

## Special Issue: Industrial Biotechnology

# **Opinion** Multiplexed Engineering in Biology

Jameson K. Rogers<sup>1</sup> and George M. Church<sup>1,\*</sup>

Biotechnology is the manufacturing technology of the future. However, engineering biology is complex, and many possible genetic designs must be evaluated to find cells that produce high levels of a desired drug or chemical. Recent advances have enabled the design and construction of billions of genetic variants per day, but evaluation capacity remains limited to thousands of variants per day. Here we evaluate biological engineering through the lens of the design-build-test cycle framework and highlight the role that multiplexing has had in transforming the design and build steps. We describe a multiplexed solution to the 'test' step that is enabled by new research. Achieving a multiplexed test step will permit a fully multiplexed engineering cycle and boost the throughput of biobased product development by up to a millionfold.

## Biomanufacturing is the Manufacturing of the Future

We are at the brink of a new era in which biology is harnessed to produce valuable new chemicals and materials. Even today, the US bioeconomy produces revenue upwards of US \$350 billion each year, with more than US\$100 billion of that figure coming from bioderived fuels and chemicals [1,2]. This is only the beginning: with cells as the chemical factories of the future, industry will no longer be restricted to compounds that are readily produced from petrochemical building blocks [3]. However, to reach this future, the long and uncertain timelines for biobased product development must be overcome.

Engineering cells for chemical production is challenging because the complexity of biology often necessitates that many designs be attempted before an optimal combination of genetic elements is discovered. The engineering process in which these designs are evaluated and iterated on is the design—build—test cycle (Figure 1 and Box 1). The rate of product development is directly related to the throughput of the design cycle, with higher throughputs resulting in reduced development times. Over the past few years, a convergence of technologies has enabled the simultaneous construction of billions of cellular variants. Each of these cellular variants tests a unique hypothesis regarding the combination of genetics elements that would result in the optimal production of a target compound. While the throughput of the design and build steps of the cycle have advanced astronomically, our ability to evaluate those designs is typically of the order of hundreds or thousands of samples per day [4] (Figure 2). As a result, the vast majority of designs go unevaluated. The impact of further advances in the design and construction of genetic variants will continue to be diminished while the evaluation bottleneck persists.

Recent advances in high-throughput evaluation of metabolic phenotypes have put a generalizable method for achieving evaluation rates of up to 1 billion cells per day within reach, providing a millionfold increase in design cycle throughput [5,6] (Figure 3). Such a dramatic leap in engineering capacity is accomplished by using genetically encoded **biosensors** (see Glossary) that enable cells to monitor their own success in producing a target compound [6,7]. Depending on

### Trends

Biotechnological development proceeds through the design-build-test cycle.

Multiplexed technologies have enabled design and build steps to achieve throughputs of up to 1 billion variants per day, far outpacing our capacity to evaluate designs.

A multiplexed strategy to evaluate designs could boost the throughput of biotechnological development by up to a millionfold.

<sup>1</sup>Wyss Institute for Biologically Inspired Engineering, Harvard University, 3 Blackfan Circle, Boston, MA 02115, USA

\*Correspondence: gchurch@genetics.med.harvard.edu (G.M. Church).





#### Trends in Biotechnology

Figure 1. The Design–Build–Test Cycle for Biological Engineering. The genes necessary for chemical production are initially placed within an organism that has not been optimized for chemical production. This microbe typically exhibits low or no productivity. The first pass through the design step of the cycle may involve varying the levels of gene expression or exploration of mutations in the enzyme active sites. These designs are implemented through DNA synthesis and subsequent genome engineering during the build step. In the test step, the newly constructed organisms are evaluated for their ability to produce the desired compound. The most productive cells are retained and used as a starting point for the next round of design. The cycle is iterated until an organism is found that meets the stated productivity requirements.

the biosensor implementation, the most productive cells either fluoresce more brightly or produce higher levels of an antibiotic-resistance gene product that allows them to outcompete their neighbors in the presence of the antibiotic.

