Programming biological operating systems: genome design, assembly and activation

Daniel G Gibson[1,2]

The DNA technologies developed over the past 20 years for reading and writing the genetic code converged when the first synthetic cell was created 4 years ago. An outcome of this work has been an extraordinary set of tools for synthesizing, assembling, engineering and transplanting whole bacterial genomes. Technical progress, options and applications for bacterial genome design, assembly and activation are discussed.

The creation of a synthetic cell in 2010 (ref. 1) presented a major paradigm shift in the field of genomics and demonstrated that it is now technically feasible to reverse the process of reading (sequencing) whole genome sequences and begin writing (synthesizing) them. These synthesized genetic instructions may then be activated to produce self-replicating biological cells with the expected functions and characteristics of that digital sequence information retrieved from a computer (**Fig. 1**). During the course of this work, several *in vitro* and *in vivo* DNA synthesis and assembly methods[2–9] have been developed and widely used in the field; however, whole-genome assembly[10] and activation[11,12] methods are not at the same level of development, and their utility has not been fully realized. Now that large DNA molecules can be constructed, the absence of well-defined genome design principles has become even more noticeable. Once these principles are better understood and defined, genomes will be rewritten and activated in a more predictable fashion. Here I discuss the state of the art of bacterial genome design, construction and activation.

REWRITING GENOMES

DNA, specifically the genome, is the software of life[13]. All of the characteristics and functions of living cells are written into the genetic code of DNA. Biologists now have the potential to act as software engineers and rewrite biological operating systems, starting from available genome sequences.

There are a variety of methods currently available for rewriting genomes. The extent to which researchers require *de novo* genome synthesis, which in turn depends on the availability of a natural DNA template, should help determine which method is used. If the desired sequence does not exist in nature, or is not readily available to them, then a synthetic approach must be taken. However, if the genome template DNA is available and does not require significant manipulation, then a complete *de novo* synthesis approach may not currently be the most efficient or cost-effective process for rewriting a genome. The constitutive parts processed by the assembly methods to produce genomes may be derived from natural and/or synthetic DNA. Several scenarios and methods for genome assembly are described below and illustrated in **Figure 2**.

Genome assembly

In 2008, the complete chemical synthesis of a bacterial genome was described for the first time[6]. It was the 583-kb *Mycoplasma genitalium* genome and represented a DNA structure that was nearly 20 times larger than any previously synthesized DNA fragment. In 2010, the 1.1-Mb *Mycoplasma mycoides* genome was completely synthesized and activated in a recipient host to produce a synthetic cell[1]. These two genomes are the largest chemically defined structures ever synthesized in a laboratory. To accomplish this goal, we (my colleagues and I at the Venter Institute) had to develop a variety of *in vitro* and *in vivo* methodologies that lead from oligonucleotides up to whole chromosomes[3,4,6–10].

***In vitro* genome assembly.** Several *in vitro* enzymatic reactions capable of assembling genome-size molecules from multiple overlapping DNA fragments have been described[3,6,8,9]. The simplest of these recombination methods can be carried out as a single isothermal reaction. The assembly reaction

[1]J. Craig Venter Institute, La Jolla, California, USA. [2]Synthetic Genomics, Inc., La Jolla, California, USA. Correspondence should be addressed to D.G.G. (dgibson@jcvi.org).

RECEIVED 11 DECEMBER 2013; ACCEPTED 22 FEBRUARY 2014; PUBLISHED ONLINE 29 APRIL 2014; DOI:10.1038/NMETH.2894

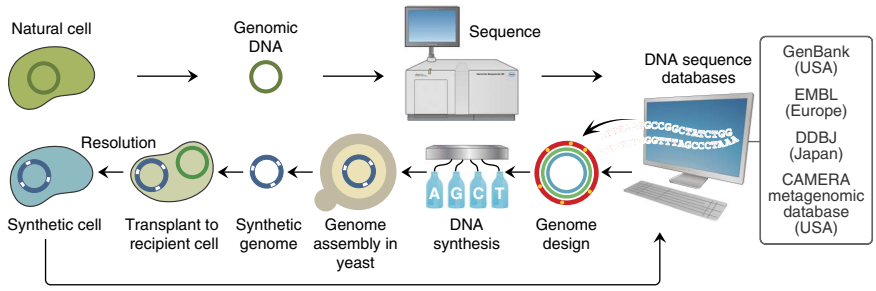

Figure 1 | Moving life into the digital world and back.

