation in the Upper School after-school program.

We continued to collaborate with Cold Spring Harbor High School (CSHHS) to deliver *Molecular and Genomic Biology*. Students devote two class periods per day to this college-level laboratory course co-instructed by DNALC staff and CSHHS faculty Jaak Raadsepp. The 2013–2014 cohort focused on hands-on experimentation and independent projects across a range of biological systems, including units on eugenics, DNA barcoding, RNAi, human and plant genomics, microbiomes, and protein structure. The 2014–2015 cohort, our tenth class, began the year with classic *DNA Science* labs in recombinant DNA technology and a survey of plant diversity at the Nature Conservancy.

The second year of partnership with St. Dominic High School in Oyster Bay progressed with the independent research component of a 5-month course, *Molecular and Genomic Biology Research*. Several students chose to develop projects using DNA barcoding to study local biodiversity. A new group of high school juniors began year three in September. The fall curriculum includes labs in bacterial genetics, genomics, bioinformatics, and RNAi.

Our newest collaboration with the Nassau BOCES Doshi STEM School entered its second year of programming with two cohorts of students in grades nine and ten. Students in *Living Environment* and research electives began the school year with DNA barcoding projects that included surveys of biological diversity in pristine and polluted environments, combined with an investigation of abiotic factors that might affect biodiversity in each location. This culminated in the presentation of posters at a small poster session and at the *Barcode Long Island* symposium.



Monthly *Saturday DNA!* sessions drew 215 participants, with parents and grandparents joining children for classes on forensics, human evolution, the central dogma, and experimental design. An additional 4183 visitors viewed films in the Multitorium, including Cablevision's multimedia presentation *Long Island Discovery*, and/or viewed the exhibition *The Genes We Share*. The annual *Great Moments in DNA Science* seminar series drew 50 Long Island high school students to three presentations by researchers. Dr. Michael Feigin, CSHL, discussed how genomic data are used to find gene targets in breast cancer; Dr. Daniel Bogenhagen, SBU, explained how the mitochondria became an essential part of eukaryotic cells; and Dr. Danielle Engle, CSHL, talked about efforts to improve survival in patients with pancreatic cancer by developing early detection strategies and identifying better biomarkers for this type of cancer.

### Graduate Training

Our collaboration with the CSHL Watson School of Biological Sciences continues to provide graduate students with the opportunity to learn skills necessary for communicating with non-biologists. During a 3-month training program, graduate students of the WSBS work together with seasoned DNALC instructors to develop effective teaching techniques. When they have completed the training, the graduate students will be prepared to quickly assess any audience and structure quality lessons accordingly.

In the first phase, the WSBS students observe a DNALC instructor teaching a laboratory and then must organize lesson plans integrating their own experiences. In the second phase of training, the graduate students co-teach with the DNALC teachers to conduct a lesson. The DNALC instructors then provide feedback, discussing strengths and appropriate preparation for an independent lesson. The final phase requires the graduate students to teach a lesson under DNALC instructor observation. When both middle and high school rotations have been completed, the WSBS students select three additional lessons to demonstrate mastery of instruction and class management skills.

Our successful collaboration with the New York Academy of Sciences (NYAS) continues to improve the science literacy among high-needs students throughout New York City. DNALC instructors provided training to graduate students and postdoctoral researchers who serve as mentors in the Academy's after-school mentoring program. Each received 6 hours of training in effective strategies to provide hands-on laboratories to middle school students.

### BioMedia Visitation and Projects

We had a record number of visitors to our suite of multimedia resources in 2014. *Google Analytics* counted 5,202,881 visits to 22 DNALC websites. Our *YouTube* videos received 873,881 views, and the *3D Brain*, *Weed to Wonder*, and *Gene Screen* apps were downloaded 668,396 times. Total website, *YouTube*, and smartphone/tablet apps visitation reached 6.75 million, a 7.7% increase over 2013.

Developed under a grant from the Dana Foundation, the *3D Brain* app continues to be successful, with 2.72 million downloads to date on Apple, Windows, and Android devices worldwide. In late 2013, we released an update for the Apple version of the *3D Brain* app that included an



Chun-hua Yang films Olympia Davison explaining how to perform a DNA extraction.

“in-app” purchase of *3D Brain HQ*, which includes higher-resolution images and a redesigned user interface for \$0.99. In 2014, there were 25,343 downloads of the *HQ* version, netting \$17,600 after Apple took its share. Proceeds from in-app purchases support the educational programs of the DNALC.

Although reduced in size by almost half, the *BioMedia* Group continued to support the DNALC’s educational programs with website development and maintenance; videography and postproduction for webinars, special events, and offsite workshops; program promotion through email campaigns; flyer, poster, and backdrop design for professional meetings; museum development; photo documentation of events and barcoding specimens; social media management; and creation of architectural concept plans for potential DNALC venues in NYC and China. The Laurie J. Landeau Multimedia Studio hosted several events, including: monthly bioinformatics webinars for *iPlant* and the CSHL *Gramene* project; mtDNA sequencing field trip follow-up webinars; remote hands-on laboratory instruction; and seminars to remote student and faculty audiences. We also routinized our traveling Livestream studio setup at several events, including meetings at CSHL, Stony Brook University, University of Nebraska, and the USDA; and for *UBP* and *UBRP* symposia.

As part of the EOT Group of *iPlant Collaborative*, we developed and implemented a new design for the *iPlant* website that launched in May. The new site incorporates several improvements over the previous *iPlant* site: increased focus on and easier access to *iPlant*’s tools; simplified structure eliminating embedded and redundant content; integration of the *iPlant* Learning Center; and modified color palette and updated imagery.

In the summer, we worked with high school student Olympia Davison to create short videos that give a student’s perspective on hands-on laboratories and research careers. Olympia interviewed CSHL neuroscientist Dr. Anne Churchland, and demonstrated “How to Load a Gel” and a “DNA Extraction” from strawberries. The videos were uploaded to YouTube, where they have had more than 1,200 views.

In September, we installed *BOLD* (named for the acronym for the Barcode of Life Data Systems), an art exhibition on DNA barcoding displayed in the front and side halls of the DNALC. Seattle-based artist Joseph Rossano partnered with biologists to engage the public around the science of DNA barcoding and



*BOLD* exhibition artwork inspired by life in coral reefs in French Polynesia.



Artist Joseph Rossano discusses his artwork with students during the installation of his show in the DNALC halls.

how it is being used to catalog the world's vast and threatened biodiversity. The work is inspired by specimens barcoded by Daniel Janzen from the Area de Conservación Guanacaste (ACG) in Costa Rica and by researchers for the University of California, Berkeley, Moorea Biocode Project being carried out at the Richard B. Gump South Pacific Research Station in Moorea. All of the pieces deliberately represent the specimens as fuzzy, obscured, or out of scale, making it difficult to identify the organism. However, on closer inspection the viewer discovers that each piece includes the DNA barcode sequence and a link to the BOLD database, where the species identity is revealed. This underscores the reality that identifying species with certainty requires more than taxonomic methods. Our educators are developing interpretive activities for visiting students, many of whom will take part in a barcoding laboratory during their visit.

As the evolution of our museum continues, plans are in place to include a one-of-a-kind reproduction of the 5300-year-old natural mummy discovered in the melting ice of the Ötztal Alps in 1991. Gary Staab, artist and life modeler, will create the "Otzi" replica (named for the region where the mummy was found) using CAT scans of the mummy and three-dimensional printing technology. Working with the South Tyrol Museum of Archeology in Italy (where the iceman's body is currently kept), the project is planned to start in early 2015 and installation is projected for late spring. The iceman's unique story will give us a platform to create a central theme on forensic biology and human origins. The publication of Ötzi's genome provides a unique opportunity to reflect on the meaning of *becoming* and *being* human. Written in each person's DNA is a record of our shared ancestry, the information that makes us both similar and unique, and information that could foreshadow each person's future health. This exhibit aims to encourage people to consider where we came from, who we are, and where we are going.

### Staff and Interns

In 2014, we said goodbye to several staff members. Although it happens rarely, two design positions were phased out due to a lack of funding. Eun Sook Jeong, Multimedia Designer, is originally from South Korea and began as an intern in 2001. With a personality and design style tagged as "Zen" by our director, she has left a beautiful legacy as lead designer on several DNALC websites, as well as making significant contributions to all of our sites. She also applied her interior design training in our unique upstairs kitchen/lounge area. Chris Weidler, Junior Designer, began as a college intern in the *BioMedia* Group in 2007. Initially assigned the tedious jobs that freed up the senior designers, he quickly proved that he could take on more intensive tasks. He soaked up knowledge in the Visual Communications major at Farmingdale State College and became a skilled Flash animator, print designer, video editor, and web designer who finalized a redesign of the *iPlant Collaborative* website just before departing.

Jermel Watkins, Ph.D., Staff Molecular Biologist and Educator, began his time with the DNALC as a high school intern from 1994 to 2000. After earning his Ph.D. in molecular and cellular pharmacology at Stony Brook University, he returned to us as a high school instructor in 2007. He pioneered new workshops, managed interns, mentored research projects, and managed prep for workshops taught at all levels. Jermel accepted a position as an Assistant Professor and a Cancer and Neuroscience Research Scientist in the Department of Biological Sciences at Hampton University.

In 2007, Jennifer Galasso joined us as a Middle School Educator, and helped produce new workshops and labs in barcoding, personalized medicine, and epigenetics. Jenn also managed middle school lab preparation and the intern staff. She is now teaching *Introduction to Research* and *Living Environment* to ninth grade students at the Doshi STEM Institute of Nassau BOCES down the road in Syosset.

Amy Nisselle, Ph.D, Multimedia and Evaluation Manager, learned about the DNALC at a 2003 congress on human genetics in Melbourne, Australia, where she saw a presentation on our multimedia content. Amy then chose to pursue a Ph.D. in multimedia genetics education. As part of her thesis, she assisted with the evaluation of our *Genes to Cognition* website in 2008. She joined the DNALC in 2010 in her dual evaluation/producer role, and we miss her evaluation abilities and her big-picture perspective on multimedia projects. She returned to Australia to be with family and so her son (just beginning to speak) will form an Australian accent. Amy continued to consult with us for several months as we searched for her replacement.

The DNALC welcomed several new staff members in 2014. Emily Harrison joined us from the CSHL campus as a middle school educator. After receiving her Bachelor's degree in Biology and a Master's in Secondary Education from Adelphi University, she became a lab technician in Greg Hannon's lab. She recently returned to Adelphi University for her Master of Science in biology and is completing her Master's thesis research on laser-activated transcription.

Ashleigh Jackobel, a native Long Islander and graduate of Northport High School, began her time at the DNALC when she was hired as a college intern. After graduating from SUNY Farmingdale with a Bachelor of Science in Bioscience and a minor in Chemistry, Ashleigh jumped at the opportunity to join our team as a middle school and high school instructor. She is hoping to pursue a Ph.D. in Virology.

Shreya Shah, a new high school educator, developed an interest in bacteria during an episode of bubonic plague near her village in India, where she volunteered in a public awareness campaign to control the spread of plague. This fascination with microorganisms led her to earn a Master of Science in Microbiology at the University of Kansas. Her recent experience at Dowling College Microbiology Research Laboratory on soil microbial diversity in Long Island ecosystems makes her an excellent choice to help lead *BLI*.

Dr. Mona Spector became the project coordinator for the DNALC's 3-year NSF project bringing RNA-Seq to undergraduate education. Mona also joins us from the CSHL campus, where her impressive contributions to elucidating the genetic causes of cancer include key discoveries in the field of programmed cell death (published in the journal *Nature*) and helping to develop a new three-dimensional culture system used to examine pancreatic cancer.

Anne Greengard started in the spring as Manager of Science Education Partnerships. Annie is a New York City native who received her B.A. from Dartmouth College. She moved on to Nashville to attend law school at Vanderbilt University, then landing a place at a large law firm in NY. Annie's family roots in science and her passion for education make her well-suited for her role working on development efforts for the *DNA Center NYC*.

We welcomed back Cornel Ghiban after a brief hiatus as a Perl programmer at Estée Lauder. He develops and maintains the expanding suite of DNALC websites and applications and plays a central role in the development of *DNA Subway*.

Christine Marizzi and Antonia Florio each received well-deserved promotions this year. First a science educator, and now Manager, DNALC *West*, Christine instructs the majority of the 3000 students who attend field trips there each year, as well as teaches offsite in local schools. Christine also played an integral role in the development of DNA barcoding programs. Since joining our staff in March 2013 as Conservation Genetics Instructor, Antonia has the significant task of managing the *UBP* and *UBRP*, and she provides instruction at all three DNALC locations. Now Manager, Research, Antonia also applies for funding, facilitates student evaluation, and plays a guiding role for *BLI*.

Since the DNALC opened, we have relied on high school and college interns to support our day-to-day operations. In addition to laboratory prep work, interns may take on independent or team projects mentored by DNALC staff members, which translate current research into classroom labs. The *BioMedia* Group also welcomes interns for summer or longer-term roles. An internship offers students the unique opportunity to gain real laboratory or design experience in an educational environment. We gathered an amazing group of interns this year and said farewell as others left for college:

**High School Interns**


---

Kayla Bianco, St. Anthony's High School	Andrew Micklos, Ward Melville High School
Abigail Buckley, Oceanside High School	Nicholas Nicolia, St. Anthony's High School
Kenneth Collado, George W. Hewlett High School	Gabrielle Ramirez, Walt Whitman High School
Liam Gensel, Harborfields High School	Gavrielle Rood, Roslyn High School
Emily Gruber, Syosset High School	Breanna Tahany, Kellenberg Memorial High School
Joshua Gruber, Half Hollow Hills High School East	Maria Urbina, Oyster Bay High School
Omotayo Ikuomenisan, Huntington High School	Scott Venturino, Huntington High School
Aiseosa Irowa, Hempstead High School	Diana Vizza, North Shore High School
Serena Jones, Portledge High School	Blakelee Woods, Glen Cove High School
Lenni Joya Amaya, Huntington High School	

**High School Interns Departing for College**


---

Gabriella Blazich, Fordham University	Jack Manzi, Binghamton University
Brittany Coscio, College of the Holy Cross	Anant Mehrotra, Rochester Institute of Technology
Michaela Lee, SUNY Oneonta	Jonathan Paris, Johns Hopkins University

**College Interns**


---

Kathryn Bellissimo, The College of New Jersey	William McBrien, Suffolk Community College
Anne Bode, University of Notre Dame	Pauline McGlone, CUNY Hunter College
Marie Jean Francois, CUNY City College	Alec Micklos, University of Tampa
Nella Hauser, SUNY Stony Brook	Arvind Rajabhathor, New York University
Ashleigh Jackobel, Farmingdale State College	Robert Scott, Sage College of Albany
Karoline Lake, Columbia University	Shenika Shah, New York Institute of Technology

**Sites of Major Faculty Workshops**

Program Key:	<i>Middle School</i>	High School	College
ALABAMA		University of Alabama, Tuscaloosa	1987–1990
		<b>Hudson Alpha Institute, Huntsville</b>	<b>2014</b>
ALASKA		<b>University of Alaska, Anchorage</b>	<b>2012</b>
		<b>University of Alaska, Fairbanks</b>	1996
ARIZONA		Arizona State University, Tempe	2009
		Tuba City High School	1988
		<b>University of Arizona, Tucson</b>	<b>2011</b>
		<b>United States Department of Agriculture, Maricopa</b>	<b>2012</b>
ARKANSAS		Henderson State University, Arkadelphia	1992
		<b>University of Arkansas, Little Rock</b>	<b>2012</b>
CALIFORNIA		<b>California State University, Dominguez Hills</b>	<b>2009</b>
		<b>California State University, Fullerton</b>	<b>2000</b>
		<b>California Institute of Technology, Pasadena</b>	<b>2007</b>
		Canada College, Redwood City	1997
		City College of San Francisco	2006
		<b>City College of San Francisco</b>	<b>2011, 2013</b>
		<b>Community College of Denver</b>	<b>2014</b>
		Contra Costa County Office of Education, Pleasant Hill	2002, 2009
		<b>Foothill College, Los Altos Hills</b>	<b>1997</b>
		Harbor-UCLA Research & Education Institute, Torrance	2003
		<b>Los Angeles Biomedical Research Institute (LA Biomed), Torrance</b>	<b>2006</b>
		Laney College, Oakland	1999
		Lutheran University, Thousand Oaks	1999
		Oxnard Community College, Oxnard	2009
		<b>Pasadena City College</b>	<b>2010</b>
		<b>Pierce College, Los Angeles</b>	<b>1998</b>
		Salk Institute for Biological Studies, La Jolla	2001, 2008

	<b>San Francisco State University</b>	1991
	<b>San Diego State University</b>	<b>2012</b>
	<b>San Jose State University</b>	<b>2005</b>
	Santa Clara University	2010
	<b>Southwestern College, Chula Vista</b>	<b>2014</b>
	<b>Stanford University, Palo Alto</b>	<b>2012</b>
	<b>University of California, Berkeley</b>	<b>2010, 2012</b>
	University of California, Davis	1986
	<b>University of California, Davis</b>	<b>2012, 2014</b>
	<b>University of California, Northridge</b>	<b>1993</b>
	University of California, Riverside	2011
	<b>University of California, Riverside</b>	<b>2012</b>
COLORADO	Aspen Science Center	2006
	Colorado College, Colorado Springs	1994, 2007
	<b>Colorado State University, Fort Collins</b>	<b>2013</b>
	<b>United States Air Force Academy, Colorado Springs</b>	<b>1995</b>
	<b>University of Colorado, Denver</b>	1998, 2009–2010, <b>2014</b>
CONNECTICUT	Choate Rosemary Hall, Wallingford	1987
DISTRICT OF COLUMBIA	<b>Howard University, Washington</b>	<b>1992, 1996, 2009–2010</b>
FLORIDA	Armwood Senior High School, Tampa	1991
	Florida Agricultural & Mechanical University, Tallahassee	2007–2008
	<b>Florida Agricultural &amp; Mechanical University, Tallahassee</b>	<b>2011</b>
	North Miami Beach Senior High School	1991
	<b>Seminole State College, Sanford</b>	<b>2013–2014</b>
	University of Miami School of Medicine	2000
	University of Western Florida, Pensacola	1991
GEORGIA	Fernbank Science Center, Atlanta	1989, 2007
	<b>Gwinnett Technical College, Lawrenceville</b>	<b>2011, 2012</b>
	<b>Morehouse College, Atlanta</b>	<b>1991, 1996–1997</b>
	<b>Spelman College, Atlanta</b>	<b>2010</b>
HAWAII	Kamehameha Secondary School, Honolulu	1990
	<b>University of Hawaii at Manoa</b>	<b>2012</b>
ILLINOIS	Argonne National Laboratory	1986–1987
	iBIO Institute/Harold Washington College, Chicago	2010
	Illinois Institute of Technology, Chicago	2009
	<b>Kings College, Chicago</b>	<b>2014</b>
	<b>University of Chicago</b>	<b>1992, 1997, 2010</b>
INDIANA	Butler University, Indianapolis	1987
	<b>Purdue University, West Lafayette</b>	<b>2012</b>
IDAHO	University of Idaho, Moscow	1994
IOWA	Drake University, Des Moines	1987
KANSAS	University of Kansas, Lawrence	1995
KENTUCKY	<b>Bluegrass Community &amp; Technical College, Lexington</b>	<b>2012–2014</b>
	Murray State University	1988
	University of Kentucky, Lexington	1992
	Western Kentucky University, Bowling Green	1992
LOUISIANA	Bossier Parish Community College	2009
	Jefferson Parish Public Schools, Harvey	1990
	John McDonogh High School, New Orleans	1993
	<b>Southern University at New Orleans</b>	<b>2012</b>
MAINE	Bates College, Lewiston	1995
	<b>Southern Maine Community College</b>	<b>2012–2013</b>
	Foundation for Blood Research, Scarborough	2002
MARYLAND	Annapolis Senior High School	1989
	<b>Bowie State University</b>	<b>2011</b>
	Frederick Cancer Research Center	1995
	McDonogh School, Baltimore	1988
	Montgomery County Public Schools	1990–1992
	National Center for Biotechnology Information, Bethesda	2002

	<i>St. John's College, Annapolis</i>	1991
MASSACHUSETTS	<b>University of Maryland, School of Medicine, Baltimore</b>	<b>1999</b>
	<b>Arnold Arboretum of Harvard University, Roslindale</b>	<b>2011</b>
	Beverly High School	1986
	Biogen Idec, Cambridge	2002, 2010
	<b>Boston University</b>	<b>1994, 1996</b>
	CityLab, Boston University School of Medicine	1997
	Dover-Sherborn High School, Dover	1989
	Randolph High School	1988
	The Winsor School, Boston	1987
	Whitehead Institute for Biomedical Research, Cambridge	2002
MICHIGAN	Athens High School, Troy	1989
	<b>Schoolcraft College, Livonia</b>	<b>2012</b>
MINNESOTA	Minneapolis Community and Technical College, Madison	2009
	<b>Minneapolis Community and Technical College, Madison</b>	<b>2013</b>
	University of Minnesota, St. Paul	2005
	<b>University of Minnesota, St. Paul</b>	<b>2010</b>
MISSISSIPPI	Mississippi School for Math & Science, Columbus	1990–1991
	<b>Rust College, Holly Springs</b>	<b>2006–2008, 2010</b>
MISSOURI	<b>St. Louis Science Center</b>	<b>2008–2010</b>
	Stowers Institute for Medical Research, Kansas City	2002, 2008
	<b>University of Missouri, Columbia</b>	<b>2012</b>
	<b>Washington University, St. Louis</b>	<b>1989, 1997, 2011</b>
MONTANA	<b>Montana State University, Bozeman</b>	<b>2012</b>
NEBRASKA	<b>University of Nebraska, Lincoln</b>	<b>2014</b>
NEVADA	University of Nevada, Reno	1992, <b>2014</b>
NEW HAMPSHIRE	Great Bay Community College, Portsmouth	2009
	<b>New Hampshire Community Technical College, Portsmouth</b>	<b>1999</b>
	St. Paul's School, Concord	1986–1987
NEW JERSEY	Coriell Institute for Medical Research, Camden	2003
	Raritan Valley Community College, Somerville	2009
NEW MEXICO	Biolink Southwest Regional Meeting, Albuquerque	2008
NEW YORK	Albany High School	1987
	American Museum of Natural History, New York	2007
	Bronx High School of Science	1987
	Canisius College, Buffalo	2007
	<b>Canisius College, Buffalo</b>	<b>2011</b>
	City College of New York	2012
	Cold Spring Harbor High School	1985, 1987
	<b>Cold Spring Harbor Laboratory</b>	<b>2014</b>
	<b>Columbia University, New York</b>	<b>1993</b>
	<b>Cornell University, Ithaca</b>	<b>2005</b>
	<i>DeWitt Middle School, Ithaca</i>	1991, 1993
	Dolan DNA Learning Center	1988–1995, 2001–2004, 2006–2009
	<b>Dolan DNA Learning Center</b>	<b>1990, 1992, 1995, 2000–2011</b>
	<i>Dolan DNA Learning Center</i>	1990–1992
	DNA Learning Center <i>West</i>	2005
	<i>Fostertown School, Newburgh</i>	1991
	<i>Harlem DNA Lab, East Harlem</i>	2008–2009, 2011–2013
	Huntington High School	1986
	Irvington High School	1986
	John Jay College of Criminal Justice	2009
	<i>Junior High School 263, Brooklyn</i>	1991
	<i>Lindenhurst Junior High School</i>	1991
	Mount Sinai School of Medicine, New York	1997
	<b>Nassau Community College, Garden City</b>	<b>2013</b>
	<b>New York Botanical Garden, Bronx</b>	<b>2013</b>
	New York City Department of Education	2007, 2012