Achieving **multiplexed** phenotype evaluation is vital because it is the final element necessary for unlocking a fully multiplexed design–build–test cycle. Here we evaluate how the design–build–test cycle has evolved over the past several years and how biological engineers can most effectively harness multiplexed phenotype evaluation. We also explore how a fully multiplexed design cycle can shift the mindset of bioengineers such that they can think in terms of whole design spaces rather than single design instances. We are excited about the next generation of biological engineering that multiplexed evaluation methods have finally enabled.

## Multiplexing Enables Next-Generation Biological Engineering

The rate of biotechnological innovation and product development is dependent on the throughput of the biological design-build-test cycle. Design cycle throughput is the product of two components: cycle speed and cycle bandwidth. Cycle speed measures how quickly each iteration of the design-build-test cycle can be completed. Cycle bandwidth reflects the number of designs that are evaluated per iteration of the cycle. Throughput is therefore the product of cycle speed and cycle bandwidth.

While either of the throughput components of the design cycle may be improved to enhance overall cycle throughput, the most impressive gains are achieved by increasing cycle bandwidth through multiplexing. The speed of the build and test steps are limited by the rate at which cells can be grown and manipulated. In contrast to speed, the physical limits on bandwidth are vast because billions of cells or trillions of DNA molecules fit in a single droplet. If each cell or molecule contains a unique design, the designs are multiplexed in space. By contrast, parallel

### Glossary

Biosensors: genetically encoded devices that monitor the intracellular concentration of a specific compound. Biosensors produce fluorescence or another readout proportional to the concentration of that compound within the cell. **Demultiplexing:** refers to the process of separating a mixture of elements into physically separate locations. This is necessary when shifting from a multiplexed process to a **singleplex** process.

Directed evolution: shifts a population of cells from one genetic state to a more desirable genetic state through a series of selections or screens performed on successive generations. Each round of a directed-evolution experiment involves subjecting a population to genetic diversification followed by enrichment for desired mutations. Fluorescence-activated cell

sorting (FACS): a high-throughput

method for screening cells based on fluorescence. Thousands of cells per second pass through a flow cell and are retained or discarded based on their fluorescence at a certain

excitation and emission wavelength. **Forward engineering:** in biology, forward engineering refers to the process of designing biological systems or parts such that they accomplish high-level design goals when implemented. Forward engineering relies on accurate models describing how DNA sequence is transduced to phenotype. This level of knowledge is often lacking, necessitating an iterative engineering process.

Multiplex: in biology, a multiplexed process operates on many distinct elements (e.g., cells, DNA molecules, metabolites) that coexist in space and time. Multiplexing enables a single process to work on millions of elements with the same effort that would be required to perform the process on a single element. Screen: in metabolic engineering, a screen is a method for identifying highly productive cells. Cells are evaluated one by one to determine which is most productive. Less productive cells are discarded. Evaluation methods are numerous and diverse.

**Selection:** in metabolic engineering, a selection is a method for identifying highly productive cells. Cells are

### Box 1. The Biological Design-Build-Test Cycle

It is no surprise that engineering biological systems is a challenging process. Product development requires the creation of a life form that behaves in a predictable and reproducible way, a process in which traditional engineering paradigms tend to be inadequate. Each component of a biological system interacts with thousands of other components within the cell. This is in stark contrast to electrical or mechanical engineering, where a given component interacts with just a handful of adjacent components. The astounding complexity that results from the high connectivity of biological systems is further exacerbated by poor characterization of components at the individual level. Consequently, finding a solution that maintains cellular viability while meeting desired design goals requires many iterations of the engineering process.

Each iteration is broken into three steps.

