Building biological foundries for next-generation synthetic biology

CHAO Ran[1,2†], YUAN YongBo[2†] & ZHAO HuiMin[1,2,3*]

[1]*Department of Chemical and Biomolecular Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA;*
[2]*Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA;*
[3]*Department of Bioengineering, Department of Chemistry, Center for Biophysics and Computational Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA*

Synthetic biology is an interdisciplinary field that takes top-down approaches to understand and engineer biological systems through design-build-test cycles. A number of advances in this relatively young field have greatly accelerated such engineering cycles. Specifically, various innovative tools were developed for *in silico* biosystems design, DNA *de novo* synthesis and assembly, construct verification, as well as metabolite analysis, which have laid a solid foundation for building biological foundries for rapid prototyping of improved or novel biosystems. This review summarizes the state-of-the-art technologies for synthetic biology and discusses the challenges to establish such biological foundries.

biofabrication, automation, metabolic engineering, biosystems design, DNA assembly, DNA sequencing, metabolite analysis

Synthetic biology is an extension of the long pursuit of ability to forward-engineer biological systems. Enabled by molecular and systems biology tools, synthetic biology emerged in the 1990s [1]. Thanks to advances in robust DNA assembly methods [2] and standardized genetic parts [3,4], this interdisciplinary field has undergone exponential growth in the first decade of the 21st century and demonstrated its tremendous potential in understanding and engineering biosystems. As a bottom-up observation oriented science, biology needs new dimensions to explore increasingly complicated systems of interest. As American physicist Richard Feynman stated, "What I cannot create, I do not understand", the capability of *de novo* designing and building genetic devices and systems empowers a top-down approach to study life sciences. For example, by constructing simplified genetic circuits, stochastic gene expression can

be modeled, which uncovers mechanisms of phenotypical variations in isogenic population [5,6]. Cryptic metabolic pathways predicted by bioinformatics can be activated by refactoring the gene expression machinery to discover novel biochemical reactions and products [7,8]. In addition, traits can be traced to their genetic background with genome scale perturbations [9,10].

By definition, synthetic biology has great potential in revolutionizing the way biosystems are engineered. In metabolic engineering, one of the most widely recognized successes is the engineered biosynthesis of anti-malarial drug artemisinin, the active component of Chinese traditional herbal medicine *Artemisia annua* [11]. Biosynthetic pathways were not only optimized for productivity but also controlled by genetic circuits to adapt to a changing environment [12]. A quorum sensing circuit was employed to program environmentally controlled cell invasion for potential anti-cancer applications [13]. Dynamic control was introduced in a biodiesel producing pathway to stabilize the level

---

†Contributed equally to this work
*Corresponding author (email: zhao5@illinois.edu)

of key intermediates [14]. On a genome scale, tools such as oligo-mediated allelic replacement [15] and CRISPR-Cas (clustered, regularly interspaced short palindromic repeats-CRISPR-associated proteins) system [16,17] are exploited to perform efficient multiplex genome editing. These tools developed in the past decade have laid the ground for synthetic biology. Nevertheless, due to the complexity of the biosystems and low success rate, a trial-and-error approach is required to survey many variations or combinations when prototyping biosystems. Therefore, it is highly desirable to develop high-throughput prototyping capabilities for synthetic biology.

In the early 2000s, successes in the construction of genetic toggle switches [18], oscillators [19], logic gates [20], and so on have initiated the first tide of synthetic biology. These early findings implied an analogy between the synthetic biology field and the electronics industry. Soon after, like in the electronic industry, efforts have been made to standardize and characterize genetic parts that can be mixed-and-matched to form progressively complex circuits, although most of them still need "hand tuning" to be functional. In early years, wiring electronic circuits was also an artisanal job until planar process was developed that allowed consistent fabrication of integrated circuits on a silicon wafer [21]. This game-changing technology streamlined the manufacturing process and accelerated the industry to the astonishing Moore's law. If the analogy holds, streamlining biofabrication process will change the landscape of synthetic biology. In the ideal situation, researchers' ideas are transformed to designs *in silico*, then submitted to automated prototyping systems and eventually tested in different organisms. With the rapid progress in this arena, we see establishing "biological foundries" no longer a fiction. This perspective reviews the state-of-the-art technologies and discusses the opportunities as well as challenges in three major key aspects of biofabrication including design, build, and test (Figure 1).

1 Design

A typical synthetic biology work flow starts from target functionalities such as compounds to be produced, substrates to be assimilated, and regulatory networks to be established. Based on prior knowledge, metabolic pathways or genetic circuits are designed while parts are selected. Usually multiple part sets are proposed due to the relatively low success rate. Then the DNA constructs are designed for appropriate assembly methods. As an example, Kim and co-workers [22] improved a xylose assimilation pathway in *Saccharomyces cerevisiae* by screening a combinatorial library. The design began with a 3-gene pathway in *Scheffersomyces stipitis* that converted xylose to xylulose-5-phosphate. Around 20 orthologs of each gene were identified by basic local alignment search tool (BLAST). A
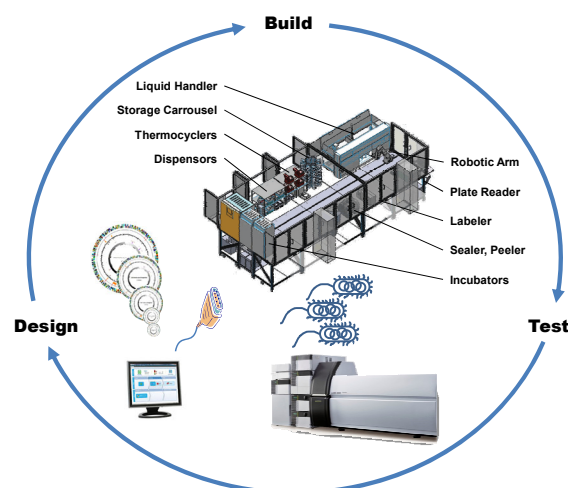

