Cancer

Cancer research has been a major focus at Cold Spring Harbor Laboratory (CSHL) for decades. Since 1987, when CSHL was designated a National Cancer Center, the field has benefited from revolutionary advances, many of which are legacies of the massive international effort to sequence the human and other genomes. CSHL scientists have had a major role in these developments, at both the conceptual and technological levels, discovering (and devising methods of discovering) oncogenes, tumor suppressor genes, and pathogenic viruses implicated in tumor development. They have also devised methods of performing comparative analyses of genomes based on microarray technology, which have revealed genomic deletions and duplications that cause changes in gene copy number, mutations that may give rise to cancers and other important illnesses. The cancer program is organized into three units. Genetics focuses on understanding the genetic basis of cancer, cancer progression, and development of resistance to chemotherapy. Gene Expression and Proliferation focuses on the regulation of gene expression, cell-division cycle control, and chromosome structure in normal and cancer cells. Signal Transduction focuses on signaling pathways and cell architecture in normal and cancer cells.

Genetics

Gregory Hannon is a pioneer in the study of RNA interference (RNAi) in mammalian systems. In RNAi, double-stranded RNA molecules induce gene silencing. Hannon is credited with the discovery of two enzymes, called Dicer and Slicer, critical in the RNAi machinery. (The work on Slicer was performed with Leemor Joshua-Tor.) His team has led the way in using RNAi to study cancer biology and genetics. They have generated a library of short-hairpin RNAs that researchers at CSHL and elsewhere apply broadly in gene-silencing studies. This year, the lab made a host of discoveries, including that of a new class of small RNAs, which was found in fruit flies to modify gene activity and suppress mobile genomic elements, thus serving as a defense mechanism. They also discovered 180 previously unrecognized microRNAs specific to monotremes in work stemming from a collaborative effort to sequence the platypus genome. The lab also explained a new way in which epigenetic information is inherited, discovering in fruit flies a class of small RNAs maternally inherited that determines an offspring’s fertility trait.

Scott Lowe’s laboratory studies cancer gene networks and determines how genetic lesions affecting these networks contribute to tumor development and resistance to therapy. This year, they continued to study cellular senescence, a potentially powerful mechanism for suppressing tumors, showing that it helps to limit wound-healing responses, a finding which suggests that it might act outside cancer to maintain tissue homeostasis following damage. Lowe also has adapted RNAi technology to produce animal models...
in which genes can be switched on and off in a spatial, temporal, and reversible manner and used this to identify and characterize new genes that modulate tumor cell responses to chemotherapy. He spearheaded an effort with other CSHL investigators to integrate genomic studies on human liver cancers with RNAi-based screening in a mouse model to identify 13 new tumor suppressor genes. These studies suggest a means of functionally annotating cancer genes and holds promise of producing new insights into the genetic basis of tumor diversity that can ultimately be exploited to tailor treatments to individual patients.

Robert Lucito, collaborating with Michael Wigler, has had an important role in developing innovative technologies, including representational difference analysis (RDA), representational oligonucleotide microarray analysis (ROMA), and comparative genome hybridization (CGH), that have proven to be valuable to cancer researchers worldwide. Also an experimentalist, Lucito has conducted studies using these techniques to detect copy-number changes in large sets of human ovarian and pancreatic cancer tissue samples. The lab also has turned its attention to epigenetics, specifically to the study of methyla-
entists to significantly change long-standing notions about both genes and genomes, revisions so sweeping that they promise to change our very definition of the gene.

The laboratory of Leemor Joshua-Tor focuses on cell regulatory mechanisms, including protein complexes involved in nucleic acid regulatory processes. The lab uses X-ray crystallography to obtain three-dimensional structures of individual proteins and atomic-level views of their interactions with other molecules. This past year, they clarified how yeast cells, through the action of genes, adjust their metabolism in response to changes in their sources of food. This research suggests how the metabolic state of a cell is linked to the expression of its genes in a way that impacts biological processes of many kinds, ranging from cancer to aging. Joshua-Tor, who in 2008 was named a Howard Hughes Medical Institute Investigator, is well known for her work on the helicase enzyme, which acts to unwind DNA strands during the DNA self-replication process, and, in research with Gregory Hannon here at CSHL, for revealing structures that help to explain the gene-silencing mechanisms of RNA interference.

Adrian Krainer and colleagues study the mechanisms of RNA splicing, the ways in which they go awry in disease, and the means by which faulty splicing can be corrected. Their approach has borne fruit in the study of spinal muscular atrophy (SMA), a genetic neuromuscular disease. Their ability to correct an mRNA splicing defect in SMA that makes a gene called SMN2 only partially functional forms the basis of a potentially powerful therapeutic approach, now being tested in mouse models. They have also studied a splicing factor, SF2/ASF, that can cause cells to become cancerous and have now shown how it acts on proteins in a group called the PI3K-mTOR (phosphoinositol-3-kinase–mammalian target of rapamycin) pathway, known for its involvement in cancers. Measurement of SF2/ASF levels could eventually lead to identification of patients who will respond well to drugs that block the pathway. With Michael Zhang, Krainer this year also demonstrated a means of identifying the many genomic targets for a particular type of splicing factor and how splicing patterns of those targets change.

