In brief

**Synthetic biologists have been** creating the genomes of

Nineteenth-century novels are typically fodder for literature conferences, not scientific gatherings. Still, at a **high-profile meeting**
organisms such as viruses and bacteria for the past 15 years. They aim to use these designer genetic codes to make cells capable of producing novel therapeutics and fuels. Now, some of these scientists have set their sights on synthesizing the human genome—a vastly more complex genetic blueprint. Read on to learn about this initiative, called Genome Project-write, and the challenges researchers will face—both technical and ethical—to achieve success.

<http://engineeringbiologycenter.org/may-2017-meeting/> of about 200 synthetic biologists in May, one presenter highlighted Mary Shelley’s gothic masterpiece “Frankenstein,” which turns 200 next year.

Frankenstein’s monster, after all, is what many people think of when the possibility of human genetic engineering is raised, said University of Pennsylvania ethicist and historian Jonathan Moreno. The initiative being discussed at the New York City meeting—Genome Project-write<http://science.sciencemag.org/content/early/2016/06/01/science.aaf6850>—has been dogged by worries over creating unnatural beings. True, part of GP-write aims to synthesize from scratch all 23 chromosomes of the human genome and insert them into cells in the lab. But proponents of the project say they’re focused on decreasing the cost of synthesizing and assembling large amounts of DNA rather than on creating designer babies.

The overall project is still under development, and the project’s members have not yet agreed on a specific road map for moving forward. It’s also unclear where funding will
What the members of GP-write do agree on is that creating a human genome from scratch is a tremendous scientific and engineering challenge that will hinge on developing new methods for synthesizing and delivering DNA. They will also need to get better at designing large groups of genes that work together in a predictable way, not to mention making sure that even larger assemblies—genomes—can function.

GP-write consortium members argue that these challenges are the very thing that should move scientists to pick up the DNA pen and turn from sequence readers to writers. They believe writing the entire human genome is the only way to truly understand how it works. Many researchers quoted Richard Feynman during the meeting in May. The statement “What I cannot create, I do not understand” was found on the famed physicist’s California Institute of Technology blackboard after his death. “I want to know the rules that make a genome tick,” said Jef Boeke, one of GP-write’s four coleaders, at the meeting.

To that end, Boeke and other GP-write supporters say the initiative will spur the development of new technologies for designing genomes with software and for synthesizing DNA. In turn, being better at designing and assembling genomes will yield synthetic cells capable of producing valuable fuels and drugs more efficiently. And turning to human genome synthesis will enable new cell therapies and other medical advances.
The Gutenberg stage

Genome writers have already synthesized a few complete genomes, all of them much less complex than the human genome. For instance, in 2002, researchers chemically synthesized a DNA-based equivalent of the poliovirus RNA genome, which is only about 7,500 bases long. They then showed that this DNA copy could be transcribed by RNA polymerase to recapitulate the viral genome, which replicated itself—a demonstration of synthesizing what the authors called “a chemical \([\text{C}_{332,652}\text{H}_{492,388}\text{N}_{98,245}\text{O}_{131,196}\text{P}_{7,501}\text{S}_{2,340}]\) with a life cycle” (Science 2002, DOI: 10.1126/science.1072266). After tinkering with a handful of other viral genomes, in 2010, researchers advanced to bacteria, painstakingly assembling a Mycoplasma genome just over about a million bases in length and then transplanting it into a host cell.

Last year, researchers upped the ante further, publishing the design for an aggressively edited Escherichia coli genome measuring 3.97 million bases long (Science, DOI: 10.1126/science.aaf3639). GP-write coleader George Church and coworkers at Harvard used DNA-editing software—a kind of Google Docs for writing genomes—to make radical systematic changes. The so-called rE.coli-57 sequence, which the team is currently synthesizing, lacks seven codons (the three-base DNA “words” that code for particular amino acids) compared with the normal E. coli genome. The researchers replaced all 62,214 instances of those codons with DNA base synonyms to eliminate redundancy in the code.