	New York Institute of Technology, New York	2006
	<b>New York Institute of Technology, New York</b>	<b>2006</b>
	<i>Orchard Park Junior High School</i>	1991
	<i>Plainview-Old Bethpage Middle School</i>	1991
	State University of New York, Purchase	1989
	<b>State University of New York, Stony Brook</b>	1987–1990, <b>2014</b>
	Stuyvesant High School, New York	1998–1999
	The Rockefeller University, New York	2003
	<b>The Rockefeller University, New York</b>	<b>2010</b>
	<i>Titusville Middle School, Poughkeepsie</i>	1991, 1993
	Trudeau Institute, Saranac Lake	2001
	Union College, Schenectady	2004
	<b>United States Military Academy, West Point</b>	<b>1996</b>
	Wheatley School, Old Westbury	1985
NORTH CAROLINA	CIIT Center for Health Research, Triangle Park	2003
	<b>North Carolina Agricultural &amp; Technical State University, Greensboro</b>	<b>2006–2007, 2009–2011</b>
	North Carolina School of Science, Durham	1987
	<b>North Carolina State University, Raleigh</b>	<b>2012</b>
NORTH DAKOTA	<b>North Dakota State University, Fargo</b>	<b>2012</b>
OHIO	Case Western Reserve University, Cleveland	1990
	Cleveland Clinic	1987
	Langston University, Langston	2008
	North Westerville High School	1990
OKLAHOMA	<b>Oklahoma City Community College</b>	<b>2000</b>
	Oklahoma City Community College	2006–2007, 2010
	Oklahoma Medical Research Foundation, Oklahoma City	2001
	Oklahoma School of Science and Math, Oklahoma City	1994
	Tulsa Community College, Tulsa	2009
	<b>Tulsa Community College, Tulsa</b>	<b>2012–2014</b>
OREGON	Kaiser Permanente-Center for Health Research, Portland	2003
	<b>Linfield College, McMinnville</b>	<b>2014</b>
PENNSYLVANIA	Duquesne University, Pittsburgh	1988
	Germantown Academy	1988
	Kimmel Cancer Center, Philadelphia	2008
RHODE ISLAND	Botanical Society of America, Providence	2010
SOUTH CAROLINA	<b>Clemson University</b>	<b>2004</b>
	Medical University of South Carolina, Charleston	1988
	<b>University of South Carolina, Columbia</b>	<b>1988</b>
TENNESSEE	NABT Professional Development Conference, Memphis	2008
TEXAS	Austin Community College–Rio Grande Campus	2000
	<b>Austin Community College–Eastview Campus–Roundrock Campus</b>	<b>2007–2009, 2013</b>
	<b>Austin Community College–Roundrock Campus</b>	<b>2012</b>
	<b>Houston Community College Northwest</b>	<b>2009–2010</b>
	J.J. Pearce High School, Richardson	1990
	Langham Creek High School, Houston	1991
	<b>University of Lone Star College, Kingwood</b>	<b>2011</b>
	Midland College	2008
	Southwest Foundation for Biomedical Research, San Antonio	2002
	Taft High School, San Antonio	1991
	<b>Texas A&amp;M University, College Station</b>	<b>2013</b>
	<b>Texas A&amp;M University, Prairie View</b>	<b>2013</b>
	Texas A&M, AG Research and Extension Center, Weslaco	2007
	<b>Trinity University, San Antonio</b>	<b>1994</b>
	<b>University of Texas, Austin</b>	<b>1999, 2004, 2010, 2012</b>
	<b>University of Texas, Brownsville</b>	<b>2010</b>
UTAH	<b>Brigham Young University, Provo</b>	<b>2012</b>
	University of Utah, Salt Lake City	1993
	<b>University of Utah, Salt Lake City</b>	<b>1998, 2000</b>



	<b>Utah Valley State College, Orem</b>	<b>2007</b>
VERMONT	University of Vermont, Burlington	1989
VIRGINIA	Eastern Mennonite University, Harrisonburg	1996
	Jefferson School of Science, Alexandria	1987
	Mathematics and Science Center, Richmond	1990
	Mills Godwin Specialty Center, Richmond	1998
	<b>Virginia Polytechnic Institute and State University, Blacksburg</b>	<b>2005, 2008–2009</b>
WASHINGTON	Fred Hutchinson Cancer Research Center, Seattle	1999, 2001, 2008
	Shoreline Community College	2011, 2012
	<b>University of Washington, Seattle</b>	<b>1993, 1998, 2010</b>
WEST VIRGINIA	Bethany College	1989
WISCONSIN	Blood Center of Southeastern Wisconsin, Milwaukee	2003
	<b>Madison Area Technical College/Madison Area College</b>	<b>1999, 2009, 2011–2014</b>
	Marquette University, Milwaukee	1986–1987
	University of Wisconsin, Madison	1988–1989
	<b>University of Wisconsin, Madison</b>	<b>2004, 2012</b>
WYOMING	University of Wyoming, Laramie	1991
PUERTO RICO	<b>Universidad del Turabo, Gurabo</b>	<b>2011, 2012, 2014</b>
	University of Puerto Rico, Mayaguez	1992
	<b>University of Puerto Rico, Mayaguez</b>	<b>1992</b>
	<b>University of Puerto Rico, Rio Piedras</b>	<b>1993</b>
	University of Puerto Rico, Rio Piedras	1994
<hr/>		
AFRICA	Godfrye Okoye University, Enugu, Nigeria	2013
AUSTRALIA	Walter and Eliza Hall Institute and University of Melbourne	1996
AUSTRIA	<b>Vienna Open Lab, Vienna</b>	<b>2007, 2012</b>
CANADA	Red River Community College, Winnipeg, Manitoba	1989
CHINA	Beijing No. 166 High School, Beijing	2013, 2014
	<b>Ho Yu College, Hong Kong</b>	<b>2009</b>
DENMARK	<b>Faroe Genome Project, Torshavn, Faroe Islands</b>	<b>2013</b>
GERMANY	Urania Science Center, Berlin	2008
ITALY	International Institute of Genetics and Biophysics, Naples	1996
	Porto Conte Research and Training Laboratories, Alghero	1993
MEXICO	ASPB Plant Biology, Merida	2008
PANAMA	<b>University of Panama, Panama City</b>	<b>1994</b>
RUSSIA	Shemyakin Institute of Bioorganic Chemistry, Moscow	1991
SINGAPORE	National Institute of Education	2001–2005
	Singapore Science Center	2013
SWEDEN	Kristineberg Marine Research Station, Fiskebackskil	1995
	Uppsala University	2000
THE NETHERLANDS	<b>International Chromosome Conference, Amsterdam</b>	<b>2007</b>
	<b>Wageningen University and Research Center, Wageningen</b>	<b>2014</b>
UNITED KINGDOM	<b>Wellcome Trust Conference Center, Hinxton</b>	<b>2012–2013</b>
	<b>University of Warwick, Coventry</b>	<b>2013</b>

## Workshops, Meetings, Collaborations, and Site Visits

Jan 6–Feb 13	Pinkerton Foundation Bi-Weekly <i>Urban Barcode Research Program</i> , “DNA Detectives,” American Museum of Natural History, New York
January 8	Site visit by Tatiana Nikolenko, RUSNANO, Moscow, Russia
January 9–14	NSF ATE Professional Development, <i>Genomic Approaches in BioSciences</i> Supplemental Workshop, Universidad del Turabo, Gurabo, Puerto Rico
January 10	Site visit by Suzhou Industrial Park and bioBAY Delegation, Suzhou, China
January 11–15	NSF <i>iPlant Collaborative</i> pre-PAG Half-Day Workshop, International Plant and Animal Genome Conference XXIII, “A Simple Bioinformatics Workflow for RNA-Seq Analysis and Distributed Genome Annotation,” San Diego, California
January 23	<i>Urban Barcode Project/Urban Barcode Research Program</i> Open Lab, <i>Harlem DNA Lab</i>

January 24	<i>Urban Barcode Project</i> “A Practical Introduction to DNA Barcoding,” University of Delaware, Newark, Delaware
January 25	<i>KidCreate</i> , O’Shea Educational Complex, New York
January 28	<i>Urban Barcode Project/Urban Barcode Research Program</i> Open Lab, <i>Harlem DNA Lab</i>
January 29	NSF <i>iPlant Collaborative</i> Webinar, “Getting Started with <i>iPlant</i> ,” DNALC
January 30	<i>Urban Barcode Project/Urban Barcode Research Program</i> Open Lab, <i>Harlem DNA Lab</i>
Jan 31–Feb 1	NSF <i>iPlant Collaborative</i> Professional Development <i>Tools and Services</i> Workshop, Linfield College, McMinnville, Oregon
Feb 3–Mar 7	<i>DNA Barcoding</i> Research Mentoring Program, “DNA Detectives,” American Museum of Natural History, New York
February 4	<i>Urban Barcode Project/Urban Barcode Research Program</i> Open Lab, <i>Harlem DNA Lab</i>
February 6	<i>Urban Barcode Project/Urban Barcode Research Program</i> Open Lab, <i>Harlem DNA Lab</i>
February 8	<i>Saturday DNA!</i> “Got Milk?” DNALC
February 8	<i>Urban Barcode Project/Urban Barcode Research Program</i> Open Lab, <i>Harlem DNA Lab</i>
February 10, 11	<i>Human Mitochondrial Sequencing</i> Laboratory Follow-up Webinar, Oceanside High School, DNALC
February 11	<i>Urban Barcode Project/Urban Barcode Research Program</i> Open Lab, <i>Harlem DNA Lab</i>
February 11	<i>Urban Barcode Project/Urban Barcode Research Program</i> Open Lab, Genspace, Brooklyn, New York
February 15	STEM Mentor Training Workshop, New York Academy of Sciences, New York
February 17–24	<i>DNA Science</i> Workshop, <i>Harlem DNA Lab</i>
February 18	<i>Urban Barcode Project/Urban Barcode Research Program</i> Open Lab, Genspace, Brooklyn, New York
February 19	PCR Workshop, Rockefeller University, New York
February 24	Pinkerton Foundation <i>Urban Barcode Research Program</i> , “Barcode of Life,” American Museum of Natural History After-School Program, New York
February 25	<i>Urban Barcode Project/Urban Barcode Research Program</i> Open Lab, <i>Harlem DNA Lab</i>
February 25	<i>Urban Barcode Project/Urban Barcode Research Program</i> Open Lab, Genspace, Brooklyn, New York
February 26	Pinkerton Foundation <i>Urban Barcode Research Program</i> , “Barcode of Life,” American Museum of Natural History After-School Program, New York
February 27	<i>Urban Barcode Project/Urban Barcode Research Program</i> Open Lab, <i>Harlem DNA Lab</i>
February 28	NSF <i>iPlant Collaborative</i> Seminar, “The <i>iPlant Collaborative</i> —Community Cyberinfrastructure for Life Science,” Genome Analysis Centre, Norwich, United Kingdom
March 3	Pinkerton Foundation <i>Urban Barcode Research Program</i> , “Barcode of Life,” American Museum of Natural History After-School Program, New York
March 3–4	NSF <i>iPlant Collaborative Tools &amp; Services</i> Workshop, Wageningen University and Research Centre, Wageningen, The Netherlands
March 4	<i>Urban Barcode Project/Urban Barcode Research Program</i> Open Lab, <i>Harlem DNA Lab</i>
March 5	Pinkerton Foundation <i>Urban Barcode Research Program</i> , “Barcode of Life,” American Museum of Natural History After-School Program, New York
March 5	“DNA Barcoding,” Queensborough Community College, Bayside, New York
March 5	Site visit by The Hearst Foundation, New York
March 5	Site visit by Dr. Jose Cordova and Dr. J. Armando Barriguete, Mexico City, Mexico
March 6	<i>Great Moments in DNA Science</i> Honors Seminar, “The Search for New Drug Targets in Breast Cancer,” Dr. Michael Feigin, CSHL
March 8	<i>Urban Barcode Project/Urban Barcode Research Program</i> Open Lab, <i>Harlem DNA Lab</i>
March 10	NSF <i>iPlant Collaborative</i> Webinar, “Getting Started with <i>iPlant</i> ,” DNALC
March 10	Pinkerton Foundation <i>Urban Barcode Research Program</i> , “Barcode of Life,” American Museum of Natural History After-School Program, New York
March 11	<i>Urban Barcode Project/Urban Barcode Research Program</i> Open Lab, <i>Harlem DNA Lab</i>
Mar 12–Apr 11	Pinkerton Foundation <i>DNA Barcoding</i> Research Mentoring Program, “DNA Barcoding,” Queensborough Community College, Bayside, New York
March 13	<i>Urban Barcode Project/Urban Barcode Research Program</i> Open Lab, <i>Harlem DNA Lab</i>
March 13–16	NSF <i>iPlant Collaborative Tools &amp; Services</i> Half Workshop, 56 <sup>th</sup> Maize Genetics Conference, Beijing, China
March 14	“Women in STEM Careers,” Panel Discussion, The Chapin School, New York, New York
March 15	<i>Saturday DNA!</i> “Silencing Genes,” DNALC
March 17	Pinkerton Foundation <i>Urban Barcode Research Program</i> , “Barcode of Life,” American Museum of Natural History After-School Program, New York
March 18	<i>Urban Barcode Project/Urban Barcode Research Program</i> Open Lab, <i>Harlem DNA Lab</i>
March 18	<i>Great Moments in DNA Science</i> Honors Seminar, “Mitochondrial DNA, Our Second Genome,” Dr. Daniel Bogenhagen, Stony Brook University, Stony Brook, New York
March 19	Pinkerton Foundation <i>Urban Barcode Research Program</i> , “Barcode of Life,” American Museum of Natural History After-School Program, New York
March 20	<i>Urban Barcode Project/Urban Barcode Research Program</i> Open Lab, <i>Harlem DNA Lab</i>

- March 22 *Urban Barcode Project/Urban Barcode Research Program* Open Lab, *Harlem DNA Lab*
- March 22 Cold Spring Harbor Schools 3<sup>rd</sup> & 4<sup>th</sup> Grade Program, DNALC
- March 24 Pinkerton Foundation *Urban Barcode Research Program*, “Barcode of Life,” American Museum of Natural History After-School Program, New York
- March 25–27 NSF *iPlant Collaborative* Train the Trainers/Tools and Services Workshop, University of California-Davis, Davis, California
- March 26 “DNA Extraction,” Presentation, Port Jefferson Science Fair, Edna Louise Spear Elementary School, Port Jefferson, New York
- March 26 Pinkerton Foundation *Urban Barcode Research Program*, “Barcode of Life,” American Museum of Natural History After-School Program, New York
- March 26–27 NSF *iPlant Collaborative* Mission Critical Bioinformatics Workshop, University of California, Davis
- March 27 *Great Moments in DNA Science* Honors Seminar, “Using Sugars to Detect Pancreatic Cancer,” Dr. Dannielle Engle, CSHL
- March 27 *Urban Barcode Project/Urban Barcode Research Program* Open Lab, *Harlem DNA Lab*
- March 28–29 Interview with Lior Patchley, “Tuxedo RNASeq,” University of California, Berkeley
- March 29 NSF ATE Professional Development, *Genomic Approaches in BioSciences* Workshop, Seminole State College, Sanford, Florida
- March 29 Tour by Renaissance Technologies, East Setauket, New York, DNALC
- March 31 Pinkerton Foundation *Urban Barcode Research Program*, “Barcode of Life,” American Museum of Natural History After-School Program, New York
- Mar 31–Apr 2 NSF Conference on Course-Based Undergraduate Research Experiences (CUREnet), Host and “Infrastructure and Data for Large-Scale Collaborations” Plenary Panel, CSHL
- April 1 *Urban Barcode Project/Urban Barcode Research Program* Open Lab, *Harlem DNA Lab*
- April 2 Pinkerton Foundation *Urban Barcode Research Program*, “Barcode of Life”, American Museum of Natural History After-School Program, New York
- April 3–4 National Science Teachers Association National Conference 2014, “Detecting Epigenetic DNA Methylation in *Arabidopsis Thaliana*,” “DNA Subway: Bringing Cutting Edge Bioinformatics into the Classroom,” and “DNA Barcoding Independent Research in the Classroom,” Boston, Massachusetts
- April 8 *Urban Barcode Project/Urban Barcode Research Program* Open Lab, *Harlem DNA Lab*
- April 10 *Urban Barcode Project/Urban Barcode Research Program* Open Lab, *Harlem DNA Lab*
- April 11 NSF *iPlant Collaborative* Webinar, “Getting Started with iPlant,” DNALC
- April 12 NSF ATE Professional Development *Genomic Approaches in BioSciences* Follow-up Workshop, Tulsa Community College, Tulsa, Oklahoma
- April 12 *Saturday DNA!* “Transcription and Translation: What Does DNA Actually Do?” DNALC
- April 12 *Urban Barcode Project/Urban Barcode Research Program* Open Lab, *Harlem DNA Lab*
- April 14–18 *Genome Science* Workshop, *Harlem DNA Lab*
- April 17–18 NSF Research Collaboration Network: Integrating Bioinformatics into Life Sciences Education, “The *iPlant Collaborative*—Community Cyberinfrastructure for Life Sciences,” Omaha, Nebraska
- April 22 “Glowing Genes,” Remote Instruction Webinar, Berean Christian Jr. Academy, Atlanta, Georgia
- April 24 *Urban Barcode Project/Urban Barcode Research Program* Open Lab, *Harlem DNA Lab*
- April 25 Site visit by Dr. Jose Angel Cordova, Mme. Lia Limon, Dr. J. Armando Barriguete, and M. Hugo Scherer, Mexico City, Mexico
- April 25 DNA Day Scavenger Hunt, DNALC and Cold Spring Harbor, New York
- April 26 *Urban Barcode Project/Urban Barcode Research Program* Open Lab, *Harlem DNA Lab*
- April 26 *Fun with DNA, Forensic Detectives and DNA Barcoding* Teacher Training Workshops, DNALC
- April 29 Pinkerton Foundation *Urban Barcode Research Program* Symposium, American Museum of Natural History, New York
- May 1 *Urban Barcode Project/Urban Barcode Research Program* Open Lab, *Harlem DNA Lab*
- May 6 Site visit by Joseph Rossano, Arlington, Washington
- May 8 *Urban Barcode Project/Urban Barcode Research Program* Open Lab, *Harlem DNA Lab*
- May 10 *Urban Barcode Project/Urban Barcode Research Program* Open Lab, *Harlem DNA Lab*
- May 12–13 NSF *iPlant Collaborative* Professional Development *Software Carpentry* Workshop, CSHL
- May 14 NSF *iPlant Collaborative* Webinar, “Getting Started with iPlant,” DNALC
- May 16 Site visit by Dr. J. Armando Barriguete, Mexico City, Mexico
- May 17 *Saturday DNA!* “Human Origins: What DNA Says About Our Ancestry,” DNALC
- May 21–22 NSF *iPlant Collaborative* Professional Development *Genomics in Education* Workshop, University of Nevada, Reno
- May 27 *Urban Barcode Project* Poster Session, American Museum of Natural History, New York
- May 27–28 NSF *iPlant Collaborative* Professional Development *Tools and Services* Workshop, Stony Brook, New York
- May 30 Site visit by The Wall Street Journal, *Harlem DNA Lab*
- June 2 *Urban Barcode Project* Symposium, American Museum of Natural History, New York