Scientists do not yet understand how dormant metastatic lesions, after they have colonized distant organs, grow into large lethal lesions. Vivek Mittal studies angiogenesis in the tumor microenvironment, which is composed of both malignant and non-malignant cells. Among the latter are bone-marrow-derived endothelial progenitor cells, or EPCs, which, Mittal has demonstrated, become involved selectively in tumor blood vessel growth. This year, Mittal and colleagues, using mouse models, showed that the levels of a protein called Id-1 increase dramatically in EPCs when tumors are present. Using RNAi to block the expression of Id-1 in living mice, the team was able to prevent mobilization of EPCs to metastasis sites, thereby inhibiting an “angiogenic switch,” a key mechanism that causes formation of blood vessels in tumors and triggers tumor growth.

David L. Spector’s lab studies the spatial organization and regulation of gene expression. Their in vivo approach is exemplified in a live cell gene expression system that has made possible examination in real time of the recruitment of members of the gene expression and silencing machineries. A current research focus is the distribution of nuclear Polycomb proteins, known to keep genes in a silent state. The team seeks to target these proteins to segments of DNA as a means of selectively silencing specific genes. Another focus is the study of noncoding RNAs retained in the nucleus. This past year, the lab discovered a nuclear mechanism that processes noncoding RNAs. MALAT1, a noncoding nuclear RNA, was observed to split into two segments, the smaller of which migrated into the cytoplasm. A longer remnant remained in the nucleus, accumulating in zones called nuclear speckles. The nuclear-retained MALAT1 transcript is up-regulated in tumors that have the propensity to metastasize. Although it is not clear what the processed parts of the original RNA do, their disparate destinations in the cell suggest that they serve different functions.
Papillomaviruses, a large viral family that induces cell proliferation at the site of infection, usually give 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. 

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. 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 at which DNA synthesis occurs. Among these proteins are those that facilitate the assembly of chromatin, the protein-DNA complexes that form the chromosomes. The prime focus of current research, however, is 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 ORC or origin recognition complex. Stillman’s research has demonstrated that ORC is also involved in the process of segregating the duplicated chromosomes in mitosis. They have found ORC at centrosomes and centromeres, structures that orchestrate chromosome separation in mitosis.

Years ago, William Tansey observed that the destruction of transcription factors is intimately tied to their ability to activate transcription, which raised the intriguing question of how the very first step in the life of a protein was functionally related to the very last step, proteolysis—the protein’s destruction by a molecular complex called the proteosome. The key link, he demonstrated, is ubiquitin, a molecule with which proteins must be tagged before they can enter the proteosome for destruction. Having linked ubiquitin and proteolysis to the activity of transcription factors, Tansey and colleagues have focused on the broader question of how ubiquitin affects transcription. This year, they discovered a means by which ubiquitin can regulate transcription without engaging the proteolytic machinery, specifically by acting as a wedge to pull apart and thereby inactivate RNA polymerase, the enzyme that orchestrates transcription. They are now exploring the biological context in which this non-proteolytic regulatory function of ubiquitin is brought to bear.

The function of DNA within cells is heavily influenced by its packaging with protein, a complex known as chromatin. The regulation of chromatin structure is essential for specifying which genes should be switched “on” or “off” in a given cell type. Alterations of chromatin structure that have been observed in cancer cells promote uncontrolled cell growth and progression of the disease. The research of CSHL Fellow Christopher Vakoc aims to understand how changes in chromatin contribute to the pathogenesis of acute leukemia, a cancer of blood cells. Using genome-wide mapping of chromatin structure together with a functional-genetics approach, Vakoc...
seeks to uncover novel chromatin pathways involved in cancer. By evaluating the reversibility of chromatin states in living organisms, he hopes to reveal novel therapeutic targets that may improve leukemia treatment.

**Signal Transduction**

Yuri Lazebnik and colleagues study cell fusion in the context of the hypothesis that viruses and other common human pathogens might cause cancer under certain conditions. They have established that massive chromosomal instability can be engendered by a transient event causing genomic destabilization without permanently affecting mechanisms such as mitosis or proliferation. The agent, in this instance, is an otherwise harmless virus that causes chromosomal disruption by fusing cells whose cell cycle is deregulated by oncogenes. The resulting cells had unique sets of chromosomes and some proved to be capable of producing aggressive epithelial cancers in mice. This year, Lazebnik’s lab developed a means of producing hybrid cells more efficiently. Rather than use live viruses to induce fusion, they perfected a method of isolating viral fusogenic proteins. Using these to induce fusion under controlled conditions, they are now exploring the consequences of cell fusion for cell viability and survival.

