The changing economics of DNA synthesis

Robert Carlson

How are the economics of synthetic biology likely to develop in the coming years?

Biological technologies come in many different guises. For millennia, humans have used selection and breeding to direct the evolution of organisms in a sort of top-down approach, a powerful but unpredictable means to achieve a desired behavior. At the opposite extreme, genes and genomes can now be written from chemical precursors, a more precise but sometimes less effective means of producing a particular biological behavior—the design rules for bottom-up engineering of biology in the vast majority of cases are still poorly understood. In between, practicing metabolic engineers use any and all tools at hand to herd and cajole organisms into producing products with market value in the many hundreds of billions of dollars.

At the core of all these approaches to biological engineering is the creation of a particular genomic sequence that produces behaviors according to human desire or need. In addition to nearly a century of evolution and selection based on early knowledge of genetics, we are already four decades into the direct manipulation of genomes through recombinant DNA technology. Synthetic oligonucleotides (oligos) have been available by mail order for the past 20 years, and synthetic genes have been built commercially from those oligos for the last ten. In that time, the number of bases a single individual can synthesize in a day using commercial instruments has increased by five orders of magnitude, whereas the per base cost of synthetic genes has dropped by nearly three orders of magnitude (**Fig. 1**).

I argue here that in the coming years, synthetic DNA manufacturers will come under increasing pressure to reduce costs and decrease turnaround times. At the same time as these DNA synthesis companies face the commoditization of their product, the commercial sector that focuses on assembling and engineering genomes (or at least chromosomes) to create novel products for use in the medical, energy or industrial sectors is likely to become increasingly profitable. Demand for synthetic DNA will consequently spread around the globe as organizations of all sizes exploit biological technologies for many different aims. That some of those aims may be less appealing than others is already prompting calls to regulate synthesis in one way or another. But the global proliferation of demand is likely to limit the effectiveness of regulations implemented on the grounds of improving safety and security.

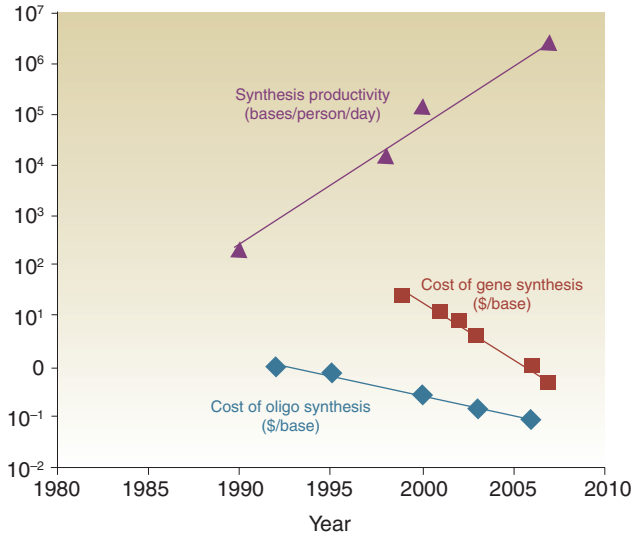

Figure 1 Productivity of oligo synthesis and cost of oligos and genes.

A nascent field

The commercial availability of synthetic DNA has clearly found a use in constructing ever longer genes and now genomes (**Fig. 2**). Recent work at the J. Craig Venter Institute (JCVI; Rockville, MD, USA) has resulted in relatively painless assembly of a 580-kb microbial genome from 101 pieces of starting material each 5–6 kb in length[1]. That this technique makes use of native recombination mechanisms in yeast suggests that it can be implemented in just about any laboratory that takes the time to learn the recipe. Combined with an existing widespread industry that regularly supplies synthetic DNA fragments of 5–10 kb in length, assembly in yeast will put the ability to build a wide range of DNA genomes in the hands of scientists, entrepreneurs and other interested parties worldwide.