Figure 1 (color online) Illustration of automated design-build-test cycle for the next-generation synthetic biology: Biosystems are designed by CAD tools, built by robotic systems, and then tested by high throughput analytical instruments. A layout for Illinois Biological Foundry for Advanced Biomanufacturing (iBioFAB), the automated biofabrication platform at the University of Illinois, USA is highlighted.

general plasmid map was designed with one set of genes as well as yeast endogenous promoters and terminators. Primer sequences with overhangs were planned to PCR amplify all the orthologs and assemble them combinatorially with the DNA assembler method [23]. Such workflow consists of two major steps, including (a) design of parts and systems and (b) design of the corresponding DNA constructs.

1.1 Design of parts and systems

Researchers' knowledge and experience played an important role in traditional biosystems design when most of the work was done by hand. However, the sizes of databases for parts, circuits, and pathways today have expanded far beyond any individual's knowledge. Even if searching for each individual part is manageable, searching for combinations of parts will end up in numerous possibilities. The tedious manual design process limits research in increasingly complex biosystems. As in the electronic industry, computer aided design (CAD) will be the ultimate solution which allows optimization and higher throughput in biosystems design. On the part level, Salis and co-workers [24] have developed a thermodynamic model to predict the protein expression level based on the ribosome binding site (RBS) sequences in *Escherichia coli*. A tool was made available to design RBSs with varying strengths, which greatly helped fine-tuning of genetic circuits and metabolic pathways. In *S. cerevisiae*, synthetic promoters were designed based on the correlation between expression level and predicted nucleosome affinity [25]. CAD has broad applications in modeling and evaluating gene circuits. Graphical user interfaces (GUIs) were also developed to

help users pick parts and simulate their designs. A number of computational tools were summarized in previous reviews [26,27]. Growing genomic and metabolomics data provide us mechanistic insights on microbial metabolic networks. Constraint-based reconstruction and analysis (COBRA) was developed to model intracellular metabolism [28]. These tools can be used to predict metabolic engineering targets for desired phenotypes. In return, high-throughput genome editing and analysis with the newly developed genome editing tools can potentially provide feedback data to further train the metabolic models for more precise predictions. Biosynthetic targets sometimes involve *de novo* composition of pathways. A few tools [29] were developed to explore the reaction and enzyme databases [30–32] for discovery of novel biochemical routes. Hatzimanikatis and co-workers [33] have developed Biochemical Network Integrated Computational Explorer (BNICE) which employed enzymatic reaction rules and thermodynamic properties to search for reactions linking the desired substrates and products *in silico*. BNICE was used to predict more than one billion polyketide structures by exploring theoretical diversities in the synthetic pathways [34]. Similarly, Hou and co-workers [35] at the University of Minnesota have developed Pathway Prediction System (PPS) to predict bio-degradation pathways.

1.2 Design of DNA constructs

Once the conceptual design of the biosystems is completed, the selected parts will be assembled with one of the available DNA assembly methods. Since most methods utilize either restriction digestion or homologous recombination,

the map of the DNA molecule that encodes the biosystem needs to be laid out nucleotide by nucleotide to ensure accurate assembly. Considering the large number of base pairs, this step is usually done in a CAD environment. The early though popular DNA design tools such as Vector NTI[TM] [36] handle the design on a continuous DNA sequence, which is competent for sequence analysis and simple cloning but becomes inefficient when designing a large number of constructs with standardized parts. New generation of DNA CAD tools suit the need of synthetic biology and adopt a modular design environment. j5 is a web-based tool that takes standard DNA part sequence files as input and designs the constructs based on user selected or software recommended assembly strategy [37]. j5 has automated the design process and is capable of processing libraries of DNA constructs in a high throughput manner. RAVEN is a newly reported tool that not only automates the design process, but also optimizes the assembly plan by self-learning algorithm with experimental results [38].

2 Build

2.1 Assembly methods

DNA assembly methods play a fundamental role in biofabrication. Recently, a number of DNA assembly methods have been developed with improved modularity, efficiency, and fidelity that fulfill the increasing demand in synthetic biology. Since most of the methods have been recapitulated in previous reviews [2,39], only the ones that can be potentially applied in the biofabrication process are briefly discussed here (Figure 2).
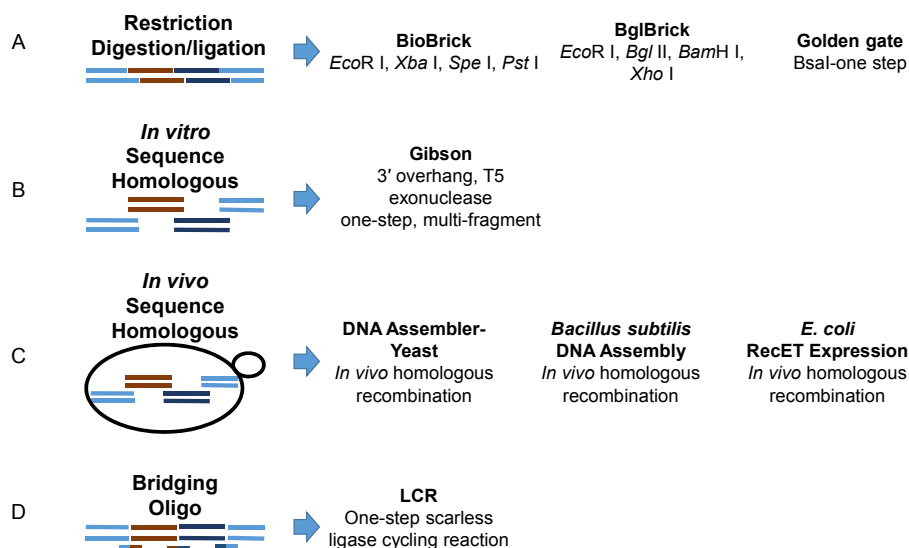