- June 2–3 Site visit by Amy Stark, University of Notre Dame, Indiana
- June 2–6 NSF ATE Professional Development *Genomic Approaches in BioSciences* Supplemental Workshop, Bluegrass Community and Technical College, Lexington, Kentucky
- June 3–4 Bio-Link Summer Fellows Forum, *Bioinformatics in the Classroom* Workshop, University of California-Berkeley, San Francisco
- June 5 Site visit by Peter Dworkin, Regeneron, Tarrytown, New York
- June 5 Ana Almeida, CSHL, Biology Career Webinar for Portuguese students
- June 6 NSF *iPlant Collaborative* New Release Previews Webinar, DNALC
- June 7 *Saturday DNA!* “Exploring Evolution,” DNALC
- June 9–13 NSF ATE Professional Development *Genomic Approaches in BioSciences* Supplemental (English and Spanish) Workshop, Bluegrass Community and Technical College, Lexington, Kentucky
- June 9–20 NSF Professional Development Infrastructure Training *Bringing Next-Generation Sequence (NGS) Analysis into Undergraduate Education* Workshop, CSHL
- June 9–23 *Forensics* Workshop, Convent of the Sacred Heart, Greenwich, Connecticut
- June 16–20 *World of Enzymes* Workshop, St. David’s School, New York
- June 19–20 *DNA Barcoding* Workshop, Cabrini College, Radnor, Pennsylvania
- June 23–24 NSF *iPlant Collaborative* Professional Development *Tools and Services* Workshop, *Understanding GWAS* Workshop, Lincoln, Nebraska
- June 23–27 NSF ATE Professional Development *Genomic Approaches in BioSciences* Supplemental Workshop, Universidad del Turabo, Gurabo, Puerto Rico
- June 27–28 NSF *iPlant Collaborative* Professional Development *Genomics in Education* Workshop, Community College of Denver, Denver, Colorado
- June 30–July 4 Pinkerton Foundation *Urban Barcode Research Program* “DNA Detectives” Workshop, *Harlem DNA Lab*
- June 30–July 4 *Fun with DNA* Workshops, DNALC (2 sessions)
- Green Genes* Workshop, DNALC
- World of Enzymes* Workshop, DNALC
- Fun with DNA* Workshop, DNA Learning Center West
- July 7–11 Pinkerton Foundation *Urban Barcode Research Program* “Barcode of Life” Workshop, Harlem DNA Lab
- July 7–11 *DNA Science* Workshops, DNALC (2 sessions)
- Forensic Detectives* Workshop, DNALC
- World of Enzymes* Workshop, DNALC
- World of Enzymes* Workshop, DNA Learning Center West
- July 7–11 NSF ATE Professional Development *Genomic Approaches in BioSciences* Supplemental Workshop, Madison Area Technical College, Madison, Wisconsin
- July 9 NSF *iPlant Collaborative* Webinar, “Getting Started with iPlant,” DNALC
- July 12–16/17–18 American Association of Plant Biologists 2014, “The iPlant Collaborative: Scalable Infrastructure for Life Science” Seminar / The iPlant Collaborative Software Carpentry Workshop, Portland, Oregon
- July 15 NSF Extreme Science and Discovery Environment (XSEDE) 2014 Conference, “*DNA Subway: Making Genome Analysis Egalitarian*,” Atlanta, Georgia
- July 14–18 *DNA Science* Workshop, DNALC
- Fun with DNA/World of Enzymes* Workshop with Beijing, China students, DNALC
- Fun with DNA* Workshop, DNALC
- Genome Science* Workshop, DNALC
- DNA Science* Workshop, DNA Learning Center West
- Fun with DNA* Workshop, Harlem DNA Lab
- July 14–18 NSF ATE Professional Development, *Genomic Approaches in BioSciences* Supplemental Workshop, Kennedy–King College, Chicago, Illinois
- July 14–25 *DNA Barcoding* Workshop, Anthony J. Genovesi Center, St. Francis College, Brooklyn, New York
- July 17–18 NSF *iPlant Collaborative* Professional Development *Software Carpentry* Workshop, Portland, Oregon
- July 19–21 *DNA Barcoding* Workshop, Marine Park at the Environmental Science Center, Brooklyn, New York
- July 21–22 *DNA Barcoding* Workshop, Rockefeller University, New York
- July 21–25 *BioCoding* Workshop, DNALC
- DNA Barcoding* Workshop, DNALC
- DNA Science* Workshop, DNALC
- Green Genes* Workshops, DNALC (2 sessions)
- Green Genes* Workshop, DNA Learning Center West
- July 23–24 “GOALS for Girls” Workshop, Rockefeller University, New York
- July 26–30 Botanical Society of America, Botany 2014, The *iPlant Collaborative Tools and Services/Genomics in Education* Workshop, Boise, Idaho
- July 28–Aug 1 *DNA Barcoding* Workshop, DNALC

- DNA Barcoding Research* Workshop, DNALC  
*Forensics* Workshop, DNALC  
*World of Enzymes* Workshop, DNALC  
*Fun with DNA* Workshop, DNA Learning Center West  
*Green Genes* Workshop, Harlem DNA Lab
- July 28–Aug 8 *DNA Barcoding* Workshop, Environmental Science Center, Brooklyn, New York  
 July 30 *DNA Barcoding* Webinar, Madison High School, San Antonio, Texas, DNALC  
 August 4–8 *DNA Barcoding Research* Workshop, DNALC  
*DNA Science* Workshop, DNALC  
*Fun with DNA* Workshop, DNALC  
*Fun with DNA/World of Enzymes* Workshop with Beijing, China students, DNALC  
*DNA Science* Workshop, DNA Learning Center West  
 Forensic Detectives Workshop, Harlem DNA Lab
- August 6 Site visit by Elizabeth Benjamin and Elena Chon, New York City Department of Education, New York,  
 Harlem DNA Lab
- August 11–15 *DNA Science* Workshop, DNALC  
*Green Genes* Workshops, DNALC (2 sessions)  
*Human Genomics* Workshop, DNALC  
*Genome Science* Workshop, DNA Learning Center West  
*DNA Science* Workshop, Harlem DNA Lab
- August 13 NSF *iPlant Collaborative* Webinar, “Getting Started with *iPlant*,” DNALC  
 August 13–14 NSF *iPlant Collaborative* Professional Development *Genomics in Education* Workshop, Hudson Alpha Institute,  
 Huntsville, Alabama
- August 17–22 10<sup>th</sup> World Congress of Genetics Applied to Livestock Production, “*DNA Subway*—An Educational  
 Bioinformatics Platform for Gene and Genome Analysis, DNA Barcoding and RNA-Seq,” Vancouver,  
 British Columbia, Canada
- August 18–22 NSF ATE Professional Development, *Genomic Approaches in BioSciences* Workshop, CSHL  
 August 18–22 *Backyard Barcoding* Workshop, DNALC  
*Fun with DNA* Workshop, DNALC  
*Silencing Genomes* Workshop, DNALC  
*World of Enzymes* Workshop, DNALC  
*DNA Science* Workshop, Delbrück Laboratory, CSHL  
*Fun with DNA* Workshop, DNA Learning Center West
- August 18–22 Pinkerton Foundation *Urban Barcode Research Program* Workshops, “Conservation Genetics,” “DNA Barcoding,”  
 Harlem DNA Lab
- August 25–29 *DNA Science* Workshop, DNALC  
*Forensic Detectives* Workshop, DNALC  
*Fun with DNA* Workshop, DNALC  
*World of Enzymes* Workshop, DNALC  
*Green Genes* Workshop, Hershey Laboratory, CSHL  
*World of Enzymes* Workshop, DNA Learning Center West
- August 25–29 *Urban Barcoding Project* Teacher Training Workshop, Harlem DNA Lab
- September 3 National Plant Genome Initiative Postdoctoral Research Fellows Annual Awardee Meeting, “Computational  
 Resources for Young Investigators: How Can *iPlant* Help?” Arlington, Virginia
- September 3 Professional Development Workshop, “Bacterial Transformation,” Progress High School, Brooklyn, New York  
 September 3–5 Banbury Meeting Presentation: “High Performance Computing (HPC) in Undergraduate Biology Education,  
 “*DNA Subway*—GUI to HPC for Undergraduates,” CSHL
- September 6 *Forensic Detectives* Teacher Training Workshop, New York Academy of Science, New York  
 September 7–10 European Conference on Computational Biology, “The *iPlant Collaborative*—Scalable Infrastructure for Life  
 Science,” Strasbourg, France
- September 15 NSF *iPlant Collaborative* Webinar, “Getting Started with *iPlant*,” DNALC  
 September 17–19 Molecular Biology Class, Marymount College, Tarrytown, New York  
 September 18–19 USDA/ARS Workshop on Arthropod Genomics, “The *iPlant Collaborative*—Scalable Infrastructure for Life  
 Science,” Beltsville, Maryland
- September 21–24 Genomic Informatics 2014 Conference, “*DNA Subway*: An HPC-Powered, User-Friendly Workflow for RNA-Seq  
 Analysis” Presentation, Churchill College, Cambridge, England
- September 25 Site visit by Datuk Seri Idris Jusoh, Kuala Lumpur, Malaysia  
 September 26 Site visit by Dr. J. Armando Barriguette, Mexico City, Mexico  
 October 8 NSF *iPlant Collaborative* Webinar, “Getting Started with *iPlant*,” DNALC  
 October 13–15 *Human Genome Science* Workshop, Beijing 166 Middle School, Beijing, China

- October 15 *Barcode Long Island* Information Session, Brookhaven National Laboratory, Upton, New York
- October 17 “Using an *Alu* Insertion Polymorphism to Study Human Populations,” Beijing 166 Middle School, Beijing, China
- October 18 *Saturday DNA!* “Bananas for DNA!” DNALC
- October 18 *Urban Barcode Project* Teacher Training Workshop, *Harlem DNA Lab*
- October 20 Site visit by Wanda Rosado, Unity Center for Urban Technologies, New York, *Harlem DNA Lab*
- October 20–22 *Human Genome Science* Workshop, Beijing 166 Middle School, Beijing, China
- October 22 21st National ATE Principal Investigators Conference, *ATE Evaluation Practice—Lessons from the Field* Pre-Conference Workshop, Washington D.C.
- October 23 *Barcode Long Island* Information Session, DNA Learning Center *West*
- October 24 “CSHL DNALC Lab Class Design and Implementation,” Beijing 166 Middle School, Beijing, China
- October 24 21st National ATE Principal Investigators Conference, *Genomic Approaches to BioSciences* Presentation, Washington D.C.
- October 30 *Barcode Long Island* Information Session, DNALC
- November 1 Metropolitan Association of College and University Biologists Annual Meeting, “Experiments in Gene and Genome Analysis,” Dowling College, Oakdale, NY
- November 4 CEI-PEA Teacher Network, “Short Fingerprinting Lab” and “Bacterial Transformation Lab,” East Harlem, New York
- November 4 *Barcode Long Island* Teacher Training Workshop, DNALC
- November 5–8 Biological Data Science Meeting, “Unleash Your Inner Data Scientist—Enabling Scalable Data Driven Collaborations with iPlant Cyberinfrastructure,” CSHL
- November 8 *Saturday DNA!* “Food for Thought,” DNALC
- November 10 NSF *iPlant Collaborative* Webinar, “Getting Started with iPlant,” DNALC
- November 13–15 2014 NABT Professional Development Conference, “DNA Barcoding: Independent Student-Driven Inquiry,” “What DNA Says About Our Human Family,” “Analyzing DNA Sequences to Understand Evolutionary Relationships,” Cleveland Convention Center, Cleveland, Ohio
- November 16 CureNET Workshop, “Scalable Experiments in Gene and Genome Analysis,” Bowie State University, Baltimore, Maryland
- November 19 *Barcode Long Island* Information Session, Midwood High School, Brooklyn, New York
- November 24 *Barcode Long Island* Information Session, Benjamin N. Cardozo High School, Bayside, New York
- December 3 NSF *iPlant Collaborative* Webinar, “Getting Started with *iPlant*,” DNALC
- December 4 *Urban Barcode Project/Urban Barcode Research Program* Open Lab, *Harlem DNA Lab*
- December 4 Site visit by Gary Staab, Staab Studios, Kansas City, Missouri
- December 5 Site visit by Dr. J. Armando Barriguete with Mexican Scientific Delegation, Mexico City, Mexico
- December 6 *Saturday DNA!* “Hairy Tales,” DNALC
- December 6 *Urban Barcode Project/Urban Barcode Research Program* Open Lab, *Harlem DNA Lab*
- December 6 *Barcode Long Island* Teacher Training Workshop, DNALC
- December 7–10 *iPlant Collaborative* Professional Development Train-the-Trainers Workshop, USDA/ARS Big Data: RNA-Seq, CSHL
- December 9 *Urban Barcode Project/Urban Barcode Research Program* Open Lab, *Harlem DNA Lab*
- December 11 Site visit to Regeneron, Tarrytown, New York
- December 12 Site visit by Ni Shuang, Wancheng Bao and Chuansen Zhao, Qilu Company, Shanghai, China
- December 13 *KidCreate*, O’Shea Educational Complex, New York, New York
- December 16 *Urban Barcode Project/Barcode Long Island* Information Session, Math for America, New York
- December 18 *Urban Barcode Project/Urban Barcode Research Program* Open Lab, *Harlem DNA Lab*





COLD SPRING HARBOR  
LABORATORY PRESS



# 2014 PRESS PUBLICATIONS

---

## Serials

- Genes & Development*, Vol. 28 ([www.genesdev.org](http://www.genesdev.org))  
*Genome Research*, Vol. 24 ([www.genome.org](http://www.genome.org))  
*Learning & Memory*, Vol. 21 ([www.learnmem.org](http://www.learnmem.org))  
*RNA*, Vol. 20 ([www.rnajournal.org](http://www.rnajournal.org))  
*Cold Spring Harbor Symposia in Quantitative Biology*, Vol. 78:  
*Immunity and Tolerance*, edited by Anne O'Garra,  
Michel Nussenzweig, Stephen Smale, David Stewart, and  
Bruce Stillman  
*Cold Spring Harbor Protocols* ([www.cshprotocols.org](http://www.cshprotocols.org))  
*Cold Spring Harbor Perspectives in Biology* ([www.cshperspectives.org](http://www.cshperspectives.org))  
*Cold Spring Harbor Perspectives in Medicine*  
([www.perspectivesinmedicine.org](http://www.perspectivesinmedicine.org))

## Laboratory Manuals

- Molecular Neuroscience: A Laboratory Manual*, edited by  
Rusty Lansford  
*Subcellular Fractionation: A Laboratory Manual*, edited  
by Paul R. Pryor

## Textbooks

- Signal Transduction: Principles, Pathways, and Processes*, edited by  
Lewis C. Cantley, Tony Hunter, Richard Sever, and  
Jeremy Thorner  
*Quickstart Molecular Biology: An Introduction for  
Mathematicians, Physicists, and Computational Scientists*,  
by Philip N. Benfey  
*Human Variation: A Genetic Perspective on Diversity, Race,  
and Medicine*, by Aravinda Chakravarti  
*Navigating Metabolism*, by Navdeep S. Chandel

## Monographs (Topic Collections from *Perspectives in Biology and Perspectives in Medicine*)

- Endocytosis*, edited by Sandra L. Schmid, Alexander Sorokin, and  
Marino Zerial  
*The Origin and Evolution of Eukaryotes*, edited by Patrick J.  
Keeling and Eugene V. Koonin  
*MYC and the Pathway to Cancer*, edited by Chi V. Dang and  
Robert N. Eisenman  
*The Skin and Its Diseases*, edited by Anthony E. Oro and Fiona  
M. Watt  
*The Genetics and Biology of Sexual Conflict*, edited by William R.  
Rice and Sergey Gavrilets  
*Tuberculosis*, edited by Stefan H.E. Kaufmann, Eric J. Rubin,  
and Alimuddin Zumla  
*Innate Immunity and Inflammation*, edited by Ruslan Medzhitov

- Human Fungal Pathogens*, edited by Arturo Casadevall, Aaron  
P. Mitchell, Judith Berman, Kyung J. Kwon-Chung, John R.  
Perfect, and Joseph Heitman  
*The Biology of Heart Disease*, edited by Margaret E. Buckingham,  
Christine L. Mummery, and Kenneth R. Chien

## History

- Father to Son: Truth, Reason, and Decency*, by James D. Watson

## Other

- Connecting with Companies: A Guide to Consulting Agreements  
for Biomedical Scientists*, by Edward Klees and H. Robert  
Horvitz  
*A Bioinformatics Guide for Molecular Biologists*, by Sarah J. Aerni  
and Marina Sirota  
*Experimental Design for Biologists*, Second Edition, by David J.  
Glass  
*Career Options for Biomedical Scientists*, edited by Kaaren Janssen  
and Richard Sever  
CSHL Annual Report 2013, Yearbook Edition  
Banbury Center Annual Report 2013

## E-books (Kindle editions)

- Father to Son: Truth, Reason, and Decency*, by James D. Watson  
*Connecting with Companies: A Guide to Consulting Agreements  
for Biomedical Scientists*, by Edward Klees and H. Robert  
Horvitz

## Websites

- Cold Spring Harbor Monographs Archive Online  
([www.cshmonographs.org](http://www.cshmonographs.org))  
Cold Spring Harbor Symposium on Quantitative Biology  
Archive ([symposium.cshlp.org](http://symposium.cshlp.org))  
*Lab Math: A Handbook of Measurements, Calculations, and  
Other Quantitative Skills for Use at the Bench, Second Edition*,  
author's blog ([www.labmath.org](http://www.labmath.org))  
*Connecting with Companies: A Guide to Consulting Agreements for  
Biomedical Scientists*, sample chapter, glossary, and additional  
information ([www.bioagreements.org](http://www.bioagreements.org))  
*Father to Son: Truth, Reason, and Decency*, timeline and excerpts  
([www.fathertoonbook.org](http://www.fathertoonbook.org))

## Services

- bioRxiv, the preprint server for biology ([www.biorxiv.org](http://www.biorxiv.org))  
BioSupplyNet, scientific supply directory ([www.biosupplynet.com](http://www.biosupplynet.com))

# COLD SPRING HARBOR LABORATORY PRESS

## EXECUTIVE DIRECTOR'S REPORT

---

Cold Spring Harbor Laboratory Press creates publications and services that help scientists achieve success, while supporting the Laboratory's reputation and goals in research, revenue, education, and communication. Rooted in the publication of the first Cold Spring Harbor Symposium on Quantitative Biology in 1933, the Press now has an internationally respected program of seven journals, 200 books, and two online services. It produces and distributes content digitally, but it also deploys long-established expertise and the latest technologies to create attractive print books when market demand warrants.

The year 2014 was another successful one for the Press in which its publications maintained or enhanced their reputation for quality, its bioRxiv preprint service made an excellent start, and its financial goals were exceeded, resulting in a significant contribution to the Laboratory. In addition, the groundwork was laid for the launch of a new research journal in 2015.

*Genes & Development* and *Genome Research* remained two of the world's four most highly cited genetics journals, and the more specialized research journals *RNA* and *Learning & Memory* continued to hold their own. There was strong growth for the newer journals, *Cold Spring Harbor Protocols*, *Cold Spring Harbor Perspectives in Biology*, and *Cold Spring Harbor Perspectives in Medicine*, each of which represents a strategic transformation of content previously available only as print manuals and monographs into successful web-based publications with all the benefits of online discovery and indexing and a sustainable business model. Total downloads of Press journal articles worldwide reached a new record level of 13 million.

In its first full year, the preprint service bioRxiv made a significant mark. More than 1000 manuscripts in 25 subject categories were posted, and authors received much feedback through on-site comments, social media, and private email. More than 300 of the posted articles were published in final form in more than 100 journals, and many journals changed their editorial policies to allow consideration of manuscripts previously made available as preprints. In this respect, and from the frequency with which posted manuscripts are revised and updated, it is clear that the existence of bioRxiv is changing the way research biologists are communicating their work.

The Press also published 21 new book titles in 2014, including Jim Watson's family memoir *Father to Son*. There was keen anticipation of the handbook *Career Options for Life Scientists* in response to the widening awareness that long-term opportunities for postdoctoral researchers in academia are lessening. The best-sellers remained the classic lab manual *Molecular Cloning* and the enormously popular career development handbooks *At the Bench* and *At the Helm*. The appetite for books in research science has shrunk substantially in recent years, but by reducing staff, cutting costs, and developing targeted marketing approaches, the Press has retained a book program with both print and electronic capabilities that continues to support its mission to the Laboratory and the scientific community at large.

### Staff

The business strategy of the Press in recent years has succeeded owing to investment in new projects with growing revenues and a strong focus on cost control, including staff numbers. At the end of the year, two more positions were eliminated, resulting in the departure of two long-time colleagues. Kaaren Janssen was one of our best developmental editors, and the long list of books her skills improved since she arrived in 1996 includes many that defined Cold Spring Harbor as a place of excellence in publishing. Geraldine Jaitin joined us in 2000 and expertly headed our

Fulfillment Department until its operations were outsourced. During the year, we also said farewell to Rena Steuer after 11 years of service as a Production Editor and Jacqueline Beggins, Assistant Production Editor. We welcomed Maria Ebbets as a replacement Assistant Production Editor. A full list of Press staff members as of December 31, 2014, is included elsewhere in this volume.

To fulfill the mission of the Press, we engage in our work some of the world's most accomplished scientists. They respect the Laboratory's uniqueness and drive for excellence, but ultimately these relationships are made possible by the skills and professionalism of the Press staff. We are fortunate to have such able and dedicated individuals in all the operations of the Press. I particularly wish to thank those who have leadership roles in our expanding and diversifying activities: Assistant Director Richard Sever, journal editors Terri Grodzicker and Hillary Sussman, and department directors Jan Argentine, Wayne Manos, Stephen Nussbaum, Marcie Siconolfi, and Linda Sussman. And I also thank my Executive Assistant, Mala Shwe Mazzullo, whose quiet efficiency and personal warmth support this organization in countless important ways.

**John Inglis**  
*Executive Director  
and Publisher*



*Genome Research* Assistant Editor Laura DeMare and Press Executive Director and Publisher John Inglis with book authors Marina Sirota and George Asimenos at the American Society of Human Genetics Annual Meeting, San Diego, California.



Press Executive Director and Publisher John Inglis presenting during a panel on *Industry Directions* at the National Press Club, Washington, D.C.