mixture in this system (commonly referred to as Gibson Assembly) contains a 5′ exonuclease, a DNA polymerase and a DNA ligase working in harmony to accomplish the seamless joining of DNA fragments without any intervening, undesired sequence during a brief 50 °C reaction. Gibson Assembly does not rely on thermocycling to bring overlapping parts together, and the assembled subfragments are covalently joined with DNA ligase; both of these features reduce the likelihood of fragmentation of large DNA molecules. In addition, ligated DNA assembly products are more efficiently electroporated into *Escherichia coli* than their unligated counterparts[9]. Other *in vitro* methods that are both seamless and sequence independent include circular polymerase extension cloning, In-Fusion (Clontech), USER (Uracil-Specific Excision Reagent; New England BioLabs) and sequence- and ligation-independent cloning (reviewed in refs. 14,15). Although these methods have been useful in the assembly of DNA up to 20 kb—the size of small genetic pathways—they have not been used in the construction of larger DNA molecules such as whole genomes.

In vitro DNA assembly methods offer greater speed and ease of use than *in vivo* approaches and are therefore generally preferred. However, the *in vitro* assembly reactions do not typically produce enough material to be useful and so must be amplified before they can be used in a subsequent round of DNA assembly or in an application. A major advantage of the Gibson Assembly method is that it produces covalently ligated constructs, and so assembled gene- and pathway-sized fragments can be readily amplified by PCR and rolling-circle amplification methods[9]. This is in contrast to cloning them into a host organism and relying on the host's DNA replication machinery to produce more of the assembled construct. In addition to speed and simplicity, the *in vitro* amplification approaches also have the advantage of not being subjected to unpredictable toxicity or stability issues within a host cell. However, it can be expected that PCR and rolling-circle amplification will produce mutations at a much higher frequency than cloning in a host. In addition, currently they cannot readily amplify genome-size fragments and will produce linear DNA molecules, which may not be useful if the goal is to produce a circular genome.

***In vivo* genome assembly.** Fully assembled bacterial genomes can be grown and engineered inside a host organism of a different species. In addition, those genomes can be assembled from overlapping parts *in vivo* using the host's natural homologous recombination machinery. Groundbreaking work by Itaya and colleagues[16] demonstrated the iterative assembly of very large nonsynthetic constructs into the *Bacillus subtilis* genome, in which they cloned almost all of the *Synechocystis* PCC6803 genome as a set of four separate 800- to 900-kb DNA segments. Also using an iterative process, Holt and colleagues[17] described the assembly of a partial *Haemophilus influenzae* genome in *E. coli* using λ Red recombination. These approaches used sequential stepwise addition of segments to reconstruct partial genomes within a recipient host strain.

The extremely efficient natural homologous recombination machinery in the yeast *Saccharomyces cerevisiae* has been leveraged for the assembly of DNA molecules extending from oligonucleotides up to complete genomes[1,2,4–7,10,18] (**Fig. 2**). Yeast was also the final destination for the two bacterial genomes that we have synthesized and has an advantage over other *in vivo* systems in that it can take up and assemble numerous overlapping DNA segments at one time. Notably, *S. cerevisiae* was used to take up and assemble 25 large DNA fragments, each ~24 kb, in a single transformation event[7]. The availability of so many yeast genetic tools also makes it an ideal host for making further modifications to the bacterial genomes once these genomes have been cloned.

Moving whole genomes into yeast

Yeast has also been developed as a host for receiving and propagating complete nonsynthetic genomes[10,19–21]. In this scenario, a yeast vector is first introduced into a bacterial cell's genome either by random transposition or by directed homologous recombination. Then the bacterial genomes, now carrying the yeast vector, are moved into yeast (**Fig. 2**). This had typically been performed using carefully purified genomes, which were immobilized and isolated in agarose to keep them intact. However, recently we demonstrated that whole genomes could be moved from bacterial cells to yeast spheroplasts without any DNA purification process (**Fig. 2**). This cell-fusion approach reduced the time required to move bacterial genomes into yeast and increased the efficiency and probability of a successful genome transfer event. The process could be made even more efficient when a restriction-free donor strain was used[21]. Restriction systems have nuclease activity and cleave DNA at specific locations. If a donor strain contains restriction systems, nucleases may enter yeast cells upon fusion and digest the yeast chromosomal DNA, resulting

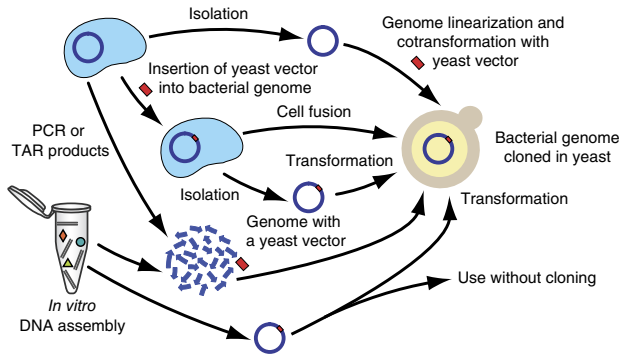

Figure 2 | Options for genome assembly. TAR, transformation-associated recombination.

in lethality to yeast. A cell-fusion approach may be especially suitable for larger bacterial genomes, which are more susceptible to *in vitro* mechanical shearing[22].