Figure 2 (color online)   Summary of key DNA assembly methods. A, DNA fragments are treated by restriction endonucleases to generate compatible ends and then ligated by ligase to create circular constructs. B, DNA fragments with overlapping regions are treated by exonuclease to generate single-stranded complementary overhangs and joined by DNA polymerase and ligase. C, Host organisms are directly transformed by DNA fragments with overlapping regions. The constructs are assembled by homologous recombination *in vivo*. D, Two neighboring fragments are annealed to a bridging oligo and then ligated by DNA ligase.

The BioBrick[TM] standard utilizes restriction endonuclease isocaudamers *Xba* I (TCTAGA) and *Spe* I (ACTAGT) to generate two compatible sticky ends (CTAG) on the receiver and donor plasmids respectively for ligation [40]. The newly generated scar sequence (ACTAGA) between the two DNA parts cannot be cut in subsequent digestions by neither enzyme, while new *Eco*R I and *Xba* I sites are introduced on the other end of the inserted fragment, which allows iterative digestion and ligation. To solve the frame shift problem of the original BioBrick[TM] standard in protein fusion applications, modified standards were proposed [41–43]. The Golden Gate method [44] relies on Type IIs restriction enzymes like *Bsa* I, which is able to cleave DNA outside of their recognition site and produce an overhang of four arbitrary nucleotides. Once ligated, the resultant fragment cannot be cut anymore unless the recognition site was purposely retained. The protocol combines digestion and ligation in one pot, which can greatly increase the efficiency by driving this irreversible reaction towards completion.

The application of restriction enzyme based methods in constructing large constructs is relatively limited since the occurrence of designated restriction sites increases with the size of fragments, which need to be removed by either site-directed mutagenesis or *de novo* synthesis. Sequence homology based methods do not need restriction digestion to generate sticky ends but use longer arbitrary overlapping regions between parts. Gibson isothermal assembly generates single-stranded complementary overhangs by digestion on the 5′ ends with T5 exonuclease [45]. The annealed ends are then jointed by Phusion DNA polymerase and *Taq* DNA ligase. All reactions happen in the same buffer at 50°C. *S. cerevisiae* has a highly efficient natural homologous recombination machinery. With this machinery, Gibson and coworkers [46] have reconstructed the *Mycoplasma genitalium* genome by assembling 25 DNA fragments. Shao and co-workers [23] have also developed DNA assembler, a robust method to assemble pathways in yeast. Similarly, homologous recombination based assembly was performed in other organisms such as *Bacillus subtilis* [47,48], plants [49,50] and engineered *E. coli* [51,52].

In 2000, Pachuk and co-workers [53] developed a method named chain reaction cloning (CRC) to assemble DNA fragments facilitated by designed single strand bridging oligos which are complementary to two ends of neighboring DNA parts. Researchers at Amyris (San Francisco, USA) systematically optimized this method and developed ligase cycling reaction (LCR) to reliably assemble thousands of constructs [54].

2.2   Assembly schemes

Robust, broadly accessible strategies for constructing multigene pathways are highly desired in synthetic biology. An iterative integration scheme for expressing multigene pathways was proposed by Wingler and Cornish [55]. Two distinct markers were used alternately to provide selective pressure for the incorporation of each fragment by homologous recombination. This sequential integration strategy improved assembly fidelity, however, at a cost of time. By contrast, one-step assembly allows faster construction of multiple fragments into one construct, although it is not easy to achieve the same level of fidelity and efficiency. Many efforts have been made to tackle this issue. Non-homologous end joining (NHEJ) of linear fragments is one of the commonly seen problems in homologous recombination based methods. To suppress the false positive clones, a counter-selective marker can be introduced at the cloning site on the vector [56]. Therefore, the host is forced to replace the adverse gene with the designated inserts. Another strategy is to separate selection marker and episome, the two essential elements on the vector to reduce the probability of false positives since at least two NHEJs are required to incorporate both elements into a self-replicable circular molecule [57]. 95% fidelity was achieved when assembling nine fragments.

