SYNTHETIC GENOMICS

BUILDING A BETTER BACTERIUM

The May 20, 2010, online edition of *Science* magazine contained pieces on Brownian motion and gravitational waves, small RNAs and drug delivery—items of interest to narrow slices of the research community. One article, though, generated instant worldwide attention. Entitled "Creation of a bacterial cell controlled by a chemically synthesized genome," the report detailed the world's first "synthetic cell," and it was at once praised and panned. Watchdog groups weighed in, as did U.S. President Barack Obama. Powered by advances in DNA synthesis and genome manipulation, the study was merely a proof-of-principle: *Mycoplasma mycoides JCVI-syn1.0* has no practical scientific or commercial value. Yet its cobalt blue colonies represent the living embodiment of an entirely new, and previously unimaginable, branch of biology. Welcome to the age of synthetic genomics.
By Jeffrey M. Perkel

"This is the control experiment. We are now at stage one."

"Synthetic genomics," reads the introduction to *Synthetic Genomics: Options for Governance,* a report by the **J. Craig Venter Institute** (JCVI), **Massachusetts Institute of Technology,** and the **Center for Strategic & International Studies,** "combines methods for the chemical synthesis of DNA with computational techniques to design it." That doesn't sound all that different from the standard molecular biology researchers have been doing for decades, and in some respects, it isn't; what's different is the incorporation of design and engineering sensibilities—not to mention the scale of the science. "These methods allow scientists to construct genetic material that would be impossible or impractical to produce using more conventional biotechnological approaches." (See report, www.jcvi.org/cms/research/projects/syngen-options/overview/)

Researchers have been making point mutations, cloning genes, and designing novel biological circuits for years. They can even transplant biological pathways, using what James Collins, a synthetic biologist and **Howard Hughes Medical Institute** investigator at **Boston University** calls "genetic engineering on steroids." (As the JCVI report notes, "There is no clear and unambiguous threshold between synthetic genomics and more conventional approaches to biotechnology.") But it can be a long, laborious process; by J. Craig Venter's estimation, **DuPont**'s development of microbes that can spin glucose into propanediol, a precursor to the company's Sorona synthetic polymer, required "10 years and well over $100 million." And that's just one pathway; rewriting a biological operating system from the ground up is a different matter entirely.

Enter synthetic genomics. Fueled by advances in gene building, metagenomics, and bio-circuitry design, researchers are coaxing microbes to do things never before possible—albeit not yet at the genomic scale. But that could soon change; in the not-too-distant future, says Collins, it may be possible to design a minimally functional genome, fold in novel or desired biochemical circuits, synthesize the DNA, and go. Venter has formed a company to do just that; **Synthetic Genomics** is using the technology to develop algae capable of cranking out faster, cheaper, and better biofuels and agricultural products, striking a $300 million deal with **ExxonMobil Research and Engineering** in 2009 to advance that aim. But they're not there yet. "To my viewpoint," says Venter, whose eponymous Institute performed the synthetic cell work, "this is the control experiment. We are now at stage one."

THE SYNTHETIC CELL

Mycoplasma mycoides JCVI-syn1.0 was the product of some 15 years and $40 million worth of effort by Venter, Clyde Hutchison, Hamilton Smith, and about two-dozen others at the JCVI. The team first sequenced and then chemically synthesized the genome of the bacterium, *Mycobacterium mycoides*, and then inserted it into a related organism, *M. capricolum*. In the parlance of synthetic biology, *M. capricolum* served as a "chassis"—a microbial shell. Loaded with the genetic operating system of its close cousin, it was then "rebooted" to produce a living synthetic cell.

