ENGINEERING LIFE: Building a FAB for Biology

BY THE BIO FAB GROUP*

*David Baker, George Church, Jim Collins, Drew Endy, Joseph Jacobson, Jay Keasling, Paul Modrich, Christina Smolke and Ron Weiss

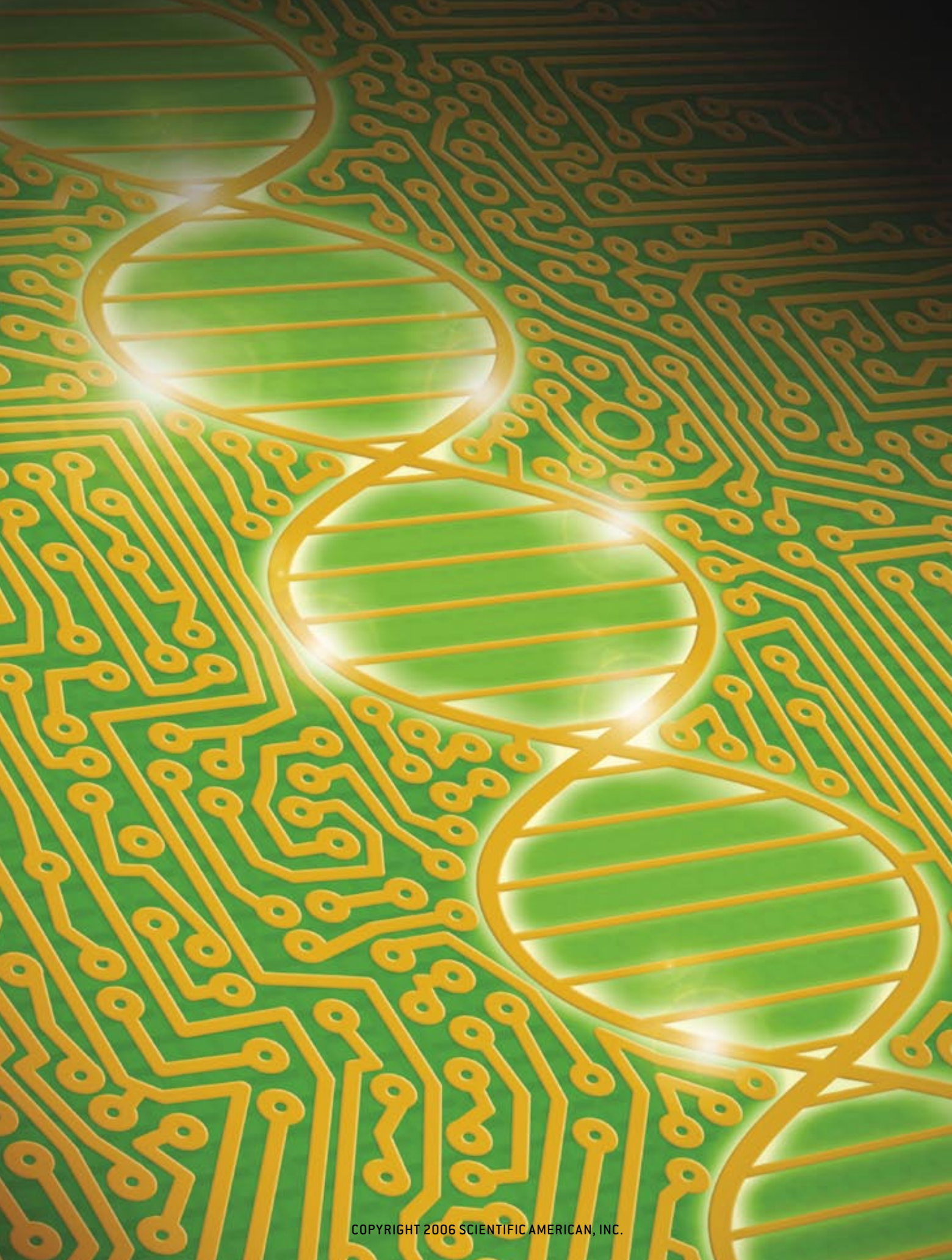
Principles and practices learned from engineering successes can help transform biotechnology from a specialized craft into a mature industry

Although the term "genetic engineering" has been in use for at least three decades, and recombinant DNA methods are now mainstays of modern research, most biotechnologists' work with living things has little in common with engineering. One reason is that the tools available for building with biological "parts" have yet to reach a level of standardization and utility equal to that in other engineering fields. Another has to do with methods and mind-sets in biology, although these, too, can be powerfully influenced by technology.

