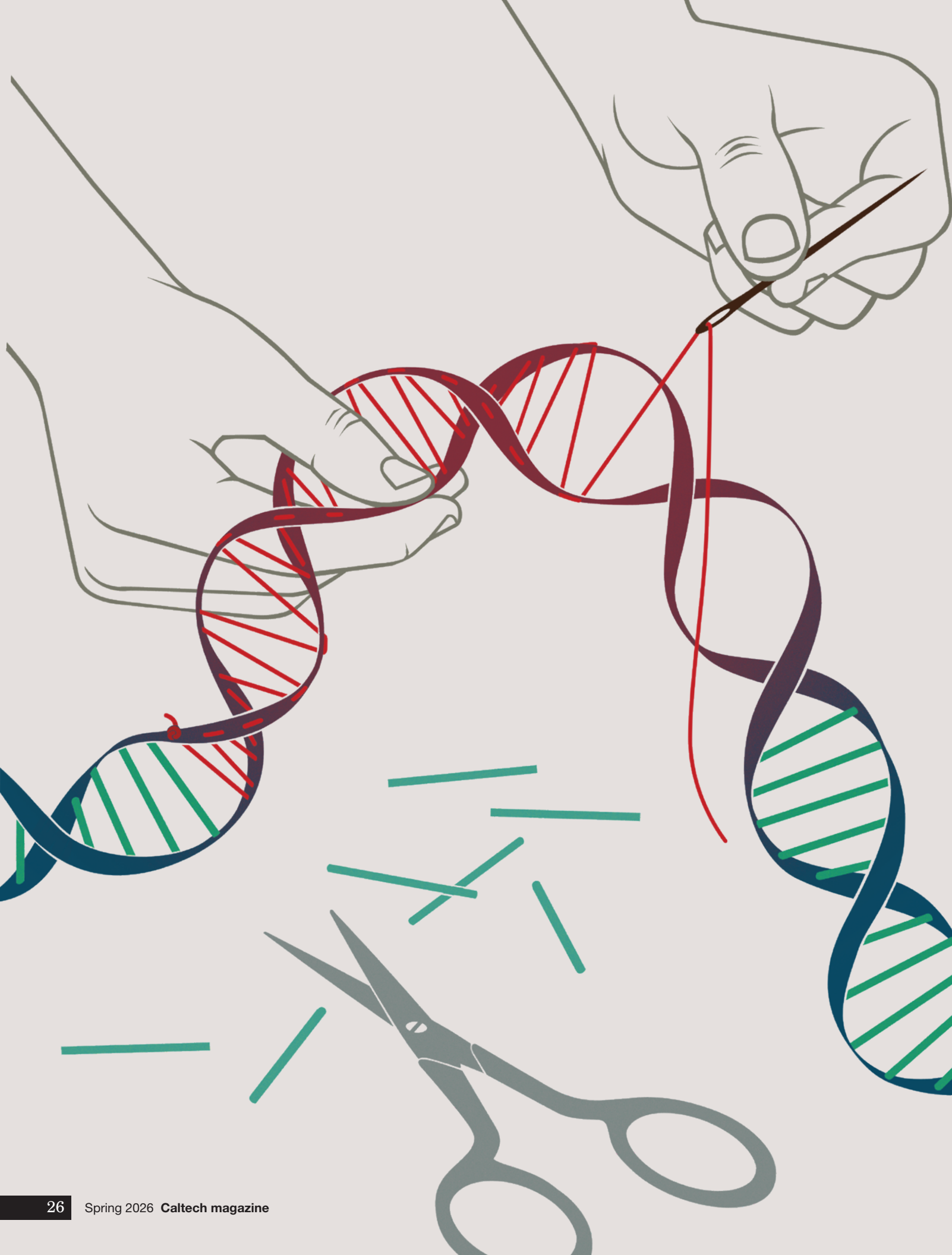




Stitching the Synthetic Biology Revolution

Caltech is building tools for and shaping approaches to the discipline, which could change the way we live.