The advantages of transposing a yeast vector into a naturally growing bacterial cell are that (i) no DNA sequence information is required and (ii) only functional genomes are recovered. However, a transformation system for the organism needs to be established. Alternatively, if the genome sequence is known, two additional approaches can be considered. In the first approach, the genomes can be purified from bacterial cells and linearized with a single-cutter restriction enzyme, if available. Yeast cells can then be cotransformed with the linearized genome and a yeast vector containing overlaps to the ends of the linearized genome, and the genomes are recombined *in vivo* (**Fig. 2**). Genomes can also be reconstructed in yeast by *in vivo* assembly of overlapping PCR products, by transformation-associated recombination (TAR) products or by a combination of both[23] (**Fig. 2**). In the TAR approach, genomic DNA is broken by mechanical means or through restriction digestion and then cotransformed into yeast with vector ends specifically 'hooking' the desired fragment[24]. A series of vector ends can be prepared to make a panel of overlapping genomic DNA fragments, which are cloned in yeast (and moved to *E. coli* for increased production when possible) and then assembled into larger and larger pieces until the complete genome is assembled in yeast. The TAR approach's advantages over PCR are a reduced likelihood of errors and the production of genome segments much larger than what can be achieved by PCR. The PCR and TAR approaches are particularly useful when a genome is being reconstructed from natural and synthetic parts, as the junctions between those parts can be readily broken at numerous defined locations. We used a combination of natural and synthetic pieces to troubleshoot a nonfunctional synthetic *M. mycoides* genome[1].

The capacity for cloning whole genomes in yeast. A variety of bacterial genomes have been cloned in yeast: these include *M. genitalium*[6,7,10], *M. mycoides*[1,10], *Mycoplasma pneumonia*[10], *Acholeplasma laidlawii*[20], *Prochlorococcus marinus*[19] and *H. influenzae*[21], ranging in size from 0.6 to 1.8 Mb, with GC contents of 24–40%. At least three factors contribute to a successful cloning event in yeast, as follows:

1. Absence of toxic gene expression. Toxic gene expression could weaken or kill the yeast host. However, because transcription[25] and translation[26] signals in eukaryotes are different from those in prokaryotes, expression of bacterial genes should be substantially reduced in yeast. In addition, mycoplasmas use the codon UGA for tryptophan rather than as a translation stop signal, which also largely reduces the risk of toxic gene product expression in yeast, as those gene products would be truncated. As an example of toxic gene expression, the complete genome of *A. laidlawii*, which uses the standard genetic code, could be cloned in yeast only after a single gene, encoding an extracellular endonuclease, was inactivated[20].

2. Genome size. The 1.8-Mb *H. influenza* genome is the largest bacterial genome to be cloned in yeast to date[21]. An upper size limit for cloning DNA constructs in yeast has yet to be determined; however, a 2.3-Mb yeast artificial chromosome has been reported[27]. Advances such as cell fusion[21] should permit cloning of large bacterial genomes, as they no longer need to be subjected to shear forces during naked DNA transfer. However, larger genomes may have an increased probability for toxic gene product expression, and replicating the large DNA structure may put more stress on the yeast cells.

3. GC content. Yeast origins of replication have reduced GC content, and these origins should be present for the initiation of DNA replication approximately every 150 kb to stably maintain a chromosome[28]. Bacterial genomes with GC content up to 40% have been cloned. As an alternative to cloning a completely contiguous 2.7-Mb *Synechococcus elongatus* PCC 7942 genome, which has a GC content of 55%, the entire genome could readily be cloned in ~100-kb sections using the PCR and TAR approaches discussed above (**Fig. 2**). Sections greater than 200 kb could be stably maintained in yeast if yeast origins were inserted[23].

