# Technology Innovation

# Three Decades of Technology Innovation at the Beckman Center

BY ASHLEY P. TAYLOR

Since the founding in 1989 of the Beckman Center for Molecular and Genetic Medicine, scientists and engineers have developed a wealth of technologies for learning about how life works at the levels of cells, molecules and genes. Scientists have figured out how to turn back the developmental clocks of ordinary, differentiated cells so that they can, like the cells in an embryo, differentiate into many different cell types. Scientists have sequenced the human genome, as well as the genomes of hundreds of other animals, not to mention those of plants, fungi and microorganisms. Now, whole-genome sequencing occurs on the level of individual patients. With the rise of the internet have come bioinformatics databases for analyzing genomic data. Beyond observing genomes, scientists have learned to edit them at precise locations using a technology called CRISPR. The list of recent developments in biotechnology could go on and on.

Many such groundbreaking technologies were developed at the Stanford Beckman Center. Ron Davis, PhD, professor of biochemistry and of genetics and director of the Stanford Genome Technology Center, and emeritus professor of biochemistry Patrick "Pat" Brown, MD, PhD, were instrumental in developing and first applying the microarray, which revolutionized molecular biology by allowing researchers to simultaneously examine the activity of thousands of genes. Also, in the 1990s, the Douglass M. and Nola Leishman Professor of Cardiovascular Disease and professor of biochemistry James "Jim" Spudich, PhD, developed a system for watching the interactions of purified muscle proteins and measuring the forces involved. In that effort, he collaborated with William R. Kenan, Jr., professor of physics and professor of molecular and cellular physiology, Steven "Steve" Chu, PhD. Chu is now developing novel optical probes to label tiny subcellular structures so that scientists can better observe them and thus learn more about how they work. Using protein engineering, professor of molecular and cellular physiology Chris Garcia, PhD, is developing new molecules, which he believes have great potential as drugs, that bind to receptors on the surfaces of cells and affect activity within. Associate professor of developmental biology, of computer science, of pediatrics and of biomedical data science Gill Bejerano, PhD, has developed computational tools to help researchers better understand the human

***Patrick Brown, MD, PhD***
*Emeritus Professor of Biochemistry*

genome and to aid clinicians in diagnosing genetic diseases. And associate professor of biochemistry and, by courtesy, of physics Rhiju Das, PhD, has developed online videogames that recruit citizen scientists to help figure out how RNAs fold, knowledge Das is using to design RNAs for diagnostics and therapeutics.

## The Microarray: A Tool for Comparing Gene-Expression Patterns and More

The genesis story of the microarray begins in the early 1990s with two professors in the Department of Biochemistry, Ron Davis and Pat Brown, as well as Stephen Fodor, PhD, who was then working at a company called Affymax in Palo Alto. Today, all three scientists have won awards for their contributions to microarray technology, among other achievements. Affymetrix, the microarray company that Fodor founded in 1993 and which ThermoFisher Scientific acquired in 2016, became one of the major commercial suppliers of microarrays. But the technology that would later be called revolutionary was at first met with resistance. Brown's involvement, for example, started with a rejected grant application.

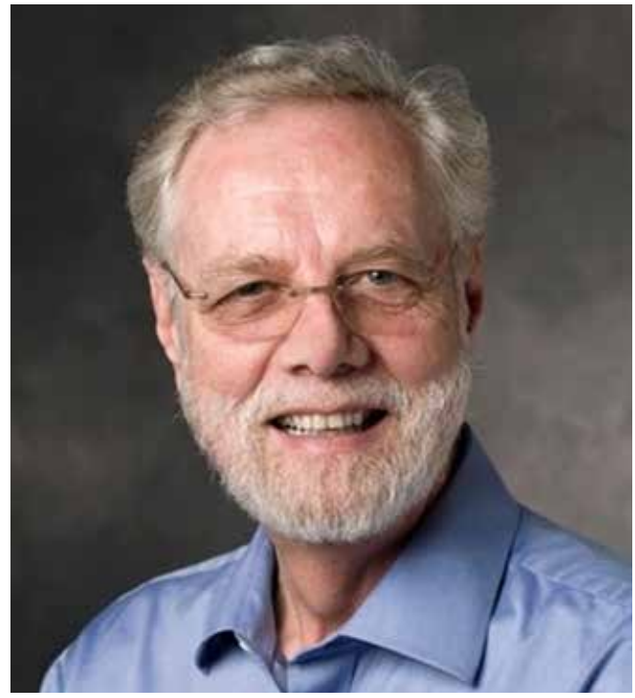
In 1992, Brown applied for federal funding to develop the microarray and for another unrelated project. The grant proposal was not well received. "It got the worst priority score I'd ever gotten or ever seen," says Brown. The National Institutes of Health (NIH) suggested Brown cut the microarray stuff and resubmit the proposal. Ignoring the NIH's advice, he developed the microarray anyway.

"What people didn't like about it was, it just seemed insanely ambitious, and they just didn't believe that it was practical, which they were wrong about," says Brown.