FINANCE

# FINANCIAL STATEMENTS

---

## CONSOLIDATED BALANCE SHEET

December 31, 2014

(with comparative financial information as of December 31, 2013)

---

	2014	2013
Assets:		
Cash and cash equivalents	\$ 56,309,959	65,258,594
Grants receivable	10,551,528	8,506,168
Contributions receivable, net	58,786,259	110,387,860
Publications inventory	1,673,595	2,187,157
Investments	442,830,529	384,070,487
Restricted use assets	5,127,815	4,504,767
Other assets	15,126,525	16,226,007
Land, buildings, and equipment, net	<u>231,650,890</u>	<u>231,988,575</u>
Total assets	<u>\$ 822,057,100</u>	<u>823,129,615</u>
Liabilities and net assets:		
Liabilities:		
Accounts payable and accrued expenses	\$ 12,510,995	10,437,314
Deferred revenue	5,509,689	5,209,891
Interest rate swap	33,623,553	18,613,481
Bonds payable	<u>97,200,000</u>	<u>97,200,000</u>
Total liabilities	<u>148,844,237</u>	<u>131,460,686</u>
Net assets:		
Unrestricted	325,723,978	294,223,173
Temporarily restricted	236,314,595	290,273,304
Permanently restricted	<u>111,174,290</u>	<u>107,172,452</u>
Total net assets	<u>673,212,863</u>	<u>691,668,929</u>
Total liabilities and net assets	<u>\$ 822,057,100</u>	<u>823,129,615</u>

# CONSOLIDATED STATEMENT OF ACTIVITIES

Year ended December 31, 2014

(with summarized financial information for the year ended December 31, 2013)

	<i>Unrestricted</i>	<i>Temporarily Restricted</i>	<i>Permanently Restricted</i>	<i>2014 Total</i>	<i>2013 Total</i>
Revenue and other support:					
Public support—contributions and nonfederal grant awards	\$ 14,658,668	20,623,012	4,001,838	39,283,518	45,095,768
Federal grant awards	27,176,257	—	—	27,176,257	26,766,266
Indirect cost allowances	23,710,599	—	—	23,710,599	23,228,697
Investment return utilized	15,695,747	801,735	—	16,497,482	14,837,880
Program fees	6,896,378	—	—	6,896,378	7,752,249
Publications sales	10,030,061	—	—	10,030,061	10,174,394
Dining services	4,322,717	—	—	4,322,717	4,715,566
Rooms and apartments	3,638,654	—	—	3,638,654	3,859,785
Miscellaneous	6,632,532	—	—	6,632,532	2,579,049
Net assets released from restrictions	<u>74,581,721</u>	<u>(74,581,721)</u>	<u>—</u>	<u>—</u>	<u>—</u>
Total revenue and other support	<u>187,343,334</u>	<u>(53,156,974)</u>	<u>4,001,838</u>	<u>138,188,198</u>	<u>139,009,654</u>
Expenses:					
Research	85,732,121	—	—	85,732,121	85,776,617
Educational programs	17,213,213	—	—	17,213,213	17,388,667
Publications	9,466,527	—	—	9,466,527	9,412,728
Banbury Center conferences	1,486,354	—	—	1,486,354	1,571,105
DNA Learning Center programs	1,973,519	—	—	1,973,519	1,623,885
Watson School of Biological Sciences programs	3,422,312	—	—	3,422,312	3,370,964
General and administrative	16,603,749	—	—	16,603,749	18,110,232
Dining services	<u>5,704,911</u>	<u>—</u>	<u>—</u>	<u>5,704,911</u>	<u>5,591,701</u>
Total expenses	<u>141,602,706</u>	<u>—</u>	<u>—</u>	<u>141,602,706</u>	<u>142,845,899</u>
Excess (deficiency) of revenue and other support over (under) expenses	45,740,628	(53,156,974)	4,001,838	(3,414,508)	(3,836,245)
Other changes in net assets:					
Investment return excluding amount utilized	770,249	(801,735)	—	(31,486)	36,766,285
Change in fair value of interest rate swap	<u>(15,010,072)</u>	<u>—</u>	<u>—</u>	<u>(15,010,072)</u>	<u>16,942,866</u>
Increase (decrease) in net assets	31,500,805	(53,958,709)	4,001,838	(18,456,066)	49,872,906
Net assets at beginning of year	<u>294,223,173</u>	<u>290,273,304</u>	<u>107,172,452</u>	<u>691,668,929</u>	<u>641,796,023</u>
Net assets at end of year	\$ <u>325,723,978</u>	<u>236,314,595</u>	<u>111,174,290</u>	<u>673,212,863</u>	<u>691,668,929</u>

# CONSOLIDATED STATEMENT OF CASH FLOWS

Year ended December 31, 2014

(with comparative financial information for the year ended December 31, 2013)

	2014	2013
Cash flows from operating activities:		
Change in net assets	\$ (18,456,066)	49,872,906
Adjustments to reconcile change in net assets to net cash provided by operating activities:		
Change in fair value of interest rate swap	15,010,072	(16,942,866)
Depreciation and amortization	13,942,830	14,816,411
Net appreciation in fair value of investments	(12,908,048)	(47,898,692)
Contributions restricted for long-term investment	(5,020,506)	(5,308,027)
Changes in assets and liabilities:		
Grants receivable	(2,045,360)	1,506,233
Contributions receivable, net of financing activities	48,392,029	16,062,486
Publications inventory	513,562	580,270
Other assets	1,909,996	(283,358)
Restricted use assets	(623,048)	(906,921)
Accounts payable and accrued expenses, net of financing activities	959,182	268,487
Deferred revenue	<u>299,798</u>	<u>(48,634)</u>
Net cash provided by operating activities	<u>41,974,441</u>	<u>11,718,295</u>
Cash flows from investing activities:		
Capital expenditures	(13,605,145)	(6,179,654)
Proceeds from sales and maturities of investments	75,649,786	119,019,919
Purchases of investments	(121,501,780)	(158,580,385)
Net change in investment in employee residences	<u>(810,514)</u>	<u>(50,707)</u>
Net cash used in investing activities	<u>(60,267,653)</u>	<u>(45,790,827)</u>
Cash flows from financing activities:		
Contributions restricted for long-term investment	4,001,838	2,121,977
Contributions restricted for investment in capital	1,018,668	3,186,050
Decrease in contributions receivable	3,209,572	18,839,280
Increase (decrease) in accounts payable relating to capital expenditures	<u>1,114,499</u>	<u>(287,585)</u>
Net cash provided by financing activities	<u>9,344,577</u>	<u>23,859,722</u>
Net decrease in cash and cash equivalents	(8,948,635)	(10,212,810)
Cash and cash equivalents at beginning of year	<u>65,258,594</u>	<u>75,471,404</u>
Cash and cash equivalents at end of year	\$ <u>56,309,959</u>	<u>65,258,594</u>
Supplemental disclosure:		
Interest paid	\$ <u>4,013,111</u>	<u>4,016,854</u>
Noncash investing and financing activity:		
Contributed property	\$ <u>—</u>	<u>641,995</u>

# FINANCIAL SUPPORT OF THE LABORATORY

Cold Spring Harbor Laboratory, Banbury Center, and the Dolan DNA Learning Center receive a substantial portion of their 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 2014.

## GRANTS January 1–December 31, 2014

### COLD SPRING HARBOR LABORATORY

<i>Grantor</i>	<i>Program/Principal Investigator</i>	<i>Duration of Grant</i>		<i>2014 Funding<sup>1</sup></i>
<b>FEDERAL GRANTS</b>				
<b>NATIONAL INSTITUTES OF HEALTH</b>				
<i>Program Project and Center Support</i>	Drs. Hannon/Krainer/Spector/Stillman	05/25/12	12/31/16	\$4,221,622.00
	Dr. Stillman—Cancer Center Core	08/17/11	07/31/16	4,608,416.00
<i>Cooperative Research Agreement Support<sup>2</sup></i>	Dr. Gingeras	09/21/12	07/31/16	2,087,649.00
	Dr. Hannon	09/26/14	06/30/17	452,400.00 *
	Drs. Osten/Albeanu/Mitra	09/26/14	06/30/17	1,245,306.00 *
	Drs. Powers/Krasnitz/Sordella/Tuveson	05/01/12	04/30/17	570,972.00
	Dr. Tuveson	06/12/14	02/28/19	258,623.00 *
<i>Research Support</i>	Dr. Albeanu	07/01/13	06/30/18	472,500.00
	Dr. Churchland	03/01/13	02/28/18	463,050.00
	Drs. Enikolopov/Koulakov	09/15/11	05/31/16	492,558.00
	Dr. Furukawa	03/01/10	02/28/15	467,775.00
	Dr. Furukawa	05/08/14	03/31/18	359,100.00 *
	Dr. Hannon	09/01/13	08/31/18	406,350.00
	Dr. Hannon	04/01/13	03/31/15	220,956.00
	Dr. Huang	07/01/11	03/31/16	565,050.00
	Dr. Huang	07/01/13	06/30/18	748,793.00
	Drs. Huang/Wigler	03/01/14	01/31/19	716,097.00 *
	Dr. Kepecs	04/01/14	03/31/19	450,250.00 *
	Drs. Kepecs/Huang	07/01/11	05/31/16	409,303.00
	Drs. Koulakov/Zador	07/01/13	03/31/18	425,250.00
	Dr. Krainer	04/01/12	03/31/17	734,358.00
	Dr. Li	07/01/10	03/31/15	564,518.00
	Dr. Li	03/01/14	02/28/19	472,500.00 *
	Dr. Martienssen	09/15/11	08/31/15	346,085.00
	Dr. Martienssen	01/20/12	11/30/15	400,837.00
	Dr. McCombie	07/23/10	02/28/15	347,879.00
	Dr. McCombie	08/16/13	06/30/16	622,428.00
	Dr. Mills	04/01/14	01/31/16	283,500.00 *
	Dr. Mills	09/16/14	08/31/19	634,545.00 *
	Dr. Mitra	03/15/13	01/31/16	250,003.00
	Dr. Osten	04/01/12	03/31/17	514,900.00
	Dr. Siepel	09/01/14	12/31/17	34,270.53 *
	Dr. D. Spector	04/01/11	03/31/15	746,284.00
	Dr. Stillman	06/01/12	05/31/16	700,725.00

<sup>1</sup>Includes direct and indirect costs

<sup>2</sup>Cooperative research agreement funding amounts include only CSHL portion of award

\*New or competing renewal grants awarded in 2014



Grantor	Program/Principal Investigator	Duration of Grant		2014 Funding <sup>1</sup>
	Dr. Tonks	07/01/10	04/30/15	\$ 745,634.00
	Dr. Trotman	07/30/14	06/30/19	392,175.00 *
	Dr. Turner	07/15/10	06/30/15	457,380.00
	Drs. Tuveson/C. Hammell/Hicks/Pappin	09/22/14	08/31/18	555,711.00 *
	Dr. Vakoc	04/01/13	03/31/18	532,576.00
	Dr. Van Aelst	08/01/13	05/31/18	524,106.00
	Dr. Van Aelst	04/01/13	03/31/18	477,000.00
	Dr. Zador	09/27/10	05/31/15	467,775.00
	Dr. Zador	09/01/12	08/31/17	401,625.00
	Dr. Zador	04/01/13	03/31/15	236,250.00
	Dr. Zador	07/01/14	04/30/19	413,438.00 *
	Dr. Zhong	01/01/14	12/31/18	401,625.00 *
<i>Research Subcontracts</i>				
NIH/Brandeis University Consortium Agreement	Dr. Mitra	09/26/14	06/30/17	301,141.00 *
NIH/Certerra, Inc. Consortium Agreement	Dr. Koulakov	01/03/13	12/31/14	124,996.00
NIH/Cornell University Consortium Agreement	Dr. Enikolopov	03/01/13	02/29/16	354,936.00
NIH/Emory University Consortium Agreement	Dr. Huang	04/01/14	02/28/18	84,746.00 *
NIH/Georgia Institute of Technology Consortium Agreement	Dr. D. Spector	09/30/06	07/31/15	200,000.00
NIH/New York University Consortium Agreement	Dr. Koulakov	06/01/14	05/31/19	176,715.00 *
NIH/Scripps Research Institute Consortium Agreement	Dr. Osten	09/30/12	08/31/17	35,923.00
NIH/Sloan-Kettering Institute for Cancer Research Consortium Agreement	Dr. Sordella	09/12/12	08/31/17	150,455.00
NIH/The Research Foundation for the State University of New York - Stony Brook Consortium Agreement	Dr. Wigler	05/01/14	04/30/19	130,599.00 *
NIH/Thomas Jefferson University Consortium Agreement	Dr. Enikolopov	03/15/14	02/28/15	17,834.00 *
NIH/University of California - Los Angeles Consortium Agreement	Dr. Wigler	09/26/14	06/30/16	626,027.50 *
NIH/University of Minnesota Consortium Agreement	Dr. Tuveson	07/01/12	06/30/15	85,397.00
NIH/University of Texas Consortium Agreement	Dr. Furukawa	09/30/11	05/31/15	18,700.00
<i>Fellowship/Career Development Support</i>				
	Dr. Anczukow	09/01/13	08/31/15	106,008.00
	Dr. Bhagwat	07/16/14	07/15/17	26,676.00 *
	S. Kelly	07/16/13	07/15/16	26,676.00
	J. Tucciarone	09/17/12	09/16/15	26,676.00
<i>Institutional Training Program Support</i>				
	Dr. Gann/Watson School of Biological Sciences	07/01/12	06/30/17	270,804.00
	Dr. Mills	09/01/11	08/31/16	178,855.00
<i>Course Support</i>				
	Advanced Immunocytochemistry: In Situ Hybridization and Live Cell Imaging	09/01/10	08/31/15	100,323.00
	Advanced Sequencing Technologies and Applications	04/10/12	03/31/15	52,816.00

<sup>1</sup>Includes direct and indirect costs

\*New or competing renewal grants awarded in 2014

Grantor	Program/Principal Investigator	Duration of Grant		2014 Funding <sup>1</sup>
	Cell and Developmental Biology of <i>Xenopus</i>	05/05/14	03/31/19	\$ 76,284.00 *
	Cellular Biology of Addiction	03/01/11	02/29/16	74,500.00
	Computational and Comparative Genomics	08/01/14	06/30/17	52,816.00 *
	Eukaryotic Gene Expression	04/01/12	03/31/17	107,205.00
	Mouse Development, Stem Cells & Cancer	04/01/12	03/31/17	106,419.00
	Neurobiology of <i>Drosophila</i>	07/15/12	06/30/17	35,000.00
	Programming for Biology	09/01/14	08/31/17	80,120.00 *
	Protein Purification and Characterization	04/01/12	03/31/17	106,289.00
	Proteomics	08/01/12	04/30/17	110,695.00
	Statistical Methods for Functional Genomics	08/05/13	07/31/17	105,657.00
	X-Ray Methods in Structural Biology	09/01/12	08/31/17	81,540.00
<i>Meeting Support</i>	Axon Guidance, Synapse Formation and Regeneration	08/01/14	07/31/15	26,694.00 *
	The Biology of Genomes	04/01/13	03/31/18	54,600.00
	Blood Brain Barrier	09/30/14	09/29/15	20,000.00 *
	Germ Cells	07/01/14	06/30/19	6,000.00 *
	Glia Health and Disease	09/30/13	08/31/14	10,000.00
	Mechanisms and Models of Cancer	08/06/14	07/31/15	7,000.00 *
	Molecular Chaperones and Stress Response	04/01/14	03/31/15	30,101.00 *
	Molecular Genetics of Aging	03/15/14	02/28/19	30,972.00 *
	Neurodegenerative Diseases: Biology and Therapeutics	09/30/14	09/29/15	10,000.00 *
	Nuclear Receptors and Disease	09/23/14	08/31/15	23,000.00 *
	PTEN Pathways and Targets	03/17/14	02/28/15	6,000.00 *
	Retroviruses	03/15/12	02/28/17	36,645.00
	Systems Biology: Global Regulation of Gene Expression	03/01/10	02/28/15	5,000.00
	The PARP Family and Friends: Gene Regulation and Beyond	03/01/14	02/28/15	5,000.00 *
	Wiring the Brain	09/30/14	08/31/15	20,000.00 *
<i>Library Support</i>	Human Genome Project Documentary History: An Annotated Scholarly Guide to the Human Genome Project	02/01/14	01/31/16	44,685.00 *
<b>NATIONAL SCIENCE FOUNDATION</b>				
<i>Multiple Project Award Support</i>	Dr. Jackson	10/01/10	09/30/15	1,162,296.00
	Drs. Lippman/Schatz	11/01/12	10/31/15	879,736.00
	Dr. Mitra	10/01/13	09/30/16	312,534.00
	Dr. Ware	06/01/12	05/31/17	2,392,521.00
<i>Research Support</i>	Dr. Albeanu	09/15/14	08/31/16	145,428.00 *
	Dr. Mitra	09/01/14	08/31/16	152,243.00 *
	Dr. Schatz	06/01/14	05/31/19	293,574.00 *
	Drs. Timmermans/M. Hammell	05/01/12	04/30/16	282,768.00
	Dr. Timmermans	08/01/14	07/31/18	302,892.00 *

<sup>1</sup>Includes direct and indirect costs

\*New or competing renewal grants awarded in 2014

<i>Grantor</i>	<i>Program/Principal Investigator</i>	<i>Duration of Grant</i>		<i>2014 Funding<sup>1</sup></i>
<i>Research Subcontracts</i>				
NSF/Cornell University Consortium Agreement	Dr. Timmermans	02/01/13	01/31/18	\$ 415,537.00
NSF/Cornell University Consortium Agreement	Dr. Ware	05/15/13	04/30/18	146,823.00
NSF/Iowa State University Consortium Agreement	Dr. Jackson	03/01/13	02/28/18	390,389.00
NSF/North Carolina State University Consortium Agreement	Dr. Martienssen	03/01/11	02/29/16	469,871.00
NSF/University of Arizona Consortium Agreement	Drs. Ware/Micklos/Schatz	09/01/13	08/31/18	1,844,316.00
<i>Fellowship Support</i>				
	C. Carlston	08/18/14	07/31/19	44,000.00 *
	D. Lambert	08/15/14	07/31/19	44,000.00 *
<i>Undergraduate Training Program Support</i>	Drs. Schatz/Churchland/Ware	04/01/12	03/31/15	118,922.00
<i>Course Support</i>				
	Advanced Bacterial Genetics	01/15/14	12/31/16	85,000.00 *
	Computational Cell Biology	08/01/12	07/31/15	70,000.00
	Frontiers and Techniques in Plant Science	05/15/12	04/30/15	103,124.00
	Yeast Genetics and Genomics	07/15/14	06/30/17	85,000.00 *
<b>UNITED STATES DEPARTMENT OF AGRICULTURE</b>				
<i>Research Support</i>				
	Dr. Jackson	01/15/11	01/14/15	130,834.00
	Dr. Jackson	12/01/13	11/30/17	123,319.00
	Dr. McCombie	09/01/14	08/31/15	500,000.00 *
<b>UNITED STATES DEPARTMENT OF THE ARMY</b>				
<i>Research Support</i>				
	Dr. Egeblad	06/15/14	06/14/19	794,485.00 *
	Dr. Hannon	09/15/12	09/14/17	1,072,791.00
	Dr. Trotman	09/22/14	09/21/17	229,412.00 *
	Dr. Tuveson	09/30/14	09/29/16	270,914.00 *
<b>UNITED STATES DEPARTMENT OF ENERGY</b>				
<i>Research Subcontracts</i>				
DOE/Brookhaven National Laboratory Consortium Agreement	Drs. Ware/Schatz	11/29/11	09/30/16	1,050,000.00
<b>UNITED STATES DEPARTMENT OF THE NAVY</b>				
<i>Course Support</i>	Synthetic Biology	07/01/14	06/30/15	50,000.00 *
<b>MISCELLANEOUS SOURCES OF FUNDING</b>				
<i>Program Project Support</i>				
David and Fanny Luke	Farm Support/Dr. Stillman	10/01/14	09/30/15	250,000.00 *
Pfizer Inc.	Dr. Stillman	01/01/12	12/31/15	2,050,000.00
Pioneer Hi-Bred International, Inc.	Drs. Jackson/Lippman/Martienssen/Timmermans/Ware	07/01/12	09/30/14	400,000.00

<sup>1</sup>Includes direct and indirect costs

\*New or competing renewal grants awarded in 2014

<i>Grantor</i>	<i>Program/Principal Investigator</i>	<i>Duration of Grant</i>		<i>2014 Funding<sup>1</sup></i>
The Simons Foundation/Autism	Dr. Wigler	01/01/12	12/31/16	\$4,000,000.00
The Simons Foundation Center for Quantitative Biology	Dr. Wigler	09/01/08	08/31/15	1,000,000.00
The Stanley Medical Research Institute	Drs. Watson/McCombie	07/01/14	06/30/16	5,000,000.00 *
St. Giles Foundation	Cancer Center Support/Dr. Stillman	01/01/14	12/31/15	100,000.00 *
<i>Research Support</i>				
Alex's Lemonade Stand	Dr. Vakoc	01/01/13	12/31/15	125,000.00
Rita Allen Foundation	Dr. C. Hammell	09/01/12	08/31/17	110,000.00
	Dr. M. Hammell	09/01/14	08/31/19	110,000.00 *
American Cancer Society	Dr. Trotman	07/01/14	06/30/18	198,000.00 *
Anonymous	Dr. Tuveson	12/01/14	11/30/16	250,000.00 *
Asothera Pharmaceuticals, Inc.	Dr. Krainer	10/24/14	04/30/15	440,997.00 *
Caroline E. Bassett	Dr. Tonks	05/19/14	05/18/15	5,000.00 *
Boehringer Ingelheim RCV GmbH & Co KG	Dr. Vakoc	11/01/13	10/31/15	228,196.00
Brain Research Foundation	Dr. Zador	01/01/14	12/31/15	75,000.00 *
The Breast Cancer Research Foundation	Drs. Wigler/Hicks	10/01/14	09/30/15	250,000.00 *
Burroughs Wellcome Fund	Dr. Furukawa	06/23/14	12/31/15	10,000.00 *
	Dr. Vakoc	09/01/11	08/31/16	151,250.00
Cedar Hill Foundation	Dr. Fearon	11/25/14	11/24/15	60,000.00 *
Columbia University/The Simons Foundation	Dr. Krainer	09/01/14	08/31/15	49,908.00 *
Dart Neuroscience LLC	Dr. Dubnau	12/01/13	11/20/14	250,000.00
Families of S.M.A.	Dr. Hua	07/01/13	06/30/15	75,000.00
The Clark Gillies Foundation	Dr. Vakoc	09/16/14	09/15/15	50,000.00 *
Global Prostate Cancer Research Foundation	Dr. Wigler	09/01/13	08/31/15	25,000.00
Glen Cove C.A.R.E.S., Inc.	Dr. Trotman	02/01/14	01/31/15	5,000.00 *
Halozyme, Inc.	Dr. Tuveson	05/20/14	05/19/16	200,000.00 *
The Irving A. Hansen Memorial Fund	Dr. Tonks	08/01/14	07/31/15	25,000.00 *
Hearts for Cancer	Dr. Diermeier	12/04/14	12/03/15	9,713.33 *
Robert and Cindy Higginson	Dr. Mills	01/15/14	01/16/15	1,000.00 *
F. Hoffman-LaRoche Ltd	Dr. Krainer	05/06/13	05/05/15	50,860.87
Howard Hughes Medical Institute - Gordon and Betty Moore Foundation	Dr. Martienssen	12/01/11	11/30/16	333,333.00
IC-MedTech Corporation	Dr. Trotman	03/12/14	10/05/15	63,658.50 *
The Esther A. & Joseph Klingenstein Fund, Inc.	Dr. Churchland	07/01/14	06/30/17	75,000.00 *
The Susan G. Komen Breast Cancer Foundation, Inc.	Dr. Hicks	10/22/13	10/21/17	254,107.00
Mara and Thomas Lehrman Charitable Fund	Dr. Tollkuhn	01/15/14	01/14/15	50,000.00 *
The Lustgarten Foundation	Dr. Egeblad	01/01/13	12/31/15	163,152.00
	Dr. Fearon	07/01/14	06/30/19	1,000,000.00 *
	Drs. Tuveson/Pappin	09/01/12	08/30/17	1,000,000.00
	Dr. Tuveson	01/01/14	01/01/15	140,357.00 *
	Dr. Tuveson	06/15/12	06/30/17	50,000.00
	Dr. Tuveson	10/16/14	10/15/15	50,000.00 *
Carol Marcincuk Fund	Dr. Tonks	09/25/14	09/24/15	9,585.00 *
The G. Harold and Leila Y. Mathers Charitable Foundation	Dr. Mitra	01/01/14	12/31/16	330,000.00 *
Breast Cancer Awareness Day in Memory of Elizabeth McFarland	Dr. Wigler	01/01/14	12/31/15	81,592.00 *

