©2016 Published in 4th International Symposium on Innovative Technologies in Engineering and Science 3-5 November 2016 (ISITES2016 Alanya/Antalya - Turkey)

Emerging Synthetic Biology Techniques: Synthetic DNA, Metabolic Pathways, & Life

[1]Daniel Holley, [2]Katherine Hooker, and [2]Jeffrey Kopszywa
*[1]University of Maryland University College, Graduate School of Biotechnology, RAF Alconbury, England
[2]University of Maryland University College, Graduate School of Biotechnology, Adelphi, Maryland, USA

Abstract

As the body of biological knowledge rapidly expands, biotechnologists are increasingly able to manipulate natural systems through techniques such as recombination and directed evolution (DE). The emerging field of synthetic biology leverages multidisciplinary knowledge in the synthesis of novel biological building blocks, optimized metabolic pathways, and even artificial organisms, theoretically toward the goal of building cellular factories that are more tractable and productive than their wild-type, recombinant, or DE predecessors. Synthetic biology encompasses two overarching themes: changing existing systems to improve overall function, and creating *de novo* systems that do not exist in nature. Most often, researchers pursue the former, since the technologies involved in systems manipulation are better understood than those involved in systems creation. However, the latter stands to grow exponentially as computational power increases and the various *-omics* disciplines (i.e., proteomics, metabolomics, etc.) mature. The authors review both themes herein, with a particular focus on synthetic DNA, synthetic metabolic pathways, and synthetic life.

Key words: biotechnology, synthetic biology, synthetic DNA, synthetic life

1. Introduction

Synthetic biology is a cutting-edge field at the intersection of genetics, biochemistry, and biophysics [1]. The ability to more easily manipulate cellular operations and the prospect of creating synthetic biological building-blocks and entirely synthetic microorganisms have major implications for industries ranging from designer-enzyme production to bioencryption. A brief examination of three synthetic biology watersheds—synthetic DNA, synthetic metabolic pathways, and synthetic life—augurs a field on the cusp of exponential growth.

2. Synthetic DNA

The basic process to construct synthetic DNA requires DNA oligonucleotides—short DNA fragments that undergo phosphoramidite chemistry to produce nucleotide chains up to 200-base-pairs (bp) long [2]. The oligonucleotides are then replicated, ligated, and hybridized to create synthetic genes [2]. Oligonucleotide synthesis can be costly, and some methods can produce errors as frequently as 1 in 300 bp [3].

*Corresponding author: Address: University of Maryland University College, Europe Campus, RAF Alconbury, England, PE28 4DA. E-mail address: daniel.holley@umuc.edu, Phone: +4401480843686

2.1. Microarray

To overcome this high error rate, researchers can use DNA microarray: a high-fidelity technology capable of synthesizing thousands of oligonucleotide sequences quickly, efficiently, and at a low cost [3]. In 2010, Harvard Medical School professor George M. Church and colleagues used microarray to synthesize a target DNA library [4]. Church's team used Roche 454, a next-generation sequencing platform, to rapidly sequence oligonucleotide pools derived from microarray [4]. Church and colleagues obtained and amplified the synthesized oligonucleotides to produce a library in which the error rate was reduced 500 fold [4]. Although microarray requires large amounts of sequences and produces few DNA sequences [5], the technology has advanced the production of synthetic DNA from hundreds of base pairs to potentially entire genomes.

2.2. DNA-Synthesis Technology

The results of the first chemical gene synthesis were published in the Journal of Molecular Biology in December 1972 [5]. Now, researchers can synthesize their own DNA or purchase sequences from multiple vendors that can ship overnight [5]. One example of a prefabricated DNA-synthesis tool is BioBrick, invented by MIT computer scientist Tom Knight [5]. BioBrick is a standardized set of rules used to define features of a DNA sequence which assemble as larger pieces *in vitro* [6]. BioBrick has limitations, as it is a sequence-driven technology that does not translate function. Hence, BioBrick is suitable for small system designs but is not practical for larger network designs [6].