At least five factors should be considered when researchers are deciding to assemble and clone an entire bacterial genome in yeast. First, can the ultimate design be achieved faster by manipulating the genome in its native environment? This will depend on the genetic tractability of that bacterial strain with respect to transformability and homologous recombination, the number of transformation and recombination iterations that are required to produce the ultimate genome design and the growth rate of the bacterial species. Second, is the native bacterial strain available? If not, then the only option is to synthesize the strain, as mentioned above. Third, is the bacterial strain a pathogen? If yes, then it may be preferable to keep that bacterial strain out of the laboratory and synthesize it without virulence factors. Once activated, it may have the potential to be used as an attenuated live vaccine. Fourth, what is the likelihood that the genome can be cloned in yeast, given its size and complexity? The guidelines above should help in that decision process. Finally, what is the likelihood that the genome can be activated to produce the engineered bacteria? It is fruitless to rewrite a genome if it cannot be activated to produce a new cell expressing those designed instructions.

GENOME ACTIVATION

The genomes of most life forms may be modified in their native environment, albeit some more efficiently than others. Church and colleagues[29–31] are using MAGE and CAGE (multiplex and conjugative automated genome engineering, respectively) methods to edit and evolve a natural *E. coli* genome template *in vivo*. MAGE is mediated by the bacteriophage λ Red ssDNA-binding protein β, which directs transformed ssDNA oligonucleotides to genetic targets where they are modified[32]. CAGE permits a hierarchical scheme to combine *E. coli* genome fragments that have individually been modified by MAGE.

Boeke and colleagues[33] are in the process of replacing the entire *S. cerevisiae* natural genome with synthetically designed DNA, one section at a time. The goal of this effort is to understand the fundamental rules governing eukaryotic genome structure and function. One of their design features includes *loxP* sites flanking each non-essential gene to enable SCRaMbLE (synthetic chromosome rearrangement and modification by *loxP*-mediated evolution) upon induction of Cre recombinase.

Biological cells, including those of higher eukaryotes such as mammals, may now be precisely edited by zinc-finger nucleases; transcription activator–like effector nucleases; and clustered, regularly interspaced, short palindromic repeats (CRISPR)-Cas9 systems (reviewed in ref. 34). These are breakthrough technologies

that will pave the way for correction of genes associated with a variety of diseases.

The approaches above are in contrast to genome activation, in which a whole genome is expressed to produce a self-replicating cell with properties defined by that genome. The term 'genome activation' includes genomes that are expressed inside a host cell, or those within a cell-free system that convert nonliving cellular machinery into a free-living cell, and can grow and divide when nutrients are provided. The challenge is to find a compatible system to activate wholly engineered genomes. Itaya and colleagues[16] cloned the *Synechocystis* genome PCC 6803 in *B. subtilis*; however, most of the *Synechocystis* genes on the chimeric genome were silent, and the cells did not take on the phenotype of the donor genome. This was presumably due to incompatibility of signals required for the initiation of transcription and translation between the two species. An operating system is an essential component for 'booting up' and working a computer. Likewise, the genome is the operating system of a cell, and the cellular milieu contains the hardware that activates and runs the genome software. The cytoplasm of a cell contains all of the parts necessary to express the information in the genome, and the genome contains all the information necessary to produce and replicate the cell. Each is ineffective without the other.

Genome transplantation

In eukaryotes, nuclear transplantation[35,36] is a well-established procedure in which the nucleus of a donor cell is introduced into an enucleated target cell. Similarly, one bacterial species can be converted into another through a process termed genome transplantation[11]. This was first demonstrated with *M. mycoides* donor genomes and *Mycoplasma capricolum* recipient cells[11]. In this approach, purified and intact chromosomes are transferred to recipient cells and grown under conditions that select for cells carrying the donor genome. Cell growth and division leads to daughter cells with either the donor genome or the recipient genome. Because the donor genome is engineered to contain an antibiotic resistance gene, only cells that contain the donor genome and express the resistance gene will grow in the presence of the antibiotic.

In further developing genome transplantation technology, we have identified several factors that should be addressed to increase the likelihood of a successful genome transplantation event.

Nuclease activity. If the recipient cell contains extracellular nucleases or restriction modification systems, these could destroy the incoming genome before it may be replicated and expressed to direct production of new cells according to that genome sequence. In the case of *M. mycoides* chromosomes isolated from their native environment and transferred to *M. capricolum* recipient cells, the restriction system present in the host did not digest the incoming donor genome because *M. mycoides* has the same restriction system and was therefore protected through methylation. However, once the donor genome was introduced into yeast, its DNA methylase was not expressed; consequently, restriction sites were no longer methylated and thus no longer protected from the recipient cell's restriction system. This restriction barrier was circumvented either by methylating the donor genome before transplantation or by constructing a restriction system–deficient recipient cell[12].