By Lori Dajose (BS '15)

Billions of years of evolution have slowly, iteratively, produced a vast library of genomes through DNA mutations, encompassing the genetic information for everything that has ever lived. Mother Nature has had a monopoly on writing in the language of life. Until now.

Earlier this year, Caltech bioengineer Kaihang Wang unveiled a new technique that can stitch together short pieces of DNA—the building blocks of living organisms—to create long unbroken strands of millions of individual genes. His invention, called Sidewinder, is the functional equivalent of “page numbers” for DNA. Think of it like assembling a book: Sidewinder enables small pieces of synthesized DNA to be labeled sequentially, like page numbers, so the strands can organize themselves in the correct order.

Wang’s DNA-weaving tool represents a major advance in synthetic biology, the practice of engineering not with electrical components or steel but with the biological components of life. Synthetic biology allows scientists to study the complex workings of living organisms by breaking apart, redesigning, and constructing new molecular parts, systems, and devices. In turn, these processes lead to the development of useful applications for treating disease, improving agriculture, and much more.

Since the field coalesced at the turn of the 21st century, Caltech researchers have led the development of tools that are integral to the practice of synthetic biology. These efforts are giving rise to a new biological frontier in which

synthetic biologists at the Institute wield cellular machinery to develop and deliver effective cancer treatments, control populations of invasive mosquitos, and manipulate plants to improve resiliency in the face of a changing climate—turning imagination, Wang says, into reality.

Origins of the Field: “A Magic Time”

Many of synthetic biology’s nascent components were developed at Caltech.

In 2003, a physicist-turned-biologist named Michael Elowitz arrived on campus as an assistant professor. As a graduate student at Princeton, Elowitz had invented a biological circuit called the repressilator: a circuit of three genes engineered into a cell where the genes interact with one another in an oscillating cycle akin to the game rock-paper-scissors. Each gene produces its own protein—think of them as rock, paper, or scissors—that physically represses the expression of one of the other genes: Rock represses scissors, scissors repress paper, and paper represses rock. Proteins degrade over time, so as concentrations of “rock” decrease, production of “scissors” shoots up, which blocks the gene for “paper.” After “scissors” eventually degrades, the “paper” proteins increase, which silences “rock,” and so on, in a perpetual cycle. This achievement marked the first time anyone managed to program a living cell to execute complex dynamic behaviors like a computer does.

When Elowitz came to Caltech, electrical engineer Richard Murray (BS '85) had not taken a biology class for two decades. However, the repressilator piqued his interest. Murray’s research focused on control and dynamical systems—systems that sense inputs in real time

and modulate themselves accordingly, like thermostats. The repressilator was an example of a dynamical system made out of biological rather than electrical parts.

“I started to get interested in the idea of engineering biological systems in the same way you would engineer electrical systems,” Murray says. “Take all the discrete components of the cell: How can we not only tweak them by adding in or repressing a gene but also engineer the entire system of biomolecular components to behave in ways that we direct?”

In the early 2000s, interdisciplinary collaborations emerged around these questions and involved researchers such as Murray, Elowitz, biophysicist Rob Phillips, chemical engineer and Nobel Laureate Frances Arnold, computer scientist Erik Winfree (PhD ’98), applied mathematician Niles Pierce, and more. “It was just a magic time in terms of crystallizing that there was an idea here that could take off,” Murray recalls.

Engineering with the components of life also helps answer open questions about life itself—a process that many researchers call “building to understand.”

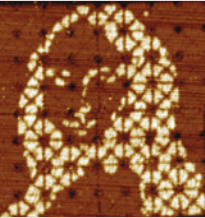
“Biology seeks to explain how complex organisms work through a kind of reverse engineering process—by discovering what genes and other molecular components are necessary for a given process and then working out how they interact with one another to provide a natural function,” Elowitz says. “Synthetic biology works in the opposite way: We take molecular components such as proteins, alter them, and redesign their interactions in order to program a new cellular behavior. What makes this interesting and challenging is that this whole engineering process unfolds inside a living cell—an environment that we did not create and do not fully understand.”