Electronic engineering, for example, was transformed beginning in 1957, when Jean Hoerni of Fairchild Semiconductor, a small company in what would later be known as Silicon Valley, invented planar technology. It was a system for layering and etching metals and chemicals within silicon wafers using templates called photomasks. This new approach allowed engineers to produce integrated circuits cleanly and consistently and to create a wide variety of circuit types just by changing the pattern on the photomask. Soon engineers could draw from libraries of simple circuits made by others

BIOLOGICAL COMPONENTS are the basis of an approach to biotechnology modeled on electronics engineering.

SLIM FILMS



and combine them in increasingly complex designs with a widening range of applications.

Until then, standard practice for manufacturing an electronic circuit had been to wire together its individual transistors one by one. It was an artisanal process, with uneven results, and a recognized bottleneck in the fledgling electronics industry. In contrast, planar technology kept steadily improving, enabling amazing advances at a rate famously commemorated by Moore's Law.

This combination of technology and methodology for designing and fabricating semiconductor chips—the "chip fab"—constitutes one of the most successful engineering paradigms of all time, and it is a valuable model for another nascent technology sector: fabrication of biological systems.

In effect, today's genetic engineers are still hand-wiring every circuit. As our colleague Tom Knight of the Massachusetts Institute of Technology artificial intelligence laboratory has observed, "The lack of standardization in assembly techniques for DNA sequences forces each DNA assembly reaction to be both an experimental tool for addressing the current research topic, and an experiment in and of itself."

Standardization of methods and components in biological engineering could give rise to design libraries of compatible parts and make outsourcing of fabrication possible. That uncoupling of concept and manufacture would free biological engineers to imagine increasingly complex devices and to use powerful engineering tools, such as computer-aided design, to manage that complexity. Toward these ends, members of our group have begun to identify and develop the equipment and techniques that could become the basis of a "bio fab." We are also trying to encourage a community that applies the best principles and practices of engineering to biotechnology.

Overview/*Fab for Life Science*

- Flexible, reliable fabrication technology along with standardized methods and design libraries gave rise to the semiconductor chip "fab" system. It enabled engineers to create extraordinarily complex and powerful electronic devices with broad applications.
- A fab approach could similarly empower biological engineers to conceive and build sophisticated devices from biological parts.
- Bio fab technologies and techniques are already being developed and used. Addressing safety issues and encouraging biologists to think more like engineers are ongoing efforts.

Quality Parts

IF INDIVIDUAL TRANSISTORS are the basic components of electronic circuits, then their biological equivalents are genes: long, carefully ordered stretches of DNA. To construct genetic circuits for advanced biological devices, therefore, we need a way to manufacture long pieces of DNA quickly, reliably and at a reasonable price.

Twenty years ago Marvin H. Caruthers of the University of Colorado at Boulder built on earlier work by others to develop a system for synthesizing single DNA strands by exploiting their natural chemistry. DNA is composed of nucleotides, which are distinguished by the type of subunit, called a base, they contain: adenine (A), cytosine (C), guanine (G) or thymine (T). Affinities between the bases cause them to pair with one another—A with T and C with G—to form the rungs of the ladderlike double-stranded DNA molecule. Chemical groups create the bonds between base pairs as well as between adjoining nucleotides along either strand.

It is a way of thinking about existing biological machines and of constructing new ones.