Changes in tissue architecture are often the first signs of cancer, but very little is known about the genes, proteins, and pathways that regulate cellular shape and polarity. **Senthil Muthuswamy** has developed a new paradigm for investigating this aspect of cancer biology. Using sophisticated model systems such as three-dimensional cell culture platforms and transgenic mice, his team recently found that a protein called Scribble normally coerces breast epithelial cells into the correct organization and shape and enforces resistance to cancer. Scribble is frequently missing in human breast cancer lesions, Muthuswamy’s team found, which suggests that it could be an appropriate target for therapies aimed at preventing precancerous lesions from becoming invasive. The lab has also identified new drug targets in patients with tumors that do not respond to chemotherapy. They identified an enzyme called Brk, which is overproduced in patients with HER2-positive tumors and helps them to become resistant to drugs such as Herceptin.

**Darryl Pappin**’s lab develops chemical and computational methods for analysis of proteins and peptides. These constitute fundamental tools for proteomics and can be applied across 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 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.

**Jacek Skowronski** and colleagues study mechanisms involved in the induction of AIDS by human and simian immunodeficiency viruses (HIV and SIV), focusing on the function of accessory proteins called Nef, Vpr, and Vpx. These virulence factors modify the cellular milieu to disrupt adaptive responses and/or innate antiviral responses and provide an environment conducive for viral replication. This year, Skowronski and colleagues discovered new details about how a simian strain of the AIDS virus replicates in macrophages, a type of immune system cell. The study revealed how Vpx enables efficient reverse transcription in the simian virus and thus overcomes an innate block that otherwise prevents viral replication. This suggests a strategy by which a future drug might interfere with the reproductive machinery of the virus to prevent or limit its ability to spread.

Despite their large variety of genetic abnormalities, cancer cells have been found to be extremely sensitive to the reversal of certain muta-
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 cells that have a particular epidermal growth factor (EGF) receptor mutation 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 are focusing on as yet unidentified mechanisms of resistance at work in other cases. They have preliminary in vitro data indicating a genetic signature predicting positive response to other treatment modalities. With colleagues at the National Institutes of Health, they are testing this hypothesis in tumor samples taken from relapsed patients.

Nicholas Tonks and colleagues study a family of enzymes called protein tyrosine phosphatases, or PTPs, that remove phosphate groups from other proteins. By changing the phosphorylation state of proteins, PTPs can profoundly affect the health of entire organisms. Tonks’ group seeks to characterize fully the PTP family, understand how their activity modifies signaling pathways, and how those pathways are abrogated in serious illnesses, from cancer to Parkinson's disease. The overall goal is to identify new targets and strategies for therapeutic intervention in human disease. This year, the lab made progress in characterizing mice in which the gene encoding JSP-1 (JNK stimulatory phosphatase 1) has been ablated, helping to define the role of this phosphatase in regulating a critical signaling pathway in Parkinson's disease. In collaboration with the Muthuswamy lab at CSHL, they also investigated new roles for PTPs in regulating signaling events in breast cancer and have identified three PTPs as potential novel tumor suppressors.

Several years ago, Lloyd Trotman discovered that the loss of a single copy of a master tumor-suppressing gene called PTEN is sufficient to permit tumorigenesis in animal models of prostate cancer. His team later found that complete loss of PTEN triggers senescence, a quiescent state that delays or blocks cancer development in affected cells. Recently, the lab has been exploring the impact of varying PTEN expression levels, alone and in conjunction with changes in other potent tumor suppressors, including p53, in mouse models for prostate cancer. They have studied how the PTEN protein is transported into and out of the cell nucleus and have sought out specific transport receptor proteins, identifying several strong candidates. They have also worked to validate their cancer progression scheme in human prostate cancer biopsy samples from the Memorial Sloan-Kettering Cancer Center, with the aim of identifying patients who have developed tumors with metastasis-favoring mutations.

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 to control cellular growth, differentiation, and morphogenesis. Alterations affecting Ras and Rho function are involved in cancer and various neurodevelopmental disorders. This year, Van Aelst’s team extended their study of mutations in a Rho-associated gene called oligophrenin-1, determining how they are linked to dysfunctions in the glutamate pathway, specifically to maturation and plasticity of excitatory glutamatergic synapses; they are probing the implications for several illnesses. Separately, they have demonstrated that a GTPase activator called DOCK7, previously shown to have a central role in axonal development and morphology, is important in neuronal migration and may be involved in defects in migration that cause developmental illness.
FACULTY LIST

Hollis Cline
Joshua Dubnau
Grigori Enikolopov
Hiroyasu Furukawa
Z. Josh Huang
Adam Kepecs
Alexei Koulakov
Bo Li
Zachary Mainen
Pavel Osten
Glenn Turner
Anthony Zador
Yi Zhong
Florin Albeanu (CSHL Fellow)

Neuroscience

Neuroscientists at CSHL are striving to trace the remarkable process by which the developing brain gives rise to immensely complex neural networks, catalog and comprehend the full range of human genes implicated in mental illness, and demonstrate with great precision—in animals and humans—how genetic mutations and a host of related cellular anomalies give rise to disease-specific pathologies. These basic science investigations promise to spawn not only unprecedented diagnostic capabilities, but also, it is hoped, important new strategies for treating and perhaps even preventing some illnesses. They are certain to help unravel the still obscure mechanisms at work in devastating neurodevelopmental illnesses such as schizophrenia, depression, bipolar illness and other mood disorders, and autism and related “spectrum” disorders, as well as neurodegenerative illnesses such as Alzheimer’s and Parkinson’s diseases. The Swartz Center for Computational Neuroscience, headed by Anthony Zador, and the Stanley Center for Cognitive Genomics are among the coordinating centers for this vital research.