Another example of DNA-synthesis technology is Gibson Assembly®, which composes sequences in parallel. Researchers used Gibson Assembly to produce the entire 16.3-kb mouse mitochondrial genome in 600 oligonucleotide segments, each of 60 bp [5]. Other assembly techniques include Golden Gate Shuffling, sequence and ligation independent cloning (SLIC), and enzymatic inverse polymerase chain reactions (EIPCR) [5]. These technological advances have driven down the cost of DNA synthesis and have encouraged research worldwide.

3. Synthetic Metabolic Pathways

Effectively being able to obtain gene coding sequences and/or produce synthetic DNA has provided scientists the capability to alter an entire metabolic pathway. Since the dawn of the biotechnology era, researchers have strived to harvest biological compounds from recombinant organisms. Medical relevance and industrial application have been the two main catalysts of innovation during the search for new biomolecules. Cellular optimizations that improve quantity, purity, and control are driving the industry forward, especially in the lucrative biofuel race. Fortunately, a vast majority of organisms already possess the ability to produce biological commodities such as ethanol and butanol. Unfortunately, these native organisms do not possess the ideal molecular mechanics to optimally produce biofuels [7].

3.1. Site-Specific Mutagenesis

Optimizing each enzymatic reaction individually and then collectively incorporating the changes into a final construct is a common means of optimizing a pathway [8]. Site-specific mutagenesis is typically achieved by polymerase chain reactions (PCR) using synthesized primer sets with intentional base-pair swaps or deletions, followed by a transformation using re-circularized plasmids containing the "new" insert. This method affords the ability to alter one site at a time, although it is not guaranteed to provide the desired results. Complicating things further, as the number of enzymatic reactions increases, the overall optimization becomes more difficult [8]. Identifying the best optimizations must take into account all products and byproducts of the pathway if the system is to intrinsically work as one.

3.2. Indirect Pathway Manipulation

There are alternative approaches to manipulating metabolic pathways other than individual enzymatic optimizations within a specific pathway. For example, the manipulation of the carbon storage regulator system within *Escherichia coli* has shown to double n-butanol while reducing unwanted byproduct accumulations of acetate and carbon dioxide [9]. The accumulation of such byproducts pulls resources away from the net desired product yield and ultimately creates a toxic environment in which the cell is cultured. Accumulation of any one product, including n-butanol, is obviously detrimental to the cell over the course of its lifetime. Fortunately, there are also methods of altering a cell to increase product tolerance.

3.3 Random Mutagenesis

Continuing the butanol-fermentation-pathway example, Mienda, Shamsir, Salleh, and Illias [10] have shown the ability to increase the tolerance of 1-butanol within *E. coli* by altering the global transcription factor cyclic adenosine monophosphate (cAMP) receptor protein (CRP) [10]. The genetic alteration achieved by Mienda and colleagues relied on error-prone PCR and DNA shuffling in contrast to a site-directed mutagenesis method [10]. Error prone PCR generates random mutations within DNA sequences by using a DNA polymerase with a high base pair error rate during successive PCR cycles. After ligation and transformations, researchers must use time-intensive screening methods to determine whether any of the mutants produce more of the desired product than their wild-type counterparts.

Another "shotgun" approach known as DNA shuffling relies on indiscriminate nuclease digestion of a DNA sequence followed by successive PCR amplifications without primers. An additional PCR cycle is then used to append complementary ends or restriction sites to assist with cloning the fragments into plasmids. As with the aforementioned CRP-altering technique, screening for productive mutants is laborious.

3.4. In Silico Metabolic Pathway Modeling

Biochemical pathway modeling facilitated by software platforms like OptFlux has shown promise in predicting parameters required for optimal system efficiency [10]. However, along

with the heuristic algorithms within these types of programs, user-defined parameters are also required and can lead to spurious results. Advances in *in silico* models will undoubtedly improve predictive screenings, which are traditionally used for site-directed and random mutagenesis. Once the fundamentals of metabolic pathway optimizations are thoroughly understood, the next challenge will be the synthesis of life itself.

4. Synthetic Life

In his seminal work, *On the Origin of Species*, the naturalist Charles Darwin concluded that "probably all the organic beings which have ever lived on this earth have descended from… one primordial form, into which life was first breathed" [11, p. 420]. Over 150 years of scientific advances in domains ranging from geography [12] to genetics [13] have supported Darwin's prescient conviction, disputes about extremely venerable microorganisms notwithstanding [14].