Status report

International teams of researchers have already synthesized six of yeast's 16 chromosomes, redesigning the organism's genome as part of the Sc2.0 project.
Bacterial genomes are no-frills compared with those of creatures in our domain, the eukaryotes. Bacterial genomes typically take the form of a single circular piece of DNA that floats freely around the cell. Eukaryotic cells, from yeast to plants to insects to people, confine their larger genomes within a cell’s nucleus and organize them in multiple bundles called chromosomes. An ongoing collaboration is now bringing genome synthesis
Researchers are building a fully synthetic yeast genome, containing 17 chromosomes that range from about 1,800 to about 1.5 million bases long. Overall, the genome will contain more than 11 million bases.

The synthetic genomes and chromosomes already constructed by scientists are by no means simple, but to synthesize the human genome, scientists will have to address a whole other level of complexity. Our genome is made up of more than 3 billion bases across 23 paired chromosomes. The smallest human chromosome is number 21, at 46.7 million bases—larger than the smallest yeast chromosome. The largest, number 1, has nearly 249 million. Making a human genome will mean making much more DNA and solving a larger puzzle in terms of assembly and transfer into cells.

Today, genome-writing technology is in what Boeke, also the director for the Institute of Systems Genetics at New York University School of Medicine, calls the “Gutenberg phase.” (Johannes Gutenberg introduced the printing press in Europe in the 1400s.) It’s still early days.

DNA synthesis companies routinely create fragments that are 100 bases long and then use enzymes to stitch them together to make sequences up to a few thousand bases long, about the size of a gene. Customers can put in orders for small bits of DNA, longer strands called oligos, and whole genes—whatever they need—and companies will fabricate and mail the genetic material.

Although the technology that makes this mail-order system possible is impressive, it’s not prolific enough to make a human genome in a reasonable amount of time. Estimates vary on how long it would take to stitch together a more than 3 billion-base human genome and how much it would cost with today’s methods. But the ballpark answer is about a decade and hundreds of millions of dollars.

Synthesis companies could help bring those figures down by moving past their current 100-base limit and creating longer DNA fragments. Some researchers and companies are moving in that direction. For example, synthesis firm Molecular Assemblies is developing an enzymatic process to write long stretches of DNA with fewer errors.
Synthesis speeds and prices have been improving rapidly, and researchers expect they will continue to do so. “From my point of view, building DNA is no longer the bottleneck,” says Daniel G. Gibson, vice president of DNA technology at Synthetic Genomics and an associate professor at the J. Craig Venter Institute (JCVI). “Some way or another, if we need to build larger pieces of DNA, we'll do that.”

---

**Genome writing: Published and drafted**

The genomes that scientists have been building for the past 15 years have increased in complexity.
**Poliovirus**

**Genome size:** about 7,500 bases

**Year synthesis was reported:** 2002

**Notes:** Researchers first built a DNA copy of the virus’s RNA genome, then translated it into RNA with enzymes.

---

**Mycoplasma genitalium**

**Genome size:** about 583,000 bases

**Year synthesis was reported:** 2008

**Notes:** Researchers disrupted one gene in the native sequence of the bacterium to prevent it from being pathogenic, added an antibiotic-resistance gene, and inserted “watermarks” to identify the genome as synthetic.
**Mycoplasma mycoides**

**Genome size:** about 1.08 million bases

**Year synthesis was reported:** 2010

**Notes:** Researchers synthesized this bacterium's genome and then transplanted it into a host cell. Later, in 2016, researchers stripped this same genome down to about 531,000 bases, deleting all genes nonessential to survival, and showed that it could keep a so-called minimal cell alive and growing.

**Escherichia coli**

**Synthetic genome size:** about 3.97 million bases

**Year the design was published:** 2016

**Notes:** Researchers radically changed the E. coli genome, replacing seven codons (the three-letter codes for amino acids) with synonyms. They are now synthesizing the entire thing from scratch.
Yeast (*Saccharomyces cerevisiae*)

**Synthetic genome size:** about 11 million bases total over 16 chromosomes, plus one “neochromosome”

**First chromosome completed:** 2011

**Expected finish:** 2020

**Smallest chromosome:** 1,800 bases

**Largest chromosome:** 1.5 million bases

**Notes:** Researchers have made widespread changes to the yeast genome, calling their modified organism Sc2.0. They are replacing one of yeast’s chromosomes at a time and then breeding cells to make an organism with all synthetic chromosomes.