Holly Cline’s lab is studying mechanisms in the brain controlling the growth of neurons, the generation of synapses, and the development of organized sensory projections between different brain regions. To understand the cellular events responsible for the stabilization of growing neuronal branches, the lab has delivered genes of interest into neurons and assessed the effects on synaptic transmission, specifically, in nerves that connect the retina and a part of the brain called the tectum involved in visual processing. Cline has learned that visual experience is a key modifier in how these circuits are built. This year, the lab demonstrated for the first time in living organisms (Xenopus tadpoles) that insulin receptor signaling in neurons regulates the maintenance of synapses, contributes to the processing of sensory information, and is involved in adjusting the plasticity of brain circuits in response to experience.

Because biological mechanisms of memory are highly conserved through evolution, many features of human memory are observed in simpler organisms such as fruit flies. Joshua Dubnau and colleagues identify genes that are important for memory and conserved across phyla. The hope is that many of these will be relevant to human memory. They recently discovered genes involved in controlling when and where specific proteins are synthesized within a neuron. These genes likely regulate neuronal communication during learning. The lab also seeks to discover how the neural circuitry of the fly brain works. They have recently shown that short- and long-term memories form in different sets of neurons: One circuit provides a memory that decays quickly and the other provides a memory that forms slowly but persists. Dubnau’s team also explores how groups of genes interact to form memories, an approach designed to shed light on complex gene networks that likely underlie human cognitive disorders.

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 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 have suggested to the team 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 also explains why the number of new neurons decreases with advancing age. Recently, Enikolopov’s group was part of a team that identified and validated the first bio-
A marker that permits neuronal progenitor cells to be tracked, noninvasively, in the brains of living human subjects. The lab is now using this discovery to reveal how neurogenesis is related to the course of diseases such as depression, bipolar disorder, Alzheimer’s, and Parkinson’s.

Hiro Furukawa’s lab is studying neurotransmission at the molecular level. They focus on two types of calcium ion channels: N-methyl-D-aspartate (NMDA) receptors and calcium homeostasis modulators (CALHM). Both are involved in the regulation of neuronal activities and in the pathogenesis of Alzheimer’s disease. NMDA receptors are very large molecules whose three-dimensional atomic structure Furukawa’s group has undertaken to solve by dividing them into several domains. They seek to understand the pharmacological specificity of ligand and modulator bindings in different subtypes of NMDA receptors in order to provide a blueprint for future drug design. A mutation of the gene coding for CALHM has been implicated in late-onset Alzheimer’s disease. The team is working to reveal the mechanism by which this calcium-specific ion channel opens by uncovering the molecule’s architecture. Such structural information promises to be useful in the design of novel drugs for treatment of Alzheimer’s.

Adam Kepecs studies the neurobiological principles by which the brain makes decisions. He and his colleagues view decisions as elementary units of behavior, from which more complex behaviors are assembled. Yet even simple decisions involve the interaction of sensory and memory information with emotional and motivational attributes, requiring the concerted action of millions of neurons across brain regions. Therefore, they take an integrative approach, combining experiments involving well-controlled rodent behavior with electrophysiology, molecular perturbations, and quantitative analysis. Their current work seeks to elucidate the neurocomputational principles of decision-making, attempting to capture more elusive attributes such as emotion, motivation, or confidence. This year, Kepecs, in a collaboration with Zachary Mainen and others here at CSHL, discovered neural signals for confidence in the rat prefrontal cortex. Their study suggests that confidence estimation is a fundamental information-processing mechanism in the brain, shared widely across species, and not strictly confined to those, such as humans, who are self-aware.

The mammalian brain is a miracle of self-assembly, a process that begins prenatally and continues well into the postnatal period. 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
ing specificity and to remember such distinctions over time. These are among the challenges that engage Koulakov and his team.

Dysfunction of excitatory, glutamatergic synapses in the brain is believed to have an important role in the pathogenesis of major psychiatric disorders, including schizophrenia and depression. But what are the causes? Where in the brain does this dysfunction occur? How does it result in the behavioral symptoms of illness? To address these issues, Bo Li and colleagues are studying normal synaptic plasticity and disease-related synaptic changes in brain circuits involved in schizophrenia and depression. Their long-term goal is to develop methods allowing the manipulation of activity in specific brain circuits in order to change disease-related behaviors. They will use 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 virus injection, 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.