### Step 1: Biological Design

Hypotheses about the sequence-phenotype relationship dictate which genetic elements should be modified. Software tools such as Rosetta and flux balance analysis help guide enzyme design and metabolic network design, respectively. Uncertainty in the sequence-phenotype relationship results in many possible designs for a single desired outcome.

#### Step 2: Genetic Construction

DNA that corresponds to each new design is synthesized and installed within the cell. Different DNA-synthesis strategies are deployed based on the type of genetic element that is being modified. The new genetic material is installed on a plasmid or the genome itself. The installation process can proceed through patching of old genetic material, complete overwriting of existing sequence, or addition of entirely new genes.

### Step 3: Phenotype Evaluation

The newly constructed organisms are evaluated for their ability to meet design goals. In biomanufacturing applications, the new cells are assessed by measuring how much of a desired compound they produce. High-performing cells are selected as the starting point for the next iteration of the design cycle.

experimentation requires the spatial separation of designs. This places a limit on the number of designs that can be processed due to constraints of space and the logistics of design manipulation.

An iteration through the design-build-test cycle is used to learn about the system at hand and provides the basis for subsequent design iterations. The simplest feedback between the test and design steps occurs as it does in nature: the most successful designs survive to become the templates for the next iteration of the cycle and the design space converges to a local maximum. More advanced feedback from the test step to the design step provides an explicit understanding of which designs function best. Multiplexed sequencing of each design followed by sequence comparison between the successful and unsuccessful designs enables the formulation of design rules that more quickly define the design space and inform subsequent engineering endeavors [8,9]. This is a form of **directed evolution** in which an engineer monitors the flow of designs and intervenes as necessary. However, such a process requires the technological capability to quickly rank and identify the best designs and to build new designs with enough speed and precision to actuate the knowledge gained. Consequently, the effectiveness of the design cycle increases substantially when multiplexing is possible.

Every step of the cycle must be multiplexed to achieve a fully multiplexed design cycle because the throughput of the full design cycle is limited by the throughput of the slowest step. While innovation in a single step results in higher step throughput, it also changes how adjacent steps are approached. Widely available gene synthesis is an example of how an innovation in the build step freed design-step engineers from the constraints of using existing DNA sequences. A more dramatic leap in how engineers approach biological design will occur once the cycle is multiplexed from start to finish.

### The Design Step: Towards Forward Engineering in Biology

A perfect design step would obviate the other steps in the design cycle. Some engineering disciplines have approached this scenario; in civil engineering, for example, it is rare to see a

exposed to an environment where productive cells have a growth advantage over less productive cells. **Singleplex:** in biology, a singleplex process operates on elements (e.g., one cell, one DNA sequence, a single chemical) separated by location or time. Singleplex operations are the opposite of multiplexed operations. Also referred to as parallel operations.





## Design throughput

Rosetta protein design	+++	
Active site mutagenesis	+++	
Computational strain design	+++	
Pathway prediction	+++	
Cis-regulatory modulation	+++	







## Build throughput

Multiplexed gene assembly	10 <sup>6</sup>	<b>→</b>	
Degenerate gene synthesis	10 <sup>7</sup>		Multiplexed gene variants
Multiplexed oligo synthesis	10 <sup>4</sup>	<b>→</b>	الالمالية المالية المال Multiplexed oligo synthesis
Error prone PCR	10 <sup>7</sup>		
Multiplexed genome engineering	10 <sup>10</sup>	→	engineering
Test throughput			
Chromatography	10 <sup>3</sup>	→	
Mass spectrometry	10 <sup>3</sup>		
Colorimetric assays	10 <sup>5</sup>	<b>→</b>	