Unique overlapping linker regions are responsible for joining and addressing DNA fragments in many assembly methods. The linker sequences can be optimized to achieve better assembly robustness. Liang and co-workers [58] experimentally optimized the 4-bp linkers in Golden Gate method for synthesizing transcription activator-like effectors nucleases (TALENs). TALENs are proteins that cleave target DNA sequences with tailored series of central repeat domains (CRDs) and are widely used in genome editing. Large-scale application of TALENS is hindered by the time-consuming synthesis protocol. The optimized linkers made it possible to assemble 13 fragments or DNA sequences coding for up to 31 CRDs in one step with 96% fidelity. Hence, the assembly scheme is simplified which enables the synthesis process readily automatable. Most homology based methods like Gibson assembly allow assembly of multiple genetic parts in one step. Nonetheless, they often suffer from unspecific annealing when dealing with highly repetitive sequences. Instead of the endogenous sequences, synthetic linkers which are optimized in parameters like orthogonality, GC content, melting temperature, and secondary structures can be used to improve the efficiency and fidelity [59–61]. Moreover, the use of pre-defined linker sets makes DNA construct design more modular. With synthetic linkers, new schemes based on Gibson method were developed to efficiently assemble large mammalian genetic circuits [59,60].

2.3   *De novo* synthesis and assembly automation

Most DNA parts used in synthetic biology today are copies of existing sequences in nature. To make the biosystems fully reprogrammable, the parts need to be synthesized *de novo*. Parts can also be customized by restriction site removal, linker addition, as well as codon optimization to

make the assembly and expression easier. Typical *de novo* synthesized DNA molecules are assemblies of chemically synthesized oligonucleotides [62]. Owing to the invention of microarray based DNA synthesis, large amount of oligos can be densely "printed" on a DNA chip, which greatly increased the throughput and lowered the cost. The newly developed assembly methods have also made it more reliable to build from oligos to a few thousand base pairs long molecules. Thus we have witnessed great breakthroughs in the past few years. In 2008, a 583 kb *M. genitalium* has been synthesized by the J. Craig Venter Institute [63]. In 2010, a *M. genitalium* cell with the chemically synthesized genome was created, of which the phenotype matched the *in silico* design [64]. On the eukaryote side, a functional designer *S. cerevisiae* chromosome III with multi-loci modifications was successfully synthesized [65]. Beside their significance in synthetic biology, these pieces of groundbreaking work have also provided more evidences to the long lasting question: can life be created?

No matter whether DNA parts were *de novo* synthesized or amplified from existing templates, DNA assembly still remains a critical and labor-intensive step in the build process. The considerable labor cost, high error rate, and insufficient throughput in manual assembly operations limit the number of designs that can be prototyped. Ideally, researchers should focus more on creative designing work instead of mechanical assembly experiments. With more and more robust DNA assembly methods, automating DNA assembly became possible. In the previously discussed high-throughput TALEN synthesis work [58], a computational tool was programmed to convert TALEN binding sequences to scripts for liquid handling stations. The RVD parts could be automatically pipetted to designated wells for TALEN synthesis. Through automation, the labor cost was greatly reduced and pipetting mistakes were avoided. By transferring the know-hows from industrial automation, the state-of-art laboratory automation systems can conduct most of the molecular biology operations. Therefore, the depth of automation in DNA assembly can be further extended to eliminate human interventions.

3 Test

3.1 Genotyping

The essence of synthetic biology is to build biosystems as designed, which requires exact control in each prototyping step. As quality control, DNA constructs are usually verified after assembly. Likewise, the mutations of extensively edited genome with desired traits need to be traced in order to precisely elucidate the phenotype. Thus genotyping tool is an important node in the biofabrication line.

DNA sequencing is the ultimate tool for construct verification because it reveals the assembly fidelity to single nucleotide resolution. However, cost and read length of Sanger sequencing, the most commonly used method for small DNA constructs, are obvious barriers for scaling up. Hence restriction digestion followed by gel electrophoresis is usually adopted on the bench side. To fulfill the requirement of high-throughput biofabrication, this process has been automated [66]. A computational tool was exploited to select restriction enzyme combinations to optimize the digestion pattern according to the designs of the constructs. Liquid handling stations followed the digestion design to pick enzymes from a stock library. The digestion patterns were analyzed with a high-throughput capillary electrophoresis system, and then compared with the prediction. Next-generation sequencing has revolutionized genotyping, which makes large-scale sequencing possible [67]. On the other hand, the cost has been greatly lowered due to the parallel sequencing capability. The smallest Illumina® sequencer can generate 25 million reads on a single run, which provides enough coverage for an entire microbial genome or hundreds of plasmids. If samples are individually barcoded, the reads can be readily aligned to the corresponding reference genome or plasmid design. Since the fixed cost on each sequencing run contributes most to the total cost, maximizing the number of samples will bring the per sample cost down to an acceptable level. When the biofabrication process scales up, next-generation sequencing may become an ideal method to verify DNA constructs.

3.2 Phenotyping

High throughput analysis of intracellular/extracellular metabolites is an essential step in prototyping biosystems. The widely used conventional methods for metabolite analysis include nuclear magnetic resonance (NMR), gas chromatography-mass spectrometry (GC-MS), liquid chromatography-mass spectrometry (LC-MS), capillary electrophoresis-mass spectrometry (CE-MS), and matrix-assisted laser desorption ionization-mass spectrometry (MALDI-MS). NMR is a powerful technique for metabolite profiling and allows direct identification and quantification of abundant analytes [68], yet the throughput is limited by the complicated sample preparation and instrument operation. GC-MS offers efficient separation of non-polar and volatile metabolites with high resolution, reproducibility, and relatively low operating cost [69]. Derivatization step is required for detection of polar and non-volatile metabolites [70]. Latest GC-MS systems feature automatic sample preparation, derivatization and injection, which reduced analytical time and human intervention. LC-MS, on the other hand, allows analysis of a wide range of metabolites without prior derivatization, especially when coupled with the electrospray ionization (ESI) method. Some recent advances significantly improved the resolution power and peak capacity [71]. For identification of unknown metabolites, tandem mass spectrometry (MS/MS) has been applied in which two or more mass analyzers are coupled by collision-induced ioni-