To better understand neuronal circuits, Josh Huang and colleagues have developed novel means of visualizing the structure and connectivity of different cell types at high resolution in living animals and of manipulating the function of specific cell types with remarkable precision. Huang is particularly interested in circuits that use GABA, the brain’s primary inhibitory neurotransmitter. The lab’s work has direct implications in neurological and psychiatric illness such as autism and schizophrenia, which involve altered development and function of GABA-ergic circuits. This past year, they discovered that neurons connect to very specific partners at very specific spots in the developing cerebellum, thanks to an underlying framework of molecular “guides” called glial cells that nudge nerve fibers to grow in the right direction and make the right contacts. The lab also made good progress in studying perturbations in the developing GABAergic system in a mouse model of Rett’s syndrome, one of the autism spectrum disorders.

Zachary Mainen studies the neural basis of mammalian behavior and cognition. He focuses on understanding the nature of the electrical activity of single neurons, the “currency” with which genes and molecules express themselves in the functioning brain. His lab also studies neural coding, or how information is represented in characteristic “spikes” of neuronal electrical discharges. This year, in collaboration with Adam Kepecs and others here at CSHL, he published results of studies in rats suggesting that the estimation of confidence that underlies decision-making may be the product of a very basic kind of information processing in the brain, shared widely across species. Separately, Mainen showed that the superior colliculus, a structure implicated in visually guided eye movements, has a much more general role in spatial decision-making than previously thought.

Pavel Osten’s lab is developing several approaches to uncover pathologies responsible for mental disorders, including autism spectrum disorders (ASDs) and schizophrenia, and neurodegenerative illnesses such as Parkinson’s and Alzheimer’s. In neuropsychiatric disorders, Osten is exploring the link between candidate genes and common “core” dysfunctions at the level of neural circuitry, specifically in mouse models of schizophrenia and autism. The ability to systematically and rapidly compare brain functions, analyze which brain regions are affected, and determine which cellular networks within these regions are altered is vital in the larger effort to understand how genes perturb biology, giving rise to symptomatic behaviors. Osten and collaborators are also creating automated systems that will take a sample of brain tissue as input, generate a three-dimensional representation of its neural circuit activity, and then map it to a brain atlas. The object is to reveal how neural circuits are
wired during development, to make possible precise characterization of neural circuit disruptions in mouse models.

When we learn an association, information from two different sensory streams is somehow bound together in the brain. For example, the smell of spoiled milk vividly evokes the taste of an injudicious gulp. How are odor and taste information represented in terms of neural spike trains, and how does learning modify those representations to form associative links among stimuli? These are among the questions that drive research in Glenn Turner's lab. His team addresses these questions by using a combination of electrophysiological, genetic, and computational approaches in the fruit fly. They directly monitor activity of neurons in the brain of an intact fly with whole-cell patch-clamp recordings. Using molecular genetics techniques, they are able to manipulate neural activity to directly test their predictions about neural coding, at the level of both spike trains and behavior. Currently, they focus on representations of smell and taste.

Anthony Zador and colleagues study how brain biology gives rise to higher-level properties such as complex behavior. They have focused on how the cortex processes sound, how that processing is modulated by attention, and how it is altered in pathology. In the lab’s “core assay,” the response of single neurons to sound stimuli is examined under distinct behavioral conditions. In animal models of autism, the team is trying to link an inability to screen out background sounds with changes in neural circuits. Separately, by showing that a very small minority of available auditory neurons in a rat cortex react strongly when exposed to a specific sound, the lab has challenged the standard model of sound representation. This year, Zador’s team generated evidence supportive of an alternative theory of information processing in the brain, showing that animals in the midst of decision-making have the ability to distinguish incoming signal spikes separated by as little as 3 milliseconds. This lends credence to a timing model of information processing as opposed to one based on the rate of signal firing.

Yi Zhong’s lab studies the neural basis of learning and memory. They work with fruit fly models to study genes involved in human cognitive disorders, including neurofibromatosis, Noonan’s syndrome, and Alzheimer’s disease. Mutations of the neurofibromatosis 1 (NFI) gene cause learning defects and neurofibromas—nerve-sheath tumors that split apart nerve fibers. The lab’s analyses of Drosophila NFI mutants have revealed how expression of the mutant gene affects a pathway crucial for learning. They have also discovered that the NFI gene and a gene called corkscrew, both implicated in Noonan’s, share a biochemical pathway. This year, they linked specific genetic defects in Noonan’s with long-term memory deficiencies that are among the symptoms of the illness. They postulate a “spacing effect,” which, if addressed with remedial learning methods, might help to address the impairment. Zhong also contributed to a study in which plaques implicated in Alzheimer’s memory loss were experimentally reduced in the fruit fly brain by overexpressing a human gene that codes for the production of an enzyme called neprilysin, or NEP.