The DNA microarray, which is now a standard part of the molecular biology toolkit, is a way of comparing gene activity in two groups of cells. Specifically, what researchers compare on a microarray are the RNAs present in the two cell groups: both their identities—which genes they come from—and their quantities. Since RNAs are produced from genes and made into proteins, they can essentially serve as proxies for both gene and protein activity. What's exciting about microarrays is that they allow researchers to do this gene-expression comparison not just for one gene, but for thousands of genes, at the same time.

The classic image of a microarray is a grid of red and green dots. Each dot represents a gene, and the color of the dot indicates the relative expression of that gene in two different cell types. But this image represents a finished experiment.

The first step is to generate the microarray itself by attaching small segments of DNA to a substrate. Brown did this by building a robot that would drop tiny samples of DNA from tens, hundreds or thousands of genes onto a glass slide. Davis started out trying to print DNA spots on nylon, but as the nylon was unstable, he joined



***Ron Davis, PhD***
*Professor of Biochemistry and of Genetics*

***Pat Brown and Ron Davis were instrumental in developing and first applying the microarray, which revolutionized molecular biology by allowing researchers to simultaneously examine the activity of thousands of genes.***

Brown in working with arrays spotted on glass. Davis also collaborated with Fodor and others at Affymetrix on experiments in which the array DNA was synthesized directly onto the slide. Next, researchers isolate RNA from the two cell types they want to compare, convert it to fluorescently labeled DNA—red for one cell type, green for the other—and expose the microarray to the two labeled samples. Analyzing the fluorescence reveals the results: if a dot is mostly red or mostly green then the gene in question produces RNA at different levels—or is differentially expressed—between the two cell types. Differential gene expression could potentially explain how the two cell groups are phenotypically different, such as why one group is healthy and the other diseased.

Both Brown and Davis originally envisioned using microarrays for other purposes, but gene-expression analysis is the application for which microarrays are known. In their first paper published about microarray experiments, in 1995, Davis and Brown used Brown's spotted arrays and RNA samples from Davis' lab to compare the expression of 45 genes between the roots and leaves of the small flowering plant and model organism *Arabidopsis thaliana*. The paper illustrated how microarrays could be used and, in its conclusion, laid out the microarray's potential for studying gene expression in different human cell types and for detecting disease-associated gene-expression patterns that could be used for diagnosis. In myriad papers that followed, Davis, Brown and others used microarrays to do just that.

In 1996, Brown and Davis reported using spotted microarrays to analyze changes in gene expression that occurred in human T cells in response to two different experimental treatments: heat shock—exposing the cells to higher temperatures—and treatment with phorbol ester, which sets off a molecular signaling cascade. The first microarray experiment done on human tissue, this was, "groundbreaking," says Davis. The experiment was also significant because the genes on the array were of unknown function, and through this experiment, the researchers deduced something about what the differentially expressed genes did—that certain number of them were involved in responding to heat shock, for instance.

The following year, Brown and Davis published the results of the first ever genome-wide microarray experiment, which was in yeast. The experiment didn't include every gene in the yeast genome, but future experiments would. There are now microarrays that cover the entire human genome.

In 1998, Brown published a paper, which has now been cited more than 17,000 times, according to Google scholar, reporting that in general, genes with similar functions tend to be expressed together—that is, at the same time and under similar conditions. Knowing this could help scientists deduce the functions of uncharacterized genes, Brown and coauthors explained. Grouping genes with similar expression patterns and functions, which is called one cluster analysis, also just gave researchers a broad-strokes view of cellular activity—here are the genes involved in cell-division; another cluster are the genes involved in making proteins—that wasn't possible through examination of individual genes. "A very

large body of systematic data is worth way more than the sum of the parts," says Brown. "Because by having the systematic data, you can really see the underlying logic in a way that you can't just with occasional snapshots."

Microarrays can be used to study how a particular process, such as sporulation in yeast, or a regulatory molecule, such as a hormone, changes gene expression. They can also be repeated over time, comparing a sample undergoing some change, such as heat shock or progression through the cell cycle, to a control sample not undergoing that change. Brown, Davis or at times both researchers, did experiments to examine all of the above.

Starting in the late 1990s, Davis began to have doubts about the accuracy of spotted arrays. It was easy to make a mistake about which DNA sample you had put in each position, Davis says. For that reason, while Brown continued to use spotted arrays, Davis pursued experiments with Affymetrix arrays. One such project involved exploring the functions of each of the yeast genome's approximately 6,000 genes.

As reported in 2002, Davis led a consortium of researchers in creating a library of yeast strains, covering the entire yeast genome, in which one gene had been deleted and replaced with two different DNA sequences, called molecular barcodes, to identify the strain. In a typical experiment, Davis would grow the yeast strains under different conditions, such as heat shock, and see how well they coped with those conditions based on how well they grew. How well they grew could be determined by making fluorescently labeled copies of the barcode sequences and hybridizing them to the microarray, which had a spot for every pair of barcodes. To take the heat-shock example, if a particular microarray spot fluoresced less, Davis could infer that the strain had failed to grow well at high temperatures and that the deleted gene must be important for coping with heat shock. Because each strain had an identifying barcode, it was possible to mix together all 6,000 strains in one tube throughout the experiment—quite a feat.

"It would be a typical experiment that a person might do, but it would be done in the past one at a time and you'd have to do 6,000 such experiments. But this allowed you to do one experiment and read out the results for 6,000 genes," says Davis. Experiments using the deletion collection are ongoing.