Other researchers, like Winfree, take a “bottom-up” approach to synthetic biology, removing biomolecules like DNA from their normal context in cellular systems in order to engineer dynamic systems at the nanoscale, such as a test tube of molecules that behave like a neural network. “For me, the next frontier is the study of information-based chemistry—the ‘magic sauce’ of biology, independent of biological cells,” Winfree says.

No matter the approach, says biophysicist Rob Phillips, at its core, synthetic biology is about “the imagination to create that which has never existed before.”

Nanoscale Engineering with DNA

DNA’s unique chemical properties, or the “magic sauce” Winfree describes, enable the work of synthetic biology researchers. The four kinds of nucleotides that make up a DNA strand—adenine (A), cytosine (C), guanine (G), and thymine (T)—pair together precisely (A with T and G with C). In biology, these base pairs form within the familiar DNA double helix of an organism’s genome.



Traditionally, genes within the genome are transcribed into RNA and then translated into proteins that function within the cell. However, researchers can take advantage of the programmability of nucleic acids like DNA and RNA to design sequences that assemble themselves into complex structures or function as dynamic devices. Researchers refer to this new discipline as molecular programming—one of the foundational tools for biological engineering.

Molecular programming occurs on the nanoscale. Designing DNA and RNA molecules to independently carry out tasks on the nanoscale within living cells and organisms exploits the very programmability that living organisms use in order to function. Programming is invaluable because researchers cannot simply take tweezers and move molecules around like you might with the wires and switches in an electrical circuit.

“Molecular-programming researchers seek to develop the principles and tools to program molecules with the ease and rigor that computers are programmed today,” says Pierce, who in 2004 developed a key technology in the field of molecular programming known as hybridization chain reaction (HCR). HCR was the first experimental demonstration of conditional nucleic acid self-assembly, opening up the possibilities for molecules to autonomously self-assemble via prescribed kinetic pathways without human intervention. Now used in thousands of labs around the world, HCR is a way to visualize and quantify biological molecules in a sample using short sequences of DNA that, in the presence of a target sequence, perform a chain reaction to produce a detectable signal, enabling imaging of individual target molecules within intact tissues. This process is central both to biological research (such as studying the development of a single cell into an adult organism) and to the creation of new therapeutics (such as measuring the effect of a candidate drug on a set of molecular targets in a cancer cell).

To enable molecular programming, Pierce and colleagues developed and support NUPACK (short for nucleic acid package), a software suite for the analysis and design of nucleic acid structures, devices, and systems used by thousands of labs worldwide. NUPACK algorithms are crucial tools to analyze and design molecular interactions that are a hallmark of the emerging fields of molecular programming, nucleic acid nanotechnology, and synthetic biology.

Such molecular interactions were on vivid display in 2016, when bioengineer Lulu Qian crafted the world’s smallest *Mona Lisa*, among other nanosized illustrations, using a technique called DNA origami developed by then Caltech research professor Paul Rothemund (BS ’94). Images captured by atomic force microscopy show the iconic smiling face of the Mona Lisa, rendered not in paint and

brushstrokes but in DNA molecules half a micrometer wide—the size of a small bacterium. These images were more than just minuscule artwork—they were proof of the accuracy and precision Qian had developed through the hierarchical self-assembly of DNA origami.

The technique involves designing DNA molecules with specific sequences so that, when tossed in a test tube, the individual sequences interact and thus assemble themselves into a desired shape. Many small shapes then assemble into a larger one like puzzle pieces forming a puzzle. The assembly process is also dynamic: The DNA sequences can process information and make decisions about whether one strand should replace another in response to signals in a molecular environment. In this way, Qian can create nanoscale computing systems such as neural networks made out of DNA instead of electrical pieces, taking advantage of the unique chemical properties of biomolecules.

Building the Bioengineer’s Toolkit

When synthetic biology is used for specific types of applications, the process follows a general pipeline. First, researchers design what they want to make in the same way an architect would create a blueprint for a building. The desired product could be a circuit of molecules to destroy cancer cells or an enzyme to make plants more drought resilient. Then, they turn their design into biological reality using synthesized DNA, proteins, and other biological building blocks. Next is lab testing, which includes delivering the synthesized material into cells. Researchers then make adjustments to improve their results.