Recently, however, scientists have disrupted the chain of common ancestry by giving rise to *Mycoplasma mycoides*—the first organism in 3.5 billion years that has no true ancestor [15]. Rather, *M. mycoides* in its current form (i.e., JCVI-syn3.0) is a product of techniques that have allowed researchers to create cellular instructions through genomic synthesis, couple those instructions with naturally evolved cellular machinery through genomic transplantation, and reduce the size of those instructions through genomic minimization.

4.1. Genomic Synthesis

A major step toward breaking *M. mycoides* from the 3.5-billion-year chain of ancestry was the chemical synthesis of a functional whole genome [16]. This entailed sequencing the DNA of a donor bacterium, *M. genitalium*; digitizing that DNA sequence to allow for computer-aided editing (e.g., the addition of watermark sequences); and chemically synthesizing the new genome from scratch [17]. Researchers used yeast artificial chromosomes of a highly transformable *Saccharomyces cerevisiae* strain as the factories in which to assemble the otherwise intractable synthetic bacterial genome [18, 19]. This heterologous approach required the use of highly purified DNA to facilitate assembly [20]. The ability to build *de novo*, double-stranded DNA sequences within *S. cerevisiae* was a major stride toward the J. Craig Venter Institute's goal of "[creating] a tool that would allow people to take their organism, clone its genome, manipulate its genome, then boot it up" [19, p. 698] by transplanting the new genome into a recipient cell.

4.2. Genomic Transplantation

For any genome to function appropriately, it must be complemented by cellular machinery capable of transcription and translation. In the case of *M. mycoides*, researchers used the bacterium *M. capricolum* as the recipient of the full-genome transplant [16]. *M. capricolum* possessed the polymerases, ribosomes, and host of other naturally evolved intracellular accoutrements necessary to express *M. mycoides'* synthetic genome [15]. Following the replacement of *M. capricolum's* native genome with *M. mycoides'* synthetic genome, the transplant recipient spawned daughter cells with no ancestral past.

4.3. Genomic Minimization

Through synthetic biology, researchers are gaining a better understanding of what makes life possible. JCVI-syn3.0, for instance, is the product of painstaking genomic minimization—the removal of non-essential genes through trial and error [21]. In the 6 years between JCVI-syn1.0 and -syn3.0, researchers reduced *M. mycoides'* genomic size from 1,079 to 531 kbp via transposon mutagenesis, leaving the synthetic bacterium with a mere 473 genes, one-third of which perform unknown vital functions [21].

A *de novo* minimal-genome organism will allow researchers to engineer a greater range of controlled cellular optimizations, leading to the increasingly efficient production of biofuels, novel proteins, and pharmaceuticals [22, 23]. Aside from serving as a template for future whole-genome syntheses, *M. mycoides'* intriguing genes of unknown function invite novel research into facets of life that we do not yet understand [21].

Conclusions

Synthetic biology is a rapidly advancing field, and those at its forefront have captured the public's imagination through their path-finding work in the 'creation of life.' The advent of synthetic DNA and the promise of malleable synthetic metabolic pathways are equally exciting and industrially promising. As interdisciplinary fields such as bioinformatics mature and reduce the cost of synthesis through better predictive modeling, and as computational power rises to the task of modeling complex pathways at a systems level, synthetic biology will become a staple of applied biological science.

Acknowledgements

The authors would like to gratefully acknowledge the support and mentorship of Dr. Rana Khan of University of Maryland University College's Graduate School of Biotechnology.