Single-step DNA assembly in yeast should prove useful as a tool to rapidly assemble metabolic pathways from many short DNA sequences. Shao and Zhao[2] demonstrated precisely this sort of application in early 2009 by assembling functional metabolic pathways

Robert Carlson is at Biodesic, Seattle, Washington, USA.
e-mail: rob@biodesic.com

from three, five and eight separated genes in a single-step process nearly identical to that implemented at the JCVI. If used with libraries of gene variants, where the sequence of every gene in a synthetic pathway might be varied at the same time, assembly in yeast could be used to rapidly and simultaneously test thousands of different mutations in different genes. This should provide a powerful tool to supplement existing metabolic engineering techniques and potentially to speed identification of useful gene and pathway variants.

New design and error correction technologies will continue reducing the cost of synthetic genes, pathways and genomes. Assembling long stretches of DNA is presently accomplished by annealing overlapping oligos, followed by ligation and PCR amplification and then an error removal step. The process typically results in a population of sequences, many of which contain at least one error consisting of base substitutions, deletions or insertions, due largely to mistakes in the source oligos and annealing errors. Proofreading of multiple individual assembly products is usually required to identify a single molecule of the desired sequence. The many kinds of assembly steps, and the many kinds of errors present in the pool before sequencing, represent a variety of opportunities to introduce improvements in gene assembly.

For instance, Integrated DNA Technologies (Coralville, IA, USA) is regularly producing and selling high-quality oligos >200 bp in length, the use of which in assembling genes should result in much lower error rates (John Havens, personal communication). At the Synthetic Biology 4.0 meeting last year in Hong Kong, Alex Borokov described a design strategy that enables using lower quality, less expensive oligonucleotides to assemble genes in one step[3]. And whereas most gene assembly techniques rely on the removal of mistakes from a sequence pool, Novici Biotech (Vacaville, CA, USA) has just released the ErrASE system, which enables true error correction of genes assembled even from unpurified oligos (Hal Padgett, personal communication; readers should note Novici is a client of my company, Biodesic). Dan Gibson has recently demonstrated that yeast can even assemble overlapping oligos *in vivo* into sequences at least one kilobase long[4]. These many examples of new technologies suggest that continued innovation will reduce the cost of gene and genome assembly, particularly as proofreading by means of DNA sequencing is a substantial fraction of the overall cost.

Practical limits of current technology

Yet it already appears that the technical ability to build large genetic circuits and genomes outstrips our ability to understand and design systems of that size. Whereas the group at the JCVI has demonstrated how to assemble nearly a megabase of DNA—a small genome's worth of genes—UC Berkeley's Jay Keasling and his colleagues[5] are working at the cutting edge of metabolic engineering while manipulating a network of just 12 genes. The cost of the latter project, based on the size of the grant that funded it, is several tens of millions of dollars to pay for infrastructure and labor. Keasling is very upfront about the extent of true design versus tinkering that the team members have been able to employ, and in his talks refers to the not inconsequential number of serendipitous "miracles" that have kept the project on schedule to commercially produce precursors to the malaria drug artemisinin next year[5].

Consequently, the 12 genes manipulated in the artemisinic acid pathway may be a practical upper limit on engineering capabilities in the near term. In the slightly longer term, the more interesting numbers for synthesis may, therefore, be only 10–50 genes and 10,000–50,000 bases. Genetic circuits of this size may represent a limit in complexity for systems with economic value for many years to come. More important questions for would-be genome engineers are, How will the costs fall for constructs of this size? When will DNA of that length be available in days or hours instead of weeks? How soon before one can buy or build a desktop box that prints synthetic DNA of this length?

Improvements in DNA synthesis and gene assembly technologies will be driven by demand. Academic researchers, particularly those attempting to model, build and test complex networks of genes, are likely to consume as much synthetic DNA as their budgets allow. The ability to experiment will be fueled in part by the availability of synthetic DNA at reasonable prices. Industrial consumers are likely to be driven more by product development projects in which timely progress is a key to producing a return on investment, which suggests an increasing market for rapid turnaround.

In the commercial world, biological technologies have been deployed largely in the service of developing drugs and transgenic plants. Those sectors are dominated by relatively large companies that earn profit margins that are the envy of businesses in other manufacturing and service sectors. With the substantial sales that support these margins, pharmaceutical and biotech industry associations advertise that their members spend a larger fraction of revenues on research than other industries. For companies in these sectors, outsourced DNA synthesis and gene assembly are technological and economic levers to reduce labor and infrastructure costs.