Human

*Not yet started*

**Genome size:** more than 3 billion bases total across 23 paired chromosomes

**Expected time to finish:** 10 years

**Smallest chromosome:** 46.7 million bases

**Largest chromosome:** 249 million bases

**Mitochondrial DNA size:** 16,569 bases

**Notes:** Researchers are still debating the ethics of this endeavor and are trying to work out exactly how this synthesis might be completed.

Credit: Will Ludwig/C&EN/Shutterstock
First words

Gibson isn’t involved with GP-write. But his research showcases what is possible with today’s tools—even if they are equivalent to Gutenberg’s movable type. He has been responsible for a few of synthetic biology’s milestones, including the development of one of the most commonly used genome-assembly techniques.

The Gibson method <http://dx.doi.org/10.1038/nmeth.1318> uses chemical means to join DNA fragments, yielding pieces thousands of bases long. For two fragments to connect, one must end with a 20- to 40-base sequence that’s identical to the start of the next fragment. These overlapping DNA fragments can be mixed with a solution of three enzymes—an exonuclease, a DNA polymerase, and a DNA ligase—that trim the 5’ end of each fragment, overlap the pieces, and seal them together.

To make the first synthetic bacterial genome <http://cen.acs.org/articles/86/i4/Bacterial-Genome-Made-Scratch.html> in 2008, that of Mycoplasma genitalium, Gibson and his colleagues at JCVI, where he was a postdoc at the time, started with his eponymous in vitro method. They synthesized more than 100 fragments of synthetic DNA, each about 5,000 bases long, and then harnessed the prodigious DNA-processing properties of yeast, introducing these large DNA pieces to yeast three or four at a time. The yeast used its own cellular machinery to bring the pieces together into larger sequences, eventually producing the entire Mycoplasma genome.

Next, the team had to figure out how to transplant this synthetic genome into a bacterial cell to create what the researchers called the first “synthetic cell.” The process is involved and requires getting the bacterial genome out of the yeast, then storing the huge, fragile piece of circular DNA in a protective agarose gel before melting it and mixing it with another species of Mycoplasma. As the bacterial cells fuse, some of them take in the synthetic genomes floating in solution. Then they divide to create three daughter cells, two containing the native genomes, and one containing the synthetic genome: the synthetic cell.

When Gibson’s group at JCVI started building the synthetic cell in 2004, “we didn’t know what the limitations were,” he says. So the scientists were cautious about overwhelming
Gibson expects that existing DNA synthesis and assembly methods haven’t yet been pushed to their limits. Yeast might be able to assemble millions of bases, not just hundreds of thousands, he says. Still, Gibson believes it would be a stretch to make a human genome with this technique.

**Model organism**

One of the most ambitious projects in genome writing so far centers on that master DNA assembler, yeast. As part of the project, called Sc2.0 (a riff on the fungus’s scientific name, *Saccharomyces cerevisiae*), an international group of scientists is redesigning and building yeast one synthetic chromosome at a time. The yeast genome is far simpler than ours. But like us, yeasts are eukaryotes and have multiple chromosomes within their nuclei.

Synthetic biologists aren’t interested in rebuilding existing genomes by rote; they want to make changes so they can probe how genomes work and make them easier to build and reengineer for practical use. The main lesson learned from Sc2.0 so far, project scientists say, is how much the yeast chromosomes can be altered in the writing, with no apparent ill effects. Indeed, the Sc2.0 sequence is not a direct copy of the original. The synthetic genome has been reduced by about 8%. Overall, the research group will make 1.1 million bases’ worth of insertions, deletions, and changes to the yeast genome (Science 2017, DOI: [10.1126/science.aaf4557](http://dx.doi.org/10.1126/science.aaf4557)).

So far, says Boeke, who’s also coleader of Sc2.0, teams have finished or almost finished the first draft of the organism’s 16 chromosomes. They’re also working on a “neochromosome,” one not found in normal yeast. In this chromosome, the designers have relocated all DNA coding for transfer RNA, which plays a critical role in protein assembly. The Sc2.0 group isolated these sequences because scientists predicted they would cause structural instability in the synthetic chromosomes, says Joel Bader, a computational biologist at Johns Hopkins University who leads the project’s software and design efforts.
The team is making yeast cells with a new chromosome one at a time. The ultimate goal is to create a yeast cell that contains no native chromosomes and all 17 synthetic ones. To get there, the scientists are taking a relatively old-fashioned approach: breeding. So far, they've made a yeast cell with three synthetic chromosomes and are continuing to breed it with strains containing the remaining ones. Once a new chromosome is in place, it requires some patching up because of recombination with the native chromosomes. “It’s a process, but it doesn’t look like there are any significant barriers,” Bader says. He estimates it will take another two to three years to produce cells with the entire Sc2.0 genome.