Presence of a cell wall. A candidate recipient cell may contain a cell wall; this may need to be removed or weakened before the cell receives donor genomes and then be regenerated afterward. The mycoplasmas, which have been successfully transplanted, do not contain a cell wall.

Genome size. Larger genomes are more susceptible to breakage than smaller ones. Even so, multimegabase chromosomes can be purified intact once they have been embedded in agarose. Polycations have been demonstrated to condense and protect large DNA molecules from shearing[27] and should be considered when handling large, purified DNA. As with moving genomes from bacteria into yeast, a cell-fusion approach could also be developed to eliminate the donor genome manipulation steps.

Similarity of donor and recipient species. A recipient cell must have the cellular machinery necessary to express and replicate the donor genome. For example, the recipient cell's replication, transcription and translation machinery would need to be compatible with replication, transcription and translation signals (for example, origin of replication, promoters, terminators and ribosome-binding sites) located on the donor genome. The donor and recipient pair should also use the same genetic code, and all the necessary tRNAs should be sufficiently available. It would be preferable that a recipient cell is deficient in homologous recombination to reduce the likelihood of recombination between the donor and recipient genomes; in this case, the donor and recipient could be from the same species.

With the knowledge gained through such work, it should be possible to design and construct a universal recipient cell that can be converted into any species designated by the transplanted donor genome. The universal recipient could be designed to be compatible with a wide variety of incoming donor genomes, deficient in nuclease and homologous recombination activities, and optimized for taking up genomic DNA. The donor genome may need to include 'self-activating' genes that are optimized for expression in the recipient cellular milieu. These genes would likely be involved in production of key components involved in transcription (such as RNA polymerase and σ factors) and translation (such as ribosomes and aminoacyl tRNA synthetases), and, once expressed, they would kick-start transcription and translation of all other genes in the donor genome. Alternatively, once genome design rules become more defined (see below), a more practical solution may be to modify the donor genome sequence such that it contains the necessary expression signals (such as promoter, terminator and ribosome binding-site sequences) to be reliably activated by the transcription and translation machinery inside the recipient cell. A major advantage of a universal recipient cell is that a standardized protocol could be developed for this process, making it more generally useful than method development on a species-per-species basis.

Cell-free genome activation

All properties of a transplanted cell are governed by the donor genome, and any residual phenotypic properties of the recipient cell are diluted out as the cells grow and divide. Using the computer analogy above, the DNA software builds its own hardware. Thus, it is expected that a cell built from only chemically synthesized cellular machinery will have identical properties to a

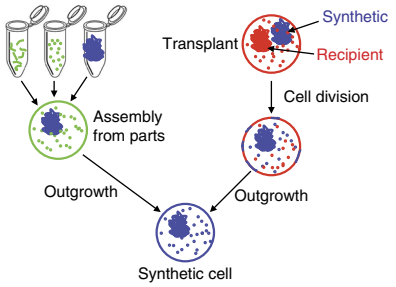

Figure 3 | A synthetic genome activated in a cell-free environment will give the same outcome as a genome activated through genome transplantation. Reproduced with permission from C. Hutchison.

cell that was transplanted with a chemically synthesized genome (**Fig. 3**). Nevertheless, the idea of creating life in a cell-free environment is intriguing, and an extraordinary amount of biological knowledge would be gained if a cell could be built from nonliving parts. Cellular parts could then be synthesized and directly tested. Church, Forster and Jewett have reviewed the *in vitro* synthesis of a minimal cell from chemically synthesized parts[37–39]. Understanding the mechanism that governs genome activation following transplantation, together with universal recipient cell development, should aid in the development of cell-free genome activation.

GENOME DESIGN

The ultimate vision in synthetic biology is to have the capacity to design and build DNA that produces a biological cell with a predictable outcome (**Fig. 4**). Significant advances have been made in DNA design at the gene and pathway level (reviewed in ref. 40) and in engineering bacteriophage genomes[41,42]. But, even with all the advances in genomics and synthetic biology, there is still not a single self-replicating cell in which the function of every one of its genes is understood.