Bioethicist Arthur Caplan, writing in *Nature*, called the work "one of the most important scientific achievements in the history of mankind." Others were more measured; *New York Times* science writer Nicholas Wade called the research "a matter of scale rather than a scientific breakthrough." The U.K.'s *Daily Mail*, in a bit of nuanced headline writing (and while simultaneously invoking the specter of global pandemic as in the Will Smith movie, *I Am Legend*), declared: "Scientist accused of playing God after

UPCOMING FEATURES

Small Devices—April 1

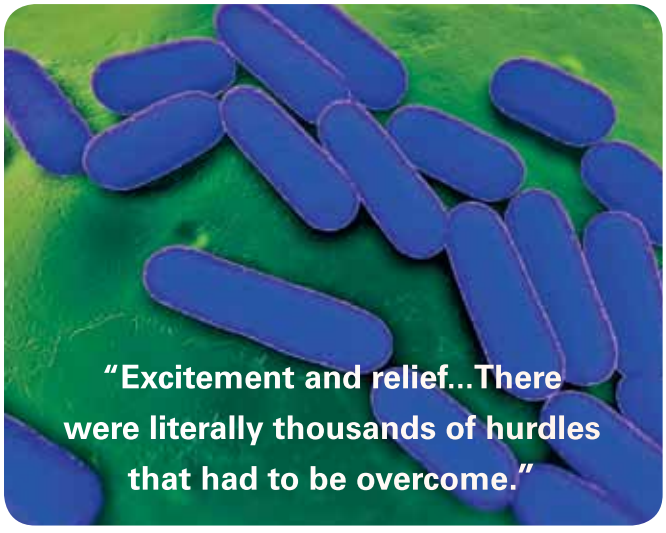
Cell Separation Technologies—May 6

Food Safety—June 24

creating artificial life by making designer microbe from scratch—but could it wipe out humanity?"

The answer to that question is undeniably no; Venter's team merely recapitulated the genome of *M. mycoides* (with the addition of a few "watermarks" and other small genetic tweaks) and transplanted it into the functioning membranes and cytoplasm of a close relative. If *M. mycoides* cannot wipe out humanity, neither can its lab-bred cousin.

To build the genome, Venter and his team turned to **Blue Heron** (acquired by OriGene Technologies in 2010). Unlike most oligonucleotide synthesis firms, which specialize in cranking out polymerase chain reaction primers by the thousands, Blue Heron (and other gene synthesis companies, including **GENEART, Gene Oracle,** and **DNA 2.0**) has mastered the art of synthesizing relatively long, entirely accurate sequences, and stringing them together to create gene-sized fragments on the order of hundreds to thousands of bases. Venter's group ordered up 1,078 1-kb "cassettes," the building blocks of the *M. mycoides* genome.



The team had already demonstrated they could assemble complete genomes, having successfully built both an intact functional virus (the 5-kb phiX174) and a bacterial genome (the 583-kb *M. genitalium*). They also showed they could transplant a natural (i.e., nonsynthetic) chromosome from one cell to another. The next step, synthesizing a genome and transplanting it, should have been simple. Yet according to Venter, the process involved "invention after invention after invention of new ways to do things"—everything from synthesis and recombination to handling bacterial restriction systems. Even DNA manipulation proved problematic. "You can't pipette whole chromosomes without just the shearing forces from pipetting tearing the DNA," he says; as a result, the team took to moving its DNA around in agarose plugs.

Using the synthetic cassettes from Blue Heron, the team assembled the genome via stepwise homologous recombination in yeast, building first 10-kb pieces, then 100-kb, and finally the complete 1,077,947-bp chromosome. Highlighting the importance of accurate DNA synthesis, a single error in the *dnaA* coding sequence set the team back three months.

In the end, a single bright blue colony signaled success. Upon receiving the news from project leader Dan Gibson, Venter says he felt "excitement and relief… There were literally thousands of hurdles that had to be overcome."

THE PROBLEM WITH BIOLOGY

Venter calls the resulting organism a "synthetic cell," and the applications of the technology used to make it run the gamut from bioengineering to basic biology. Chief Scientific Officer Richard Roberts of **New England Biolabs,** which supplies reagents supporting synthetic biology, suggests one possible use: designing organisms in which one of the 64 triplets is reassigned to some novel, non-natural amino acid. That would require a complete genomic rewrite, as well as the insertion of additional machinery, such as new aminoacyl-tRNA synthetases. "That's not something you could do by mutagenesis or by any sort of simple genetic engineering methods," he says.