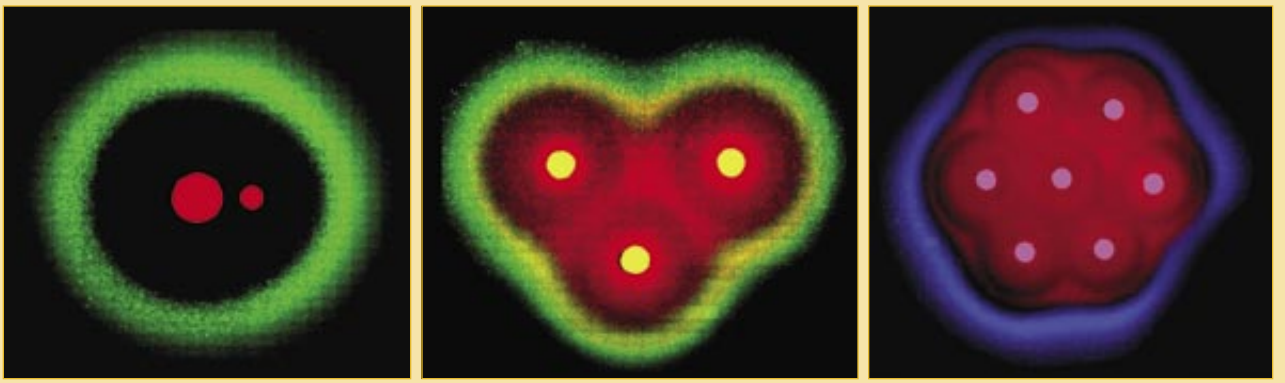
Caruthers's method, known as solid phase phosphoramidite chemistry, is still the basis of most commercial DNA synthesis. It begins with a single nucleotide attached to a solid support, such as a polystyrene bead, suspended in liquid. When exposed to an acid, the nucleotide's base becomes open to forming a bond with a new nucleotide added to the solution. That second nucleotide is then exposed to acid, and another nucleotide is joined to it, contributing to the growing chain. Repeating this cycle makes it possible to synthesize any desired nucleotide sequence with an error rate of approximately one base in 100.

Unfortunately, many of the genetic constructs that biological engineers wish to build are far longer than that. A simple network of genes may be several thousand bases long; the genome of even a small organism such as a bacterium can run to several million bases. Several of us working on finding synthesis methods with higher yields and lower error rates have therefore looked to nature for clues.

In living organisms, biological machinery composed of enzymes such as polymerase is able to manufacture and repair DNA molecules at speeds of up to 500 bases a second, with error rates of about one base in a billion. That represents a trillionfold performance improvement in yield throughput (output divided by error rate) over the best DNA synthesis machines, which add a base every 300 seconds. Moreover, multiple polymerases work in parallel when copying a long piece of DNA, such as a bacterial genome, so they are able to churn out about five million bases in 20 minutes.

One of us (Church) set out to emulate that parallelism by adapting the existing technology of microarrays. These are

Living Devices



Programmed bacteria form ring patterns in response to signals from colonies of "sender" cells that appear red (*left*), yellow (*center*) or pink (*right*). Lawns of engineered *Escherichia coli* "receiver" cells detect chemicals emitted by senders placed in their midst. The receivers then produce different-colored fluorescent proteins depending on their distance from the senders, which they judge by the chemicals' concentration.

Increasingly complex patterns can be created by varying the initial placement of sender colonies. This example of an artificial multicellular system can be used in research to understand signaling and pattern formation in natural cell systems, such as during an organism's development. The same technology could also be applied in sensors, three-dimensional tissue engineering, and biofabrication of materials by programmed organisms.

large slides dotted with short, single DNA strands known as oligonucleotides, or oligos, about 50 to 70 bases in length. They are manufactured simultaneously right on the microarray surface using phosphoramidite chemistry, anchored in a grid pattern that approaches densities of one million dots per square centimeter. To the traditional technology, we added cuttable linkers that allow specific oligos to be released from the microarray. Each dot in our experimental microarray is about 30 microns wide and contains some 10 million oligo molecules.