This experiment illustrates what Davis sees as the overall importance of the microarray: it changed the way biologists approached their research. "Up until microarrays were developed, people were doing things one at a time," says Davis. They might examine the expression of one gene, realize they needed to do the same for another gene or two, and end up doing several experiments one after the next. Microarrays allow researchers to do thousands of experiments at the same time. "I think the biggest impact is teaching people that you could do things in parallel, as opposed to serially," Davis continues. DNA is not the only molecule scientists put on microarrays. There are arrays of proteins, of sugars, of antibodies. In all of these experiments, scientists are doing, in parallel, experiments they might otherwise have done one at a time. That approach to science paved the way for other large-scale experimental

***Microarray technology also ushered in the era of precision medicine. "To do precision medicine you need lots of data on an individual, and you have to do it inexpensively or it won't be developed," says Davis.***

*Microarrays have become a standard way of determining which subtype of cancer a patient has and, accordingly, whether and how to treat it, says Brown.*

efforts, such as automated DNA sequencing, Davis adds.

Much microarray work also has medical applications. Microarrays can help scientists doing basic research learn about gene-expression changes that occur in various diseases. They also, Davis says, ushered in the era of precision medicine or tailoring treatments to individual patients. "To do precision medicine you need lots of data on an individual and you have to do it inexpensively or it won't be developed," says Davis. Microarrays were the technology needed to gather those data.

A prime example of the use of microarrays for precision medicine is in the area of cancer diagnostics and treatment planning. Brown's lab, and others, characterized gene-expression patterns in different cancers and, in the process, learned that some cancers previously thought of as being monolithic groups had subtypes that were distinguished by their gene-expression patterns. "By recognizing those subtypes as distinct kinds of tumors, they could be treated differently," says Brown. Microarrays have become a standard way of determining which subtype of cancer a patient has and, accordingly, whether and how to treat it, says Brown.

Davis's work has also had precision-medicine applications. For example, in 2012, Davis contributed to a paper in which microarray analysis was used to identify five genes whose expression in patients' blood cells changed depending on whether or not a kidney transplant was rejected. Further, in a test group of blood samples from transplant patients, expression of these five genes could be used to diagnose transplant rejection.

Davis also collaborated on a project in which scientists identified gene-expression changes in the blood that could distinguish, 12 hours after a trauma, such as a car accident, between patients likely to have a smooth recovery and those likely to go into multiple-organ failure after a few days. As reported in January 2019, this work has led to development of a test that can predict, within 24 hours after a trauma, which patients will have organ dysfunction and other complications in order to better treat them.

The microarray technology developed in the 1990s is now being replaced by DNA sequencing, according to Davis, who says that sequencing is now a more efficient way to do many of the tests for which microarrays were previously employed. "But that's not surprising. That's what happens with technology," says Davis. "It exists for a while and it's replaced. But it had a big impact on the community."

And whether the experiments are done with microarrays, sequencing or other methods, the underlying concept, which originated with microarrays, of simultaneously screening for thousands of different molecules and using the information to learn about biology and human disease is of enduring value, says Davis.

## The Dual-Beam Laser Trap: A Tool for Characterizing Molecular Motors

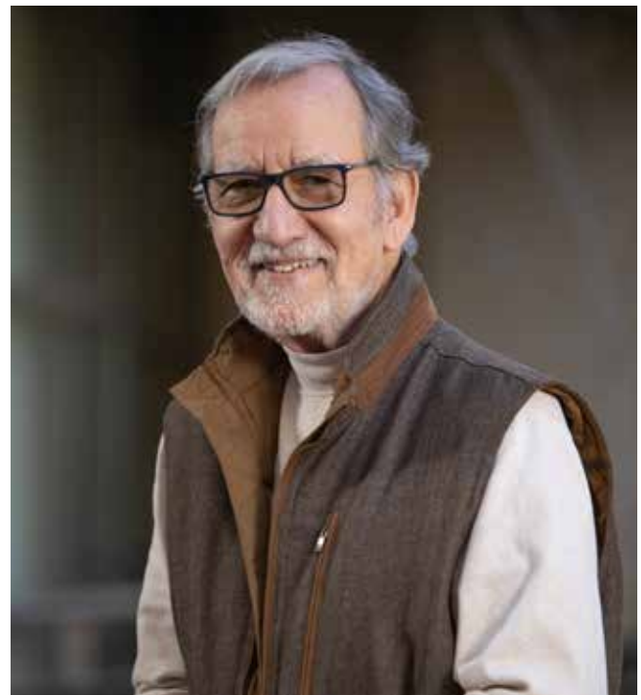
Around the same time that Brown and Davis were developing the microarray, Jim Spudich, then a professor of structural biology, was trying to understand how two essential muscle proteins, actin and myosin, interact to make muscles contract.

Scientists now know, from Spudich's work and that of others, that muscles contract when myosin molecules reach across and attach to actin filaments, pull the actin filaments back by 5 to 10 nanometers, then recock and repeat the cycle, like sailors pulling on a rope in a "heave ho" motion. Each cycle, which moves the actin filament by about one one-hundredth the width of a human hair, is powered by consumption of one molecule of the cellular fuel ATP. This results in muscle contraction because within the functional unit of the muscle, called the sarcomere, there are two sets of myosin molecules, both of which are pulling the actin filaments toward the sarcomere's center, shortening it. Evidence that actin and myosin interact this way was provided by a paper that Spudich and colleagues published in 1994. But in the 1980s, it was just a model.