### Trends in Biotechnology

Figure 2. Available Technologies and Corresponding Throughput at Each Stage of the Design Cycle. Colored circles illustrate throughput: test throughput in red, build throughput in green, and design throughput in blue. Design throughput refers to the universe of potential designs that an engineer may want to explore and is potentially unlimited in scope. In practice, however, an upper limit on design space complexity is influenced by the number of designs that can be built and evaluated in the subsequent steps of the cycle. Recent technologies have increased build throughput by several orders of magnitude. Multiplexed DNA synthesis has opened the door to very large libraries of individually designed sequences while multiplexed genome engineering has enabled vast numbers of genomic edits to be made rapidly. Test throughput remains the primary bottleneck in the design cycle. Available technologies require physical separation of each design, conflicting with the multiplexed technologies of the build step. Protein illustration used under Creative Commons License with attribution to David S. Goodsell and the RCSB PDB.

bridge constructed repeatedly until it works as desired because the design step is so highly refined. However, our knowledge of biological design principles is imperfect. For example, sequence determinants of transcription are not fully characterized, epigenetic regulation remains fuzzy, codon usage is often cryptic, and the stability of folded protein products is still mysterious. For these reasons, the accuracy of the biological design step lags behind that of more traditional engineering disciplines and necessitates continued innovation.

Fluorescence based screening



Antibiotic based selections



### Trends in Biotechnology

Figure 3. Multiplexed Phenotype Evaluation Eliminates the Primary Bottleneck of the Biological Design-Build-Test Cycle. Biosensor-based phenotype evaluation is implemented through screens or selections. Fluorescencebased screening relies on linking the chemical productivity of a cell to the expression of a fluorescent reporter protein. Cells that produce more of the product compound fluoresce more brightly and are quickly separated from less productive cells through flow cytometry at rates of up to 1 million cells per minute. Antibiotic-based selections are enabled by linking the chemical productivity of a cell to the expression of an antibiotic-resistance gene. Cells that produce more of the product compound are able to survive when exposed to the antibiotic while less productive cells die. Using selections, billions of cells can be evaluated simultaneously. Such high throughput in the test step enables a fully multiplexed design cycle and enables engineers to take on previously insurmountable design challenges.



The effectiveness of the design step has risen meteorically with DNA sequencing technology. In the burgeoning era of biotechnology, locating or amplifying a gene was hindered by a lack of sequence knowledge. This made even simple design endeavors, such as recombinant protein expression, an arduous task. Once a gene's sequence was known, exploration of new biological designs was restricted to random walks in adjacent sequence space because hypotheses about functional regions had yet to be formulated. Later, as the amount of sequenced DNA exploded, so did the potential for designing novel biological systems. The vast repertoire of sequenced genomes has enabled the formulation of design rules, in turn enabling engineers to hone in on the active sites of enzymes, borrow homologous sequences from distant species, and locate regulatory elements for transcription and translation.

Modern-day design tools build on these sequence–function relationships and enable the **forward engineering** of biological systems. Precise prediction of ribosomal binding site strength and promoter activity is now possible with tools like the Salis RBS calculator [9–11]. New proteins can be constructed *in silico* before being implemented *in vivo* [12]. The metabolism of entire organisms can be modeled mathematically, revealing which genes should be turned up or down to accomplish a given metabolic goal [13–15].

While these tools provide guidelines for designing new functionality within organisms, their design predictions still fall within the ambiguity of biology. Consequently, many designs must be evaluated before an optimal result is obtained. Take the modification of an enzyme's substrate specificity as an example. If our design tools allow us to identify an active site of seven amino acids, that active site would exist in a design space of 1 billion potential active sites. If we know that those amino acids should be positively charged, our design space would comprise just 2000 possible proteins. But how do engineers construct such a large number of designs?