Wang’s Sidewinder is a revolutionary new tool for synthesizing biological designs. To help people understand Sidewinder, Wang likes to point to the invention of the printing press in the 1440s, which enabled the creation of thousands of pages each day from a single printer. But assembling these individual pages into books was an arduous process, requiring book binders to scrutinize each page for contextual clues and place them in the correct order. Pagination was not invented until decades later and did not become commonplace for centuries.

The practice of DNA synthesis encountered a similar bottleneck. Until Wang’s Sidewinder invention, it was challenging to synthesize long pieces of DNA, though it was possible to synthesize short pieces that are a few hundred base pairs long, called oligonucleotides (or oligos), as used in the HCR and DNA origami processes.



Kaihang Wang, center, together with grad students Noah Robinson, left, and Weilin Zhang, right, developed the Sidewinder technology.

► The Qian group has shown that these DNA neural networks can use heat as an energy source and can learn to perform pattern classification tasks on their own.

But to create entire genes, and ultimately genomes, and thereby enable the realization of biotechnologies from personalized cancer vaccines to synthetic crops, longer DNA sequences up to billions of bases in length would need to be stitched together with perfect accuracy from an ever-increasing number of oligos.

Sidewinder enables the synthesis of new genes and even entire genomes within just a few days, if not hours. “We can now make long DNA, regardless of complexity and sequence, and do this faster, more easily, and more cheaply than has been possible,” Wang says.

The final stage of the synthetic biology pipeline, delivery, focuses on how molecular components get where they need to go. One method for delivery draws inspiration from naturally occurring viruses. Viruses, much like a Trojan horse, are adept at sneaking past cellular barriers to get their genetic material into cells. While real viruses can make you sick, scientists such as biologist Viviana Gradinaru (BS ’05) are testing whether viral packaging could instead be filled with therapeutic molecules or genes and designed to make its way into the intended cells.

Gradinaru uses a class of virus known as adeno-associated viruses (AAVs) and alters their capsids, or shells, based on their ability to penetrate specific cells. For example, some AAVs can be designed to target specific brain cells called neurons. These would be useful when developing treatments for brain disorders such as Huntington’s disease, which is caused by a mutation in a single gene. Gradinaru’s team modifies the AAV to custom applications whether the cargo is intended for a neuron or another kind of cell.

As with AAVs, delivery tools often need to be customized for specific applications. In particular, many of the tools developed to work with human cells do not translate to plant cells. To address this, chemical engineer Gözde Demirer is developing tailored tools for plant engineering, a major frontier in agricultural sustainability.

A World of Applications

As the synthetic biology toolkit expands, so does its potential applications. Twenty-five years after unveiling the repressilator, Elowitz is now applying synthetic circuit design to cancer therapeutics. The two leading forms of cancer treatment, chemotherapy and drug-based targeted therapies, each have their own drawbacks. Chemotherapy is great at killing cells but lacks specificity, leading to debilitating side effects like liver toxicity. Targeted therapies, on the other hand, can be very specific for diseased cells but often lead to resistance.

Projects in Elowitz's lab, led by graduate students Andrew Lu and Lukas Moeller, provide foundations for an alternative paradigm. Their therapeutic circuits are sets of engineered proteins that detect whether a cell is expressing the mutant *RAS* oncogene, the most prevalent cancer-related gene in human cancers. If mutant *RAS* is detected in the cell, the circuit initiates a cell-death program that destroys it. But the same circuit has no effect on healthy cells. Along with collaborators, the team has analyzed the livers of mice with cancer that had received the treatment through an injection and some, the control group, that had not. The control group mice had livers riddled with tumors while the livers of the mice treated with the circuits were cleared. An exciting aspect of the system is that it is programmable. By swapping out sensor modules in the proteins, it should be possible to adapt the system to other cancers or diseases.

In addition to treating individual organisms, synthetic biology research can be applied to entire populations. Biologist Bruce Hay spends his time focused on swarms of pesky insects. Mosquitoes, for example, carry infectious diseases like malaria. However, they are also a major food source for birds, so wiping them out entirely would be damaging to the ecosystem. Hay's team is engineering ways to deliver genes into mosquito populations that would only kill the insect if it were carrying malaria.