We call these strands construction oligos because they are designed to overlap with one another in sequence so that they can later be assembled to form longer DNA constructs, such as a whole gene. But any oligos containing sequence errors must be weeded out. For that purpose, we have pursued two different error-correction systems.

The first uses the same microarray synthesis method to produce what we call selection oligos, with sequences complementary to the construction oligos. We then release the selection oligos from their slide and wash them across the construction oligo array. The selection oligos will follow base-pairing rules and bind, or hybridize, with their complementary construction oligos to form double DNA strands. We can then identify any unmatched construction oligo strands or gross imperfections in bound pairs as containing errors and release the bad oligos from the array. Interestingly, although the selection oligos are just as likely to bear some mistakes—having been made in the same manner as the construction oligos—the probability that erroneous sequences in either set will find a perfect complement is very low. Thus, using one set of oligos to proofread the other is an effective approach that allows us to create oligos with an average of only one error in every 1,300 bases.

As one might expect, biological systems have an interest in copying themselves accurately, and our second method of error correction is borrowed directly from nature. One of us (Modrich) first worked out the details of the process 10 years ago and dubbed it "MutS, L, H." When two DNA strands hybridize but their A-T and C-G base pairing is not perfect, the double-stranded molecule will not assume a helix shape at the location of the mismatch. MutS is a naturally occurring protein that recognizes and binds to such imperfections and eventually recruits other proteins, MutL and MutH, to correct the error. One of us (Jacobson), with Peter Carr of M.I.T., has employed this system to achieve error rates of just one in 10,000 bases of synthetic DNA, which is sufficient fidelity to produce small networks of genes.

These technologies—releasable parallel synthesis and error correction—permit us to assemble long, relatively error-free DNA constructs far more rapidly and inexpensively than has been possible to date. They can therefore constitute the

GREETINGS FROM BACTERIA glow in this device entered in the 2004 iGEM competition by a team from the University of Texas at Austin. The group incorporated multiple light-receiving and color-producing genetic parts into *Escherichia coli* to turn a biofilm into biological film that displays images inscribed by light on its surface. In a nod to computer programming tradition, the machine's first message was "Hello World."

COURTESY OF RON WEISS *Princeton University (top)*; AARON CHEVALIER/ UNIVERSITY OF TEXAS AT AUSTIN/UNIVERSITY OF CALIFORNIA, SAN FRANCISCO *(bottom)*

THE ABSTRACTION ADVANTAGE

Biological engineers can benefit from methods that made very large scale integrated (VLSI) electronics practical for the semiconductor industry. Standardization of technologies allowed chip engineers to specialize in circuit design or fabrication and to thereby manage complex problems at different levels of abstraction. Bio fab engineers can also cope with complexity by using abstraction hierarchies to hide unnecessary information. Thus, a bio fab designer working at the level of whole systems need worry only about which devices to include and how to connect them to perform the desired function without having to manufacture each device from scratch. Similarly, a device-level designer should know the functions and compatibility of individual parts within a device, whereas a parts-level engineer should understand how each part works internally but need not be able to synthesize its DNA raw material.