In the '80s, Spudich and a graduate student, Stephen Kron, MD, PhD, now a professor of molecular genetics and cell biology at the University of Chicago, developed a method for watching purified actin and myosin proteins interact under the microscope. "When you added ATP, these actin filaments moved around on the surface in a directed fashion, being propelled along by something that these myosin molecules were doing to make the actin move," says Spudich. Spudich and Kron also showed that movement required just actin, myosin and ATP; no additional accessory sarcomeric proteins were needed. These experiments, "established that in fact, the model that actin moved along myosin, which was in all the textbooks, was correct," says Spudich.

But there were still open questions. Spudich's data suggested that myosin moved actin by 5 to 10 nanometers when it burned one ATP, but others had data suggesting that myosin moved actin by 100 nanometers or more at a time, which would not be consistent with the model favored by muscle investigators for actin-myosin interactions, that heave-ho motion, called the swinging cross bridge model.
"And that's sort of where we were in the late 1980s," says Spudich. "It was clear that somebody was going to have to design a way to watch just one myosin molecule interact with a single actin filament and see how far it moved." That's exactly what, with help from a graduate student, a sabbatical visitor and fellow Stanford faculty member Steve Chu, he was able to do in the early '90s. Just as the endeavor to understand how actin and myosin interact depended on work that Spudich had done in the '80s, so it relied on work that Chu had done during that period, when he was a scientist at Bell Labs in New Jersey.

The idea of holding onto biomolecules grew out of the work that Chu did on how to hold onto atoms in a vacuum with light. The idea had been proposed by Chu's Bell Labs colleague Arthur Ashkin in 1978, but the optical forces that light could exert on atoms was far too feeble to hold onto fast-moving atoms. In 1984, Chu showed that a set of counterpropagating laser beams, which he called, "optical molasses," could be used to slow atoms from the speed of a supersonic



***James Spudich, PhD***
*Professor of Biochemistry*

*The idea of holding onto biomolecules grew out of the work that Chu did on how to hold onto atoms in a vacuum with dual-beam lasers, for which he won the 1997 Nobel Prize in Physics.*

jet plane to that of a scurrying cockroach. At those speeds, the atoms were cooled to less than 0.0001 degree above absolute zero, but once that cold, the atoms were easily trapped. For his laser cooling and atom-trapping work, Chu won the 1997 Nobel Prize in Physics.

The same physics applies to any size particle, but the optical forces are proportional to the volume of the particle. For a micron-sized particle, water at room temperature could substitute for the exotic optical molasses. In the same year atoms were trapped, Ashkin and Chu showed that the optical tweezers could be used to hold onto polystyrene spheres cooled by water. In the following year, Ashkin showed that optical tweezers could be used to trap bacteria and viruses. For the invention and demonstration of optical tweezers for biological applications, Ashkin received the 2018 Nobel Prize in Physics.

In the fall of 1987, Chu moved from Bell Labs to Stanford to join the Departments of Physics and Applied Physics. Biological molecules were too small to be held directly with optical tweezers at room temperature, but Chu reasoned that if a polystyrene sphere could be glued to an individual biomolecule, the laser trap could be used to directly manipulate molecules in aqueous solutions. In 1988, he began to search for someone to teach him enough biochemistry to glue the tiny plastic spheres to a single molecule of DNA and found Kron who was working in the laboratory of Spudich, observing actin filaments moving on a bed of myosin molecules in an optical microscope. While continuing to work full time on his PhD thesis with Spudich, Kron began to moonlight with Chu. By 1989, they showed for the first time that biomolecules could be directly manipulated with optical tweezers.

Independent of this work, Spudich's graduate student Jeffrey Finer, MD, PhD, had become interested in using laser traps to see if they could observe one myosin molecule interacting with a single actin filament. By then, Kron had graduated and ironically, it was Robert Simmons, PhD, a professor from King's College London on sabbatical in Spudich's lab, who pointed out that Chu was working nearby on the second floor of the physics building. In a modification of the laser trap apparatus used to hold onto DNA, Finer and Simmons glued a single bead to each actin filament and tugged on it with optical tweezers, watching under a microscope, as myosin molecules pulled in the other direction. "It's like you're pulling on one end of the rope and then the other person pulls on the other end of the rope to keep the rope stationary," says Chu. In this way, they determined that myosin was exerting a force in the range of 1 to 5 piconewtons—equivalent to 1 to 5 trillionths the force that your average apple exerts under the pull of Earth's gravity—on the actin filament, as they reported in a 1993 paper. Such a force is perhaps too tiny to conceptualize, but the combined forces of many myosin molecules pulling on many actin filaments drive the movements of cells, muscles and animals.