zation chambers (e.g., Q-TOF). CE-MS is considered one of the most versatile analytical techniques in proteomics and metabolomics [72]. CE separates analytes based on their charge and size, which is capable of analyzing a diverse range of chemical compounds, charged or polar molecules with minimal sample preparation. Therefore, it has good potential for high throughput metabolite analysis and even single cell metabolomics and subcellular structure study [73]. The resolution of CE, in many cases, is higher than LC-MS because the capillary in CE usually provides better plug flow of analytes hence narrower peaks than that of the chromatography column in LC [74]. Compared to other MS-based analytical methods, MALDI-MS has higher sensitivity and smaller sample volume. Moreover, MALDI-MS is a high throughput technique due to its ability to analyze a large number of metabolites in one shot with direct ionization mechanism [75]. The choice of matrix is critical in MALDI-MS analysis. Other than conventional matrices like DHB (2,5-dihydroxybenzoic acid) and CHCA (alpha-cyano-4-hydroxy-cinnamic acid), 9-aminoacridine (9-AA) has been reported as a better matrix in quantitative detection of metabolites [76]. Based on 9-AA matrix, a MALDI-MS-based approach was reported by Yukihira and co-workers which enables high sensitivity and high throughput metabolite analysis for intracellular metabolic dynamics [75]. This high throughput method potentially enables large-scale analysis that deals with tens of thousands of samples and even real-time monitoring of intracellular metabolism. The recent MS advances even enable the analysis of a single yeast cell [77,78]. Cells were first deposited on hydrophilic reservoirs in a microarray chip and ionized by laser irradiation. The concentrations of central intermediates of energy metabolism were measured by MALDI-MS.

4   Conclusion

In its relatively short history, synthetic biology has made great progress in accelerating the design-build-test cycle. A number of advanced tools in each process of the cycle have sketched a preliminary blueprint for future automated biofabrication. Some steps have been already automated for a few specific types of biosystems. These steps include but are not limited to DNA assembly design and construct verification. Steps like pathway design, genetic circuit design, DNA assembly, and metabolite analysis are also partially automated. Given the short time span all these advances happened, we can be confident to foresee that a "prototype-on-demand" capability will become available in the near future.

The future for high throughput synthetic biology is promising yet challenging. In the design process, genetic circuit can be designed and simulated in a CAD environment, but the simulation fidelity will drop with increasing circuit complexity due to the incomplete characterization data of parts. The synthetic routes recommended by retro-biosynthesis algorithms are still selected in an artisanal approach. Once a route is selected, a set of enzymes that is most likely to catalyze the reactions are usually handpicked. Although there are tools available to assist the selection based on structure comparison against the databases, the total possibilities are still vast [29]. The assembly methods for building DNA molecules need higher efficiency and fidelity for robust automation. The modularity needs to be further improved for easier design. Most analytical methods for phenotyping lack multiplexing capacity to achieve desired throughput. Lastly, integrating existing technologies in a fully automated biofabrication facility still requires extensive engineering effort to follow up.

Cheerfully, significant resources are being invested from both governments and industries around the world that should take synthetic biology to a new era. Among the pioneers are the Defense Advanced Research Projects Agency (DARPA) in the USA and Biotechnology and Biological Sciences Research Council (BBSRC) in UK. With an increasing number of researchers attempting to develop high throughput synthetic biology tools, we believe that establishing industrialized "biological foundries" is just a matter of time.

1   Cameron DE, Bashor CJ, Collins JJ. A brief history of synthetic biology. Nat Rev Microbiol, 2014, 12: 381–390
2   Chao R, Yuan Y, Zhao H. Recent advances in DNA assembly technologies. FEMS Yeast Res, 2015, 15: 1–9
3   Shetty RP, Endy D, Knight TF, Jr. Engineering biobrick vectors from biobrick parts. J Biol Eng, 2008, 2: 5
4   Partregistry. Registry of standard biological parts. 2014, http://www.webcitation.org/6O8Ha2b2B
5   Elowitz MB, Levine AJ, Siggia ED, Swain PS. Stochastic gene expression in a single cell. Science, 2002, 297: 1183–1186
6   Blake WJ, KÆrn M, Cantor CR, Collins JJ. Noise in eukaryotic gene expression. Nature, 2003, 422: 633–637
7   Luo YZ, Huang H, Liang J, Wang M, Lu L, Shao ZY, Cobb RE, Zhao HM. Activation and characterization of a cryptic polycyclic tetramate macrolactam biosynthetic gene cluster. Nat Commun, 2013, 4: 2894
8   Yamanaka K, Reynolds KA, Kersten RD, Ryan KS, Gonzalez DJ, Nizet V, Dorrestein PC, Moore BS. Direct cloning and refactoring of a silent lipopeptide biosynthetic gene cluster yields the antibiotic taromycin a. Proc Natl Acad Sci USA, 2014, 111: 1957–1962
9   Warner JR, Reeder PJ, Karimpour-Fard A, Woodruff LBA, Gill RT. Rapid profiling of a microbial genome using mixtures of barcoded oligonucleotides. Nat Biotechnol, 2010, 28: 856–862
10   Cong L, Ran FA, Cox D, Lin SL, Barretto R, Habib N, Hsu PD, Wu XB, Jiang WY, Marraffini LA, Zhang F. Multiplex genome engineering using CRISPR/Cas systems. Science, 2013, 339:

819–823

11   Paddon CJ, Westfall PJ, Pitera DJ, Benjamin K, Fisher K, McPhee D, Leavell MD, Tai A, Main A, Eng D, Polichuk DR, Teoh KH, Reed DW, Treynor T, Lenihan J, Fleck M, Bajad S, Dang G, Dengrove D, Diola D, Dorin G, Ellens KW, Fickes S, Galazzo J, Gaucher SP, Geistlinger T, Henry R, Hepp M, Horning T, Iqbal T, Jiang H, Kizer L, Lieu B, Melis D, Moss N, Regentin R, Secrest S, Tsuruta H, Vazquez R, Westblade LF, Xu L, Yu M, Zhang Y, Zhao L, Lievense J, Covello PS, Keasling JD, Reiling KK, Renninger NS, Newman JD. High-level semi-synthetic production of the potent antimalarial artemisinin. Nature, 2013, 496: 528–532

12   Holtz WJ, Keasling JD. Engineering static and dynamic control of synthetic pathways. Cell, 2010, 140: 19–23

13   Anderson JC, Clarke EJ, Arkin AP, Voigt CA. Environmentally controlled invasion of cancer cells by engineered bacteria. J Mol Biol, 2006, 355: 619–627

14   Zhang F, Carothers JM, Keasling JD. Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids. Nat Biotechnol, 2012, 30: 354–359

15   Wang HH, Isaacs FJ, Carr PA, Sun ZZ, Xu G, Forest CR, Church GM. Programming cells by multiplex genome engineering and accelerated evolution. Nature, 2009, 460: 894–898

16   Jiang WY, Bikard D, Cox D, Zhang F, Marraffini LA. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. Nat Biotechnol, 2013, 31: 233–239

17   Bao Z, Xiao H, Liang J, Zhang L, Xiong X, Sun N, Si T, Zhao H. Homology-integrated CRISPR-Cas (HI-CRISPR) system for one-step multigene disruption in *Saccharomyces cerevisiae*. ACS Synth Biol, 2014, doi: 10.1021/sb500255k

18   Gardner TS, Cantor CR, Collins JJ. Construction of a genetic toggle switch in *Escherichia coli*. Nature, 2000, 403: 339–342

19   Elowitz MB, Leibler S. A synthetic oscillatory network of transcriptional regulators. Nature, 2000, 403: 335–338

20   Guet CC, Elowitz MB, Hsing W, Leibler S. Combinatorial synthesis of genetic networks. Science, 2002, 296: 1466–1470

21   Baker D, Group BF, Church G, Collins J, Endy D, Jacobson J, Keasling J, Modrich P, Smolke C, Weiss R. Engineering life: building a fab for biology. Sci Am, 2006, 294: 44–51

22   Kim B, Du J, Eriksen DT, Zhao HM. Combinatorial design of a highly efficient xylose-utilizing pathway in *Saccharomyces cerevisiae* for the production of cellulosic biofuels. Appl Environ Microb, 2013, 79: 931–941

23   Shao ZY, Zhao H, Zhao HM. DNA assembler, an *in vivo* genetic method for rapid construction of biochemical pathways. Nucleic Acids Res, 2009, 37: e16

24   Salis HM, Mirsky EA, Voigt CA. Automated design of synthetic ribosome binding sites to control protein expression. Nat Biotechnol, 2009, 27: 946–950

25   Curran KA, Crook NC, Karim AS, Gupta A, Wagman AM, Alper HS. Design of synthetic yeast promoters via tuning of nucleosome architecture. Nat Commun, 2014, 5: 4002

26   MacDonald JT, Barnes C, Kitney RI, Freemont PS, Stan GB. Computational design approaches and tools for synthetic biology. Integr Biol (Camb), 2011, 3: 97–108

27   Brophy JA, Voigt CA. Principles of genetic circuit design. Nat Methods, 2014, 11: 508–520

28   Lewis NE, Nagarajan H, Palsson BO. Constraining the metabolic genotype-phenotype relationship using a phylogeny of in silico methods. Nat Rev Microbiol, 2012, 10: 291–305

29   Prather KL, Martin CH. *De novo* biosynthetic pathways: rational design of microbial chemical factories. Curr Opin Biotechnol, 2008, 19: 468–474

30   Kanehisa M, Goto S, Kawashima S, Okuno Y, Hattori M. The kegg resource for deciphering the genome. Nucleic Acids Res, 2004, 32: D277–280

31   Ellis LBM, Roe D, Wackett LP. The University of Minnesota biocatalysis/biodegradation database: the first decade. Nucleic Acids Res, 2006, 34: D517–521

32   Schomburg I, Chang A, Schomburg D. Standardization in enzymology—data integration in the world's enzyme information system brenda. Persp Sci, 2014, 1: 15–23

33   Hatzimanikatis V, Li C, Ionita JA, Henry CS, Jankowski MD, Broadbelt LJ. Exploring the diversity of complex metabolic networks. Bioinformatics, 2005, 21: 1603–1609

34   Gonzalez-Lergier J, Broadbelt LJ, Hatzimanikatis V. Theoretical considerations and computational analysis of the complexity in polyketide synthesis pathways. J Am Chem Soc, 2005, 127: 9930–9938

35   Hou BK, Ellis LB, Wackett LP. Encoding microbial metabolic logic: predicting biodegradation. J Ind Microbiol Biotechnol, 2004, 31: 261–272