<sup>1</sup>Includes direct and indirect costs

\*New or competing renewal grants awarded in 2014

<i>Grantor</i>	<i>Program/Principal Investigator</i>	<i>Duration of Grant</i>		<i>2014 Funding<sup>1</sup></i>
The McKnight Endowment Fund for Neuroscience	Dr. Churchland	07/01/12	06/30/15	\$ 75,000.00
	Dr. Kepecs	02/01/14	01/31/17	100,000.00 *
The John Merck Fund	Dr. Churchland	06/01/11	05/31/15	75,000.00
The Don Monti Memorial Research Foundation	Drs. Stillman/Tonks	03/01/08	02/28/16	500,000.00
Louis Morin Charitable Trust	Dr. Tuveson	12/01/14	11/30/15	115,000.00 *
State of New York Department of Health/ Empire State Stem Cell Fund	Dr. Enikolopov	06/01/14	05/31/17	358,948.00 *
Nimbus Discovery, Inc.	Dr. Tuveson	12/15/14	12/14/15	75,000.00 *
In Memory of Jack O'Brien	Dr. Hicks	08/01/14	07/31/15	10,010.00 *
Omicia, Inc.	Dr. G. Lyon	11/17/14	11/16/15	10,000.00 *
Pancreatic Cancer UK	Dr. Tuveson	10/23/12	10/22/15	65,000.00
Mr. Joe Paternoster	Dr. Vakoc	09/16/14	09/15/15	10,000.00 *
Mr. Paul J. Paternoster	Dr. Vakoc	09/16/14	09/15/15	10,000.00 *
Pershing Square Sohn Cancer Research Alliance	Dr. Trotman	07/01/14	06/30/17	200,000.00 *
The Pew Charitable Trusts	Dr. Albeanu	08/01/12	07/31/17	60,000.00
	Dr. Churchland	08/01/14	07/31/19	60,000.00 *
	Dr. Hannon	01/01/11	12/31/15	355,063.00
Phi Beta Psi Sorority	Dr. Egeblad	08/15/14	08/14/15	54,130.00 *
Pioneer Hi-Bred International, Inc.	Dr. Ware	03/15/12	06/30/15	54,873.00
The Hazen Polsky Foundation	Dr. Vakoc	01/01/12	12/31/15	50,000.00
Precision Medicine Research Associates, LLC	Dr. Tuveson	03/20/14	03/19/15	157,836.00 *
	Dr. Tuveson	12/01/14	11/30/15	268,390.00 *
Prostate Cancer Foundation	Dr. Hicks	09/01/13	08/31/15	184,325.00
Christina Renna Foundation Inc.	Dr. Van Aelst	02/01/14	01/31/15	20,000.00 *
Marie Robertson Memorial Fund	CSHL Neuroscience Program Support	01/01/14	12/31/14	34,375.00 *
Diane Emdin Sachs Memorial Fund	Dr. Sordella	09/01/12	08/31/15	1,233.00
Eleanor Schwartz Charitable Foundation	Dr. Tuveson	08/01/14	07/31/17	200,000.00 *
Mr. Doug Shorestein	Dr. Tuveson	03/01/13	02/28/15	25,000.00
The Simons Foundation	Dr. Churchland	07/01/14	06/30/17	200,000.00 *
	Dr. Huang	08/01/14	07/31/17	82,473.00 *
	Dr. Mills	08/01/14	07/31/15	212,100.00 *
	Dr. Shea	01/01/14	12/31/14	81,970.00 *
The Sontag Foundation	Dr. Zheng	10/01/13	09/30/17	150,000.00
Starr Cancer Consortium	Drs. Hannon/Atwal	01/01/14	12/31/15	396,000.00 *
	Dr. Tuveson	01/01/14	12/31/15	192,000.00 *
The Swartz Foundation	Dr. Zador	01/01/14	12/31/14	75,000.00 *
Friends of the TJ Foundation Inc.	Dr. Vakoc	09/16/14	09/15/15	50,000.00 *
Joan & Sanford I. Weill Medical College	Dr. Fearon	07/01/14	06/30/19	165,061.90 *
Whitehall Foundation, Inc.	Dr. Albeanu	01/01/12	12/31/14	75,000.00
	Dr. Shea	09/01/11	08/31/15	70,740.00
Women's Partnership for Science	Dr. Joshua-Tor	01/01/14	12/31/14	174,631.64 *
	Dr. Timmermans	01/01/14	12/31/14	162,659.44 *
Elisabeth R. Woods Foundation Inc.	Dr. Sordella	09/27/13	09/26/15	9,000.00
The WTFC Organization, Inc.	Dr. Tonks	07/15/14	07/14/15	5,500.00 *
The Bradley Zankel Foundation, Inc.	Dr. Zheng	01/22/14	01/21/15	15,000.00 *
<i>Fellowship Support</i>				
Agricultural Biotechnology Center	Dr. Benkovics	05/01/13	07/04/15	18,862.00
American Cancer Society	Dr. Ardito	07/01/13	06/30/16	50,000.00

<sup>1</sup>Includes direct and indirect costs

\*New or competing renewal grants awarded in 2014

<i>Grantor</i>	<i>Program/Principal Investigator</i>	<i>Duration of Grant</i>		<i>2014 Funding<sup>1</sup></i>
Boehringer Ingelheim Fonds	M. Delas Vives	08/01/13	07/31/15	\$ 60,000.00
	A. Hohmann	08/01/13	07/31/15	60,000.00
	J. Kebschull	08/01/13	07/31/15	60,000.00
Brain and Behavior Research Foundation	Dr. Ahrens	01/15/14	01/14/16	30,000.00 *
	Dr. Horev	01/15/14	01/14/16	30,000.00 *
	Dr. Y. Kim	01/15/14	01/14/16	30,000.00 *
	Dr. K. Krishnan	01/15/14	01/14/16	30,000.00 *
	Dr. J. Lu	01/15/13	01/14/15	30,000.00
Terri Brodeur Breast Cancer Fund	Dr. Paul	01/15/13	01/14/15	30,000.00
	Dr. Anczukow	01/01/13	12/31/14	50,000.00
Marie Curie International Outgoing Fellowships for Career Development	Dr. Hangya	01/01/13	12/31/14	100,718.98
	Dr. Benkovic	05/01/13	04/04/15	84,460.00
European Molecular Biology Organization	Dr. Claeys	07/01/14	04/30/16	43,765.00 *
Genentech Foundation	Watson School of Biological Sciences	10/01/12	09/30/15	58,751.00
German Research Foundation	Dr. Knauer	05/01/14	04/30/16	27,000.00 *
	Dr. Kuettner	09/01/14	08/31/16	12,160.00 *
Lola A. Goldring	Dr. Stillman	10/01/14	09/30/15	85,000.00 *
Hope Funds for Cancer Research	Dr. Knott	07/01/12	06/30/15	50,500.00
Human Frontier Science Program	Dr. Elyada	04/01/14	03/31/17	61,080.00 *
	Dr. Huilgol	04/01/14	03/31/17	51,720.00 *
	Dr. Livneh	01/01/14	12/31/16	52,500.00 *
	Dr. Malinauskas	04/01/14	03/31/17	51,720.00 *
	Dr. He	02/01/13	08/15/14	27,083.61
International Rett Syndrome Foundation	Dr. Ballouz	06/04/14	06/03/15	900.00 *
International Society of Computational Biology	Dr. Verleyen	06/04/14	06/03/15	900.00 *
Annette Kade Charitable Trust	Watson School of Biological Sciences	10/15/14	10/14/15	25,000.00 *
Mr. and Mrs. John Phelan	Watson School of Biological Sciences	09/01/10	08/31/15	100,000.00
Damon Runyon Cancer Research Foundation	Dr. Chio	10/01/13	09/30/17	52,000.00
	Dr. Sabin	01/01/12	12/31/15	52,000.00
Sass Foundation for Medical Research Inc.	Dr. Roe	06/01/13	05/31/15	50,000.00
Lauri Strauss Leukemia Foundation	Dr. Roe	12/01/14	11/30/15	12,500.00 *
The Swartz Foundation	Drs. Albeanu/Anselmi	01/01/14	12/31/14	55,000.00 *
	Drs. Churchland/Kaufman	01/01/14	12/31/14	55,000.00 *
	Drs. Koulakov/Ferrete	01/01/14	12/31/14	55,000.00 *
	Dr. Zador	01/01/14	12/31/14	15,000.00 *
Umea University	Dr. Olund	12/01/12	11/30/14	12,520.00
Vasterbottens Lans Landsting/Swedish Research Council	Dr. Olund	01/01/14	12/31/14	22,140.00 *
Vasterbottens Lans Landsting	Dr. Olund	01/01/14	12/31/14	19,581.85 *
<i>Training Support</i>				
New York State Department of Economic Development	Dr. Mills, Postdoctoral Cancer Research Training	11/01/12	10/31/17	144,301.00
The Lita Annenberg Hazen Foundation	Undergraduate Research Support	05/01/08	04/30/18	10,000.00
William Townsend Porter Foundation	Undergraduate Research Support	04/01/14	03/31/15	13,500.00 *
University of Maryland	Undergraduate Research Support	06/01/14	05/31/15	8,700.00 *

<sup>1</sup>Includes direct and indirect costs

\*New or competing renewal grants awarded in 2014

<i>Grantor</i>	<i>Program/Principal Investigator</i>	<i>Duration of Grant</i>		<i>2014 Funding<sup>1</sup></i>	
<i>Meeting Support</i>					
Anonymous	Plasmids: History & Biology	07/08/14	07/07/15	\$ 15,000.00	*
Bethyl Laboratories, Inc.	The Cell Cycle	05/20/14	05/19/15	13,000.00	*
Calico Life Sciences LLC	Molecular Genetics of Aging	07/24/14	07/23/15	5,000.00	*
The Company of Biologists LTD	Avian Model Systems	02/01/14	01/31/15	6,448.40	*
Dr. Gerald J. Cordani	Meeting Support	03/26/14	03/25/15	1,000.00	*
Forest Research Institute, Inc.	79th CSHL Symposium on Quantitative Biology Cognition	04/24/14	04/23/15	50,000.00	*
Gilead Sciences Inc.	Retroviruses	05/01/14	04/30/15	12,500.00	*
GlaxoSmithKline, LLC	Molecular Genetics of Aging	11/11/13	11/19/14	5,000.00	
Illumina, Inc.	Avian Model Systems	02/01/14	01/31/15	3,000.00	*
The Lalor Foundation, Inc.	Germ Cells	08/06/14	08/05/15	8,000.00	*
Life Sciences Foundation	Plasmids: History & Biology	07/08/14	07/07/15	2,500.00	*
March of Dimes Foundation	Germ Cells	05/28/14	05/27/15	7,500.00	*
Medimmune, LLC	Plasmids: History & Biology	07/08/14	07/07/15	2,500.00	*
Merck Sharp & Dohme Corp.	Neurodegenerative Diseases: Biology & Therapeutics	10/15/14	10/14/15	5,000.00	*
Monsanto Company	Plasmids: History & Biology	07/08/14	07/07/15	2,500.00	*
Mr. Richard P. Novick	Plasmids: History & Biology	07/08/14	07/07/15	5,000.00	*
Regeneron Pharmaceuticals Inc	Personal Genomes: Discovery, Treatment & Outcomes	10/15/14	10/14/15	10,000.00	*
Society for Developmental Biology	Avian Model Systems	02/25/14	02/24/15	4,000.00	*
Society for General Microbiology	Plasmids: History & Biology	07/08/14	07/07/15	10,155.54	*
The Swartz Foundation	79th CSHL Symposium on Quantitative Biology Cognition	05/01/14	04/30/15	25,000.00	*
University of Arizona	Avian Model Systems	02/01/14	01/31/15	1,000.00	*
<i>Course Support</i>					
American Brain Tumor Association	Brain Tumors	02/01/14	01/31/15	25,000.00	*
Howard Hughes Medical Institute	Course Support	08/01/11	07/31/15	600,000.00	
Illumina, Inc.	Advanced Sequencing Technologies and Applications	11/01/13	10/31/14	10,000.00	
Nancy Lurie Marks Foundation	Genetics and Neurobiology of Language	07/01/12	06/30/16	50,000.00	
<i>Library Support</i>					
The Ellen Brenner Memorial Fund		12/15/14	12/14/15	2,000.00	*
Howard Hughes Medical Institute		10/01/14	09/30/15	110,812.00	*
The New York State Education Department		07/01/14	06/30/15	3,804.00	*
<i>Library Meeting Support</i>					
Lexogen GmbH	Messenger RNA: From Discovery to Synthesis and Regulation in Bacteria and Eukaryotes	03/18/14	03/17/15	778.20	*
Life Sciences Foundation	Messenger RNA: From Discovery to Synthesis and Regulation in Bacteria and Eukaryotes	03/18/14	03/17/15	2,500.00	*
New England Biolabs, Inc.	Messenger RNA: From Discovery to Synthesis and Regulation in Bacteria and Eukaryotes	03/18/14	03/17/15	10,000.00	*
The Rockefeller University	Messenger RNA: From Discovery to Synthesis and Regulation in Bacteria and Eukaryotes	03/18/14	03/17/15	40,000.00	*

<sup>1</sup>Includes direct and indirect costs

\*New or competing renewal grants awarded in 2014

## DNA LEARNING CENTER GRANTS

Grantor	Program	Duration of Grant	2014 Funding <sup>†</sup>
<b>FEDERAL GRANTS</b>			
National Institutes of Health	<i>Barcode Long Island</i>	7/14–3/19	48,628
National Science Foundation	<i>Infrastructure and Training to Bring Next-Generation Sequence (NGS) Analysis into Undergraduate Education</i>	9/13–8/16	198,670
National Science Foundation	Gramene, Cereal Comparative Genomics	6/12–9/14	28,460
National Science Foundation, University of Arizona	<i>iPlant Collaborative: Cyberinfrastructure for the Life Sciences</i>	9/13–8/18	873,877
National Science Foundation	Advanced Technology Education (ATE) <i>Genomic Approaches in BioSciences</i>	4/11–3/15	239,397
<b>NONFEDERAL GRANTS</b>			
Beijing No. 166 High School	Chinese Collaboration Agreement	5/14–4/15	16,780
Alfred P. Sloan Foundation	<i>DNA Center NYC start-up</i>	12/13–11/16	138,993
Alfred P. Sloan Foundation	<i>High-Performance Computing in Undergraduate Biology Education</i>	5/14–9/14	11,207
National Grid Foundation	Scholarships for Minority and Underserved Students in the Central Islip Union Free School District	10/14–10/15	15,000
Pinkerton Foundation	<i>Urban Barcode Research Program</i>	1/13–5/15	156,535
The Bank of America Charitable Foundation	<i>Genetics and Biotechnology Lab</i>	11/14–11/15	20,000
William Townsend Porter Foundation	<i>Harlem DNA Lab for Underprivileged Students</i>	4/14–3/15	13,500

The following schools and school districts each contributed \$1,000 or more for participation in the *Curriculum Study* program:

Bellmore–Merrick Central High School District	1,500	Levittown Union Free School District	1,500
East Meadow Union Free School District	1,500	Long Beach City School District	1,500
East Williston Union Free School District	2,750	North Shore Central School District	1,500
Elwood Union Free School District	1,500	North Shore Hebrew Academy	1,500
Fordham Preparatory School	1,500	Oceanside Union Free School District	1,500
Garden City Union Free School District	1,500	Oyster Bay–East Norwich Central School District	1,500
Great Neck Union Free School District	1,500	Plainedge Union Free School District	1,500
Green Vale School	1,500	Plainview–Old Bethpage Central School District	1,500
Half Hollow Schools Central School District	1,500	Portledge School	1,500
Harborfields Central School District	1,500	Port Washington Union Free School District	1,500
Herricks Union Free School District	1,500	Ramaz Upper School	1,500
Huntington Union Free School District	1,500	Roslyn Union Free School District	1,500
Island Trees Union Free School District	3,000	Syosset Central School District	1,500
Jericho Union Free School District	3,000		

The following schools and school districts each contributed \$1,000 or more for participation in the *Genetics as a Model for Whole Learning* program:

Bais Yaakov Academy for Girls	1,200	Locust Valley Central School District	9,582
Bellmore Union Free School District	2,400	Merrick Union Free School District	3,900
Bellmore–Merrick Union Free School District	4,200	M.S. 447, Brooklyn	1,800
Berkeley Carroll School	1,300	North Bellmore Union Free School District	2,600
Cold Spring Harbor Central School District	12,600	North Shore Central School District	1,050
East Meadow Union Free School District	3,828	Oceanside Union Free School District	1,625
East Williston Union Free School District	2,900	Oyster Bay–East Norwich Central School District	3,175
Elwood Union Free School District	3,725	Portledge School	2,400
Floral Park–Bellerose Union Free School District	7,500	Port Washington Union Free School District	10,300
Franklin Lakes Public Schools, NJ	1,300	Redeemer Lutheran School	1,300
Friends Academy	3,300	Rockville Centre Union Free School District	6,240
Friends Seminary of New York	1,100	Roslyn Union Free School District	5,250
Garden City Union Free School District	11,655	Saint Dominic Elementary School	4,200
Great Neck Union Free School District	8,250	Saint Joseph School	2,400
Half Hollow Hills Union Free School District	1,284	Saint Patrick School	5,000
Herricks Union Free School District	2,275	Scarsdale Union Free School District	6,300
Hicksville Public Schools	1,400	Smithtown Union Free School District	1,680
Hofstra STEP	1,500	Syosset Union Free School District	38,800
Holy Child Academy	1,600	Three Village Central School District	2,800
Huntington Union Free School District	8,275	Valley Stream Union Free School District 13	1,050
Jericho Union Free School District	9,450	YBH of Passaic	1,400
Lawrence Union Free School District	2,100	Yeshiva Darchei Torah	2,775
Lindenhurst Union Free School District	1,000		

<sup>†</sup>Includes direct and indirect costs.



**BANBURY CENTER GRANTS**

<i>Grantor</i>	<i>Program</i>	<i>Duration of Grant</i>	<i>2014 Funding</i>
<b>FEDERAL SUPPORT</b>			
National Institute of Mental Health	Brain Camp VI	2014	\$ 26,500
National Science Foundation	High Performance Computing in Undergraduate Biology Education: Scanning the Landscape	2014	29,000
<b>NONFEDERAL SUPPORT</b>			
American College of Medical Genetics and Genomics (ACMG)	Genetics Training in the Genomic Era	2014	33,920
Boehringer Ingelheim Fonds	Science—Get It Across!	2014	60,967
Bring Change 2 Mind	The College Toolbox Project: Eliminating Stigma	2014	7,500
Cold Spring Harbor Laboratory Corporate Sponsor Program	Interpreting Personal Genomes: How Are We to Set Appropriate Statistical Standards for Identifying Pathogenic Variants?	2014	52,362
Cold Spring Harbor Laboratory Corporate Sponsor Program	Epigenetics and Agriculture	2014	59,398
Cold Spring Harbor Laboratory Office of Development	Defeating Ovarian Cancer	2014	10,000
Dana Foundation	Lewy Body Dementia: Current Status, Future Directions	2014	10,000
Friends of TJ Foundation	Rhabdomyosarcoma: A Critical Review of Research and What It Means for Developing Therapies	2014	20,000
Nelson DeMille	Lewy Body Dementia: Current Status, Future Directions	2014	5,000
Jonathan Gray	Defeating Ovarian Cancer	2014	25,000
Indiana University Foundation	The College Toolbox Project: Eliminating Stigma	2014	7,500
Indiana University	The College Toolbox Project: Eliminating Stigma	2014	6,360
The Margaret Clark Morgan Foundation	The College Toolbox Project: Eliminating Stigma	2014	2,120
The Mayday Fund	Genetics of Pain and Pain Inhibition	2014	56,250
Oliver Grace Cancer Fund	The Immune System and Cancer	2014	40,852
Oliver Grace Cancer Fund	ROS in Biology and Cancer	2014	53,801
Oliver Grace Fund	Lewy Body Dementia: Current Status, Future Directions	2014	20,700
Michelle Paternoster Foundation for Sarcoma Research	Rhabdomyosarcoma: A Critical Review of Research and What It Means for Developing Therapies	2014	20,000
Mr. & Mrs. Howard Phipps	Interdisciplinary Symposium on Creativity	2014	15,000
Prothema Biosciences Inc.	Lewy Body Dementia: Current Status, Future Directions	2014	10,000
The Daniel & Joanna S. Rose Fund	Interdisciplinary Symposium on Creativity	2014	15,000
Alfred P. Sloan Foundation	High Performance Computing in Undergraduate Biology Education: Scanning the Landscape	2014	15,500
Society of Biological Psychiatry	Brain Camp VI	2014	710
The Swartz Foundation	Connections and Communications in the Brain	2014	44,337

# CORPORATE SPONSOR PROGRAM FOR MEETINGS SUPPORT

---

The Corporate Sponsor Program continues to provide critical funding for the vigorous meetings program held at Cold Spring Harbor Laboratory, whether at Grace Auditorium on the main Laboratory campus or at the Banbury Center. Without the strong foundation provided by the Program, we could neither plan with confidence for the year's meetings nor introduce new and unusual topics.

Raising funds for this program continues to be challenging even as the economy improves. We are thus especially grateful to the companies that joined us in 2014.

The members of the Program receive special privileges in acknowledgment of their contributions. We waive all on-site fees for eight representatives of the Corporate Sponsors at our meetings. Three scientists from Sponsors may attend meetings at the Banbury Center, where attendance is otherwise only by invitation of the organizers. Member companies also receive gratis copies of Cold Spring Harbor Laboratory Press publications, including the journals *Genes & Development*, *Genome Research*, and *RNA*.