The boundaries continue to be tested for building a minimal cell—that is, a cell that has only the machinery necessary for independent life. This is an effort that began in 1995 following sequencing of the genome from *M. genitalium*[43], a species that has a genome that is naturally close to a minimal cell. We were not, however, able to activate the *M. genitalium* genome, maybe

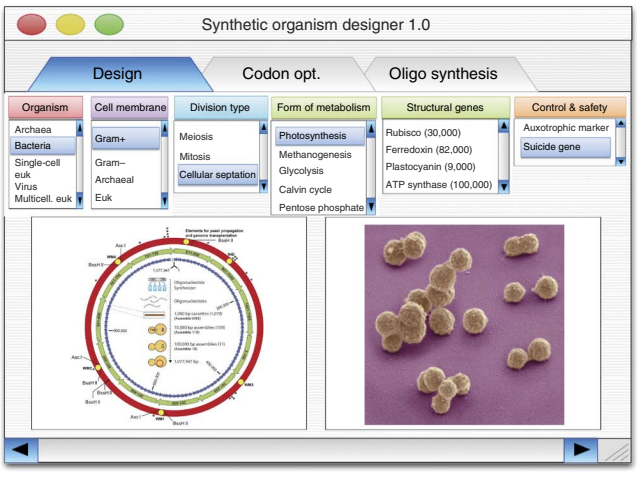

Figure 4 | A futuristic vision for designing synthetic organisms on demand.

partly owing to its extremely slow growth rate; but now, after 15 years, we have the genome synthesis and activation tools to build and test any genome design starting from the sequence of the *M. mycoides* genome. We now use *M. mycoides* and not *M. genitalium* as the experimental model system to produce a minimal cell. The design-build-test cycle has never been easier than now at the whole-genome level with this system. Any new genome sequence can be designed and readily tested for functionality by assaying for viable cells following genome transplantation. Scientists at the J. Craig Venter Institute and Synthetic Genomics, Inc. are whittling away at the *M. mycoides* genome and testing hundreds of new genome combinations to produce a minimal cell. The ultimate goal is to synthesize a cell in which every single gene's role is understood such that the minimal cell can be computationally modeled. This would significantly help scientists understand how cells work and how to design them from scratch. Recently, Covert's team[44] exceptionally demonstrated computational modeling of natural *M. genitalium*.

Once minimized, cellular life becomes easier to understand and predict. Minimal cellular systems will become test beds for understanding the rules governing genome design, including the capacity to predictably (i) group genes with similar functions and produce genetic modules that can be swapped, (ii) change the genetic code, (iii) modify the genome such that it is easier to chemically synthesize and assemble in a host organism, (iv) modify the genome such that it can be activated in a recipient cell or in a cell-free environment, (v) change regulatory elements to alter expression profiles and (vi) add genes or genetic pathways to the minimal chassis that function once activated. This last point is intriguing, as the minimal cell could be used as a launching pad for learning to make more complex organisms. For example, can one add a 'cell-wall genetic module' to produce a minimal cell now containing a cell wall? The capacity to build and test any new design is extremely powerful. It will establish genome design rules that will help researchers better understand biology and produce synthetic cells that carry out their programmed functions.

CONCLUDING REMARKS

Scientists have been mastering the techniques to read the genetic code and now are taking the next step to write the genetic code as synthetic biologists and to produce organisms with new and improved biological functions that do not already exist in nature. Although synthetic biologists are still learning how to predictably design DNA, particularly large genetic pathways or genomes, the technology is available to build them. A very useful set of DNA synthesis and assembly tools emerged during the quest to build a synthetic cell. These methods have been applied to the construction of synthetic influenza vaccine viruses for rapid response to pandemics[45], genetic pathways[46,47], organellar genomes[8,48], viral genomes[41,42,49] and bacterial genomes[1,6].

The synthetic genomics vision is to take synthetic biology to an entirely new level where complete genomes are designed, written and activated, in a fully automated pipeline, to create new biological operating systems of cells that predictably carry out their programmed instructions. To accomplish this goal, improved tools for genome design, synthesis and assembly, and activation must be developed. DNA synthesis costs must fall to below pennies per base pair, as has been seen with the dramatic drop in DNA sequencing costs[29]. This will ultimately be enabled

by new automated DNA technologies that can reliably convert low-cost oligonucleotides (for example, those produced on a microchip in small quantities) into accurate synthetic DNA fragments. Thousands of genome combinations can then be built and tested at an affordable price. Once the synthetic genomics vision is realized, cells with extraordinary properties will be produced. The applications are endless with these new technologies, and there is now great potential for transforming carbon dioxide, plant biomass and coal into high-value products such as medicines, foods, biofuels, plastics or chemicals, and for creating new vaccines.

ACKNOWLEDGMENTS

I would like to thank past and present members of the synthetic biology groups at the J. Craig Venter Institute and Synthetic Genomics, Inc., for making all of these technologies I discuss here possible, especially C. Venter, H. Smith and C. Hutchison, who had the vision to make a synthetic cell as early as 1995. I also thank M. LaPointe, T. Richardson, J. Ward and J. Eads for their contributions to **Figures 1**, **2** and **4** and C. Hutchison for producing **Figure 3**.