### The Build Step: Billions of Designs per Day

The build step of the design cycle describes the process in which potential designs are constructed out of DNA and integrated into a cell. The cost of gene synthesis has decreased to the point where ordering several genes (e.g., GeneArt Strings, IDT gBlocks) is trivial for most laboratories [16,17]. Combined with technologies that enable seamless plasmid construction (e. g., Gibson [18] and Golden Gate assemblies [19]) and simple methods for modifying the genome (lambda red recombineering [20] and multiplexed automated genome engineering [21]), parallel construction is a robust process. However, achieving a meaningful increase in build bandwidth by multiplexing with synthesized genes will remain cost prohibitive as long as the construction of those genes is a parallel process itself. Thus, engineers are limited to either random or sitedirected mutagenesis to achieve multiplexed construction of genetic elements. This has been a serious constraint for building many designs in a multiplexed manner. Recently, microarraybased oligonucleotide synthesis has enabled multiplexed construction of precisely designed sequences of up to 200 base pairs, allowing the evaluation of hundreds of thousands of complex biological hypotheses simultaneously [8,9,22]. Modification of gene expression in trans provides additional versatility for the build step. Both small regulatory RNAs [23] and CRISPR interference [24] can be multiplexed to rapidly fine-tune gene expression. While genome engineering has the benefit of providing permanent changes to metabolism, multiplexing in trans enables further leverage of recent advances in multiplexed oligo synthesis. Regardless of the specific method, the capability to multiplex the build step is transforming how engineers approach the design cycle: experiments that were previously infeasible are now within reach.

Simultaneous advances in genome engineering have made the construction of billions of genomic variants a routine process [6,21,25–27]. Multiplexed genome engineering allows specified or degenerate mutations to be targeted anywhere in the genome [21]. Such facile genome engineering enables new classes of experiments. As an example, the entire set of



metabolic modifications suggested by *in silico* analytical techniques such as flux balance analysis can now be explored simultaneously. When optimizing the production of a target compound, this type of analysis identifies genes that are important to modulate but does not accurately specify what their level of expression should be [13]. Multiplexed genome engineering enables the combinatorial exploration of gene expression levels for each of the genes of interest [6,21]. If ten genes are targeted with mutations corresponding to ten levels of expression, the resulting design space comprises 10 billion genomes.

## The Test Step: Evaluation Rates Still Lag Behind

Despite the success in multiplexing the design and build steps, evaluating 10 billion designs with current technology would take decades because the test step of the engineering cycle remains a parallel process. Biological designs constructed in multiplex must first be **demultiplexed** before analysis, negating the value of multiplexed construction in the first place. One reason phenotype evaluation lacks an adequate multiplexed solution is that analytical methods differ dramatically for different phenotypes. When cells are engineered to produce fuels or chemicals, design success is often determined by the amount of compound produced. To measure this concentration using chromatography or mass spectrometry, cells must be separated into individual designs (the demultiplexing step) and cultured in parallel in small volumes, such as in 96-well plates. Next, either the supernatant or cell lysates are prepared such that the concentration of the molecule of interest can be determined [4]. In this workflow, throughput is typically limited to hundreds or thousands of design evaluations per day [4].

## Enabling a Fully Multiplexed Design Cycle

Allowing cells to report their own progress in making a specific chemical enables a multiplexed solution to the test step. Rather than assaying individual designs, engineers should be able to define a design goal and immediately isolate cells that meet a specified level of performance. If cells keep track of their own progress, the time required for separation of productive cells from unproductive cells is no longer proportional to the number of cells evaluated. **Selections** are an example of such a multiplexed evaluation method. In a selection, only cells that have a certain phenotype survive. This decouples the time required for evaluation from the number of cells evaluated. However, selections are typically based on an *ad hoc* link between a phenotype of interest and a necessary cell function. For instance, selecting for increased utilization of a new sugar is possible if all other energy sources are withheld, but selecting for the increased production of a novel chemical is not so simple. A general method for multiplexed phenotype evaluation is the last step required for a fully multiplexed design–build–test cycle.