First, though, researchers will have to bone up on their biology. Genome sequencing and metagenomics efforts have filled databases to overflowing with novel genes, but researchers simply don't know what many of them do. Even less well understood are the regulatory layers controlling those activities. Venter's study, says Raik Grünberg, a postdoctoral fellow at the **Centre for Genomic Regulation-European Molecular Biology Laboratory (CGR-EMBL)** Systems Biology Unit in Barcelona who develops synthetic biological circuits, highlights not only a technological development, but also researchers' biological ignorance. "It shows that we can now write genomes. But at the same moment everyone is realizing … we don't really know what to write."

Another problem is that it's one thing to draw a straightforward pathway on paper; it's quite another to make it work in practice. Unlike the electrical circuits on which those drawings are based, biology simply isn't binary, but stochastic. Promoters aren't 100 percent on or off, for instance, and operator sequences are not all the same. "It can take only a matter of few days or weeks to design a synthetic gene circuit to look like the schematic," says Collins, "but it can take many months to try to actually construct it so that it functions as desired." What inevitably follows is a long period of what Collins calls "post-hoc tweaking."

"That's where most of us spend most of our time," he says.

The effort can pay big dividends, however, as with the bioengineering of microbes that can synthesize artemisinic acid, a precursor to the antimalarial drug artemisinin. Artemisinin is a terpenoid normally extracted from wormwood, a lengthy and expensive process. **University of California-Berkeley** Professor Jay Keasling led that effort, which took the better part of a decade, to provide a rapid, reliable, and low-cost source of the drug. Microbially derived artemisinin, he says, could eventually cost just a tenth of the native material. "We might be able to save half a million children a year," Keasling says.

Keasling's team started by transplanting a yeast mevalonate isoprenoid pathway and a synthetic (codon-optimized) amorphadiene synthase gene into *E. coli*, creating a strain capable of turning sugar into amorphadiene, a precursor to artemisinic acid. The next biosynthetic step is a series of oxidation reactions, all of which require a cytochrome-P450. There the team hit a stumbling block, as that enzyme's identity was unknown. But with luck, and some comparative genomics, the team cloned the necessary gene and transferred it into yeast, giving a strain that could produce artemisinic acid. The final step was to migrate the entire pathway back into *E. coli*.

According to Keasling, this work—supported by $42 million from the **Bill and Melinda Gates Foundation**—represents the culmination of years of genetic tinkering with promoters and ribosome binding sites, RNA stabilization elements continued »

and transcription factor operators. One key problem, he says, was that one of the intermediates (hydroxymethylglutaryl-CoA) is actually toxic to *E. coli*. Once the team identified that step, they tweaked it by both suppressing the biosynthetic enzyme and activating the utilization enzyme. They also constructed a synthetic protein scaffold—a kind of biological assembly line—to "channel" metabolic intermediates from enzyme to enzyme and prevent them from accumulating, increasing output an additional 75-fold. The whole process, Roberts says, represents "probably … the most complicated genetic engineering feat to date."

Keasling licensed the work to a spin-off company called **Amyris Biotechnologies,** which in turn licensed it to **Sanofi Aventis.** "Right now, they are scaling up the process, and we should have artemisinin out late this year or early next year," he says.

RNA SOLUTIONS

Such feats of bioengineering highlight the power of synthetic biology. Yet their nearly universal reliance on protein-mediated regulation underscores one of its shortcomings, as well. "There's a design gap right now in synthetic biology," says Christina Smolke, assistant professor of bioengineering at **Stanford University.** Natural biological systems, she says, "have very complex regulatory strategies in play. And they're layering different mechanisms—not just transcription, but RNA-based mechanisms and posttranslational mechanisms. So everything is very tightly regulated."