Importantly, rather than catching and injecting every mosquito with the therapy—an impossible and absurd task—Hay is engineering the gene therapy to be passed along in the population. These so-called “selfish” gene systems take advantage of the Darwinian concept of survival of the fittest to ensure the inheritance and spread of the desired anti-malaria gene among the offspring. Hay designs these systems to be reversible, so that the effects of synthetic elements can be counteracted if need be.

▶ One tool uses self-assembled protein nanoparticle platforms to deliver useful small molecules and proteins into plant cells, leading to more productive crops.

Synthetic biology can also provide alternatives to working with real organisms. For example, researchers have limited access to donated embryos for research, and the embryos may only be grown for a set period of time (up to 14 days, according to international policies, or 12 days in California). However, there are many unanswered scientific questions about this important stage of development, one at which most pregnancies fail.

Developmental biologist Magdalena Zernicka-Goetz is a pioneer in creating synthetic human embryos for research. Her lab uses mouse stem cells to make embryo-like structures that look and behave identically to a mouse embryo, complete with a primitive heart and brain cells. These structures do not come from a sperm and egg; they are the result of chemically coaxing mouse stem cells to interact and self-organize into structures that develop like an embryo, enabling the study of the earliest stage of life without using real embryos. Zernicka-Goetz has used these models to discover how defects in certain genes impact brain development, among other discoveries.

Phillips notes there is still much to learn about biological systems. “For 60 percent of genes in the *E. coli* bacteria family—not to mention 99 percent of genes in mammals—we don't know how they are turned on or off at all,” Phillips says. “We can build a genome, but we don't know how it works or how to control it. It's like sailors navigating on flawed maps from 400 years ago: We will find land, but it may take time. Synthetic biology is a major field of opportunity that could improve the human condition. There's so much to be done.”

Much of the research mentioned in this story was made possible by federal funding from the National Science Foundation and the National Institutes of Health, among others. For a full list of funding sources, scan the QR code at right.



Gözde Demirer is the Clare Boothe Luce Assistant Professor of Chemical Engineering.

Michael Elowitz is the Roscoe Gilkey Dickinson Professor of Biology and Bioengineering and a Howard Hughes Medical Institute Investigator.

Viviana Gradinaru is the Lois and Victor Troendle Professor of Neuroscience and Biological Engineering and a Howard Hughes Medical Institute Investigator. She is also the Allen V. C. Davis and Lenabelle Davis Leadership Chair and director of the Richard N. Merkin Institute for Translational Research.

Bruce Hay is a professor of biology.

Richard Murray is the Thomas E. and Doris Everhart Professor of Control and Dynamical Systems and Bioengineering.

Rob Phillips is the Fred and Nancy Morris Professor of Biophysics, Biology, and Physics.

Niles Pierce is the John D. and Catherine T. MacArthur Professor of Applied and Computational Mathematics and Bioengineering, and the executive officer for biology and biological engineering.

Lulu Qian is a professor of bioengineering.

Kaihang Wang is an assistant professor of biology and biological engineering.

Erik Winfree is a professor of computer science, computation and neural systems, and bioengineering.

Magdalena Zernicka-Goetz is a Bren Professor of Biology and Biological Engineering.

Bringing Financial Vision to Life

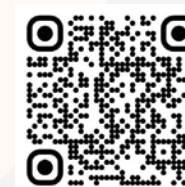
CEFCU exclusively serves the extended Caltech community and their families. Membership is open to Caltech, Jet Propulsion Laboratory, and The Huntington employees, contractors and volunteers, as well as Caltech undergraduate and graduate students, the Caltech Associates, Caltech Alumni Association, CARA, TMT, Polytechnic School, and immediate family (by blood or marriage) of current members.



800/592-3328 • www.cefcu.org



Join today!



Must qualify for CEFCU membership to join.
Minimum \$5 deposit and one-time \$5 membership fee due upon opening any CEFCU share account.
Federally insured by NCUA.