36   Lu G, Moriyama EN. Vector NTI, a balanced all-in-one sequence analysis suite. Brief Bioinform, 2004, 5: 378–388

37   Hillson NJ, Rosengarten RD, Keasling JD. J5 DNA assembly design automation software. ACS Synth Biol, 2012, 1: 14–21

38   Appleton E, Tao JH, Haddock T, Densmore D. Interactive assembly algorithms for molecular cloning. Nat Methods, 2014, 11: 657–662

39   Ellis T, Adie T, Baldwin GS. DNA assembly for synthetic biology: from parts to pathways and beyond. Integr Biol (Camb), 2011, 3: 109–118

40   McLennan A. Building with Biobricks: Constructing a Commons for Synthetic Biology Research. Cheltenham: Edward Elgar, 2012. 176–201

41   Grünberg R, Arndt K, Müller K. Fusion protein (freiburg) biobrick assembly standard [OL]. [2009-04-18]. http://hdl.handle.net/ 1721.1/ 45140

42   Phillips I, Silver P. A new biobrick assembly strategy designed for facile protein engineering [OL]. [2006-04-20]. http://hdl.handle.net/ 1721.1/32535

43   Anderson JC, Dueber JE, Leguia M, Wu GC, Goler JA, Arkin AP, Keasling JD. Bglbricks: A flexible standard for biological part assembly. J Biol Eng, 2010, 4: 1

44   Engler C, Kandzia R, Marillonnet S. A one pot, one step, precision cloning method with high throughput capability. PLoS One, 2008, 3: e3647

45   Gibson DG, Young L, Chuang RY, Venter JC, Hutchison CA, Smith HO. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Methods, 2009, 6: 343–345

46   Gibson DG, Benders GA, Andrews-Pfannkoch C, Denisova EA, Baden-Tillson H, Zaveri J, Stockwell TB, Brownley A, Thomas DW, Algire MA, Merryman C, Young L, Noskov VN, Glass JI, Venter JC, Hutchison CA, 3rd, Smith HO. Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. Science, 2008, 319: 1215–1220

47   Itaya M, Fujita K, Kuroki A, Tsuge K. Bottom-up genome assembly using the *Bacillus subtilis* genome vector. Nat Methods, 2008, 5: 41–43

48   Yonemura I, Nakada K, Sato A, Hayashi JI, Fujita K, Kaneko S, Itaya M. Direct cloning of full-length mouse mitochondrial DNA using a *Bacillus subtilis* genome vector. Gene, 2007, 391: 171–177

49   Zhu CF, Naqvi S, Breitenbach J, Sandmann G, Christou P, Capell T. Combinatorial genetic transformation generates a library of metabolic phenotypes for the carotenoid pathway in maize. Proc Natl Acad Sci USA, 2008, 105: 18232–18237

50   Farre G, Naqvi S, Sanahuja G, Bai C, Zorrilla-Lopez U, Rivera SM, Canela R, Sandman G, Twyman RM, Capell T, Zhu CF, Christou P. Combinatorial genetic transformation of cereals and the creation of metabolic libraries for the carotenoid pathway. Trans Plants Methods Mol Biol, 2012, 847: 419–435

51   Zhang YM, Muyrers JPP, Testa G, Stewart AF. DNA cloning by homologous recombination in *Escherichia coli*. Nat Biotechnol, 2000, 18: 1314–1317

52   Fu J, Bian XY, Hu SB, Wang HL, Huang F, Seibert PM, Plaza A, Xia LQ, Muller R, Stewart AF, Zhang YM. Full-length rece enhances linear-linear homologous recombination and facilitates direct cloning for bioprospecting. Nat Biotechnol, 2012, 30: 440–446

53   Pachuk CJ, Samuel M, Zurawski JA, Snyder L, Phillips P, Satishchandran C. Chain reaction cloning: a one-step method for

directional ligation of multiple DNA fragments. Gene, 2000, 243: 19–25

54   De Kok S, Stanton LH, Slaby T, Durot M, Holmes VF, Patel KG, Platt D, Shapland EB, Serber Z, Dean J, Newman JD, Chandran SS. Rapid and reliable DNA assembly via ligase cycling reaction. Acs Synth Biol, 2014, 3: 97–106

55   Wingler LM, Cornish VW. Reiterative recombination for the *in vivo* assembly of libraries of multigene pathways. Proc Natl Acad Sci USA, 2011, 108: 15135–15140

56   Anderson PR, Haj-Ahmad Y. Counter-selection facilitated plasmid construction by homologous recombination in saccharomyces cerevisiae. Biotechniques, 2003, 35: 692–694

57   Kuijpers NG, Solis-Escalante D, Bosman L, van den Broek M, Pronk JT, Daran JM, Daran-Lapujade P. A versatile, efficient strategy for assembly of multi-fragment expression vectors in saccharomyces cerevisiae using 60 bp synthetic recombination sequences. Microb Cell Fact, 2013, 12: 47

58   Liang J, Chao R, Abil Z, Bao Z, Zhao H. Fairytale: a high-throughput tal effector synthesis platform. ACS Synth Biol, 2014, 3: 67–73

59   Guye P, Li Y, Wroblewska L, Duportet X, Weiss R. Rapid, modular and reliable construction of complex mammalian gene circuits. Nucleic Acids Res, 2013, 41: e156

60   Torella JP, Boehm CR, Lienert F, Chen JH, Way JC, Silver PA. Rapid construction of insulated genetic circuits via synthetic sequence-guided isothermal assembly. Nucleic Acids Res, 2014, 42: 681–689