But the second floor of the physics building wasn't the ideal place to watch a single myosin molecule move an actin filament—too many vibrations. "It didn't have the stability we needed to see these 5 to 10 nanometer strokes and furthermore, we

*"It was clear, in the late 1980s, that somebody was going to have to design a way to watch just one myosin molecule interact with a single actin filament and see how far it moved," said Spudich.*

envisioned building a dual-beam laser system to trap an actin dumbbell that could be lowered onto a pedestal with a single myosin molecule on top," says Spudich.

It was around that time that Spudich, then chairman of Stanford's Department of Structural Biology, was recruited to join the Beckman Center, where Spudich found his ideal lab space for setting up a better microscope. This was a room in the Beckman Center basement: "a very quiet space with very little vibration, which was needed to do this very sensitive experiment. And they allowed me to use that room to build what we call the dual-beam laser trap for measuring one myosin molecule."

In the new setup, Finer and Simmons glued one bead to each end of the actin filament, creating what they call an actin dumbbell. They used two sets of optical tweezers, one at each end of the actin filament, to hold it above a glass slide covered in little bumps sparsely coated with purified myosin molecules. To do the experiment, Spudich says, they used the optical tweezers to lower the actin dumbbell onto the bumps in search of a myosin molecule resting on top. "Most of the bumps didn't have anything, so nothing happened. But if we found a bump that had a myosin molecule on it, then the myosin grabbed onto the actin and pulled on the actin filament. And we could see that because the dumbbell moved," says Spudich.

"What Jeff found was it did not move by a hundred nanometers or more; it moved by somewhere between 5 and 10 nanometers," says Spudich. Finer, Simmons and he had, in 1994, provided decisive evidence in favor of the swinging cross bridge model. The Spudich lab firmly established the model with further studies. In that same experiment, they reported the force generated by a single myosin molecule pulling on actin and using one molecule of ATP, and they found it to be 3 to 4 piconewtons—similar to the force measurement that Spudich, Chu, Finer and Simmons had made previously.

Now, Spudich is studying how problems with cardiac myosin result in hypertrophic cardiomyopathy, a disease in which the heart muscle gets too thick, and single molecule force measurements are essential for these studies. Indeed, since the early studies in the '90s, single molecule biology has exploded into a field of its own, and researchers everywhere are characterizing their favorite biological molecules at the single molecule level. By examining a cellular component one molecule at a time, one can truly ascertain how it works to perform its particular function in the cell, Spudich says. Hence, there are now hundreds of investigators meeting regularly at single molecule biology conferences around the world.

## Rare-Earth Nanoprobes: Tools for Labeling and Visualizing Molecules

When Chu was collaborating with Spudich on laser trapping of biological molecules, his affiliation was with the physics and applied physics departments. But when, in 2013, he returned to Stanford from his time as President Obama's secretary of energy, he joined the Beckman Center's Department of Molecular and Cellular Physiology, while also retaining his affiliation with the Department of Physics. Recently,

***Steven Chu, PhD***
*Professor of Molecular and Cellular Physiology and of Physics*

he has been greatly improving rare-earth nanocrystals, a new class of fluorescent particles, as probes to label specific molecules and cellular structures. His goal is to be able to visualize proteins at work, in real time, at nanometer resolution—the resolution of molecules—but without the drawbacks of the current arsenal of optical probes.

"Fluorescent labels such as green fluorescent proteins (GFPs) and organic dyes have revolutionized molecular and cellular biology," says Chu, "but they fluoresce for about one second before they stop emitting light." This photo-instability makes it difficult to track individual molecules over time. Quantum dots are much more photostable, but are chemically toxic to cells, says Chu, adding that the wavelengths of light needed to excite quantum dots are phototoxic. In sum, using quantum dots for experiments can damage the cells under investigation.

Chu's rare-earth nanoprobes, while comparable in size with quantum dots, are much larger that fluorescent proteins or organic dyes. Their larger size (with diameters of greater than 10 nanometers) is a disadvantage, but they are completely photostable, and the infrared light used to excite the particles has no known phototoxicity. "Instead of looking for seconds with a limited number of photons, we can look for hours, weeks or months," says Chu.

By putting these nanoparticles into the vesicles that move neurotransmitters around in neurons, Chu can track vesicle movement with nanometer precision and millisecond time resolution in order to study fundamental questions in molecular transport and signaling, and to track movement of stem cells, cancer cells and immune cells.

Beyond basic research, such probes could help surgeons to achieve clean margins when removing tumors. Currently, surgeons find tumor edges by feel, Chu says. They then cut away some of what they presume to be healthy tissue around the tumor and send it off to be analyzed. Sometimes, unfortunately, they end up detecting cancer cells in that sample and have to operate again. Chu's very bright nanoparticles emit green light and when used to label cancer cells, could allow surgeons to see with their naked eyes a small number of remaining cancerous cells. While operating-room applications could take years to be approved by the Food and Drug Administration (FDA), they have a chance, since they are not chemically toxic, says Chu. Any particles not rinsed out during surgery will slowly dissolve over many months, Chu adds.

## Synthekines' and Other Engineered Ligands for Tinkering with Cell Fate and Function

Chu's nanoparticle probes have to bind specifically to molecules of interest in the same way that molecular messages, such as growth factors, drugs and many hormones, bind to receptors on the outsides of cells and thereby affect what happens within. How these molecular messages,

called ligands, bind to their receptors is the domain of Chris Garcia. Using structural biology and protein engineering, Garcia is developing novel ligands that tweak this signaling system in ways that are helping him learn about biology—and could lead to the development of new drugs.