We acknowledge our Sponsors in all relevant publications, including the books of abstracts given to each of the 7,000 participants who come to the meetings each year. The names of the sponsoring companies are listed on the poster describing the meetings, and this is mailed to approximately 17,000 scientists throughout the world. In addition, the companies are listed on the Cold Spring Harbor Laboratory website on the Meetings Office and Banbury Center pages. Members in 2014 were as follows.

## CORPORATE SPONSORS

Agilent Technologies, Inc.  
Bristol-Myers Squibb Company  
Genentech, Inc.  
Life Technologies  
(part of Thermo Fisher Scientific)  
New England BioLabs, Inc.

## PLANT CORPORATE SPONSORS

Monsanto Company

## FOUNDATIONS

Hudson-Alpha Institute for Biotechnology

## DEVELOPMENT

---

In celebration of the 125th anniversary of Cold Spring Harbor Laboratory (CSHL), we reflect on the many people who have shaped this great institution and helped it become one of the most revered scientific organizations in the world. As Vice President for Development, I have had the privilege to work closely with our dedicated donors and friends, including Marilyn and Jim Simons, Cathy Soref, Ginny Knott, Mary Lindsay, Kristina Davison, and so many others.

The Laboratory not only has a long history of world-class science, it also has a legacy of involvement among on the part of and their families. One example is the Davenport/Spire family. In 1969, newly appointed Director of Cold Spring Harbor Laboratory Jim Watson expressed his hope to find an “angel” donor to support CSHL’s new cancer research program in an article in the *Long Island Press*. John and Edna Davenport responded to his plea, donating more than \$100,000 of Pfizer stock for building an extension to James Laboratory. This extraordinary gift provided critical support for the tumor virus program, enabling the Laboratory to establish itself as a world leader in cancer research. Over the years, John and Edna’s family has stayed involved, supporting the Laboratory’s education and research programs. In 2014, through a friendship with John’s grandson Bill, the Laboratory received a \$3 million pledge to name “the Edna and John Davenport Genetic Screening Facility,” an invaluable shared resource critical to many of our scientists’ work.

In 2014, we raised almost \$6.7 million in unrestricted funding, largely due to the overwhelming success of the Double Helix Medal Dinner. This year, we proudly honored Matt Meselson, Andrew Solomon, and Marlo Thomas for their uncompromising passion and hard work toward improving human health and changing the world for the better.

This is truly a year to celebrate both Cold Spring Harbor Laboratory and our generous supporters. Thank you for always believing in the work we do and partnering with us on this incredible journey toward a healthier world.

*Charles V. Prizzi, Vice President for Development and Community Relations*



Edith Seligson, Kate Friedman, and Kristina Perkin Davison with Alan Alda at the Double Helix Medal Dinner



Bill and Heather Spire at CSHL’s President’s Council reception

## Cold Spring Harbor Laboratory Association

The Cold Spring Harbor Laboratory Association (CSHLA) continues to be instrumental in its support of CSHL. Members continue to generously host outreach events and introduce friends and colleagues to the Laboratory to raise awareness of CSHL's cutting-edge research and education programs. The year 2014 saw more unrestricted funding raised than ever before. With events spearheaded by directors of the Association with new president Frank O'Keefe at the helm, more than \$7 million were raised. Some of this important funding resulted from events that were supported by the Association, its directors, community members, major donors, foundations, and corporate sponsors, including the 13th annual Women's Partnership for Science luncheon, the 21st annual golf tournament at Piping Rock Club, and the 9th annual Double Helix Awards dinner honoring Matt Meselson, Andrew Solomon, and Marlo Thomas for their efforts in bringing awareness to disease research.

### Directors

#### 2014 Directors

Frank O'Keefe, *President*  
James A. Woods, *Treasurer*  
Virginia Knott, *Secretary*

Elizabeth Ainslie  
Lori Bahnik  
Hans E.R. Bosch  
Edward A. Chernoff  
Susan A. Cohen  
Maria de Lesseps  
Nelson DeMille  
Carolyn Gulotta  
Nancy Lippman Israeli, M.D.  
Joel J. Kassimir, M.D.  
Kassahun Kebede  
Laureen Stanton Knutsen  
Suzanne Leeds  
Jeffrey A. Nagel

Timothy Q. O'Neill  
Olga Pharr  
Whitney F. Posillico  
Charles S. Ryan, Ph.D.  
Caroline Monti Saladino  
Tracey DeMatteis Serko  
Lonnie Shoff  
Hope Geier Smith  
Heather Lane Spehr  
Nancy Szigethy  
Peter S. Tilles  
Lara S. Trafelet  
Marjorie J. van de Stouwe, M.D.

#### Honorary Directors

Mary D. Lindsay  
Anne R. Meier  
Cathy Cyphers Soref  
Pearl F. Staller



2014 CSHLA Directors

## Cold Spring Harbor Laboratory Corporate Advisory Board

The Corporate Advisory Board (CAB) is comprised of prominent business leaders from the tristate community and is a vital source of funding and outreach for Cold Spring Harbor Laboratory. Board members are the driving force behind the Laboratory's annual golf outing at Piping Rock Club, which raises critical unrestricted funding for research and education programs. CAB president Eddie Chernoff graciously chaired his 10th CSHL outing, which honored CSHL Trustee, CAB member, and longtime friend to the Lab, Ed Travaglianti, and more than \$175,000 was raised.

### Corporate Advisory Board

Edward A. Chernoff, MARS, *Chairperson*

Michael Aboff, Aboff's Inc.  
 Paul Amoruso, Oxford & Simpson Realty  
 Rocco S. Barrese, Dilworth & Barrese  
 Edward Blaskey, TD Bank  
 Thomas J. Calabrese, Daniel Gale Sotheby's International Realty  
 Christopher Callaghan, GFI Group  
 Jeffrey Carstens, Capital One  
 Richard A. Catalano, KPMG, LLP  
 Jonathan Connors, Wells Fargo  
 Marian Conway, New York Community Bank Foundation  
 J. Daniel Decker, Eppendorf North America  
 Robert Dickstein, Ph.D., Pall Corporation  
 Kevin Foley, Fidelity  
 Brian Fox, McKinsey & Company  
 Brian Goldstein, M.D.  
 Lawrence Goodman, Curtis, Mallet-Prevost,  
 Colt & Mosle LLP  
 Mary Jane Helenek, Luitpold Pharmaceuticals  
 Diana Hoadley, JP Morgan  
 Richard W. Humann, H2M

Robert Isaksen, Bank of America  
 M. Patricia Janco, Wells Fargo  
 Neil M. Kaufman, Abrams Fensterman  
 John C. Kean III, Kean Development Corporation  
 Norman Kelker, Enzo Life Sciences  
 Robert Keller, National Grid Foundation  
 Amy Koreen, M.D.  
 Laurie J. Landeau, V.M.D.  
 Victoria Meagher  
 Rob Palatnick, DTTC  
 John G. Passarelli, M.D.  
 David Peikon  
 Patricia Petersen, Daniel Gale Sotheby's International Realty  
 Stephen Ross, Ph.D., Nikon Instruments  
 Raju Sarwal, M.D.  
 Nancy Szigethy, NMS Management  
 Edward Travaglianti, TD Bank  
 Andrew Whiteley, Enzo Life Sciences  
 Joseph M. Willen, Advantage Title  
 Hans Zobel



CSHL Board Member and Golf Honoree Ed Travaglianti (*left*) with Paul Amoruso and Chris Giamo



CAB Chairman Eddie Chernoff (*right*) with Bryan McMillen, Pat Petersen, and Brando Casalicchio

## Honor Roll of Donors

### Contributions of \$5 million+ lifetime

Anonymous  
 Estate of Donald E. Axinn  
 The Arnold and Mabel Beckman Foundation  
 Jamie Nicholls and Fran Biondi  
 David Boies, Chairman of Boies, Schiller & Flexner LLP  
 The Breast Cancer Research Foundation  
 Mr. and Mrs. Landon T. Clay  
 The Dana Foundation  
 Mr. and Mrs. Norris W. Darrell  
 Dart Neurosciences LLC  
 Estate of Kathryn W. Davis  
 DeMatteis Family Foundation  
 Mr. and Mrs. Charles F. Dolan—Dolan Family Foundation  
 The William Stamps Farish Fund  
 Estate of Robert B. Gardner, Jr.\*  
 Gen\*NY\*Sis grant from New York State Dormitory Authority  
 Kate Medina Guthart and Leo A. Guthart  
 Howard Hughes Medical Institute  
 W.M. Keck Foundation  
 Mr. and Mrs. David H. Koch  
 Laurie J. Landeau, V.M.D. and the Laurie J. Landeau Foundation  
 Mr. and Mrs. Robert D. Lindsay and Family  
 Estate of Elisabeth Sloan Livingston  
 The Lustgarten Foundation  
 Lucille P. Markey Charitable Trust  
 Nancy Abeles Marks  
 The Don Monti Memorial Research Foundation  
 New York State Empire State Development Corporation  
 Pershing Square Foundation  
 The Honorable and Mrs. Leon Polsky—Hazen Polsky Foundation  
 The Quick Family  
 The Robertson Family—Banbury Fund  
 The Seraph Foundation  
 Drs. Marilyn and James Simons  
 The Simons Foundation  
 Alfred P. Sloan Foundation  
 St. Giles Foundation  
 Theodore R. Stanley—Stanley Medical Research Institute  
 The Starr Foundation  
 Thompson Family Foundation  
 Dr. and Mrs. James D. Watson  
 Mr. and Mrs. Henry Wendt III  
 The Roy J. Zuckerberg Family Foundation

### Contributions of \$1 million+ lifetime or \$100,000 in 2014

Paul G. Allen Foundation  
 Angel Foundation

Anonymous (4)  
 Gonzalo Río Arronte Foundation  
 Lalit and Kavita Bahl  
 BioBay of Suzhou Venture Capital Group  
 Booth Ferris Foundation  
 Bristol-Myers Squibb Company  
 Burroughs Wellcome Fund  
 The Mary K. Chapman Trust  
 Cutler Family Education and Animal Welfare Fund  
 Michel David-Weill  
 Ira W. DeCamp Foundation  
 Edward P. Evans Foundation  
 Charitable Lead Annuity Trust—Will of Louis Feil  
 Samuel Freeman Charitable Trust  
 Coleman Fung Foundation  
 Robert L. Garland\*  
 Mr. and Mrs. Robert A. Gay  
 Genentech, Inc.  
 Kenyon Gillespie  
 Jacob Goldfield  
 Lola Goldring  
 The Florence Gould Foundation  
 The Grace Family  
 The Harrison Foundation and The Mary Anderson Harrison Foundation  
 Thomas Hartman Foundation for Parkinson's Research  
 Janet Strauss and Jeff Hawkins  
 Jo-Ellen and Ira Hazan  
 William Randolph Hearst Foundation  
 The Karches Foundation  
 Walter B. Kissinger  
 Mr. and Mrs. John Klingenstein—The Esther A. and Joseph Klingenstein Fund, Inc.  
 Elise H. Knight  
 Mr. and Mrs. David M. Knott

The Forrest C. and Frances H. Lattner Foundation  
 Betsy and Bryan H. Lawrence  
 Mara and Thomas Lehrman  
 The Lehrman Institute  
 Mr. and Mrs. Stephen M. Lessing  
 Mary D. Lindsay  
 The Litchfield Charitable Trust  
 Louis Morin Charitable Trust  
 Lisa Lourie  
 Dr. Robert W. Lourie  
 Mr. and Mrs. David L. Luke III\*  
 Josiah Macy, Jr. Foundation  
 Manhasset Women's Coalition Against Breast Cancer  
 The G. Harold and Leila Y. Mathers Charitable Foundation  
 Gillian and Eduardo Mestre  
 Mr. and Mrs. William R. Miller  
 Dr. and Mrs. Howard L. Morgan  
 Estate of Dr. Thomas C. Nelson  
 OSI Pharmaceuticals, LLC  
 David and Lucile Packard Foundation  
 The Pew Charitable Trusts  
 Amy and John Phelan  
 William and Maude Pritchard Charitable Trust  
 Fannie E. Rippel Foundation  
 The Robertson Foundation  
 Mr. and Mrs. David M. Rubenstein  
 Raymond and Beverly Sackler Foundation  
 Eleanor Schwartz Charitable Foundation  
 Edith and Alan Seligson  
 Jeanne and Herbert J. Siegel  
 Mr. and Mrs. Howard Solomon  
 Mr. and Mrs. Erwin P. Staller  
 Dr. and Mrs. James M. Stone  
 Dr. Jerome Swartz—The Swartz Foundation



CSHL President Bruce Stillman with the 2014 Double Helix Medals honorees Matt Meselson, Marlo Thomas, and Andrew Solomon



Masthead Cove Yacht Club presents a check to CSHL scientist Dr. Nick Tonks for the Carol Marcincuk Fund

Waclaw T. Szybalski, D.Sc.  
Mr. and Mrs. Paul J. Taubman  
Whitehall Foundation, Inc.

**Contributions of \$50,000+**

The Achelis Foundation  
Anonymous  
Cedar Hill Foundation  
Mr. and Mrs. Edward A. Chernoff—  
MARS  
Ellen and Casey Cogut  
Claire Friedlander Family Foundation  
Genentech Foundation  
The Clark Gillies Foundation  
Joni Gladowsky Breast Cancer Foundation  
Mr. and Mrs. Jeffrey E. Kelter  
Breast Cancer Awareness Day in Memory  
of Elizabeth McFarland  
Mr. and Mrs. Douglas P. Morris  
Nimbus Discovery, Inc.  
The Sass Foundation for Medical  
Research  
Hope Geier Smith and Theodore  
Smith III  
Swim Across America Nassau/Suffolk  
The Friends of T.J. Foundation  
Wodecroft Foundation

**Contributions of \$25,000+**

Annette Kade Charitable Trust  
Ashley and Frank O'Keefe  
Dr. and Mrs. Mark S. Ptashne—The Joe  
Lewis Jefferson Foundation, Inc.  
Dr. and Mrs. Stuart T. Weisbrod  
Dr. and Mrs. Thomas P. Maniatis  
Drs. Maria and Charles Ryan—Forest  
Laboratories, Inc.  
Drs. Nancy and Ron Israeli  
Francis Goelet Charitable Lead Trusts  
Hearts for Cancer  
Ike, Molly and Steven Elias Foundation  
John and Laura Desmarais  
Jon and Mindy Gray Family Foundation

Life Technologies  
Long Island Real Estate Group, Inc.  
Loren Eng and Dinakar Singh  
Louise Parent and John Casaly  
Mr. and Mrs. Claudio Del Vecchio  
Mr. and Mrs. Douglas W. Shorenstein  
Mr. and Mrs. Francis J. Ingrassia  
Mr. and Mrs. Geoffrey de Lesseps—  
Alpha/Omega Charitable  
Foundation  
Mr. and Mrs. Harry Slatkin  
Mr. and Mrs. James F. Stebbins  
Mr. and Mrs. Louis Moore Bacon  
Mr. and Mrs. Oliver R. Grace, Jr.—  
OBX, Inc.  
Mr. and Mrs. Ronald A. Gottlieb  
Mr. and Mrs. Thomas A. Saunders III  
Pall Corporation  
Pamela Hurst-Della Pietra and Stephen  
Della Pietra  
Paul J.R. Paternoster  
Pfizer, Inc.

Pien and Hans Bosch  
Susan T. Harris  
Susan Wasserstein and George Sard  
The Bahnik Foundation  
The Carson Family Charitable Trust  
The Irving A. Hansen Memorial  
Foundation  
The Marc Haas Foundation  
The Morningside Foundation  
The Rockefeller University Banbury  
Meeting  
Thomas C. Quick  
Wilhelmus Johannes Van der Goes  
William Townsend Porter Foundation

**Contributions of \$10,000+**

Elizabeth and Lee Ainslie  
The Bank of America Charitable  
Foundation  
Caroline E. Bassett  
Blavatnik Family Foundation  
Broad Hollow Bioscience Park, Inc.  
David P. Brown  
Mr. and Mrs. George W. Carmany III  
Dr. and Mrs. Bayard D. Clarkson  
Dr. Stanley N. Cohen—Banbury  
Meeting  
Susan and Glenn Cohen  
Samara P. Cohen  
Patricia Petersen—Daniel Gale Sotheby's  
International Realty  
Barbara Amonson and Vincent Della  
Pietra  
Sandy and Nelson DeMille  
Miriam and David Donoho  
Eppendorf North America  
F.A.C.T. Breast Cancer Foundation  
The Fascitelli Family Foundation  
Fortunato Sons Contracting, Inc.  
Mr. and Mrs. Stephen C. Freidheim



Nelson DeMille, Bruce Stillman, Ed Travaglianti, and Sandy DeMille at CSHL's annual Golf Tournament



Barbara Buckley of Manhasset Women's Coalition Against Breast Cancer presents a check to Dr. Sarah Diermeier for her breast cancer research. Also pictured are CSHL President Dr. Bruce Stillman (*left*) and Director of Research Dr. David Spector (*right*)

Dr. Alexander Gann  
 Goddard Family Charitable Foundation  
 Kappa Alpha Order-University of Maryland  
 Dr. and Mrs. Joel Kassimir  
 Thomas L. Kempner, Jr. and Katheryn C. Patterson  
 Andrea B. and Peter D. Klein  
 Lauren and Ragnar M. Knutsen  
 Mitchell and Karen Kuflik Charitable Foundation  
 Long Island 2 Day Walk to Fight Breast Cancer  
 Bridget and John Macaskill  
 Robert L. Marcus  
 Edward E. Matthews  
 Marcia Kramer Mayer  
 Mirus Bio LLC  
 National Grid Foundation  
 North Shore-LIJ Health System  
 Susanne C. Olin  
 Mr. and Mrs. George D. O'Neill  
 Wendy O'Neill  
 Robert and Lauryl Palatnick  
 Joe Paternoster  
 Michael Peluso  
 Mrs. Richardson Pratt, Jr. and Family—Spionkop Charitable Trust  
 Mr. and Mrs. Preston V. Pumphrey  
 Rauch Foundation, Inc.  
 Mr. and Mrs. John R. Reese  
 Christina Renna Foundation  
 Lonnie and Michael Shoff  
 Deborah K. Solbert  
 St. Jude Children's Research Hospital  
 Grace and Bruce Stillman  
 Nancy M. Szigethy  
 TD Bank  
 Mr. and Mrs. Peter S. Tilles  
 Joan P. Tilney  
 Mr. and Mrs. Edward Travaglianti

Sandy and Howard Tytel  
 Dr. and Mrs. Roy Vagelos  
 Dr. and Mrs. Fredrick Valauri  
 West Islip Breast Cancer Coalition for Long Island  
 Elisabeth R. Woods Foundation  
 Ann Eden Woodward Foundation  
 The Bradley Zankel Foundation  
 Karen E. Zoller

**Contributions of \$5,000+**

Arlene and Alan Alda  
 Mr. and Mrs. W. Dillaway Ayres, Jr.  
 Fred H. Baron  
 Geoffrey Beene Foundation  
 Frank and Nancy Bynum—The Bynum Family Foundation  
 Trudy and Tom Calabrese  
 Mr. and Mrs. Christopher C. Callaghan  
 Mr. and Mrs. Frank Castagna  
 Michele and Jim Celestino  
 Centerbrook Architects and Planners, LLC  
 Rita Cleary  
 Nancy Parmet Cook and Jim Cook  
 Milton Cooper  
 Ide and David Dangoor  
 Dau Family Foundation  
 Panera Bread  
 Dr. Lee MacCormick Edwards Charitable Foundation  
 Diane and Bob Fagiola  
 Stephen F. Melore, Esq.—Farrell Fritz, P.C.  
 Michael Focazio-Carissa Maringo Fund  
 Mr. and Mrs. Roger Fradin  
 The Deane A. and John D. Gilliam Foundation  
 Glen Cove C.A.R.E.S.  
 H2M Architects + Engineers  
 The Randall and Mary Hack Foundation

Robin Hadley  
 Harvest Real Estate Services, Inc.  
 Mr. and Mrs. Mark Hoffman  
 Joan M. Hutchins  
 J. Fallon's Tap Room  
 Kean Development Company, Inc.  
 The Kebede Family Foundation  
 The Kleinknecht Family Foundation  
 Drs. Amy and Roger Koreen  
 Dan Levy  
 Mr. and Mrs. Thomas H. Lister  
 Luitpold Pharmaceuticals, Inc.  
 Mrs. David J. Mahoney  
 Carol Marcincuk Fund  
 Mr. and Mrs. Edward C. Meagher  
 The Elena Melius Foundation  
 Mickie and Jeff Nagel  
 New York Community Bank Foundation  
 Indra K. Nooyi  
 Oxford & Simpson Realty Services  
 David P. Pearson  
 Amy Falls and Hartley Rogers  
 Pamela and Richard Rubinstein Foundation  
 Mr. and Mrs. Phillip Satow  
 Israel Scheut  
 Mr. and Mrs. Jonathan R. Serko  
 Mary F. Smith Family Foundation  
 Mr. and Mrs. Douglas S. Soref  
 Mr. and Mrs. Richard A. Spehr  
 Stand Up for Suzanne  
 The Hastings Foundation  
 Thermo Fisher Scientific  
 Drs. Marjorie J. van de Stouwe and Scott Ratner  
 Dr. and Mrs. George F. Vande Woude  
 Mr. and Mrs. Paul A. Vermynen, Jr.  
 J.B. Vermynen  
 W.F. Whitman Family Foundation  
 Deedee and Michael Wigler  
 Mr. and Mrs. Richard H. Witmer, Jr.  
 The WTFC Organization



Olivia Flatto, Frank Wisner, Fran Biondi, and Judy Cormier at the 2014 Double Helix Medals





CSHL President Bruce Stillman, CSHL Chairman Jamie Nicholls, Gabrielle Bacon, CSHL scientist Dr. Ann Churchland, and Kristina Perkin Davison at the Women's Partnership for Science lunch

**Contributions of \$2,500+**

Mr. and Mrs. Robert J. Adams  
 Mr. and Mrs. Robert H.B. Baldwin  
 Dominic Bencivenga  
 Liz and Rod Berens  
 Dr. Steven Brenner  
 Lisa and Tim Broadbent  
 Mr. and Mrs. William J. Candee III  
 Annie and Peter Cannell  
 Carol and Stephen Canter  
 Keri and Eric Carlstrom  
 Mrs. Robert L. Clarkson, Jr.  
 Julia Pershan and Jonathan Cohen  
 Mr. and Mrs. Jonathan Connors  
 Cottage Pharmacy & Boutique  
 Andrew Darrell  
 Dilworth & Barrese, LLP  
 Fidelity Investments  
 G & M Earth Moving, Inc.  
 Liebe and Seth Gadinsky  
 Gateway Realty  
 Chris Giamo  
 Mr. and Mrs. Mark W. Hamer  
 Cheryl Henson and Edwin A. Finn, Jr.  
 James B. Hicks, Ph.D.  
 Sonia S. Jauhar, M.D.  
 Neil Winter—Jefferies LLC  
 Johnston Family Foundation  
 Mr. and Mrs. Andrew L. Kahn  
 Neil M. Kaufman-Abrams, Fensterman  
 KPMG LLP  
 Gordon E. Lamb—Jefferson Family  
 Charitable Foundation  
 Fay J. Lindner Foundation  
 Mr. and Mrs. Christopher B. Mailman  
 Mr. and Mrs. Timothy Q. O'Neill  
 Onyx Management Group, LLC  
 Dr. and Mrs. John G. Passarelli  
 Mr. and Mrs. Simon Pharr  
 The Litwin Foundation  
 Whitney F. Posillico  
 Nicole and Charlie Prizzi

Anne and Bernard Reynolds  
 Alexandra Munroe and Robert  
 Rosenkranz  
 Mrs. Miriam K. Rothenberg  
 Lari and Francis Russo  
 Susan and Charles Sawyers  
 Scribner  
 Robert D. Selvaggio  
 Debbie Parmet Sondock and Cliff  
 Sondock  
 Samuel L. Stanley, Jr. and Ellen Li  
 Mr. and Mrs. Vincent Striano—Striano  
 Electric Co.  
 Carl J. Studer and Michele Miroff  
 Nancy B. Taylor  
 Erich Theophile and Steven Yee  
 Barbara and Donald Tober Foundation  
 Mr. and Mrs. Stanley S. Trotman, Jr.  
 United HealthCare Services, Inc.