COMPETING FINANCIAL INTERESTS

The author declares competing financial interests: details are available in the online version of the paper.

1. Gibson, D.G. *et al.* Creation of a bacterial cell controlled by a chemically synthesized genome. *Science* **329**, 52–56 (2010).
2. Gibson, D.G. Oligonucleotide assembly in yeast to produce synthetic DNA fragments. *Methods Mol. Biol.* **852**, 11–21 (2012).
3. Gibson, D.G. Enzymatic assembly of overlapping DNA fragments. *Methods Enzymol.* **498**, 349–361 (2011).
4. Gibson, D.G. Gene and genome construction in yeast. *Curr. Protoc. Mol. Biol.* **94**, 3.22 (2011).
5. Gibson, D.G. Synthesis of DNA fragments in yeast by one-step assembly of overlapping oligonucleotides. *Nucleic Acids Res.* **37**, 6984–6990 (2009).
6. Gibson, D.G. *et al.* Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. *Science* **319**, 1215–1220 (2008).
7. Gibson, D.G. *et al.* One-step assembly in yeast of 25 overlapping DNA fragments to form a complete synthetic *Mycoplasma genitalium* genome. *Proc. Natl. Acad. Sci. USA* **105**, 20404–20409 (2008).
8. Gibson, D.G., Smith, H.O., Hutchison, C.A. III., Venter, J.C. & Merryman, C. Chemical synthesis of the mouse mitochondrial genome. *Nat. Methods* **7**, 901–903 (2010).
9. Gibson, D.G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).
10. Benders, G.A. *et al.* Cloning whole bacterial genomes in yeast. *Nucleic Acids Res.* **38**, 2558–2569 (2010).
11. Lartigue, C. *et al.* Genome transplantation in bacteria: changing one species to another. *Science* **317**, 632–638 (2007).
12. Lartigue, C. *et al.* Creating bacterial strains from genomes that have been cloned and engineered in yeast. *Science* **325**, 1693–1696 (2009).
13. Venter, J.C. *Life at the Speed of Light: From the Double Helix to the Dawn of Digital Life* (Viking, 2013).
14. Ma, S., Tang, N. & Tian, J. DNA synthesis, assembly and applications in synthetic biology. *Curr. Opin. Chem. Biol.* **16**, 260–267 (2012).
15. Ellis, T., Adie, T. & Baldwin, G.S. DNA assembly for synthetic biology: from parts to pathways and beyond. *Integr. Biol. (Camb.)* **3**, 109–118 (2011).
16. Itaya, M., Tsuge, K., Koizumi, M. & Fujita, K. Combining two genomes in one cell: stable cloning of the *Synechocystis* PCC6803 genome in the *Bacillus subtilis* 168 genome. *Proc. Natl. Acad. Sci. USA* **102**, 15971–15976 (2005).
17. Smailus, D.E., Warren, R.L. & Holt, R.A. Constructing large DNA segments by iterative clone recombination. *Syst. Synth. Biol.* **1**, 139–144 (2007).
18. Ma, H., Kunes, S., Schatz, P.J. & Botstein, D. Plasmid construction by homologous recombination in yeast. *Gene* **58**, 201–216 (1987).
19. Tagwerker, C. *et al.* Sequence analysis of a complete 1.66 Mb *Prochlorococcus marinus* MED4 genome cloned in yeast. *Nucleic Acids Res.* **40**, 10375–10383 (2012).
20. Karas, B.J., Tagwerker, C., Yonemoto, I.T., Hutchison, C.A. III. & Smith, H.O. Cloning the *Acholeplasma laidlawii* PG-8A genome in *Saccharomyces cerevisiae* as a yeast centromeric plasmid. *ACS Synth. Biol.* **1**, 22–28 (2012).
21. Karas, B.J. *et al.* Direct transfer of whole genomes from bacteria to yeast. *Nat. Methods* **10**, 410–412 (2013).
22. Schwartz, D.C. & Cantor, C.R. Separation of yeast chromosome-sized DNAs by pulsed field gradient gel electrophoresis. *Cell* **37**, 67–75 (1984).
23. Noskov, V.N. *et al.* Assembly of large, high G+C bacterial DNA fragments in yeast. *ACS Synth. Biol.* **1**, 267–273 (2012).
24. Kouprina, N. & Larionov, V. Selective isolation of genomic loci from complex genomes by transformation-associated recombination cloning in the yeast *Saccharomyces cerevisiae*. *Nat. Protoc.* **3**, 371–377 (2008).