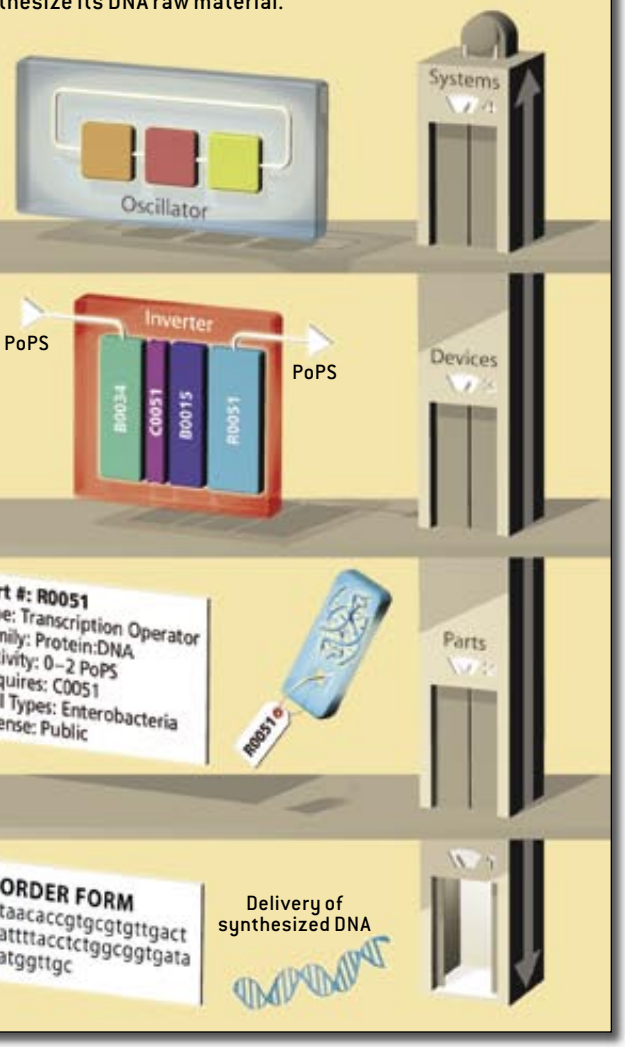
ABSTRACTION HIERARCHY

Systems
Combinations of biological devices that perform functions encoded by humans. A system of three inverters, for example, can operate as an oscillator.

Devices
Combinations of parts that perform discrete tasks. One inverter can take an input signal—for example, "HIGH"—and convert it to the opposite output signal, "LOW." A common signal carrier standard, polymerase per second (PoPS), allows devices to more easily be combined into systems.

Parts
Genetic material encoding biological functions. A transcription operator such as part #R0051, for example, is a piece of DNA that works with a matching binding protein (#C0051 in this case) to regulate gene activity. Off-the-shelf parts with clear specifications can be combined in a variety of devices.

DNA
Sequences for genetic parts. These can be specified by parts designers, manufactured off-site, then delivered. Fast synthesis technologies with low error rates make fabrication of custom DNA quick and reliable.



basis of a bio fab, and much like semiconductor chip lithography, these processes can be expected to keep steadily improving over time. That frees us to think about what we will build in the fab.

Nature Revised and Improved

AMONG OUR EARLIEST OBJECTIVES is using the bio fab platform to explore new ways of combating disease. Two of us (Keasling and Baker) run laboratories involved in creating cures for two of the most insidious diseases plaguing humanity: malaria and HIV. Although we are pursuing different types of therapeutics, the work of both our groups relies heavily on the ability to synthesize long, accurate pieces of DNA. Our projects thus offer examples of how the fab approach will considerably change the way biomedical scientists can go about developing new cures.

In the case of malaria, a treatment already exists that can eradicate the causative parasite from the body of an infected person. It is the small molecule C-15 sesquiterpene, commonly known as artemisinin, a natural compound made by the sweet wormwood plant, found mostly in northern China. The trees produce too little of the substance for the drug to be widely deployed at an affordable cost, however. That is why Keasling's group has been working for the past five years to copy the collection of genes, known as a genetic pathway, responsible for manufacturing artemisinin in the tree and

SLIM FILMS

inserting them into yeast to mass-produce the compound.

Once inside the yeast, this pathway can also be modified to operate much more efficiently than it does in the native plant. So far we have been able to redesign key subsets of the genes, known collectively as the mevalonate pathway, to produce an artemisinin precursor called amorphadiene at yields 100,000 times greater than the original pathway produces in bacteria. Increasing yields still further, to the point of making the drug widely available, will require us to reengineer the entire artemisinin pathway in an integrated way.

The full pathway consists of nine genes, each with an average length of about 1,500 DNA bases. Every new version of the pathway that we construct therefore contains approximately 13,000 bases. It would also be useful for us to be able to make variants of each gene in the pathway so that we could see which combinations worked most effectively. Manufacturing just two variants of each gene would mean synthesizing 2^9, or 512, constructs, for a total of about six million nucleotide bases. That is an extremely challenging goal using conventional DNA synthesis techniques, but that amount of DNA can fit on a single microarray chip.