One such method is based on genetically encoded biosensors. Biosensors provide a general framework for linking intracellular chemical concentration to cell function and provide a generalizable method for multiplexing design evaluation in cases of metabolic engineering. Allosterically regulated transcription factors, riboswitches, domain-inserted proteins, and ligand-dependent dimerization or stability schemes are all methods of genetically encoding biosensors. Biosensors based on allosteric transcription factors allow expression of a target gene when bound by a specific small molecule. Transcription factors cluster into more than 20 major families [28]. Currently, the *lacl* family contains 29 000 sequenced members, while the *gntR* family contains 49 000 members [29]. There are over 200 000 sequences available for members of the *tetR* family of transcriptional repressors [28]. These naturally occurring transcription factors bind to an impressive range of compounds. If a microbe has an incentive to either consume or avoid a compound, there is likely to be a transcription factor that has evolved to bind it. Recent advances in protein engineering and directed evolution have produced designer transcription factors that bind compounds for which natural transcription factors have yet to be discovered [30,31].

Biosensors based on allosterically regulated transcription factors enable **screens** and selections for a vast repertoire of compounds by providing a transcriptional readout for intracellular metabolite concentration. When the transcriptional output of the biosensor is an antibioticresistance gene, biosensor activation confers antibiotic resistance. Treating a population of cells with the appropriate antibiotic allows cells to survive only if they produce the required amount of product. Alternatively, if the transcriptional readout of the biosensor is a fluorescent protein, cells with more effective designs will fluoresce more brightly. Fluorescent biosensors enable millions of cells to be screened per minute with high-throughput methods such as **fluorescenceactivated cell sorting (FACS)**.

Combining multiplexed phenotype evaluation with next-generation sequencing enables a deep understanding of the design space being explored [8,9]. Sorting cells into bins based on their fluorescence is a multiplexed method for assigning each cell a rank that is based on the quality of the design it contains. Deep sequencing of the bin contents provides a list of designs for each bin. Ranks are assigned to each design based on which bin they were found in. Each iteration of the design cycle provides millions of design–rank pairs. This wealth of information allows design rules to be developed much more rapidly than would otherwise be possible.

## Concluding Remarks: A New Approach for Metabolic Engineering

Routine incorporation of biosensor-based multiplexing will change how metabolic engineers approach the design cycle. Engineers will explore complete design spaces overnight and the design cycle bottleneck will shift to data analysis and experimental design. Several groups have already adopted biosensor-based multiplexing in metabolic engineering applications. Success has been demonstrated for biosensors implemented as both selections [5,6,32] and screens [5,33–36]. See Table 1 for a summary of biosensor-mediated metabolic engineering outcomes.

Simple design rules have been formulated for constructing and deploying metabolite-responsive biosensors that transform screens and selections from *ad hoc* solutions to well-characterized methodologies [6,7]. These design rules make multiplexing simple for other engineers to implement. Small and large biotechnology firms are beginning to evaluate how these multiplexing technologies will fit into their product development pipelines while academic laboratories are further expanding biosensor-based multiplexing capacity and developing novel applications. Advances in *de novo* biosensor construction further expand the range of compounds for which multiplexed evaluation is possible [37].

An entirely high-throughput design-build-test cycle will allow bioengineers to address design challenges that were previously out of reach. Applications span agricultural products, drop-in replacements for fuels and chemicals, novel chemical products, and previously unattainable

Table	1.	Examples	of	Biosensor-Media	ated	High-	Throughput	Metabolic	Engineering	g
-------	----	----------	----	-----------------	------	-------	------------	-----------	-------------	---

Molecule	Biosensor Mode	Titer Fold Improvement	Throughput	Year	Refs
Naringenin	Selection	36	10 <sup>9</sup>	2014	[6]
Glucarate	Selection	22	10 <sup>9</sup>	2014	[6]
Lysine	Screen	37	10 <sup>8</sup>	2013	[38]
Histidine	Screen	>40	10 <sup>8</sup>	2013	[38]
Arginine	Screen	87	10 <sup>8</sup>	2013	[38]
Triacetate lactone	Screen	20	10 <sup>4</sup>	2013	[36]
Mevalonate	Screen	3.8	10 <sup>5</sup>	2011	[30]
Butanol	Screen	1.4	10 <sup>3</sup>	2010	[4]