It's important to be able to target drugs to particular cell types and areas of the body, Garcia explains. You want a drug to have a specific function to solve a particular problem. For a while, it's been possible to design drugs that target what are called G protein-coupled receptors in this cell-specific way, Garcia says. In part for this reason, says Garcia, G protein-coupled receptors are the most common target for drug developers. They're also common in the clinic: more than 30% of drugs approved by the FDA target G protein-coupled receptors; these include treatments for allergies, hypertension and asthma. "But the class of receptors that I work on, which is an equally large class of receptors, nobody had really been thinking about the concepts for these receptors," says Garcia. Receptors in this class, which are called single-pass transmembrane receptors, and which include growth and cytokine receptors, "mediate a huge swath of human biology," Garcia says, yet he adds that there is virtually no drug development happening with these types of receptors.

People overlook this receptor class because, unlike the ligands that target G protein-coupled receptors, those that target cytokine and growth receptors tend to have effects in multiple cell types, a phenomenon called pleiotropy. "When you dose somebody with a cytokine, let's say, for example, interleukin-2 or gamma interferon or stem cell factor, you not only hit the cell that you want to activate, but you hit a bunch of other cells," says Garcia. Sometimes, this can result in toxicity. Sometimes, these multiple effects even cancel each other out. But using protein engineering, Garcia is creating versions of these receptors that can do all sorts of interesting and pharmacologically relevant things—without the pleiotropic effects.

"We have engineered tuned versions of these cytokines that basically gives them a whole new life as drugs; we have created a cytokine pharmacology. But even more than that, it opens up this incredibly important class of receptors for drug discovery again," says Garcia.

The process starts with atom-level three-dimensional crystal structures of the ligand and its receptor, which serve as "blueprints" for the Garcia group's ligand engineering efforts, much as a carpenter uses the blueprints of a house to guide remodeling. Garcia's lab has developed crystal structures—a notoriously difficult task—for many of the most therapeutically important classes of receptors bound to their ligands. The next step is to create a collection of ligands with structural variations near the receptor binding sites. Finally, Garcia and his group test all of the ligands to find the ones that bind receptors as intended and have the desired signaling output—for example stimulating, versus tamping down, the immune system.

***K. Christopher Garcia, PhD***
*Professor of Molecular and Cellular Physiology, of Structural Biology, and Howard Hughes Medical Institute Investigator*

For instance, the cytokine gamma interferon (IFNy), which Garcia calls, "a holy grail drug for immunotherapy," can bind to two different receptors: one that stimulates the immune system; another that has the opposite effect. Garcia engineered a version of gamma interferon in which only the first function remains. "We biased it towards immuno-stimulation," says Garcia.

Sometimes, Garcia's engineered ligands can alter cell fate. For example, when hematopoietic stem cells begin to differentiate, they can take one of two paths to become either lymphoid cells, such as T cells, or myeloid cells, such as red blood cells. One of Garcia's ligands stalls differentiation at a stage that appears to Garcia to be somewhere between the two cell fates. "It doesn't look like a stem cell or a myeloid progenitor. It's something intermediate," says Garcia. Another possibility, he adds, is that the novel ligand caused the stem cells to differentiate into a cell type that doesn't exist in nature all.

Creating receptors that don't exist in nature is, in fact, another area of Garcia's research. The receptors Garcia studies have two components, which the ligand draws together into what's called a dimer when it binds, like a Lego block that sticks to two others, bringing them together. Garcia is developing novel ligands that bring together naturally existing receptor components in combinations that do not naturally occur. These novel ligands, called synthekines (for synthetic cytokine), will likely have novel biological functions. "Synthekines induce new activities in natural cells by just taking advantage of different receptors that are expressed on the cell surface," says Garcia. "Looking beyond the menu of cytokines encoded in our genome to creating completely synthetic activities is going to be the future of cytokine drug discovery," Garcia believes.

By tinkering with naturally occurring ligands and by designing ligands that bring their receptors' Lego-like components together in new combinations, Garcia hopes to realize the potential of single-pass transmembrane receptors as drug targets.

## Computational Tools for Decoding the Human Genome

Gill Bejerano wants to understand, in his words, "how the human genome does its `thing.'" To that end, since joining Beckman Center faculty in 2007, Bejerano has pursued three approaches to genomics research: comparative genomics, disease genomics and functional genomics. All three approaches help him correlate genotype, the DNA letters of the genome, to traits, including disease, which are referred to as phenotypes.

One approach, which Bejerano developed and dubbed Phenotree, correlates particular phenotypes and their emergence in or disappearance from different branches of the mammalian evolutionary tree with the appearance or disappearance of genetic changes in mammalian genomes in order to discover the genes and genomic regions that control phenotypes of interest. It's called a "comparative genomics" approach because it involves trying to learn more about the human genome by comparing it to the genomes of other species, Bejerano explains. Phenotree compares the genomes of humans, mice, dolphins, giraffes, monkeys and other mammals—over 100 species in total. The Phenotree approach

***Using structural biology and protein engineering, Garcia is developing novel ligands that tweak the signaling system in ways that are helping him learn about biology and could lead to the development of new drugs.***

continues to be refined and is available for any researcher to use.