The same technology that makes wholesale synthesis of gene networks possible can also be employed to generate novel proteins, such as new catalysts for synthetic chemistry reactions or environmental waste remediation and highly specific enzymes for gene therapy or pathogen destruction. Baker's group is developing computational methods for designing such new protein structures, including two that mimic essential features on the surface of HIV, which are already being tested as potential vaccines.

The trouble is that computer models are not sufficiently advanced to guarantee that each newly designed protein will have the desired function, but the computers can generate tens or hundreds of promising candidate structures to try. Turning all those into genetic sequences would require synthesis of hundreds of thousands of DNA bases—a difficult and expensive proposition using current technology but one well within the reach of the first generation of bio fab techniques.

These DNA and protein synthesis projects targeting malaria and HIV illustrate an approach, enabled by bio fab technology, that could be applied to a wider range of diseases, including newly emerging threats. For example, by combining high-speed, low-cost DNA sequencing methods [see "Genomes for All," by George M. Church; Scientific American, January] with the synthetic capabilities of the fab, a novel virus such as SARS or a new flu strain could be characterized, and protein-based vaccines against them could be readied far faster than is currently possible.

Of course, a bio fab is more than a collection of speedier synthesis technologies. It is a way of thinking about existing biological machines and of constructing new ones, which borrows both language and methodology from engineering.

BioBricks

In 2000 Michael Elowitz and Stanislas Leibler, then at Princeton University, as well as one of us (Collins), with colleagues Tim Gardner and Charles Cantor of Boston University, built the first basic circuit elements—a ring oscillator and a toggle switch—from biological parts. Scientists had known for some 25 years that natural organisms employ this type of circuitry to regulate their own genes, but the separate efforts of our two teams represented the first successes in manufacturing functional artificial biological circuitry.

Much of the future of biological technology will require many different groups to contribute subsystems.

What we mean by that term is well illustrated by Elowitz and Leibler's ring oscillator, which they began as an attempt to build a synthetic biological clock, hoping that it would provide insight into the clocks that exist naturally in biological systems. Their basic circuit consisted of a DNA ring called a plasmid containing three genes: *tetR*, *lacI* and λ *cI*, which encode the proteins TetR, LacI and λ cI, respectively. For any gene to be translated into a protein, the enzyme polymerase must first bind to a region of the DNA strand called a promoter that lies upstream of the gene. Polymerase then transcribes the gene into messenger RNA, which in turn is translated into a protein. If polymerase cannot bind the promoter, the gene is not translated and the protein is not made.

Elowitz and Leibler arranged for the protein products of the three genes in their circuit to selectively bind to one an-

THE AUTHORS

THE BIO FAB GROUP members are David Baker of the University of Washington, George Church of Harvard Medical School, Jim Collins of Boston University, Drew Endy and Joseph Jacobson of the Massachusetts Institute of Technology, Jay Keasling of the University of California, Berkeley, Paul Modrich of Duke University, Christina Smolke of the California Institute of Technology and Ron Weiss of Princeton University. They are friends, colleagues and sometime collaborators who wrote this article as a group because the diversity of their expertise, and hence of their contributions to the bio fab effort, embodies the interdisciplinary nature of biological engineering. All the authors are also scientific advisers to Codon Devices in Cambridge, Mass., the first commercial enterprise launched to apply engineering principles to synthetic biology. Church, Endy, Jacobson and Keasling are among its founders. Endy is also a founder of the not-for-profit BioBricks Foundation, and Keasling founded Amyris Biotechnologies.

Better Safety through Synthesis

Exploration of the many new opportunities a bio fab would offer for medicine, manufacture of new materials, sensors, waste remediation and energy production is just beginning. But like any worthwhile undertaking, it involves risk. A hallmark of biological systems is their ability to evolve and replicate, prompting understandable concerns that biological "devices" might cause unintentional or deliberate harm.