### **Outstanding Questions**

Efficient engineering relies on modular systems that can be independently designed and reused in various applications. Such a system allows engineers to specialize in the design of one type of system and collaborate with other engineers who specialize in the design of another. Modular engineering also enables efforts expended for one application to be applied towards subsequent applications. Is it possible to advance the biological design-build-test cycle to the point where each step of the cycle is independent and modular? Will it be possible to have evolutionary engineers (who design selections and screens) who operate independently from metabolic engineers (who design bioproduction systems)?

The initial work that goes into developing the biosensor for a new compound provides an activation energy that will deter some engineers from using a multiplexed approach. What is necessary to reduce the activation energy for biosensor use? Some ideas include: laboratories and companies that focus on custom biosensor development; a large library of predeveloped biosensors; and clear design rules for biosensor construction and characterization.

A fully multiplexed design-build-test cycle enables millions of designs to be evaluated per cycle. Sequencing the entire design pool at each iteration provides an unprecedented amount of data linking phenotypes to DNA sequence. What kind of data-analysis pipeline will be required to develop design rules based on sequence-phenotype relationships? How can this information be used to automatically inform the next iteration of design?



pharmaceuticals. The increase in design cycle throughput that is enabled by multiplexing will embolden bioengineers to go after lofty targets with immediate and global impact.

#### References

- 1. Carlson, R. (2011) Biodesic 2011 Bioeconomy Update, www. biodesic.com/library/Biodesic 2011 Bioeconomy Update.pdf
- 2. Carlson, R. (2014) The U.S. Bioeconomy in 2012 Reached \$350 Billion in Revenues, or about 2.5% of GDP. www.synthesis.cc/ 2014/01/the-us-bioeconomy-in-2012.html
- 3. White House
- 4. Dietrich, J.A. et al. (2010) High-throughput metabolic engineering: advances in small-molecule screening and selection. Annu. Rev. Biochem. 79, 563-590
- 5. Dietrich, J.A. et al. (2013) Transcription factor-based screens and synthetic selections for microbial small-molecule biosynthesis. ACS Synth. Biol. 2, 47-58
- 6. Raman, S. et al. (2014) Evolution-guided optimization of biosynthetic pathways. Proc. Natl. Acad. Sci. U.S.A. 111, 17803–17808
- 7. Rogers, J.K. et al. (2015) Synthetic biosensors for precise gene control and real-time monitoring of metabolites. Nucleic Acids Res. 43, 7648-7660
- 8. Goodman, D.B. et al. (2013) Causes and effects of N-terminal codon bias in bacterial genes. Science 342, 475-479
- 9. Kosuri, S. et al. (2013) Composability of regulatory sequences controlling transcription and translation in Escherichia coli. Proc. Natl. Acad. Sci. U.S.A. 110, 14024-14029
- 10. Salis, H.M. et al. (2009) Automated design of synthetic ribosome binding sites to control protein expression. Nat. Biotechnol. 27. 946-950
- 11. Mutalik, V.K. et al. (2013) Precise and reliable gene expression via standard transcription and translation initiation elements. Nat. Methods 10, 354-360
- 12. Rohl, C.A. et al. (2004) Protein structure prediction using Rosetta. Methods Enzymol, 383, 66-93
- 13. Orth, J.D. et al. (2010) What is flux balance analysis? Nat. Biotechnol. 28, 245-248
- 14. Varma, A. and Palsson, B.O. (1994) Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type Escherichia coli W3110. Appl. Environ. Microbiol. 60, 3724-3731
- 15. Ng, C.Y. et al. (2015) Advances in de novo strain design using integrated systems and synthetic biology tools. Curr. Opin. Chem. Biol. 28, 105-114
- 16. Carr, P.A. and Church, G.M. (2009) Genome engineering. Nat. Biotechnol. 27, 1151-1162
- 17. Carlson, R. (2014) Time for New DNA Synthesis and Sequencing Cost Curves. www.synthesis.cc/2014/02/time-for-new-costcurves-2014.html
- 18. Gibson, D.G. et al. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat. Methods 6, 343-345
- 19. Engler, C. and Marillonnet, S. (2013) Combinatorial DNA assembly using Golden Gate cloning. Methods Mol. Biol. 1073, 141-156