Changing economics

But the bio-economy is changing rapidly. Revenues from industrial applications of

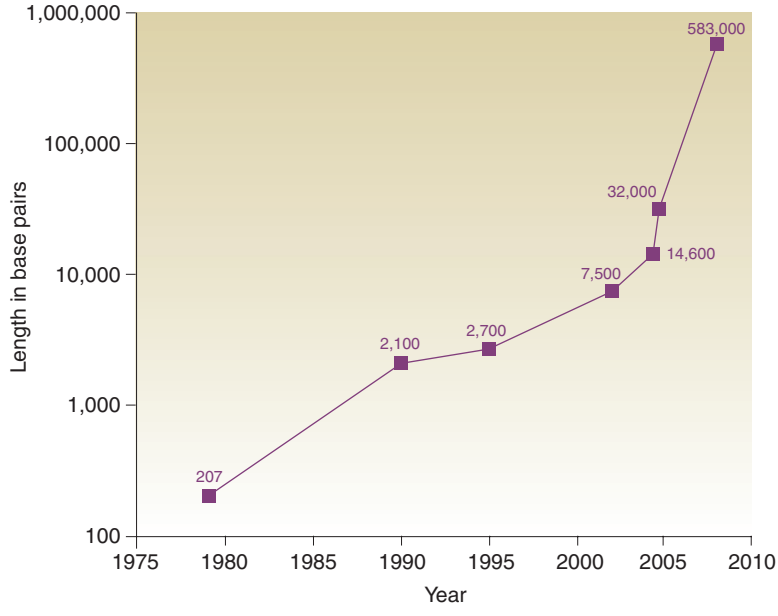

Figure 2 Longest published synthetic DNA[1,12–17] (http://www.synthesis.cc/2009/06/data-and-references-for-longest-published-sdna.html).

biotech appear to have overtaken US revenues from biotech drugs (biologics) and transgenic crops[6], each in the neighborhood of $70 billion, whereas the combination of biofuels, enzymes and materials has reached about $85 billion[7,8].

The relative size of revenue numbers is important for the future of biological engineering for several reasons. Because of lower regulatory burdens, industrial applications are likely to be faster moving and may rapidly become more competitive in their respective markets than biologics or transgenic crops. Green chemistry based on biological processing is already competing with more traditional synthetic chemistry based on petroleum feedstocks in markets worth many hundreds of billions of dollars worldwide[7,8]. It seems likely that in markets that are simultaneously less regulated and closer to the consumer, smaller organizations—operating on smaller budgets and shorter timelines—will have plenty of room to enter markets with products derived from biological processing.

This leads to the question of just which parts of the infrastructure for biological engineering hold the greatest economic value. Is it in the design (bits), or in the objects (atoms)? The synthetic biology market—the ecology of companies that produce and consume products and services related to building genes and genomes—still isn't very big. A generous estimate would put the market for synthetic genes in the neighborhood of $100 million in 2009 (ref. 9). Thus, the revenues for any given synthesis firm are (optimistically) probably no more than a few tens of millions of dollars. Compare the value of the gene synthesis market with the many tens of billions of dollars in revenues from industrial biotech in the United States, and an interesting perspective presents itself.

It is possible that the value of gene synthesis as a service is transitory. Although the assembly of large DNA circuits is presently a technological challenge, and is therefore valuable, the relative value of that assembled DNA is quite small. Of much greater value are the molecules or behaviors specified by those sequences: networks that enable computation or fabrication, enzymes that facilitate processing plants into fuels or fine chemicals, proteins and other molecules that serve as therapeutics and antibiotics.

DNA is cheap, and getting cheaper. Given that the maximum possible profit margin on assembling synthetic genes is falling exponentially (roughly the difference between the cost of genes and the cost of oligos; **Fig. 1**), it would seem that finding value in those particular atoms will become ever more difficult. The design of genetic circuits (resulting in bits) definitely costs more in labor than obtaining the physical sequence by express delivery (resulting in atoms). Moreover, because the electronic specification of DNA sufficient to reproduce the molecule, maintaining proprietary control over sequences of value during the design phase will become ever more important to commercial viability.