Thirty-one years ago a conference convened at Asilomar in California to address similar worries about the then new technology of recombinant DNA. For the first time, scientists could extract an individual gene from one organism and insert it into another, producing genetic combinations that might not exist naturally. That ability is now an essential tool in virtually every molecular biology laboratory in the world, in part because the governance that came from Asilomar eased fears about the use of recombinant DNA.

In a sense, then, the issues surrounding the new technologies that make up a fab for biology are not themselves new, but our community is committed to keep discussing them. A panel of scientists and ethicists tackled the implications of synthetic genomics at the Synthetic Biology 2.0 conference held in May. Its conclusions as well as results from a 15-month study of risks, benefits and potentially necessary safeguards, funded by the Alfred P. Sloan Foundation, will be available online at www.syntheticbiology.org in the coming months.

Right now scientists can certainly take the same precautions, such as working in secure biosafety laboratories, and observe the same ethical codes that have served us well for 30 years. Of course, ensuring that responsible investigators behave responsibly is easy. But the possibility that one day widespread access to DNA synthesis capability could allow malefactors to create deadly new pathogens, for example, is also a concern. It has prompted one of us (Church) to propose a compliance monitoring system that might include registration of synthetic biology workers—much as researchers working with so-called select agents are currently registered with the U.S. government—as well as surveillance of purchases of designer organisms, equipment and precursor materials for synthetic biology.

BIOBRICK VIALS contain DNA parts for biological engineering. These are safe for use in the lowest-level biosafety laboratories.

Another intriguing prospect is that the bio fab itself could represent the ultimate safe system because of the exquisite control it will permit. Most of the applications we have described would not require synthetic organisms ever to be exposed to the environment, but, just in case, those organisms could be created with genetic encoding different from any in nature, making it impossible for them to exchange genes with other life-forms. A synthetic biological device might be designed to self-destruct after a certain number of cell divisions or to be dependent for life on chemicals not present in normal environments. Genetic watermarks could be inscribed in every BioBrick to identify and track fabricated organisms. In other engineering disciplines, the ability to construct devices with higher precision affords greater safety—for example, in triply redundant flight-control systems in aircraft. We think the same may prove true of synthetic biological systems built in the bio fab.

—The Bio Fab Group

other's promoter regions. Thus, the LacI protein would bind the *tetR* promoter, whereas the λ cI protein would bind the *lacI* gene's promoter, and TetR would bind the promoter of the λ *cI* gene. These interrelations enable the protein product of one gene to block polymerase from binding to the promoter of another gene. Manufacture of the three proteins consequently happens in an oscillatory cycle: an abundance of LacI protein represses *tetR* gene activity; the absence of TetR protein then allows the λ *cI* gene to be turned on, which has the effect of repressing LacI production, and so on.

When one of the protein products in this cycle is also linked to a gene for making a green fluorescent protein and the entire circuit is inserted into bacteria, the oscillation of this device can be observed as the bacteria blink on and off like holiday lights. Similarly, the latest version of the Collins group's genetic toggle switch can be used to program bacteria to detect cellular DNA damage and then report their findings by arranging themselves into a green fluorescent "lawn" known as a biofilm.

Perhaps the most striking thing about these synthetic biological circuits is that they are identical in function to the first types of circuits that electrical engineers build when they want to test a new process for manufacturing semiconductor chips. Engineers know that basic components, such as an oscillator or a switch, are logically complete. Being able to build these simple parts reliably and accurately makes it possible to design and fabricate much more complex circuitry. Once biological engineers, too, can take such basic building blocks for granted, they can move on to more complicated projects such as multicellular systems, two- and three-dimensional designs, and devices whose function is not biological.

One of us (Weiss) recently produced a prototype for a multicellular system that could be used, for example, to detect

SAM OGDEN

explosives or other chemicals and then report them with a visible signal [*see box on page 47*]. This biological machine allows us to program millions of bacterial cells with instructions and protocols both for communicating with one another while carrying out their orders and for outputting light signals in a variety of patterns.