How does the brain encode stimuli from the outside world to generate specific perceptions that, in turn, trigger complex behaviors? How is the brain shaped by sensory experience and what modifications occur in neuronal circuits that allow us to learn and remember? These are questions guiding the work of CSHL Fellow Florin Albeanu, who is using the olfactory bulb in living mice as the subject of his current studies. Rodents depend on olfaction for finding food and mates and avoiding predators. Airborne chemicals, translated into neuronal signals by specific receptors in the nose, are sent directly to the olfactory bulb. Although readily accessible, little is known about the neuronal processing that occurs there. Technological advances in optical imaging enable Albeanu to monitor patterns of activity at unprecedented synaptic resolution, in real time, as animals “behave.”
Bioinformatics and Genomics

Bioinformaticians in literally dozens of labs are devising new capabilities that their colleagues are putting to novel uses in their experiments. These technologies, it is universally acknowledged, are prerequisites for satisfactory progress in diverse areas of research, from cancer to neurological diseases such as Alzheimer’s and Parkinson’s to neuropsychiatric illnesses such as schizophrenia and autism to all aspects of plant biology and genetics. At the Lita Annenberg Hazen Genome Sequencing Center, directed by W. Richard McCombie, CSHL scientists have been among the leaders in devising, refining, and updating techniques that enabled the sequencing of human and other “model” genomes earlier in this decade. They are currently generating the technical and conceptual means for labs worldwide to use a new generation of sequencers that are paving the way to low-cost individual genomes and an era of personalized medicine.

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 at the beginning of the decade. Having brought online a new generation of Solexa sequencers at CSHL this past year, McCombie’s team optimized their function to a level at which 10 billion DNA bases (single “rungs” on the double-helical DNA “ladder”) can be sequenced in a typical day, and on some days as many as 20 billion. McCombie has been a leader of efforts to sequence the flowering plant Arabidopsis thaliana, the fission yeast Schizosaccharomyces pombe, and Homo sapiens. New projects are under way to sequence genes of special interest, including DISCI, a strong candidate gene for schizophrenia, and genomic regions likely implicated in bipolar illness. With the Memorial Sloan-Kettering Cancer Center, they are using a method called hybrid resequencing, developed with Greg Hannon here at CSHL, to look at mutations in samples collected from patients with prostate cancer.

Jonathan Sebat’s laboratory is studying the role of genetic variation, and particularly gene copy-number variation, in schizophrenia, autism, and other neuropsychiatric illnesses. This past year, Sebat and collaborators discovered that a significantly increased rate of rare structural mutations exists in the genomes of individuals with schizophrenia as compared with healthy controls. They found, moreover, that the mutations were powerful, increasing individual risk 10–20 times, and that in people with schizophrenia, the genes disrupted by the
mutations nearly half the time were involved in pathways known to be important in brain development. One implication is that the universe of genetic risk factors for schizophrenia consists of many different rare mutations, each one present in comparatively few individuals or even a single one. Thus, the paradox: Despite the large constellation of rare mutations contributing to the disease, one must sift through thousands of cases before any one of them is likely to appear more than once. In prior work, Sebat, in collaboration with Michael Wigler at CSHL, discovered that spontaneous mutations—genetic errors in children that do not occur in either parent—are far more common in autism than previously thought.

Lincoln Stein’s lab is developing databases, data analysis tools, and user interfaces to organize, manage, and visualize the vast body of information being generated by biologists. One of the unsolved mysteries of the genome is how its genes are precisely regulated to promote orderly growth and development and respond to changes in the environment. The modENCODE project (model organism encyclopedia of DNA elements) is an international consortium organized to find and characterize the elements that regulate the genomes of the fruit fly and the roundworm. Last year, the Stein lab established the modENCODE Data Coordination Center, which is responsible for collecting, integrating, and publishing the information collected by the consortium in a form that can be extended and combined with information from other human and model organism genome databases.

The HapMap project is another focus of Stein’s lab, a database of human single-base-pair variations. Last year, the HapMap database was expanded from information on the genetic variability of 4 populations to 11 human populations, including individuals from Africa, Asia, Europe, Mexico, and the Indian subcontinent. The variability information is integrated with recently published information on more than 50 susceptibility regions in 20 common human diseases, including diabetes, rheumatoid arthritis, Crohn’s disease, coronary artery disease, and bipolar disorder.

The Stein lab also manages and curates the WormBase and Reactome databases. WormBase, a database of the genome and biology of the roundworm Caenorhabditis elegans, gives users quick access to the large C. elegans literature. It has been integrated by Stein’s group with an open-source online publication called WormBook, which now has 140 monograph “chapters” that link to gene entry data in WormBase. Reactome, an interactive database of fundamental biological pathways in humans, integrates the peer-reviewed literature with high-throughput genomic information, including a growing database of protein–protein interactions (now covering more than half of the annotated genome). Finally, Stein’s group has begun work on the iPlant Collaborative, an effort to better enable plant biology researchers to collaborate in cyberspace.