In the race to generate new products with new value, firms that develop and sell engineering infrastructure like DNA synthesis and gene assembly will be under pressure to reduce costs and decrease turnaround times. The question is whether customers for DNA of a specific sequence will continue to order it from centralized facilities, or whether economic, technical and regulatory factors might contribute to a decentralization of synthesis. New technologies could enable desktop instruments that provide rapid and secure gene synthesis. Similar technological transitions have resulted in profound transformations of the infrastructure we use for computing, printing and communicating, all of which can now fit in a pocket. At Biodesic, my engineering and design company, our experience is that electrical engineering projects that once required five to ten people working for several years, at a cost of several million dollars, can now be accomplished by one person in less than six months using open source tools. To be sure, there is no guarantee that biological technologies will follow the same route. But neither is there any a priori reason to think that route implausible or unlikely.

We will certainly be hearing more about regulating access to synthesis over the coming months and years.

Market and regulatory forces

To be completely clear, it is not my argument that there will be a gene or genome synthesizer in every small business or even every home, but rather that relatively soon the technology is likely to work well enough to be marketed and used that way. In practice, DNA will be assembled in whatever locale and at whatever scale is demanded by the market and allowed by regulation. And we will certainly be hearing more about regulating access to synthesis over the coming months and years. I suspect, however, that regulations intended to shape the DNA synthesis market will be short-lived and ineffective.

Implementing restrictions on biological technologies that might improve safety and security should always remain among our options. But implementing regulations without a careful examination of possible consequences is unwise. Instituting security measures, and maintaining auditable records of both security and access, will incur costs for producers, users, governments and society as a whole. Understanding the potential costs of restricting access to synthesis first requires examining proposed mechanisms of regulation in greater depth.

To facilitate the control of access to DNA synthesis, Garfinkel *et al.*[10] propose the option of establishing a registry of DNA synthesizers, service providers and certified users. Such requirements would allow DNA synthesis only in what amounts to secure facilities, where security is defined by monitored operation of DNA-synthesis technology either through licensed or permitted ownership of instruments or through licensing of "legitimate users," or both. Sequences submitted to these secure facilities might be kept on file for some number of years to facilitate any forensic efforts. Screening software would examine submitted sequences to identify potential threats in the form of genes and pathways that code for toxins or genomes that code for pathogens.

With respect to the costs of this registry, Garfinkel *et al.*[10] note that "if a review mechanism were too burdensome, small startup firms might shift to in-house DNA synthesis instead." Thus, one of the immediate *social* costs of implementing a registry might be that some otherwise legitimate users opt out of participating due to the *monetary* costs of compliance, thereby limiting the utility of the registry. Given a choice—or if forced by regulatory action to make a choice—some designers of new DNA circuits will inevitably conduct business with synthesis providers who do not maintain an archive of design files. Those who choose to drop off the grid by synthesizing genes in-house could be monitored only if reagents and instruments were strictly controlled. As a result, one potential outcome of restricting access to synthesis might mirror the problem encountered by the US Drug Enforcement Agency when it cracked down on domestic methamphetamine production: information on activities the agency wished to monitor and suppress became much harder to obtain, whereas methamphetamine use continued to rise[7,11]. Similarly, regulatory actions that motivate users to pursue synthesis outside the registry may reduce knowledge

of what is being synthesized, and by whom. Thus, it is not at all clear that regulation will limit access to synthesis technology by users who may be considered a threat. Restricting access to DNA synthesis may motivate some consumers—including those most deserving of scrutiny—to seek access to producers who are either not bound by restrictions or who are willing to ignore them.