61   Casini A, MacDonald JT, De Jonghe J, Christodoulou G, Freemont PS, Baldwin GS, Ellis T. One-pot DNA construction for synthetic biology: the modular overlap-directed assembly with linkers (modal) strategy. Nucleic Acids Res, 2014, 42: e7

62   Kosuri S, Church GM. Large-scale *de novo* DNA synthesis: technologies and applications. Nat Methods, 2014, 11: 499–507

63   Gibson DG, Benders GA, Andrews-Pfannkoch C, Denisova EA, Baden-Tillson H, Zaveri J, Stockwell TB, Brownley A, Thomas DW, Algire MA, Merryman C, Young L, Noskov VN, Glass JI, Venter JC, Hutchison CA, Smith HO. Complete chemical synthesis, assembly, and cloning of a mycoplasma genitalium genome. Science, 2008, 319: 1215–1220

64   Gibson DG, Glass JI, Lartigue C, Noskov VN, Chuang RY, Algire MA, Benders GA, Montague MG, Ma L, Moodie MM, Merryman C, Vashee S, Krishnakumar R, Assad-Garcia N, Andrews-Pfannkoch C, Denisova EA, Young L, Qi ZQ, Segall-Shapiro TH, Calvey CH, Parmar PP, Hutchison CA, Smith HO, Venter JC. Creation of a bacterial cell controlled by a chemically synthesized genome. Science, 2010, 329: 52–56

65   Annaluru N, Muller H, Mitchell LA, Ramalingam S, Stracquadanio G, Richardson SM, Dymond JS, Kuang Z, Scheifele LZ, Cooper EM, Cai Y, Zeller K, Agmon N, Han JS, Hadjithomas M, Tullman J, Caravelli K, Cirelli K, Guo Z, London V, Yeluru A, Murugan S, Kandavelou K, Agier N, Fischer G, Yang K, Martin JA, Bilgel M,

Bohutski P, Boulier KM, Capaldo BJ, Chang J, Charoen K, Choi WJ, Deng P, DiCarlo JE, Doong J, Dunn J, Feinberg JI, Fernandez C, Floria CE, Gladowski D, Hadidi P, Ishizuka I, Jabbari J, Lau CY, Lee PA, Li S, Lin D, Linder ME, Ling J, Liu J, London M, Ma H, Mao J, McDade JE, McMillan A, Moore AM, Oh WC, Ouyang Y, Patel R, Paul M, Paulsen LC, Qiu J, Rhee A, Rubashkin MG, Soh IY, Sotuyo NE, Srinivas V, Suarez A, Wong A, Wong R, Xie WR, Xu Y, Yu AT, Koszul R, Bader JS, Boeke JD, Chandrasegaran S. Total synthesis of a functional designer eukaryotic chromosome. Science, 2014, 344: 55–58

66   Dharmadi Y, Patel K, Shapland E, Hollis D, Slaby T, Klinkner N, Dean J, Chandran SS. High-throughput, cost-effective verification of structural DNA assembly. Nucleic Acids Res, 2014, 42: e22

67   Metzker ML. Sequencing technologies—the next generation. Nat Rev Genet, 2010, 11: 31–46

68   Coen M, Holmes E, Lindon JC, Nicholson JK. NMR-based metabolic profiling and metabonomic approaches to problems in molecular toxicology. Chem Res Toxicol, 2008, 21: 9–27

69   Fiehn O. Extending the breadth of metabolite profiling by gas chromatography coupled to mass spectrometry. Trac-Trend Anal Chem, 2008, 27: 261–269

70   Khakimov B, Motawia MS, Bak S, Engelsen SB. The use of trimethylsilyl cyanide derivatization for robust and broad-spectrum high-throughput gas chromatography-mass spectrometry based metabolomics. Anal Bioanal Chem, 2013, 405: 9193–9205

71   Allwood JW, Goodacre R. An introduction to liquid chromatography-mass spectrometry instrumentation applied in plant metabolomic analyses. Phytochem Anal, 2010, 21: 33–47

72   Mischak H, Coon JJ, Novak J, Weissinger EM, Schanstra JP, Dominiczak AF. Capillary electrophoresis-mass spectrometry as a powerful tool in biomarker discovery and clinical diagnosis: an update of recent developments. Mass Spectrom Rev, 2009, 28: 703–724

73   Lapainis T, Rubakhin SS, Sweedler JV. Capillary electrophoresis with electrospray ionization mass spectrometric detection for single-cell metabolomics. Anal Chem, 2009, 81: 5858–5864

74   Khakimov B, Bak S, Engelsen SB. High-throughput cereal metabolomics: current analytical technologies, challenges and perspectives. J Cereal Sci, 2014, 59: 393–418

75   Yukihira D, Miura D, Saito K, Takahashi K, Wariishi H. MALDI-MS-based high-throughput metabolite analysis for intracellular metabolic dynamics. Anal Chem, 2010, 82: 4278–4282

76   Vaidyanathan S, Goodacre R. Quantitative detection of metabolites using matrix-assisted laser desorption/ionization mass spectrometry with 9-aminoacridine as the matrix. Rapid Commun Mass Spectrom, 2007, 21: 2072–2078

77   Yanes O. Metabolomics playing pinata with single cells. Nat Chem Biol, 2013, 9: 471–473

78   Rubakhin SS, Romanova EV, Nemes P, Sweedler JV. Profiling metabolites and peptides in single cells. Nat Methods, 2011, 8: S20–29