So far, even with these significant changes to the chromosomes, the yeast lives at no apparent disadvantage compared with yeast that has its original chromosomes. “It’s surprising how much you can torture the genome with no effect,” Boeke says.

Boeke and Bader have founded a start-up company called Neochromosome that will eventually use Sc2.0 strains to produce large protein drugs, chemical precursors, and other biomolecules that are currently impossible to make in yeast or E. coli because the genetic pathways used to create them are too complex. “With synthetic chromosomes we’ll be able to make these large supportive pathways in yeast,” Bader predicts.

**Design within reach**

Whether existing genome-engineering methods like those used in Sc2.0 will translate to humans is an open question.

Bader believes that yeast, so willing to take up and assemble large amounts of DNA, might serve as future human-chromosome producers, assembling genetic material that could then be transferred to other organisms, perhaps human cells. Transplanting large human chromosomes would be tricky, Synthetic Genomics’ Gibson says. First, the recipient cell must be prepped by somehow removing its native chromosome. Gibson expects physically moving the synthetic chromosome would also be difficult: Stretches of DNA larger than about 50,000 bases are fragile. “You have to be very gentle so the chromosome doesn’t break—once it’s broken, it’s not going to be useful,” he says. Some researchers are working on more direct methods for cell-to-cell DNA transfer, such as getting cells to fuse with one another.
Once the scientists solve the delivery challenge, the next question is whether the transplanted chromosome will function. Our genomes are patterned with methyl groups that silence regions of the genome and are wrapped around histone proteins that pack the long strands into a three-dimensional order in cells’ nuclei. “If the synthetic chromosome doesn’t have the appropriate methylation patterns, the right structure, it might not be recognized by the cell,” Gibson says.

Biologists might sidestep these epigenetic and other issues by doing large-scale DNA assembly in human cells from the get-go. Ron Weiss, a synthetic biologist at Massachusetts Institute of Technology, is pushing the upper limits on this sort of approach. He has designed methods for inserting large amounts of DNA directly into human cells. Weiss endows human cells with large circuits, which are packages of engineered DNA containing groups of genes and regulatory machinery that will change a cell’s behavior.

In 2014, Weiss developed a “landing pad” method to insert about 64,000-base stretches of DNA into human and other mammalian cells. First, researchers use gene editing to create the landing pad, which is a set of markers at a designated spot on a particular chromosome where an enzyme called a recombinase will insert the synthetic genetic material. Then they string together the genes for a given pathway, along with their regulatory elements, add a matching recombinase site, and fashion this strand into a circular piece of DNA called a plasmid. The target cells are then incubated with the plasmid, take it up, and incorporate it at the landing site (Nucleic Acids Res. 2014, DOI: 10.1093/nar/gku1082<http://dx.doi.org/10.1093/nar/gku1082>).

This works, but it’s tedious. “It takes about two weeks to generate these cell lines if you’re doing well, and the payload only goes into a few of the cells,” Weiss explains. Since his initial publication, he says, his team has been able to generate cells with three landing pads; that means they could incorporate a genetic circuit that’s about 200,000 bases long.

Weiss doesn’t see simple scale-up of the landing pad method as the way forward, though, even setting aside the tedium. He doesn’t think the supersized circuits would even function in a human cell because he doesn’t yet know how to design them.
“The limiting factor in the size of the circuit is not the construction of DNA, but the design,” Weiss says. Instead of working completely by trial and error, bioengineers use computer models to predict how synthetic circuits or genetic edits will work in living cells of any species. But the larger the synthetic element, the harder it is to know whether it will work in a real cell. And the more radical the deletion, the harder it is to foresee whether it will have unintended consequences and kill the cell. Researchers also have a hard time predicting the degree to which cells will express the genes in a complex synthetic circuit—a lot, a little, or not at all. Gene regulation in humans is not fully understood, and rewriting on the scale done in the yeast chromosome would have far less predictable outcomes.