In this context, one must also keep in mind the already intrinsically international nature of the DNA-synthesis market[9]. Effective restrictions of access to synthesis must therefore be international in scope and must track the flow of valuable design information through electronic networks, often across borders. This raises the most important vulnerability in DNA synthesis registries and archives, one inherent in the inevitable and increasing reliance on information technology. Whether in the form of electronic signatures, databases of 'legitimate users', screening software or a design tool, this information can be viewed, copied and even altered. Moreover, it is subject to a growing number of security threats that cover the range from simple human mistakes, to fraud, to interception of information during transmission, to complex software attacks on other complex software. Under any international regulatory regime that required screening, individual firms would be faced with exposing their designs to multiple sets of eyes, which would threaten their economic security.

If the policy decision is made by either industry or government to adopt electronic records, another question that must be addressed is, Who will indemnify the security of sequence archives? That is, as the archives by definition hold information that customers deem to be economically valuable, the archives will increasingly be targets for industrial espionage. Who pays for securing the archives long-term? How much does insurance cost? Who is ultimately responsible in the event of a breach?

The future

In the preceding article, I have argued that DNA assembly technologies will become increasingly important tools in specifying synthetic genes and genomes. Those genomes will be used to produce products with revenues much higher than are presently associated with producing the genome itself. As the value of the products generated from synthetic genes and genomes increases, so will the demand to produce DNA faster and cheaper. In that environment, any barrier, whether economic, technological or regulatory, that slows down access to synthetic genes and genomes will be subject to pressure. We are likely to see continued innovation that enables ever more rapid genome construction.

COMPETING INTERESTS STATEMENT
The author declares competing financial interests: details accompany the full-text HTML version of the paper at http://www.nature.com/naturebiotechnology/.

ACKNOWLEDGMENTS
Portions of this text are drawn from *Biology is Technology*, scheduled for publication in 2010. I would like to thank J. Mulligan, J. Minshull, J. Koschwanez, L. Stewart, D. Endy and R. Wehbring for illuminating and entertaining discussions.

1. Gibson, D.G. *et al. Science* **319**, 1215–1220 (2008).
2. Shao, Z. & Zhao, H. *Nucleic Acids Res.* **37**, e16 (2009).
3. Borokov, A. SynBuild, an accurate high-throughput gene assembly platform that uses microarray synthesized oligos. Poster at Synthetic Biology 4.0, Hong Kong, China, October 10–12, 2008.
4. Gibson, D.G. *Nucleic Acids Res.* published online, doi:10.1093/nar/gkp687 (10 September 2009).
5. Keasling, J.D. Synthetic biology in pursuit of low-cost, effective, anti-malarial drugs. Presentation at the Institute for Systems Biology Seventh Annual International Symposium, Seattle, April 21, 2008.
6. Carlson, R. *Nat. Biotechnol.* **27**, 984 (2009).
7. Carlson, R. *Syst. Synth. Biol.* **1**, 109–117 (2007).
8. Carlson, R. *Biology is Technology* (Harvard University Press, Cambridge, in the press).
9. Newcomb, J., Carlson, R. & Aldrich, S. *Genome Synthesis and Design Futures: Implications for the US Economy* (Bio Economic Research Associates, Cambridge, Massachusetts, 2007).
10. Garfinkel, M.S., Endy, D., Epstein, G.L. & Friedman, R.M. *Synthetic Genomics: Options for Governance* (J. Craig Venter Institute, The Center for Strategic and International Studies, Massachusetts Institute of Technology, October 2007). http://www.jcvi.org/cms/fileadmin/site/research/projects/synthetic-genomics-report/synthetic-genomics-report.pdf
11. Carlson, R. *Biosecur. Bioterror.* **1**, 203–214 (2003).
12. Khorana, H.G. *Science* **203**, 614–625 (1979).
13. Mandecki, W., Hayden, M.A., Shallcross, M.A. & Stotland, E. *Gene* **94**, 103–107 (1990).
13. Stemmer, W.P.C., Crameria, A., Hab, K.D., Brennan, T.M. & Heynekerb, H.L. *Gene* **164**, 49–53 (1995).
15. Cello, J., Paul, A.V. & Wimmer, E. *Science* **297**, 1016–1018 (2002).
16. Tian, J. *et al. Nature* **432**, 1050–1054 (2004).
17. Kodumal, S.J. *et al. Proc. Natl. Acad. Sci. USA* **101**, 15573–15578 (2004).