Inspired by these early examples, one of us (Endy), along with Knight and our M.I.T. colleague Randy Rettberg, is developing a library of biological components similar to the libraries available to chip designers. This Registry of Standard Biological Parts should facilitate a wide range of biological building projects, and our hope is that others will contribute new entries. So far the registry contains more than 1,000 individual BioBricks, as we call them, including many parts analogous to electronics, such as inverters, switches, counters, amplifiers, and components that can receive input or output a display. We have also defined a standard signal carrier—polymerase per second, or PoPS—akin to the current in a wire connecting two electronic components, so that bio fab engineers can more easily combine and reuse genetic devices.

To demonstrate the power of the fab approach and to seed this new field, the M.I.T. group offered the first course in fab-style engineering with biological parts in 2003. That class quickly evolved into an annual competition, which will draw teams from more than 30 universities this summer. In its short existence, the International Genetically Engineered Machine (iGEM) contest has already generated a number of amazing cellular devices, including biofilms that can record and display a photograph as well as programmed cells able to sense and respond like switches to small-molecule inputs, such as caffeine.

Still another iGEM entry that three of us (Smolke, Collins and Church) developed was a device capable of digital counting using a series of DNA segments. Twenty such DNA bits would be enough to count and report up to one million (2^{20}) cell states. This technology can be incorporated into sensors, which in turn might be connected to engineered metabolic pathways, such as the Keasling group's optimized version for artemisinin production. That would allow increased manufacture of the desired drug at literally the flip of a switch.

Constructing Synthetic Biology

WHEN WE AUTHORS of this article began our efforts to build a bio fab, no clear approach existed to making long DNA constructs accurately, rapidly and inexpensively. Today that is one among several technologies in an expanding toolbox for biological engineering. We are progressing toward first designing and modeling biological devices in computers, then "cutting" them into biological form as the final step—much as silicon chips are planned, then etched.

As with semiconductor circuitry, this approach has the added benefit of allowing us to optimize interactions between parts and to anticipate bugs. This ability grows increasingly useful as the constructed systems become increasingly complex. Yet another advantage of designing in the abstract is that a biological engineer does not need to actually build every part from scratch or even know how every one of them works internally—only that they do so reliably.

The students participating in iGEM may represent the first generation of biologist-engineers trained from the beginning of their careers to think of themselves as both. An important challenge going forward, however, will be to get more biologists to think like silicon engineers (and lure more engineers into biology)—particularly when it comes to sharing parts. Until now, biotechnology has been characterized by self-contained teams working to develop single-purpose applications, such as one drug compound. Much of the future of biological technology will require many different groups to contribute subsystems. Our hope is that building a fab for biology will facilitate that progression and help to spur advances as revolutionary as those achieved in the semiconductor industry. SA

Bio Fab Beginnings

A handful of companies and organizations are already applying engineering principles and tools to commercial biological manufacturing, bringing the fab closer to reality.

COMPANY	FOCUS
BioBricks Foundation Cambridge, Mass.	Promoting open tools, standards and parts for biological engineering
Blue Heron Biotechnology Bothell, Wash.	DNA synthesis
Amyris Biotechnologies Emeryville, Calif.	Engineering metabolic pathways for drug production in microbes
Codon Devices Cambridge, Mass.	Building biological devices
Foundation for Applied Molecular Evolution Gainesville, Fla.	Generating novel proteins and materials
Synthetic Genomics Rockville, Md.	Engineering microbes to produce fuels

MORE TO EXPLORE

Synthetic Life. W. Wayt Gibbs in *Scientific American,* Vol. 290, No. 5, pages 74–81; May 2004.

Foundations for Engineering Biology. Drew Endy in *Nature*, Vol. 438, pages 449–453; November 24, 2005.

Let Us Go Forth and Safely Multiply. George Church in *Nature,* Vol. 438, page 423; November 24, 2005.

Adventures in Synthetic Biology. Drew Endy, Isadore Deese, the M.I.T. Synthetic Biology Working Group and Chuck Wadey. A comic book available online at **http://openwetware.org/wiki/Adventures**

Registry of Standard Biological Parts: **http://parts.mit.edu/**