Using multidisciplinary approaches that combine computational analysis, modeling, and prediction with experimental verification, Doreen Ware’s lab seeks a deeper understanding of genome organization in plants. By looking comparatively across the genomes of plants in the same lineage, Ware and colleagues seek answers to such questions as, How are genes conserved and lost over time? What are the fates of duplicated genes? Her team also studies gene regulation in plants, specifically looking at cis-regulatory elements and characterizing microRNA genes and their targets. This past year, the lab has been responsible for providing annotations of the maize genome for an international consortium. Within their own lab, they have also brought new, fully sequenced genomes into an existing integrated data framework, to enhance the power of their comparative studies. This framework now encompasses the genomes of Arabidopsis, maize, rice, sorghum, grape, and poplar. They have devoted special attention to examining diversity within species, particularly maize and grape, in assays that make use of powerful new genotyping technology.
Quantitative Biology

Mathematical and statistical insights are at the very heart of the technologies that have made possible comparative genomic studies pioneered at CSHL. Thanks to the foresight of two major private donors, James Simons and Henry “Hank” Greenberg, 2008 marked the unofficial launch of a new Center for Quantitative Biology (CQB), which in the coming year will take up residence in a new building in the Hillside Research Complex. This new research focus at CSHL, under the interim leadership of Michael Wigler, reflects the importance of math and statistics in the forging of new approaches to problems and new ways of understanding the results of biological experiments. The program is bringing together under one roof some of the world’s most gifted mathematical minds, who will apply their insights in the formulation of research hypotheses pertinent to the study of molecular biology. Beyond individual research applications, the work of CSHL’s quantitative biologists has the potential to revise fundamental notions of how we think about data itself.

Gurinder S. “Mickey” Atwal and colleagues are applying insights from the physical sciences to the study of biological phenomena. Specifically, they develop and use mathematical and computational tools to address quantitative principles governing the behavior of correlated “many-body” biological systems. Such systems range from molecular interactions in a single cell to the evolution of Homo sapiens. They are now seeking to understand evolutionary forces acting on the genome in the context of human diseases. In collaborations with colleagues at the Institute for Advanced Study at Princeton University, Atwal has modeled the process by which genetic variants, or alleles, evolve. This has recently led to surprising insights about the role that p53, a master tumor suppressor gene, has in reproduction. This work also bears on the question of demonstrating recent selective pressures acting on our genomes.

Partha Mitra seeks to develop an integrative picture of brain function, incorporating theory, informatics, and experimental work. His theoretical interests are primarily in formalizing the treatment of biological function using ideas and methods from engineering. In informatics work, his lab is developing computational tools for integrative analysis of neurobiological data, spanning electrophysiological, neuroanatomical, and, more recently, genomic data, from multiple species pertaining to the brain. Mitra has organized the Brain Architecture Project, a multi-institutional effort to curate information from the literature about human neuroanatomical connectivity that will also advocate for large-scale studies of connectivity in model organisms. This past year, the lab launched http://brainarchitecture.org, a website featuring full-text search for more than 55,000 papers. Separately, Mitra and colleagues progressed in their efforts to develop experimental methodologies for postmortem human brain tract tracing, a means of tracing neuronal connections in brain areas more rapidly and over greater distances than previously possible. Mitra also recently presented results of neurobiology research demonstrating the de novo evolution of song culture in zebra finches, starting from an isolate population.

Michael Wigler’s group uses methods for comparative genome analysis to study cancer and human genetic disease. These methods (called representational oligonucleotide microarray
analysis [ROMA] and methylation-specific oligonucleotide microarray analysis [MOMA]) evolved from an earlier technique called RDA (representational difference analysis), used to find tumor suppressors, oncogenes, and pathogenic viruses. Current microarray-based techniques, including comparative genome hybridization (CGH), reveal changes in the numbers of copies of sections of the genome and regions of deletion and duplication, mutations that may underlie the evolution of cancers. Wigler’s group focuses on breast cancer and leukemias, and they are engaged in clinical studies with major research hospitals to discover mutation patterns predicting treatment response and outcome. In collaboration with Jim Hicks and Gregory Hannon at CSHL, the lab has applied microarray techniques and hybrid selection to explore the role of epigenetics in cancer. With Jonathan Sebat, they have made headway in the discovery of the causative mutations in autism. Their results show that spontaneous mutation has a far greater role in autism than previously suspected. They have developed a new theory of autism’s genetic basis that explains otherwise bewildering patterns of inheritance and are testing the new genetic model in other disabling genetic disorders. Wigler also has spearheaded the development of the Center for Quantitative Biology at CSHL, with initial funding from the Simons and Starr foundations.