- 20. Datsenko, K.A. and Wanner, B.L. (2000) One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proc. Natl. Acad. Sci. U.S.A. 97, 6640-6645
- 21. Wang, H.H. et al. (2009) Programming cells by multiplex genome engineering and accelerated evolution. Nature 460, 894-898
- The White House (2012) National Bioeconomy Blueprint, The 22. Quan, J. et al. (2011) Parallel on-chip gene synthesis and application to optimization of protein expression. Nat. Biotechnol. 29, 449-452
  - 23. Na, D. et al. (2013) Metabolic engineering of Escherichia coli using synthetic small regulatory RNAs. Nat. Biotechnol. 31, 170-174
  - 24. Wu, J. et al. (2015) Enhancing flavonoid production by systematically tuning the central metabolic pathways based on a CRISPR interference system in Escherichia coli. Sci. Rep. 5, 13477
  - 25. Wang, H.H. et al. (2012) Genome-scale promoter engineering by coselection MAGE. Nat. Methods 9, 591-593
  - 26. Isaacs, F.J. et al. (2011) Precise manipulation of chromosomes in vivo enables genome-wide codon replacement. Science 333, 348-353
  - 27. Lajoie, M.J. et al. (2013) Genomically recoded organisms expand biological functions. Science 342, 357-360
  - 28. Cuthbertson, L. and Nodwell, J.R. (2013) The TetR family of regulators. Microbiol. Mol. Biol. Rev. 77, 440-475
  - 29. Finn, R.D. et al. (2014) Pfam: the protein families database. Nucleic Acids Res. 42, D222-D230
  - 30, Tang, S.Y. and Cirino, P.C. (2011) Design and application of a mevalonate-responsive regulatory protein. Angew. Chem. Int. Ed. Engl. 50, 1084-1086
  - 31. Raman, S. et al. (2014) Engineering allostery. Trends Genet. 30, 521-528
  - 32. Lee, S.W. and Oh, M.K. (2015) A synthetic suicide riboswitch for the high-throughput screening of metabolite production in Saccharomyces cerevisiae. Metab. Eng. 28, 143-150
  - 33. Binder, S. et al. (2012) A high-throughput approach to identify genomic variants of bacterial metabolite producers at the singlecell level. Genome Biol. 13, R40
  - 34. Jha, R.K. et al. (2014) Engineering an Acinetobacter regulon for biosensing and high-throughput enzyme screening in E. coli via flow cytometry. Nucleic Acids Res. 42, 8150-8160
  - 35. Siedler, S. et al. (2014) SoxR as a single-cell biosensor for NADPHconsuming enzymes in Escherichia coli. ACS Synth. Biol. 3, 41-47
  - 36. Tang, S.Y. et al. (2013) Screening for enhanced triacetic acid lactone production by recombinant Escherichia coli expressing a designed triacetic acid lactone reporter. J. Am. Chem. Soc. 135, 10099-10103
  - 37. Taylor, N.D. et al. (2015) Engineering an allosteric transcription factor to respond to new ligands. Nat. Methods Published online December 21, 2015. http://dx.doi.org/10.1038/nmeth.3696
  - 38. Schendzielorz, G. et al. (2014) Taking control over control: use of product sensing in single cells to remove flux control at key enzymes in biosynthesis pathways. ACS Synth. Biol. 3, 21-29