RNA-based regulators, for instance, have different kinetics and are more malleable than protein, with relatively simple folding rules and a research-friendly modular architecture. They also exact less of an energetic burden on the cell. "As we start to think about genome design," Smolke says, "issues [such as] the energetic cost of the entire system and how much energy it requires to run all the programs you want to actually get in there, become significant."

Smolke's lab, which builds microbes capable of synthesizing benzylisoquinoline alkaloids (another class of pharmacologically interesting plant-derived compounds), is developing regulatory RNA modules to try to incorporate some subtlety into its synthetic circuitry. In a report published last November in *Science*, her team described synthetic mini-genes with built-in RNA modules that would, upon sensing the presence of one or more cell-signaling proteins, induce an alternative splicing event that up- or down-regulates the expression of either a fluorescent reporter or a pro-apoptotic gene.

According to Smolke, the regulatory modules comprise three elements—a sensor, an actuator, and an information processor that links the two—all contained within a three-exon, two-intron synthetic construct encoding the output gene. The approach is completely generalizable, she says. Her team used the approach to make cells responsive to signaling through disease pathways, but it could be used, for instance, to keep toxic metabolites in check; all the researcher needs do is exchange one sensing element for another. Even the actuator is modular; Smolke's lab has used microRNAs, antisense RNAs, and even ribozymes.

Such regulators could help researcher exert finer control over synthetic systems. But they also add another layer of complexity for those who would design novel genomes. Enter **University of California, San Francisco** biologist Christopher Voigt. Voigt has been engineering logic circuits, like NOR and XOR gates, from synthetic DNA and *E. coli*.

Yet circuitry represent just half of the problem of programming cells, Voigt says; the other half is software. Just as computer programmers would rather code in high-level languages like C++ than in the 1s and 0s of the computer, so too is it easier to instruct DNA synthesizers in a high level language than in the language of As, Cs, Gs, and Ts—especially when writing sequences the size of a genome.

Voigt is now working with **Life Technologies** to develop a "genetic compiler" and language for programming synthetic genomes. The compiler would reduce human-readable instructions to a series of fundamental components, which could then be strung together *in silico* and synthesized in vitro. "The idea is for Life Technologies to have it where you write your genetic code like C++ and [the software] converts it into a DNA sequence that they synthesize for you," he says.

In the short term at least, most such work will continue to be done at the level of individual circuits and pathways. But technology evolves, and with it, science itself. Already, new biological vistas are opening. Says Luis Serrano, head of the CRG-EMBL Systems Biology Unit (and Grünberg's advisor), "If you can make a genome from scratch, now people can play. And by playing you can learn. And by learning we will be able in the future to engineer [genomes] better, or even to design them from scratch."

FEATURED PARTICIPANTS

Amyris Biotechnologies
www.amyrisbiotech.com

Bill and Melinda Gates Foundation
www.gatesfoundation.org

Blue Heron
www.blueheronbio.com

Boston University
www.bu.edu

Center for Strategic and International Studies
www.csis.org

Centre for Genomic Regulation
pasteur.crg.es/portal/page/portal/Internet

DNA 2.0
www.dna20.com

DuPont
www.dupont.com

European Molecular Biology Laboratory
www.embl.org

ExxonMobil Research and Engineering
www.exxonmobil.com

Gene Oracle
www.geneoracle.com

GENEART
www.geneart.com

Howard Hughes Medical Institute
www.hhmi.org

J. Craig Venter Institute
www.jcvi.org

Life Technologies
www.lifetechnologies.com

Massachusetts Institute of Technology
www.mit.edu

New England Biolabs
www.neb.com

Sanofi Aventis
www.sanofi-aventis.us

Stanford University
www.stanford.edu

Synthetic Genomics
www.syntheticgenomics.com

University of California, Berkeley
www.berkeley.edu

University of California, San Francisco
www.ucsf.edu

Jeffrey M. Perkel is a freelance science writer based in Pocatello, Idaho.

DOI: 10.1126/science.opms.p1100053