One way that Phenotree identifies genes or groups of genes that control a given trait is by looking for genes that are absent or dysfunctional in species that have lost the trait. For example, Bejerano and colleagues tried using Phenotree to identify a gene that controls the ability to synthesize Vitamin C from food, an ability that is present in many mammals, but that has been lost in several branches of the mammalian evolutionary tree, such that multiple bat species, some primates, including humans, and guinea pigs cannot make Vitamin C. Bejerano explains the experimental approach, "We went across the genomes of all of these species and we asked, 'Can we find a gene that looks perfectly healthy in all the species that can synthesize Vitamin C, but looks like it's falling apart exactly in the species that cannot?'" Phenotree did just that and identified the gene, called GULO, that researchers already knew was essential for Vitamin C synthesis. This experiment told the researchers that their system was working. Since then, Bejerano's group has expanded their approach to answer a set of related questions, including, of course, those for which they do not already know the answers.

Correlating genotype and phenotype can also help diagnose human patients: "When you get the genomes, for example, of a few patients that have a mysterious disease and nobody knows where that disease starts in the genome, you can ask, 'Do these patients share anything that the healthy population of people that we've sequenced do not share?' And you often get very precise answers," says Bejerano. This is medical genomics. The Bejerano group has developed over a dozen computational components that together can greatly accelerate both patient diagnosis and disease-gene discovery. One such tool, which Bejerano created to compare human genomic data and phenotypes, is called Phrank.

Phrank relies on the increasing ease with which clinicians can sequence a patient's genome in order to diagnose diseases with a genetic basis. There are currently over 5,000 diseases that can be traced back to mutations in single genes, and researchers identify hundreds of new genes that can cause diseases every year. For a physician to go through the list of the hundred or so disease genes in which a patient has mutations can take many hours of a physician's valuable time. Phrank aims to save physician time by automating a portion of this process.

To use Phrank, a clinician takes the patient's sequencing results, makes a list of the disease genes that contain suspicious mutations, and gives that list to Phrank, along with a list of the patient's symptoms, or phenotypes, translated into codes Phrank can understand. The program then compares the patient's data to genotype-phenotype relationships in its database to generate a list of genes that could explain the patient's symptoms, ordered from most likely to least likely to be responsible. Phrank also generates a ranked list of genetic diseases that the patient might have.

According to a study published this past July, Phrank is better at ranking candidate

***Gill Bejarano, PhD***
*Associate Professor of Developmental Biology, of Computer Science, of Pediatrics, and of Biomedical Data Science*

*The Bejerano group has developed over a dozen computational components that together can greatly accelerate both patient diagnosis and disease-gene discovery.*

genes and diseases than two other tools that aim to produce similar results. In approximately one out of four cases, Phrank puts the gene responsible for the patient's disease at the top of the list; in over half of cases, the gene responsible for the disorder is one of the top five on the list. The same is true for the list of potential diseases.

In 2017, Bejerano put Phrank online for clinicians to use, free of charge, as part of the web portal AMELIE, which includes Phrank plus some other tools from the Bejerano lab. A special feature of AMELIE (which stands for Automatic Mendelian Literature Evaluation) is that it mines the scientific literature to find papers relevant to an individual patient's case. Its output is not only a list of candidate genes, but related scientific papers, also ranked and organized by gene. The code for Phrank alone has also been available for clinicians to download and incorporate into their own systems since July 2018.

Perhaps because of its accuracy, Phrank, via AMELIE, is very popular. AMELIE has received over 2,000 submissions of patient data per month for the last year, Bejerano says.

"We have evidence from the clinicians we talked to that we really are helping patients and their families by empowering the clinicians to very quickly find the causal gene in as many cases as they possibly can," says Bejerano.

A third tool, called GREAT correlates gene-regulatory datasets—such as those obtained by measuring transient marks on the histone molecules that pack our DNA—and phenotypes. Bejerano and his team first put GREAT, a functional genomics tool, online in 2010 and are continually improving it. Scientists can use GREAT to study how epigenomic modifications, such as histone marks, differ between healthy and diseased cells, for example, or to observe how epigenomic modifications change as cells differentiate during development. "[There are] many different questions you can ask both in the developmental context and in the disease context," Bejerano says of GREAT. "And that's why, I think, we have had over one million job submissions to this tool that we put out there. That's pretty cool for an academic lab."

## Eterna: An Online Videogame for Designing RNA Molecules

Rhiju Das develops algorithms for designing RNA molecules. The goal is to be able to write sequences of RNA letters—A, U, C and G—that will fold up into particular structures and perform desired functions, many of which could be medically relevant. "There are numerous emerging therapeutics and vaccines, and diagnostics that are based on designer RNA molecules," says Das. But creating such molecules is easier said than done. "Some of these design problems are beyond the capabilities of our best computer simulations; they're beyond our best people, experts in RNA science, including folks who have won Nobel prizes for their discoveries in RNA. So, designing molecules that can fit at the heart of a new molecular therapy is a hard, hard problem," says Das.