Mr. and Mrs. Douglas A. Warner III  
 Deborah Norville and Karl Wellner  
 Wells Fargo Private Bank  
 Wolfensohn Family Foundation

**Contributions of \$1,000+**

A.L.L. Caring Foundation  
 Aboff's Inc.  
 Advantage Title Agency, Inc.  
 Chandni and Harry Anand  
 Erica and Henry Babcock  
 William W. Bancroft and Thomas  
 M. Bancroft Families—Norwood  
 Foundation  
 Nicholas Bartilucci—D&B Engineers and  
 Architects, P.C.  
 Mr. and Mrs. William B. Beekman  
 Dr. and Mrs. Arthur M. Bernhang  
 Kathy and Gene Bernstein  
 Giuseppe and L. Elizabeth Bertani  
 Jan and Don Bettencourt  
 Elizabeth B. Borden—Katharine Bradford  
 Foundation  
 Drs. David Botstein and Renee A. Fitts  
 Dr. William Braden III—Porter Braden  
 Endowment Fund  
 Ronald J. Brecht  
 Frank Brosens and Deenie Brosens  
 Foundation  
 Mary Jo Otsea and Richard H. Brown  
 C. Wonder LLC  
 C.M.S. Electrical Services Inc.  
 Maria B. Campbell  
 Daniel Carbone—Syosset Limousine  
 Michelle Carlson, M.D.  
 Muffy Carlstrom  
 Mary and Richard Catalano  
 Jim Childress and Ann Thompson  
 Jane H. Choate  
 Chubb Commercial Insurance  
 Dorcas O. Cochran



Cliff Sondock, Debbie Parmet Sondock, Frank O'Keefe, Mona Spector, David Spector, Scott Ratner, Marjorie van de Stouwe, Rob Palatnik, and Lauryl Palatnik at the 2014 Annual Meeting



Danielle Faria, Caroline Monti Saladino, and Jackie Lomtevas present a check from the Don Monti Memorial Research Foundation to CSHL President Bruce Stillman

Cold Spring Harbor Wine Shoppe  
 Cosel-Pieper Family Foundation  
 Susan and Curtis Cushman  
 Mrs. George W. Cutting, Jr.  
 Mr. and Mrs. Henry P. Davison II  
 Guy L. de Chazal and Caroline de Chazal  
 Meleanor L. Deming  
 Design Development Consulting Ltd.  
 Dr. Lester Dubnick  
 Brenda Earl  
 Electronix Systems  
 Frederick W. Engel  
 Enzo Life Sciences  
 Robert J. Feeny  
 Frederick Fogelman

Putney Cloos and Brian Fox  
 Mr. and Mrs. Douglas B. Fox  
 E. Maxwell Geddes, Jr.  
 Mr. and Mrs. Albert Gherlone  
 Joan and Alfred Goldberg  
 Mr. and Mrs. William H. Grover  
 Fran and John Gutleber  
 Mr. and Mrs. Frederick C. Haas  
 Reverend Linda Peyton Hancock  
 William A. Herzog  
 Cindy and Robert Higginson  
 Theodora W. Hooton  
 Indie By Jenna Lauren, Inc.  
 Frederic Ingraham  
 Lisa Jonas  
 Mr. and Mrs. Barclay G. Jones III  
 Mr. and Mrs. James M. Large  
 Janet and Gary Levy  
 MacLean Family Charitable Fund  
 Jane N.P. Mallinson  
 Marsh  
 Mr. and Mrs. Neil T. McGoldrick  
 Mr. and Mrs. Howard R. Merriman, Jr.  
 Mr. and Mrs. Clarence F. Michalis  
 Leo Model Foundation  
 Kathleen Mooney  
 Mr. and Mrs. Stephen V. Murphy  
 Mr. and Mrs. Francis W. Murray  
 Richard P. and Faye Nespola Charitable  
 Foundation  
 Northville Industries  
 Norwood Foundation, Inc.  
 Hugh and Patricia O'Kane Family  
 Foundation  
 One Beacon Charitable Trust  
 Mrs. William H. Osborn, Jr.  
 Mr. and Mrs. Richard S. Overton

Rhoda Parmet  
 Betsy and Jeffrey Pash  
 Mr. and Mrs. Nicholas B. Paumgarten  
 Mr. and Mrs. David Peikon  
 Carol Perry  
 Mr. and Mrs. Robert Pilkington  
 Emily K. Rafferty  
 Mr. and Mrs. William D. Roche  
 Mr. and Mrs. Arthur M. Rogers, Jr.  
 Ronald K. Rothman  
 Dr. and Mrs. Gerald M. Rubin  
 Mr. and Mrs. J. Wright Rumbough, Jr.  
 Wendy Ruzicka  
 George and Jayne Sabbagh  
 Diane Emdin Sachs Memorial Fund  
 Dr. and Mrs. Raju Sarwal  
 Lisa and Lee Schalop  
 Robin and Enrique Senior  
 Zachary, Sean, and Jordan Shallat  
 Mr. and Mrs. Trevor A. Smith  
 Mr. and Mrs. James L. Spingarn  
 The Society of St. Johnland  
 The James Stewart Family  
 Judith and Irwin Tantleff  
 Pamela M. Thye  
 Mr. and Mrs. Bart T. Tiernan  
 Mr. and Mrs. Pascal Franchot Tone  
 Helene P. Victor  
 Richard John Walk  
 Mr. and Mrs. Thomas J. Walsh  
 Walter B. James Fund No. 2  
 Gerald I. White  
 Diane and Howard Wohl Family  
 Foundation  
 Andrew Wylie  
 Dr. Charles Yanofsky

## CORPORATE AND FOUNDATION MATCHING GIFT CONTRIBUTORS

American Express Charitable Fund  
Bank of America Matching Gifts  
Bristol-Myers Squibb Foundation  
Causecast Foundation  
Dow Jones & Co.

Honeywell Matching Gift Program  
iStar Financial, Inc.  
The J.P. Morgan Chase Foundation  
Merck Partnership for Giving  
Microsoft Matching Gifts Program

Pfizer Foundation Matching Gifts  
Program  
Wells Fargo Foundation  
United Way of New York City

### IN HONOR OF

Keri Balmer  
Timothy Broadbent  
Edward Chernoff  
Susan Cohen  
Susan P. Connors  
Norris W. Darrell  
Kristina Perkin Davison  
Marjorie Feeny  
Robin Hadley

Mary D. Lindsay  
Robert Lindsay  
Bob and Sandy Messer  
Alea Mills  
Dr. Herbert Perr  
Dr. Mark Reimers  
Martin Ronemus  
Beatrice Rutkovsky  
Susan Wasserstein and George Sard

Charles Schmall  
Josh Silverman  
Andrew Solomon  
Freddie Staller  
Dr. Bruce Stillman  
Sylvia Thomas  
Edward Travaglianti  
John W. Ussery

### IN MEMORY OF

Pamela S. Banker  
Dr. Melvin Belsky  
Jack Billingsley  
Jon Brandt  
Ellen Brenner  
Carol A. Buonaiuto  
Ernest W. Caspari  
Susan Chiat  
Josephine and Ted Chodkowski  
John P. Cleary  
Jack Curry  
Vilma Diana  
Kerri Douglas  
Tina Duque  
Lee Edwards  
Ardell Bierman Fox  
Irwin Frish  
Nancy Slaughter Gay  
Joni Gladowsky  
Dr. Christopher G. Goff  
Charles Harris

Dr. John F. Heggie  
Thomas A. Hunt  
Jill Kipnes  
Judy Lettow  
Allan Lisse  
David L. Luke III  
Fanny Luke  
Cathryn MacInnes  
Richard Maguire  
Carol Marcincuk  
Joyce Marcus  
Carissa Maringo  
Elizabeth McFarland  
Rose M. Mooney  
Monica Mowery  
Arnold E. "Hack" Noyes  
Jack O'Brien  
Francis O'Connor  
Jeffrey Palmer  
Dr. Cecile Pickart  
G. Jeffrey Poletti

Christina Renna  
Jack Richards  
Elizabeth Roberts  
George Roberts  
Karl and Mona Runkle  
Beatrice Rutkovsky  
Diane Edmin Sachs  
Charles Schmall  
Thomas Schmitz  
Rhoda Schrom  
Martha Seidel  
Suzanne Shea  
Margaret Sheppard  
Fred Sileo  
Jason Stewart  
Mr. and Mrs. Carl Joseph Studer  
Sandy van Rensselaer  
Michael J. Wallace  
Caroline Wienclaw  
Dr. Mark Zoller

# LABORATORY MANAGEMENT

---

Bruce W. Stillman, Ph.D., President & Chief Executive Officer

James D. Watson, Ph.D., Chancellor Emeritus

W. Dillaway Ayres, Jr., Chief Operating Officer

## RESEARCH

**David L. Spector, Ph.D.**

*Director of Research*

**Sydney Gary, Ph.D.**

*Director, Research Operations*

**Walter L. Goldschmidts, Ph.D.**

*Executive Director,*

*Sponsored Programs*

## Professors

Douglas Fearon, M.D.

Thomas Gingeras, Ph.D.

*Professor and Head,*

*Functional Genomics*

Gregory Hannon\*, Ph.D.

Z. Josh Huang, Ph.D.

David Jackson, Ph.D.

Leemor Joshua-Tor\*, Ph.D.

Alexei Koulakov, Ph.D.

Adrian Krainer, Ph.D.

*Program Chair, Cancer*

*and Molecular Biology*

Robert Martienssen\*, Ph.D.

*Program Chair, Genomics*

*and Plant Biology*

W. Richard McCombie, Ph.D.

Alea Mills, Ph.D.

Partha Mitra, Ph.D.

Adam Siepel

*Program Chair, Simons*

*Center Quantitative*

*Biology*

Marja Timmermans, Ph.D.

Nicholas Tonks, Ph.D.

*Deputy Director,*

*CSHL Cancer Center*

David Tuveson, M.D., Ph.D.

*Deputy Director,*

*CSHL Cancer Center*

Linda Van Aelst, M.D., Ph.D.

Michael Wigler, Ph.D.

Anthony Zador, Ph.D.

*Program Chair, Neuroscience*

Yi Zhong, Ph.D.

## Associate Professors

Anne Churchland, Ph.D.

Joshua Dubnau, Ph.D.

Mikala Egeblad, Ph.D.

Grigori Enikolopov, Ph.D.

Hiroyasu Furukawa, Ph.D.

Adam Kepecs, Ph.D.

Bo Li, Ph.D.

Zachary Lippman, Ph.D.

Pavel Osten, M.D., Ph.D.

Darryl Pappin, Ph.D.

Michael Schatz, Ph.D.

Raffaella Sordella, Ph.D.

Arne Stenlund, Ph.D.

Lloyd Trotman, Ph.D.

Glenn Turner, Ph.D.

## Assistant Professors

Dinu Albeanu, Ph.D.

Gurinder Atwal, Ph.D.

Jesse Gillis, Ph.D.

Christopher Hammell, Ph.D.

Molly Hammell, Ph.D.

Ivan Iossifov, Ph.D.

Justin Kinney, Ph.D.

Alexander Krasnitz, Ph.D.

Je Hyuk Lee, M.D., Ph.D.

Dan Levy, Ph.D.

Gholson Lyon, M.D., Ph.D.

Stephen Shea, Ph.D.

Jessica Tollkuhn, Ph.D.

Christopher Vakoc, M.D., Ph.D.

Hongwu Zheng, Ph.D.

## Research Professor

James Hicks, Ph.D.

Scott Powers, Ph.D.

## Research Assistant Professors

Kenneth Chang, Ph.D.

Scott Lyons, Ph.D.

Michael Ronemus, Ph.D.

Seungtae Yoon, Ph.D.

## Senior Computer Scientist

Boris Yamrom, Ph.D.

## Computer Scientist

Swagatam Mukhopadhyay, Ph.D.

Ashwin Prakash, Ph.D.

## Senior Research Investigator

Carrie Davis, Ph.D.

## Research Investigators

Joan Alexander, M.D.

Eric Antoniou, Ph.D.

*Research Investigator/*

*Project Manager, CCG*

Isabel Aznarez Da Silva, Ph.D.

Camila Dos Santos, Ph.D.

Diane Esposito, Ph.D.

Yimin Hua, Ph.D.

Simon Knott, Ph.D.

Yoon-Ha Lee, Ph.D.

Shane Edward McCarthy, Ph.D.

Tatyana Michurina, Ph.D.

Youngkyu Park, Ph.D.

Natalia Peunova, Ph.D.

Michael Regulski, Ph.D.

Yi-Jun Sheu, Ph.D.

Zihua Wang, Ph.D.

## Adjunct Professors

Scott Lowe, Ph.D.

Anil Malhotra, M.D.

Larry Norton, M.D.

## Adjunct Associate Professor

Doreen Ware, Ph.D.

*USDA Adjunct Associate*

*Professor*

## Visiting Scientists/ Collaborative Scientists

Rhys Adams, Ph.D.

Gyula Bencze, Ph.D.

Wan-Jin Chen, Ph.D.

Ebru Demir, Ph.D.

Santiago Jaramillo Escobar, Ph.D.

Duda Kvitsiani, Ph.D.

Hysell Oviedo, Ph.D.

Shilpi Paul, Ph.D.

Lisa Prazak, Ph.D.

Min Zhang, M.D.

## Visiting Professors

Fritz Henn, M.D., Ph.D.

Matthew Ridley, Ph.D.

Kathleen Rockland, Ph.D.

Stuart Weisbrod, Ph.D.

## Visiting Scholar

Bhubaneswar Mishra, Ph.D.

## Director, Antibody and Phage Display

Bryce Nelson, Ph.D.

## Director, Gene Targeting Shared Resources

Leyi Li, Ph.D.

## Database Manager, Genomics

Christopher Zaleski

## Senior Fellow

Olga Anczukow-Camarda, Ph.D.

## Clinical Fellow

Mariano Ponz Sarvise, M.D., Ph.D.

Kenneth Yu, M.D.

## CSH Fellow

Lingbo Zhang, Ph.D.

## Postdoctoral Fellows

Sandra Ahrens, Ph.D.

Martin Akerman, Ph.D.

Xu An, Ph.D.

Francesca Anselmi, Ph.D.

Gayatri Arun, Ph.D.

\*Employee of the Howard Hughes Medical Institute

- Lindsey Baker, Ph.D.  
 Rachappa Balkunde, Ph.D.  
 Sara Ballouz, Ph.D.  
 Jan Bergmann, Ph.D.  
 Yael Berstein, Ph.D.  
 Jennifer Beshel, Ph.D.  
 Sonali Bhattacharjee, Ph.D.  
 Giulia Biffi, Ph.D.  
 Dario Bressan, Ph.D.  
 Martin Brill, Ph.D.  
 Hong Goo Chae, Ph.D.  
 Cristina Chen, Ph.D.  
 Iok In Chio, Ph.D.  
 Monika Chugh, Ph.D.  
 Hannes Clacys, Ph.D.  
 Vincenzo Corbo, Ph.D.  
 Katie Creasey, Ph.D.  
 Benjamin Czech, Ph.D.  
 Filipe de Sousa Borges, Ph.D.  
 Edgar Demesa Arevalo, Ph.D.  
 Sarah Diermeier, Ph.D.  
 Max Doerfel, Ph.D.  
 Mark Donoghue, Ph.D.  
 Dennis Eckmeier, Ph.D.  
 Elad Elkayam\*, Ph.D.  
 Bradford Elmer, Ph.D.  
 Ela Elyada, Ph.D.  
 Dannielle Engle, Ph.D.  
 Christopher Faeahnle, Ph.D.  
 Delphine Fagegaltier, Ph.D.  
 Iliaria Falcatori, Ph.D.  
 Gaofeng Fan, Ph.D.  
 Michael Feigin, Ph.D.  
 Daniel Ferrante, Ph.D.  
 Astrid Desiree Haase,  
 M.D., Ph.D.  
 Jong-Jin Han, Ph.D.  
 Balazs Hangya, M.D., Ph.D.  
 Anna Hangyane Benkovic, Ph.D.  
 Rowan Herridge, Ph.D.  
 Toshihide Hige, Ph.D.  
 Junya Hirokawa, Ph.D.  
 Guy Horev, Ph.D.  
 Manzar Hossain, Ph.D.  
 Dhananjay Huilgol, Ph.D.  
 Aman Husbands, Ph.D.  
 Change-Il Hwang, Ph.D.  
 Jonathan Ipsaro, Ph.D.  
 Yannick Jacob\*, Ph.D.  
 Byoung Il Je, Ph.D.  
 Yinping Jiao, Ph.D.  
 Daniel Jones III, Ph.D.  
 Erkan Karakas, Ph.D.  
 Matthew Kaufman, Ph.D.  
 Yongsoo Kim, Ph.D.  
 Stefan Klingler, Ph.D.  
 Steffen Knauer, Ph.D.
- Keerthi Krishnan, Ph.D.  
 Navasona Krishnan, Ph.D.  
 Victoria Kuettner, Ph.D.  
 Billy Lau, Ph.D.  
 Seung Cho Lee, Ph.D.  
 Seung-Jae Lee\*, Ph.D.  
 Young Koung Lee, Ph.D.  
 Zachary Lemmon, Ph.D.  
 Jayon Lihm, Ph.D.  
 Christopher Liseron-  
 Monfils, Ph.D.  
 Uri Livneh, Ph.D.  
 Jiangteng Lu, Ph.D.  
 Michael Ludwig, Ph.D.  
 Tomas Malinauskas, Ph.D.  
 Luis Mejia, Jr., Ph.D.  
 Olivia Mendivil Ramos, Ph.D.  
 Faith Mercan, Ph.D.  
 Mehrab Modi, Ph.D.  
 Adam Mohamed Naguib, Ph.D.  
 Almudena Molla-Morales, Ph.D.  
 Farzaneh Najafi, Ph.D.  
 Saikat Nandi, Ph.D.  
 Arun Narasimhan, Ph.D.  
 Gal Nechooshtan, Ph.D.  
 Shuang Ni, Ph.D.  
 Christos Noutsos, Ph.D.  
 Jaclyn Novatt, Ph.D.  
 Dawid Nowak, Ph.D.  
 Erik Olund, Ph.D.  
 Kin Fan On, Ph.D.  
 Gonzalo Hugo Otazu  
 Aldana, Ph.D.  
 Jean-Sebastien Parent, Ph.D.  
 Juwon Park, Ph.D.  
 Soonju Park, Ph.D.  
 Mario Penzo, Ph.D.  
 Katherine Petsch, Ph.D.  
 Hyun Jae Pi, Ph.D.  
 Arnaud Pommier, Ph.D.  
 Ileng Ramachandran, Ph.D.  
 Sachin Ranade, Ph.D.  
 Clare Rebbeck, Ph.D.  
 Michael Regan, Ph.D.  
 Ashlan Reid, Ph.D.  
 Jie Ren, Ph.D.  
 Benjamin Roche, Ph.D.  
 Jae Seok Roe, Ph.D.  
 Nikolay Rozhkov, Ph.D.  
 Elena Rozhkova, Ph.D.  
 Leah Sabin, Ph.D.  
 Ana Margarida Santos De  
 Almeida, Ph.D.  
 Juergen Scharner, Ph.D.  
 Hillary Schiff, Ph.D.  
 Andrea Schorn, Ph.D.  
 Serif Senturk, Ph.D.
- Mario Shields, Ph.D.  
 Meng-Fu Shih, Ph.D.  
 Atsushi Shimada, Ph.D.  
 Yichun Shuai, Ph.D.  
 Damianos Skopelitis, Ph.D.  
 Sebastian Soyk, Ph.D.  
 Carlos Stahlhut Espinosa, Ph.D.  
 Marcus Stephenson-Jones, Ph.D.  
 Yilin Tai, Ph.D.  
 Nami Tajima, Ph.D.  
 Oliver Tam, Ph.D.  
 Julian Taranda, Ph.D.  
 Shaolei Teng, Ph.D.  
 Herve Tiriack, Ph.D.  
 Ante Tocilj, Ph.D.  
 Vasily Vagin, Ph.D.  
 Alexander Vaughan, Ph.D.  
 Wim Verleyen, Ph.D.  
 Elvin Wagenblast, Ph.D.  
 Bo Wang, Ph.D.  
 Bor-Shuen Wang, Ph.D.  
 Chia-Lin Wang, Ph.D.  
 Minghui Wang, Ph.D.  
 Kaja Wasik, Ph.D.  
 John Wilson, Ph.D.  
 Sze Wong, Ph.D.  
 Qingyu Wu, Ph.D.  
 Qiaojie Xiong, Ph.D.  
 Cao Xu, Ph.D.  
 Yu-Ting Yang, Ph.D.  
 Kai Yu, Ph.D.  
 Yang Yu, Ph.D.  
 Wei Zheng, Ph.D.
- Graduate Students**
- Robert Aboukhalil  
 Brinda Alagesan  
 Pedro Alexandre Garcia Da Silva  
 Naishitha Anaparthi  
 Arkarup Bandyopadhyay  
 Giorgia Battistoni  
 Philippe Batut, Ph.D.  
 Anand Bhagwat  
 Emilis Bruzas  
 Colleen Carlston  
 Brittany Cazakoff  
 Sudipto Chakraborty  
 Mana Chandhok  
 An-Yun Chang  
 Kung-Chi Chang  
 Lital Chartarifsky  
 Ke Chen  
 Derek Cheng  
 Sam Chiappone  
 Mohan Chitra  
 Na Hyun Cho
- Hsiang-Chen Chou  
 Maria Joaquina Delas Vives  
 Kristen Delevich, Ph.D.  
 Steffen Dickopf  
 Noah Dickler  
 Nicolas Erard  
 Han Fang  
 Miriam Fein  
 Silvia Fenoglio  
 Angeliki Field-Pollatou  
 Dea Filippini  
 Talitha Forcier  
 Nicholas Gallo  
 Tyler Garvin  
 Sanchari Ghosh  
 Hamza Giaffar  
 Michael Giangrasso  
 Jacqueline Giovanniello  
 Siong Wee Goh  
 Jiaqi Gu  
 Priyanka Gupta  
 Michael Gutbrod  
 Abram Handly-Santana  
 Yu-Jui Ho  
 Anja Hohmann  
 Elizabeth Hutton  
 Dong-Woo Hwang  
 Hyun-Yong Jeon  
 Yibin Jia  
 Laura Jimenez Barron  
 Yagmur Kan  
 Justus Kebschull  
 Sean Kelly  
 Daniel Kepple  
 Matthew Koh  
 Lisa Krug  
 Devinn Lambert  
 Yan Ha Lee  
 Jesse Levine  
 Shanshan Li  
 Laura Maiorino  
 Fred Marbach  
 Paul Masset  
 Monika Moissidis  
 Maria Nattestad  
 Tomoki Nomakuchi  
 David Raposo Nunes  
 Onyekachi Odoemene  
 Tobiloba Oni  
 Jason O'Rawe  
 Debjani Pal  
 Ian Peikon, Ph.D.  
 Jean Pie  
 Sashank Pisupati  
 Mathangi Ramesh  
 Ricardo Raudales  
 Priyadarshini Ravichandran