Besides being willing to take up and incorporate DNA, yeast is relatively simple. Upstream from a yeast gene, biologists can easily find the promoter sequence that turns it on. In contrast, human genes are often regulated by elements found in distant regions of the genome. That means working out how to control large pathways is more difficult, and there’s a greater risk that changing the genetic sequence—such as deleting what looks like repetitive nonsense—will have unintended, currently unpredictable, consequences.

Gibson notes that even in the minimal cell, the organism with the simplest known genome on the planet, biologists don’t know what one-third of the genes do. Moving from the simplest organism to humans is a leap into the unknown. “One design flaw can change how the cell behaves or even whether the cells are viable,” Gibson says. “We don’t have the design knowledge.”

**Learn by doing**

Many scientists believe this uncertainty about design is all the more reason to try writing human and other large genomes. “People are entranced with the perfect,” Harvard’s Church says. “But engineering and medicine are about the ‘pretty good.’ I learn much more by trying to make something than by observing it.”

Others aren’t sure that the move from writing the yeast genome to writing the human genome is necessary, or ethical. When the project to write the human genome was made public in May 2016, the founders called it Human Genome Project-write. They held the first organizational meeting behind closed doors, with no journalists present. A backlash ensued.
In the magazine *Cosmos*, Stanford University bioengineer Drew Endy and Northwestern University ethicist Laurie Zoloth in May 2016 warned of unintended consequences of large-scale changes to the genome and of alienating the public, potentially putting at risk funding for the synthetic biology field at large. They wrote that “the synthesis of less controversial and more immediately useful genomes along with greatly improved sub-genomic synthesis capacities ... should be pursued instead.”

GP-write members seem to have taken such criticisms to heart, or come to a similar conclusion on their own. By this May’s conference, “human” was dropped from the project’s name. Leaders emphasized that the human genome would be a subproject proceeding on a conservative timescale and that ethicists would be involved at every step along the way. “We want to separate the overarching goal of technology development from the hot-button issue of human genome writing,” Boeke explains.

Bringing the public on board with this kind of project can be difficult, says Alta Charo, a professor of law and bioethics at the University of Wisconsin, Madison, who is not involved with GP-write. Charo cochaired a National Academy of Sciences study on the ethics and governance of human gene editing, which was published in February.

She says the likelihood of positive outcomes, such as new therapies or advances in basic science, must be weighed against potential unintended consequences or unforeseen uses of genome writing. People see their basic values at stake in human genetic engineering. If scientists achieve their goals—making larger scale genetic engineering routine and more useful, and bringing it to the human genome—major changes are possible to what Charo calls “the fabric of our culture and society.” People will have to decide whether they feel optimistic about that or not. (Charo does.)

Given humans’ cautiousness, Charo imagines in early times we might have decided against creating fire, saying, “Let’s live without that; we don’t need to create this thing that might destroy us.” People often see genetic engineering in extreme terms, as a fire that might illuminate human biology and light the way to new technologies, or one that will destroy us.
Charo says the GP-write plan to keep ethicists involved going forward is the right approach and that it’s difficult to make an ethical or legal call on the project until its leaders put forward a road map.

The group will announce a specific road map sometime this year, but it doesn’t want to be restrictive ahead of time. You know when you’re done reading something, Boeke said at the meeting in May. But “writing has an artistic side to it,” he added. “You never know when you’re done.”

Katherine Bourzac is a freelance science writer based in San Francisco.

Chemical & Engineering News
ISSN 0009-2347
Copyright © 2017 American Chemical Society

**Comments**

**Rama Viswanathan  (July 10, 2017 3:37 PM)**
I am probably being naive, but why was the first total synthesis of a genome that of the poliovirus, presumably the same virus causing the disease that humans in a different century worked hard to eradicate by vaccination? There was no commentary in the article about this red flag.

» Reply

**Gordon Robinson  (July 13, 2017 4:08 PM)**
Nice review of plans and progress in a challenging and controversial set of projects. It does seem that starting with the mouse genome instead of human would yield many of the same insights and tools, and an easy test-bed for assessing genome-level changes w/o provoking a predictable ethical backlash; but perhaps the latter response was part of these creators’ plan to stimulate an open discussion?

» Reply