Das and others have developed computer algorithms to predict how a given RNA sequence will fold and, conversely to write RNA sequences to form desired structures, but the algorithms are often wrong. To improve them, Das created an online videogame, called Eterna (http://

***Rhiju Das develops algorithms for designing RNA molecules. The goal is to write sequences of RNA letters that will fold up into particular structures and perform desired functions.***

eternagame.org), where the objective is to design RNA molecules with particular structures.

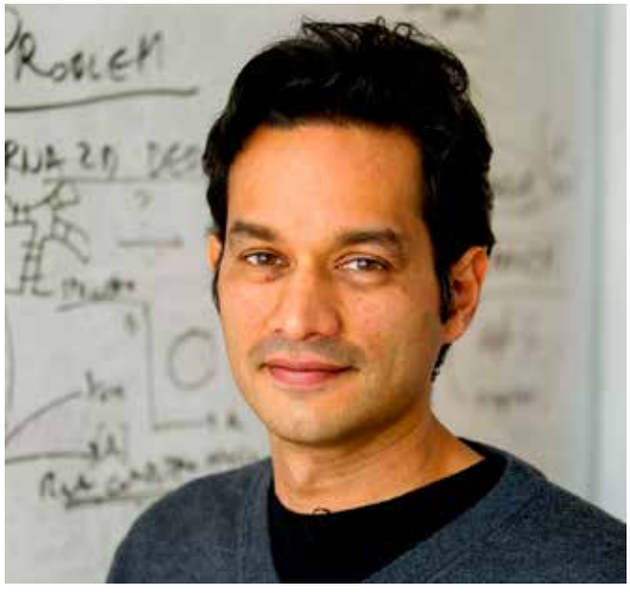
During early levels of the game, players' RNAs just fold up according to the algorithm of the day, but eventually, players get to a level called "lab," where researchers in Das's laboratory synthesize the RNA molecules players have designed and determine whether they fold as predicted. If not, the players learn that they need to change something about their RNA-design strategies and Das's laboratory learns that it needs to change its algorithms.

Eterna went online in 2011. Within a few months, tens of thousands of players were designing RNA molecules. "From this virtuous cycle of gameplay and experimental feedback, they became the best in the world at designing RNA molecules. Their efforts outperformed any algorithm that I or others in the RNA-folding field had prepared," says Das. The Das lab also started using artificial intelligence methods for developing their algorithms. With these methods, he says, "we do better than prior automated algorithms, but we are still having trouble doing as well as human players."

Now Das's group is using their algorithms—and working with Eterna players—to design two types of RNA molecules with medical applications. The first is a tuberculosis (TB) test to determine whether latent TB—present in one in four humans—has become active.

In 2016, another Beckman Center faculty member, Purvesh Khatri, PhD, associate professor of medicine and of biomedical data science and assistant professor at the Institute for Immunity, Transplantation and Infection, reported that he and his team had discovered three RNA molecules whose presence in the blood is indicative of an active TB infection. Soon thereafter, Khatri asked Das if he could develop a designer RNA molecule that could detect the three TB-associated RNAs and effect a color change on a piece of paper—something like a pregnancy test. At first, Das says, it sounded like a fairly easy problem to solve. Then Khatri told him that the RNA sensor on the test would have to detect not just the presence or absence of these three RNA molecules, but the ratios of their concentrations in the blood. The designer RNA would have to be a molecular calculator. That, Das thought, sounded impossible. But Das's team of Eterna developers and players themselves were more optimistic. "They said, 'Hey, if you could make an RNA that could fold up into two different structures depending on the concentrations of these input RNAs, then maybe you could make a molecular calculator and it might work,'" Das recalls.

Late in 2016, Das set up the OpenTB Challenge, which asked Eterna players to design such a molecule. He also worked with Beckman Center colleague Will Greenleaf, PhD, associate professor of genetics and, by courtesy, of applied physics, to design a specialized microscope that would allow them to test all of the candidate molecules that Eterna players submitted. Now, about two years and 30,000 candidate RNA molecules later, Das and Greenleaf have a few dozen molecules that perform as molecular calculators, folding one way when the ratio of the three RNAs is above a certain threshold, another way when that ratio is below the threshold. Last year, Das says, the Bill & Melinda



***Rhiju Das, PhD***
*Associate Professor of Biochemistry*

Gates Foundation heard about the OpenTB Challenge and Das had a meeting with Bill Gates. "Essentially, right after that meeting, the Gates Foundation cut us a check to then develop these OpenTB molecules designed by Eterna players and to try to translate these molecules into a point-of-care test," says Das. That's what the Das lab is working on now. The goal is that public health workers visiting communities in areas where tuberculosis is still a big problem could use such a test to quickly screen anyone who thought he or she might be infected with active TB. Anyone who tested positive could then visit a larger facility for a more thorough test and, if necessary, treatment.

Now Das and Eterna players are working on the OpenCRISPR Challenge in which the goal is to design RNA molecules that would improve the therapeutic potential of the gene-editing technology CRISPR, which depends on a so-called guide RNA to lead the gene-editing machinery to its target. "The problem is, when you put the CRISPR machines in and you get them going, you can't turn them off. It's kind of like giving someone an iPhone that only has an on button," says Das. But with the right guide RNA, which Das's lab and Eterna players are working to design, that could change. ■