Michael Zhang’s laboratory develops mathematical and computational methods that can be combined with advanced experimental technologies to transform data into biological knowledge about transcription and gene expression, work that has manifold implications for the study of cancer and many other diseases. Their tools, used by investigators throughout CSHL and beyond, are designed to identify functional genetic elements within molecular sequences as well as pathways that control and regulate gene expression. Zhang’s group has developed a series of computational tools that make use of statistical pattern-recognition techniques to identify exons, promoters, and posttranslational modification signals in large genomic DNA sequences. They also study alternative splicing of exons and collaborate with other labs to characterize splicing enhancers and silencers. This past year, Zhang’s team published widely cited studies on optimization of sequencing with next-generation Solexa equipment, using software to predict promoter regions at high resolution, and, in collaboration with Adrian Krainer at CSHL, defining the regulatory networks of two related tissue-specific (brain and muscle) splicing factors, Fox-1 and Fox-2.

Rapid growth of the research literature in genomics and molecular biology threatens the ability of scientists to digest and make use of valuable data. CSHL Fellow Ivan Iossifov seeks to develop automated methods for extracting, structuring, and interpreting research data. Building on text-mining methods that he developed in his Ph.D. work, losiifov is now focusing on using these to build representations of molecular networks for humans and for two model organisms, Mus musculus (lab mouse) and Saccharomyces cerevisiae (baker’s yeast). In collaboration with Partha Mitra at CSHL, he is also working on a means of organizing information about connections among brain regions, work that could prove valuable in efforts to trace the etiology of common neurodevelopmental disorders such as autism, schizophrenia, and bipolar disorder.
Plant genetics has long had a major role in the advance of all areas of research at CSHL, epitomized by the work of Nobel Laureate Barbara McClintock, who discovered transposable genetic elements—which she called “jumping genes”—in the 1940s. The contemporary effort to understand the genetic and cell biological mechanisms that enable plants to grow and develop has obvious implications not only for agriculture, but also for human health, as well as the development of next-generation biofuels and related alternative energy sources. Plant geneticists at CSHL have been among the leaders of efforts to sequence the first plant genomes, of *Arabidopsis* and maize, and continue to be at the center of sequencing and genome-annotation projects involving a host of cereal crops that feed the growing population of our planet. CSHL scientists have been among the pioneers in the study of RNAi, a gene-regulatory mechanism unknown only a decade ago, which has been shown, first in plants, then in other organisms, to have a vital and evolutionarily ancient genome-defending role.

David Jackson and colleagues study genes and signals that regulate plant growth and architecture. This year, they built on past success in identifying a gene that controls communication between plant cells via small channels called plasmodesmata. These channels, which direct the flow of nutrients and signals through growing tissues, are regulated during development. Jackson’s team performed a genetic screen and found some mutants that are affected in transport, one of which was found to affect redox signaling. This suggests a mechanism through which plant cells can adjust trafficking in the channels in different developmental stages. The lab also cloned a gene called *tasselsheath*, whose activity they correlated with the shutting down of leaf growth when flowering gets under way. Jackson’s past success in developing fluorescent-protein-tagged maize lines bore fruit when a team that included Jackson this year identified a gene called *sparse inflorescence1* that proves to be vital in the formation of corn “ears.”

Zachary Lippman’s research focuses on a universal growth habit in plants called sympodial growth, seen in widely varying species from orchids to tomatoes to trees. It is characterized by a renewal of growth that follows a cessation of growth following prior flowering, resulting in an array of stems. Lippman’s lab seeks to understand how sympodial growth affects the vigor and reproductive success of a plant, specifically its ability to produce vegetative shoots and flower-bearing shoots called inflorescences. In the tomato plant and other members of the plant family to which it belongs, called *Solanaceae*, they study genes that control the branching of flowering shoots and how those genes function to control flower production. Their
aim is to decipher the molecular basis of the natural variation for inflorescence branching seen in that important family, with an eye to increasing crop yields.

Epigenetic mechanisms of gene regulation—chemical and conformational changes to chromatin bundles of DNA and protein—have 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, associated with plant color diversity and caused by inactivation of a gene positioned near densely packed chromosomal material called heterochromatin. They have found that heterochromatin is programmed by small RNA molecules arising from repeating genetic sequences. This past year, they addressed a nagging conundrum: Because genes contained within heterochromatin are silenced, how can they give rise to RNA molecules that help to modify the histone proteins around which DNA is spooled? They solved the puzzle by tracking cells through their cycle of growth and division. They also mapped changing epigenetic modifications enabling mobile genetic elements to run amok within a genome.

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 arise. They also produce signals important for the determination and patterning of lateral organs. Marja Timmermans and colleagues are using a genomic approach to study genes active in the meristem. They have used mutational analyses to identify a protein complex that suppresses stem cell fate during organ development. This complex includes the chromatin remodeling factor HIRA, an epigenetic regulator that helps to control stem cell identity. This year, the lab made progress in elucidating the molecular mechanism through which HIRA acts, demonstrating its role in specifying stem cell niches during embryogenesis and, later on, after commitment to differentiation, in shutting off stem cells. They also identified regulatory mechanisms that allow for the precise spatial accumulation of developmentally important small RNAs in plants. This work revealed that polarity in leaves is established via mobile small RNAs; a description of the mobile small RNA, tasiR-ARF, in pattern formation suggests possible roles of other small RNAs as morphogen-like signals during development in plants and animals.