\*Employee of the Howard Hughes Medical Institute

Annabel Romero Hernandez  
Ioana Rus  
Eugene Jun Wen Seah  
Chen Shen  
Jia Shen  
Junwei Shi  
Nitin Shirole  
Justin Snider  
Dongyan Song  
Colin Stoneking  
Eric Szelenyi  
Sophie Thomain  
Jason Tucciarone  
Charles Underwood  
Jack Walleshauser III  
Jue Xiang Wang  
Yiyang Wu  
Robert Wysocki  
Yali Xu  
Ming Yang  
Georgi Yordanov  
Allen Yu  
Jia-Ray Yu  
Cinthya Zepeda Mendoza, Ph.D.  
Anqi Zhang  
Siwei Zhang  
Xin Zhou

### Computational Science

Peter Andrews, Ph.D.  
Alexander Dobin, Ph.D.  
Evan Ernst  
Elena Ghiban  
James Gurtowski  
Yuan Hao, Ph.D.  
Gregor Havkin  
Ying Jin, Ph.D.  
Jude Kendall  
Melissa Kramer  
Vivek Kumar, Ph.D.  
Sunita Kumari, Ph.D.  
Anthony Leotta  
Zhenyuan Lu  
Steven Marks  
Ferenc Mechler, Ph.D.  
David Molik  
Marcela Monaco, Ph.D.  
Joseph Mulvaney  
Andrew Olson  
Kith Pradhan, Ph.D.  
Srividya Ramakrishnan  
Fritz Sedlazeck, Ph.D.  
Vijay Kumar Sreenivasa  
Gopalan, Ph.D.  
Joshua Stein, Ph.D.  
James Thomason III  
Kannan Umadevi Venkataraju

Liya Wang, Ph.D.  
Xuehong Wei

### Research Associates

Carmelita Bautista  
Barry Burbach  
Hilary Cox  
Jorg Drenkow  
Cristina Fernandez Marco, Ph.D.  
Carla Gauss, Ph.D.  
Diana Gizatullina  
Hyun Soo Kim, Ph.D.  
Lisa Manche  
Anirban Paul, Ph.D.  
Keith Rivera  
Linda Rodgers  
Patricia Wendel  
Caizhi Wu, Ph.D.  
Melody Wu, Ph.D.  
Scilla Wu

### Laboratory Facility Managers

Sabrina Boettcher\*  
Patricia Brady  
Robert Gerdes  
Stephanie Goldsmith\*  
Sara Goodwin, Ph.D.  
Stephen Hearn  
Pamela Moody  
Timothy Mulligan  
Joseph Simorowski\*  
Spencer Teplin  
Jennifer Troge  
Ming Wang

### Scientific IT Support

David Trimboli  
Peter Van Buren

### Scientific Administrators

Sadie Arana\*  
Lorraine Baldwin  
AnnMarie Ballew  
Christine Bedell  
Patricia Bird  
Becky Dong  
Irene Gill  
Carol Gundersen  
Idee Mallardi  
Sandra Michelsen  
Susan Milham  
Sandra Moran  
Elizabeth Panagot  
Kylie Parker  
Andrea Potopsingh

Dorothy Tsang Hu  
Madeline Wisnewski

### Laboratory Technicians

Sai Dipikaa Akshinthala  
Alexandra Ambrico  
Nabil Azamy  
Nicholas Baltera  
Eleni Boikoglou, Ph.D.  
Amy Brady  
Christopher Brooks  
Jacqueline Cappellani  
Yon Chang  
Joshua Cohen  
Martin Davis  
Anthony DeSantis  
Panchajanya Deshpande  
Sean D'Italia  
Katja Doerfel  
Amanda Epstein\*  
Margaret Falkowski  
Neil Franciotti  
Inessa Hakker  
Ga-Ram Hwang  
Meagan Keane  
Sun Kim  
Chantal LeBlanc\*  
Emily Lee  
Eunjung Lee  
Jingjing Li  
Allison Liberato  
Angela Licata  
Ying Hsiu Liu  
Maria Lorenc  
Beicong Ma  
Senem Mavruk Eskipehlivan  
Jean McIndoo  
Nicoletta Memos  
Carolyn Mende  
Joseph Milazzo  
Jessica Minder  
Patricia Mocombe  
Stephanie Muller  
Victoria Nelson  
Hien Nguyen  
Jie Ouyang  
Ramesh Palaniswamy  
Ami Patel  
Umamaheswari Ramu  
Michael Riggs  
Julie Rosenbaum  
Nikita Saha Turna  
Stephen Savoia  
Alexandra Scavelli  
Stephen Schuck  
Lei Hoon See  
Noriko Simorowski

Tara Skopelitis  
Young-Eun Son, Ph.D.  
WeiJia Tan  
Alexander Tolpygo  
Gaurang Trivedi  
Eric Wang  
Robert Wappel  
Kaitlin Watrud  
Kevin Weber  
Kevin Wright  
Lei Zhang  
Lifang Zhang, Ph.D.  
LiGuo Zhang

### Laboratory Facility Staff

Alexander Aguilera  
Damir Begonja  
Fred Bencivenga  
Lisa Bianco  
Oral Bryan  
Amy Burrows  
Michael Cahn  
Manuel Canas  
Jodi Coblentz  
Kristin Cox  
Charlene De Poto  
Maria Demott  
Jared Downing  
Eileen Earl  
Wilson Garciaguirre  
Jill Habel  
Denise Hoppe  
Xiao Ling Long  
Maria Mosquera  
Gustavo Munoz  
Aigoul Nourjanova  
Jason Pellegrini  
Raisa Puzis  
Sherry Salgado  
Amanda Sendkewitz  
Ann Smith  
Walter Vigil Garcia  
Stephanie Woods, D.V.M.

### Media Makers/Farm Assistants

Nancy Bolanos  
Terrance Chisum  
Adriana Hincapie  
Kelly Ann Holfester  
Amy Ji  
Vincent Meschan  
Michele Murrell  
Mariana Neves Dos Santos Leite  
Hong Jie Shen

**Research Administration**

Maureen Bell

**EDUCATION****Banbury Center****Jan A. Witkowski, Ph.D.***Executive Director*

Patricia Iannotti

Barbara Polakowski

Janice Tozzo

**CSHL Press****John R. Inglis, Ph.D.***Executive Director*

Barbara Acosta

Janet Argentine

Cynthia Blaut

Carol Brown

Kathleen Bubbeo

Margaret Calicchia

Kathleen Cirone

Heather Cosel-Pieper

Marie Cotter

Laura DeMare, Ph.D.

Maryliz Dickerson

Maria Ebbets

Corrisa Farmer

Bibiane Garite

Michael Henigman

Nancy Hodson

Catherine Kelly

Delia King

Tara Kulesa

Wayne Manos

Mala Mazzullo

Joanne McFadden

Mary Mulligan

Stephanie Novara

Stephen Nussbaum

Keriaki Pollakis

Robert Redmond

Theodore Roeder

Mary Sabala

Eric Sawey, Ph.D.

Diane Schubach

Richard Sever, Ph.D.

Inez Sialiano

Marcie Siconolfi

Maria Smit, Ph.D.

Hillary Sussman, Ph.D.

Linda Sussman

Denise Weiss

**DNA Learning Center****David A. Micklos***Executive Director*

Kayla Bianco

Abigail Buckley

Elna Carrasco

Kapeel Chougule

Kenneth Romi Collado

Lauren Corrieri

Antonia Florio, Ph.D.

Liam Gensel

Cornel Ghiban

Joshua Gruber

Emily Harrison

Nella Hauser

Omotayo Ikuomenisan

Aiseosa Irowa

Ashleigh Jackobel

Marie Jean Francois

Lenni Joya Amaya

Mohammed Khalfan

Karoline Lake

Mary Lamont

Susan Lauter

Melissa Lee

Christine Marizzi

Kathleen McAuley

Amanda McBrien

Pauline McGlone

Colleen McKechnie

John Messina

Valerie Meszaros

Gabrielle Nagel

Eric Nash, Ph.D.

Nicholas Nicolia

Gabrielle Ramirez

Carolyn Reid

Brooke Roeper

Gavrielle Rood

Shenika Shah

Shreya Shah

Mona Spector, Ph.D.

Breanna Tahany

Maria Urbina

Scott Venturino

Jason Williams

Blakelee Woods

Chun-hua Yang

Xiaoqun Zhang

**Library and Archives****Ludmila T. Pollock***Executive Director*

Paula Abisognio

Thomas Adams

Clare Clark

Matthew Covey, Ph.D.

Stephanie Satalino

Gail Sherman

Rhonda Veros

Judith Wieber, Ph.D.

**Meetings and Courses****David J. Stewart, Ph.D.***Executive Director*

Stephen Alexander

Alfred Baman

Katherine Bradley

Edward Campodonico

Joseph Carrieri

Reginald Cayo

Katiria Ceballos

Maoyen Chi, Ph.D.

Wendy Crowley

William Dickerson

James Duffy

Bradley Frey

Constance Hallaran

Julia Kuhl

Charla Lambert, Ph.D.

Andrew Mendelsohn

Diana Monno

Nicholas Moore

Maureen Morrow

Demetria Mudar

Michele Navarro

Andrea Newell

Kenneth Orff

Valerie Pakaluk

Jonathan Parsons

Stephen Silver

Mary Smith

Margaret Stellabotte

Samuel Stewart

James Whitaker

Barbara Zane

**Watson School of  
Biological Sciences****Alexander Gann, Ph.D.***Dean*

Carrie Cowan, Ph.D.

Kimberly Creteur

Kimberly Geer

Alyson Kass-Eisler, Ph.D.

**Terri I. Grodzicker, Ph.D.***Dean of Academic Affairs***ADMINISTRATIVE  
DEPARTMENTS****Comptrollers Office  
and Accounting****Lari C. Russo***Chief Financial Officer*

Sharon Bense

Karine Boyce

Damian Desiderio

Maryanne Detweiler

Karen Filasky

Mary Ellen Goldstein

Karen Jones

Patricia Maroney

Mary Mastropaolo

Alison McDermott

Kristine Murphy

Katherine Ortiz

Patricia Penner

Nicole Solomon

Annette Tommaso

Patricia Urena

Ann Williams

**Central Administration**

Maureen Berejka

Bridgette Brady

Carol Link

Stephanie Peragallo

Denise Roberts, Ph.D.

**Culinary Services****James Hope***Director*

Doris Abrego

Inolka Abreu

Alan Aguirre

Jeannette Amato

Eliseo Amaya

Ingrid Amaya

Yamic Ambra

Francis Bowdren

Maryam Brinkwart

Elmer Canales

Giuditta Carino

William Carmona

Francis Cayamcela

James Charvis

Marvin Chavez

Mildred Contreras Pinto

Julie Ehrlich

Suzanne Esposito

Bartola Garcia

Luis Garcia

Andres Gonzalez

Silvia Gonzalez  
 Barbara Gordon  
 Jennifer Gordon  
 James Hynter  
 Pierre Jean-Baptiste  
 Deborah Koslosky  
 Pamela Lancellotti  
 Jennifer LoSordo  
 Michael LoSordo  
 Juan Magana  
 Jose Maradiaga Gonzalez  
 Thomas McIlvaine  
 Alan Minicozzi  
 Laurie Moller  
 Frederik Munter  
 Jorge Ramirez  
 Jose Reyes  
 Susana Roman  
 Marlenis Romero  
 Robert Scherer  
 Karen Scola  
 Shannon Silva  
 Jose Velasquez  
 Hector Vigil Garcia  
 Maria Villatoro  
 Maria Vilorio  
 Tracy Woolfson  
 Dennis Zelaya  
 Jennifer Ziegler  
 Lidia Zuleta

#### Daycare Center

**Deborah Keller**  
*Director*  
 Michele Bauman  
 Suzanne Chitwood  
 Margie Davis  
 Stephanie Dennis  
 Danielle Duryea  
 Karen Flore  
 Sara Genovese  
 Christina Godfrey  
 Mary Ellen Hagenmiller  
 Jamie John  
 Stefanie Levitz  
 Cynthia Lyon  
 Tracie Martinetti  
 Deirdre McDermott  
 Paula Messinger  
 Christina Meyer  
 Lucinda Mow  
 Meryl Rosen-Wells  
 Kara Scholz  
 Judith Smith  
 Shannon Smith

#### Development

**Charles V. Prizzi**  
*Vice President*  
 Jean DeForest  
 Diane Fagiola  
 Anne Greengard, J.D.  
 Karen Orzel  
 Wendy Schneps  
 Bailey Stark  
 Jill Stone

#### Facilities

**Arthur Brings**  
*Vice President,  
 Chief Facilities Officer*  
 Andres Alarcon  
 Robert Albinson  
 Brian Allen  
 Leslie Allen  
 Edward Anderson  
 Peter Anderson  
 Jerry Armstrong  
 Richard Beehler  
 Elizabeth Begonja  
 Matthew Bird  
 William Bishop  
 Margaret Brock  
 Keith Brown  
 Dessie Carter  
 Daniel Chapman  
 Christopher Clark  
 Robert Collins  
 Drew Comer  
 Michael Cook  
 Gabriela Cruz  
 Matthew D'Amico  
 Peter Dale  
 John Damico  
 Nicholas Dassau  
 James Decker  
 Robert Dickerson  
 Glen DiMaria  
 Steven Dlugolonski  
 Kevin Donohue  
 Michelle Doran  
 Wesley Dreusike  
 Paul Edwards  
 Robert Eifert  
 Rafael Estrella  
 Timothy Foley  
 Robert Gensel  
 Jeffrey Goldblum  
 Robert Gorney  
 Erick Greene  
 Leo Greene

Stephen Gregorovic  
 Wilman Gutierrez  
 Wayne Hamilton  
 Gerard Holler  
 Joseph Houser  
 Matthew Hoyer  
 Christopher Hubert  
 Louis Hunter  
 Michael Hutchinson  
 Nicholas Jarett  
 Randal Jones  
 Daniel Jusino  
 Kevin Keenan  
 Jesse Kitzen  
 Jeffrey Klaverweiden  
 Oscar Lastra  
 Walter Leute  
 Jian Lin  
 Charles Mandell  
 Gregory Marotta  
 Patricia McAdams  
 Joseph McCoy  
 Tahid McCoy  
 Christopher McEvoy  
 Kevin McHugh  
 Daniel Miller  
 Douglas Morrow III  
 Guido Mosquera  
 Louis Muno  
 Mary Muno  
 Michael Murphy  
 Kenneth Nastri  
 Herbert Neazer  
 George Newell  
 Michael Nolasco  
 Michael Noone  
 Wayne Pav  
 Jose Pena Corvera  
 John Pisciotta  
 David Pollock  
 Manuel Ramones  
 Wilson Ramones  
 Steven Randall  
 Alfredo Rivera  
 Ronald Romani  
 Marlene Rubino  
 Frank Russo  
 Richard Rutherford  
 Steven Sardone  
 John Schindler  
 Claudia Schmid  
 Lauren Schmidt  
 Peter Schwind  
 Cesar Sisalima  
 Michael Spichalsky  
 Clifford Sutkevich

Steven Tang  
 Hock Yew Teo  
 Albaro Trinidad  
 Leigh Tyrrel  
 Steven Ulloa  
 James Van Wie  
 Fredy Vasquez  
 Trinidad Velasquez  
 Wilfredo Velasquez  
 Benjamin Veneable III  
 Alvin Watson  
 Kyle Welton  
 Leslie Wenzel  
 Graham Wildt  
 Curtis Williams  
 Fred Zier  
 Daniel Zimmer

#### Human Resources

**Katherine G. Raftery**  
*Vice President*  
 Wendy Alexander  
 Michele Borriello  
 Camille Cava-Belluscio  
 Patricia Crawford  
 Jessica Heckel  
 Lesley Inglis  
 Michelle Krul  
 Marygrace Navarra  
 Victoria Panebianco  
 Melissa Rowsell  
 Marjorie Ryan  
 Patricia Schoeffler

#### Information Technology

**Hans Erik Aronson**  
*Vice President Information  
 Technology/CIO*  
 Alex Barton  
 William Burns  
 Roy Reeves Crowe  
 Darian Defalco  
 Peter DiLauro  
 Daniel Drayzen  
 Osama El Demerdash  
 Frank Falco  
 Kelvin Flores  
 Todd Heywood, Ph.D.  
 Alex Jerez  
 Thomas Karow  
 Sean Kelly  
 Louis Malfi  
 Gerald McCloskey  
 Bhargavkumar Patel



Barbara Peters  
 Vincent Reale  
 Leila Sawicki  
 Brian Scholer  
 Salvatore Serafino  
 Michael Skuthan  
 David Jonathan Torres  
 Steven Zielinski

**Procurement**

**Philip Lembo**

*Vice President*  
 Deborah Aufiero  
 Francis Bella  
 Kimberly Bronson  
 Carol Brower  
 Susan De Angelo  
 Jeffrey DuPree  
 Elizabeth Janow  
 Christopher Oravitz  
 Barbara Purcell  
 Krystyna Rzonca  
 Theresa Saia

**Public Affairs**

**Dagnia Zeidlickis**

*Vice President,  
 Communications*  
 Robert Als  
 Margot Bennett  
 Jessica Giordano  
 Deborah Gordon  
 Jaclyn Jansen, Ph.D.  
 Philip Renna  
 Peter Tarr, Ph.D.

**Sponsored Programs**

**Walter L. Goldschmidts, Ph.D.**

*Executive Director*  
 Irina Armstrong  
 Maureen Davies  
 Rachel Desrochers  
 Carol DuPree  
 Diane Errico  
 Donna Gavin  
 Arthur Gunston  
 Jill Hemish, Ph.D.  
 Joanne Janelli  
 Ruta Ledins  
 Stacy Leotta  
 Dave Neale  
 Catherine Perdikoylis

**Technology Transfer**

**Teri Willey**

*Vice President, Business  
 Development/Technology  
 Transfer*  
 Debra Arenare, J.D.  
 Elizabeth Cherian-Samuel  
 Vladimir Drozdoff, J.D., Ph.D.  
 Shiu Yie Hsieh  
 Lorraine McInerney  
 Alexander Sosnovski

**General Counsel**

**John P. Maroney, J.D.**

*Vice President,  
 General Counsel*  
 Amanda Reiss, J.D.  
 Julie Sutherland

**RESEARCH STAFF  
 DEPARTURES DURING 2014**

**Research  
 Investigators**

Emily Hodges  
 Kandasamy Ravi

**Adjunct Assistant Professor**

Senthil Muthuswamy

**Visiting Scientists/  
 Collaborative  
 Scientists**

Megan Cleary

**Director, Gene Targeting  
 Shared Resources**

Ping Jiang

**Postdoctoral  
 Fellows**

Christine Ardito-Abraham  
 Huyen Bui  
 Alexander Campbell  
 Marcela Dotto  
 Cheryl Eifert  
 Andrea Eveland  
 Pascal Grange  
 Baofeng Guo  
 Miao He  
 Michael Huebner  
 Ke Jiang

Shimpei Kawaoka  
 Sukant Khurana  
 Claus-Dieter Kuhn  
 Duda Kvitsiani  
 Haohong Li  
 Jinyu Li  
 Li Li  
 Cora MacAlister  
 Shoshana Marcus  
 Adaobi Mofunanya  
 Amit Mukherjee  
 June Hee Park  
 Roberto Perales  
 Jonathan Preall  
 Joshua Sanders  
 Naoki Takada  
 Yi Wei  
 Huiqing Zhan  
 Ling Zhang

**Computational  
 Science**

Han Fang  
 Ewa Grabowski  
 Giuseppe Narzisi  
 Eric Paniagua  
 Vadim Pinskiy  
 Chenghai Xue  
 Charles Youens-Clark  
 Yingbin Zhang

**Research  
 Associate**

Shiqiang Dai  
 Leopold Puzis