References

[1] Haseloff J, Ajioka J. Synthetic biology: history, challenges and prospects. Journal of the Royal Society 2009;6(4). doi:10.1098/rsif.2009.0176.focus
[2] Hranueli D, Starcevic A, Zucko J, Diego Rojas J, Diminic J, Baranasic D, et al. Synthetic biology: A novel approach for the construction of industrial microorganisms. Food Technology & Biotechnology 2013;51(1):3-11.
[3] Matzas M, Stähler P, Kefer N, Siebelt N, Boisguérin V, Leonard J, et al. High-fidelity gene synthesis by retrieval of sequence-verified DNA identified using high-throughput pyrosequencing. Nature Biotechnology 2010;28:1291-4. doi:10.1038/nbt.1710
[4] Lee B, Cho S, Song Y, Chang Kim S, Cho B. Emerging tools for synthetic genome design. Molecules and Cells 2013;35:359-370. doi:10.1007/s10059-013-0127-5

[5] Baker M. The next step for the synthetic genome. Nature 2011;473(7347): 403-8. doi:10.1038/473403a
[6] Alterovitz G, Muso T, Ramoni, M. The challenges of informatics in synthetic biology: From biomolecular networks to artificial organisms. Briefing in Bioinformatics 2009;11(1):80-95. doi:10.1093/bib/bbp054
[7] Na D, Kim T, Lee S. Construction and optimization of synthetic pathways in metabolic engineering. Current Opinion in Microbiology 2010;13:363-370. doi:10.1016/j. mib.2010.02.004
[8] Holtz W, Keasling J. Essay: Engineering static and dynamic control of synthetic pathways. Cell 2010;140:19-23. doi:10.1016/j.cell.2009.12.029
[9] McKee A, Rutherford B, Chivian D, Baidoo E, Juminaga D, Kuo D, et al. Manipulation of the carbon storage regulator system for metabolite remodeling and biofuel production in Escherichia coli. Microbial Cell Factories 2012;11(1):79-90. doi:10.1186/1475-2859-11-79
[10] Mienda B, Shamsir M, Salleh F, Illias R. In silico prediction of Escherichia coli metabolic engineering capabilities for 1-butanol production. Current Science 2015;108(4):686-694.
[11] Darwin C. On the origin of species by means of natural selection [U.S. edition's annotated draft]. New York: D. Appleton and Company; 1861.
[12] Wegener A. The origin of continents and oceans. Quarterly Journal of the Royal Meteorological Society 1925;51(215):287. doi:10.1002/qj.49705121512
[13] Ijdo J, Baldini A, Ward D, Reeders S, Wells R. Origin of human chromosome 2: An ancestral telomere-telomere fusion. Proceedings of the National Academy of Sciences of the United States of America 1991;(20):9051.
[14] Theobald D. On universal common ancestry, sequence similarity, and phylogenetic structure: The sins of P-values and the virtues of Bayesian evidence. Biology Direct 2011;6:60. http://doi.org/10.1186/1745-6150-6-60
[15] Gowrishankar J. Craig Venter, and the claim for 'synthetic life'. Current Science 2010; 99(2): 152.
[16] Gibson D, Glass I, Lartigue C, Noskov V, Chuang R, Algire M, et al. Creation of a bacterial cell controlled by a chemically synthesized genome. Institute of Medicine (US) Forum on Microbial Threats 2011. Retrieved on 17 July 2016, from http://www.ncbi.nlm.nih.gov/books/ NBK84435/.
[17] Levin M. Creation of a bacterial cell controlled by a chemically synthesized genome. Journal of Scientific Exploration 2010;24(4):783-4.
[18] Gibson D. Synthesis of DNA fragments in yeast by one-step assembly of overlapping oligonucleotides. Nucleic Acids Research 2009;37(20):6984–6990. http://doi.org/10.1093/nar/ gkp687
[19] Rusk N. Engineering bacteria in yeast. Nature Methods 2009;6(10):698. doi:10.1038/ nmeth1009-698
[20] Rusk N. Advent of synthetic life. Nature Methods 2010;7(7):487. doi:10.1038/nmeth0710-487
[21] Hutchison C, Chuang R, Noskov V, Assad-Garcia N, Karas B, Ma L, et al. Design and synthesis of a minimal bacterial genome. Science 2016;351(6280). doi:10.1126/science.aad6253
[22] Dorit R. Making life from scratch: Artificial intelligence is not human intelligence, nor is synthetic life the same as life with evolutionary history. American Scientist 2013;(5):342.
[23] Riis S. The ultimate technology: The end of technology and the task of nature. Artificial Life 2013;19(3-4):471-485. doi:10.1162/ARTL_a_00119