25. Kornberg, R.D. Eukaryotic transcriptional control. *Trends Cell Biol.* **9**, M46–M49 (1999).
26. Kozak, M. Initiation of translation in prokaryotes and eukaryotes. *Gene* **234**, 187–208 (1999).
27. Marschall, P., Malik, N. & Larin, Z. Transfer of YACs up to 2.3 Mb intact into human cells with polyethylenimine. *Gene Ther.* **6**, 1634–1637 (1999).
28. van Brabant, A.J., Buchanan, C.D., Charboneau, E., Fangman, W.L. & Brewer, B.J. An origin-deficient yeast artificial chromosome triggers a cell cycle checkpoint. *Mol. Cell* **7**, 705–713 (2001).
29. Carr, P.A. & Church, G.M. Genome engineering. *Nat. Biotechnol.* **27**, 1151–1162 (2009).
30. Wang, H.H. *et al.* Programming cells by multiplex genome engineering and accelerated evolution. *Nature* **460**, 894–898 (2009).
31. Isaacs, F.J. *et al.* Precise manipulation of chromosomes *in vivo* enables genome-wide codon replacement. *Science* **333**, 348–353 (2011).
32. Ellis, H.M., Yu, D., DiTizio, T. & Court, D.L. High efficiency mutagenesis, repair, and engineering of chromosomal DNA using single-stranded oligonucleotides. *Proc. Natl. Acad. Sci. USA* **98**, 6742–6746 (2001).
33. Dymond, J.S. *et al.* Synthetic chromosome arms function in yeast and generate phenotypic diversity by design. *Nature* **477**, 471–476 (2011).
34. Gaj, T., Gersbach, C.A. & Barbas, C.F. III. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol.* **31**, 397–405 (2013).
35. Gurdon, J.B. & Byrne, J.A. The first half-century of nuclear transplantation. *Proc. Natl. Acad. Sci. USA* **100**, 8048–8052 (2003).
36. Gurdon, J.B. & Wilmut, I. Nuclear transfer to eggs and oocytes. *Cold Spring Harb. Perspect. Biol.* **3**, a002659 (2011).
37. Forster, A.C. & Church, G.M. Towards synthesis of a minimal cell. *Mol. Syst. Biol.* **2**, 45 (2006).
38. Forster, A.C. & Church, G.M. Synthetic biology projects *in vitro*. *Genome Res.* **17**, 1–6 (2007).
39. Jewett, M.C. & Forster, A.C. Update on designing and building minimal cells. *Curr. Opin. Biotechnol.* **21**, 697–703 (2010).
40. Medema, M.H., van Raaphorst, R., Takano, E. & Breitling, R. Computational tools for the synthetic design of biochemical pathways. *Nat. Rev. Microbiol.* **10**, 191–202 (2012).
41. Chan, L.Y., Kosuri, S. & Endy, D. Refactoring bacteriophage T7. *Mol. Syst. Biol.* **1**, 2005.0018 (2005).
42. Jaschke, P.R., Lieberman, E.K., Rodriguez, J., Sierra, A. & Endy, D. A fully decompressed synthetic bacteriophage oX174 genome assembled and archived in yeast. *Virology* **434**, 278–284 (2012).
43. Fraser, C.M. *et al.* The minimal gene complement of *Mycoplasma genitalium*. *Science* **270**, 397–403 (1995).
44. Karr, J.R. *et al.* A whole-cell computational model predicts phenotype from genotype. *Cell* **150**, 389–401 (2012).
45. Dormitzer, P.R. *et al.* Synthetic generation of influenza vaccine viruses for rapid response to pandemics. *Sci. Transl. Med.* **5**, 185ra68 (2013).
46. Guye, P., Li, Y., Wroblewska, L., Duportet, X. & Weiss, R. Rapid, modular and reliable construction of complex mammalian gene circuits. *Nucleic Acids Res.* **41**, e156 (2013).
47. Temme, K., Zhao, D. & Voigt, C.A. Refactoring the nitrogen fixation gene cluster from *Klebsiella oxytoca*. *Proc. Natl. Acad. Sci. USA* **109**, 7085–7090 (2012).
48. O'Neill, B.M. *et al.* An exogenous chloroplast genome for complex sequence manipulation in algae. *Nucleic Acids Res.* **40**, 2782–2792 (2012).
49. Smith, H.O., Hutchison, C.A. III., Pfannkoch, C. & Venter, J.C. Generating a synthetic genome by whole genome assembly: phiX174 bacteriophage from synthetic oligonucleotides. *Proc. Natl. Acad. Sci. USA* **100**, 15440–15445 (2003).