Bridging the gap between systems biology and synthetic biology

Di Liu, Allison Hoynes-O'Connor and Fuzhong Zhang*

Department of Energy, Environmental and Chemical Engineering, Washington University, St. Louis, MO, USA

Systems biology is an inter-disciplinary science that studies the complex interactions and the collective behavior of a cell or an organism. Synthetic biology, as a technological subject, combines biological science and engineering, allowing the design and manipulation of a system for certain applications. Both systems and synthetic biology have played important roles in the recent development of microbial platforms for energy, materials, and environmental applications. More importantly, systems biology provides the knowledge necessary for the development of synthetic biology tools, which in turn facilitates the manipulation and understanding of complex biological systems. Thus, the combination of systems and synthetic biology has huge potential for studying and engineering microbes, especially to perform advanced tasks, such as producing biofuels. Although there have been very few studies in integrating systems and synthetic biology, existing examples have demonstrated great power in extending microbiological capabilities. This review focuses on recent efforts in microbiological genomics, transcriptomics, proteomics, and metabolomics, aiming to fill the gap between systems and synthetic biology.

Keywords: systems biology, synthetic biology, microbial engineering, metabolic engineering, cell factory

INTRODUCTION

Recent advances in genetics and molecular biology, including genome sequencing, bioinformatics, and high-throughput experimentation, have enabled the collection of large sets of data. These data provide a comprehensive understanding of complex biological systems, boosting the development of both systems biology and synthetic biology.

Systems biology aims to develop novel methodologies to study the functionality of the biological system as a whole. When studying microorganisms, these methodologies not only help to understand how microbes adapt, evolve, and interact with other organisms (Navid et al., 2009; Zaneveld et al., 2011), but also reveal the profile and the dynamics of RNAs, proteins, and metabolites, elucidate their intracellular interactions, and uncover complex regulatory networks. Several review articles have discussed the role of systems biology in the study of microbes (Brul et al., 2008; de Lorenzo and Galperin, 2009; Heinemann and Sauer, 2010; Kohlstedt et al., 2010).

Synthetic biology focuses on constructing artificial tools to achieve particular functions. Microbes are excellent hosts for many important applications such as bioremediation, biodegradation, bioconversion, and bioproduction. Particularly, engineered microbes have been extensively used to produce therapeutic proteins, industrial enzymes, small molecular pharmaceuticals, chemicals, biofuels, and materials. Review articles focusing on engineering microbes as cell factories are also available (Picataggio, 2009; Gowen and Fong, 2011).

Although systems biology and synthetic biology focus, respectively, on science and technology, knowledge of systems biology guides the design of better synthetic biology tools, which can in turn provide insights to systems biology. Here we review the recent development of systems and synthetic biology methodologies for the understanding and control of genomics, transcriptomics, proteomics, and metabolomics in microbes. We also discuss the further possibilities for these two fields to benefit from each other.

GENOMICS

As the basic unit of heredity, a gene is transcribed to an mRNA, which is translated to a protein, and all these molecules work together to perform complex functions within a living organism. Systems biology methodologies, such as whole genome sequencing, enable better understanding of gene function, which in turn allows the development of synthetic biology tools to manipulate genetics.

At the systems biology level, deciphering genetic codes provides valuable information on the structure and function of genes. Initiated in the 1970s, sequencing techniques have gone through striking development, and now fully automated DNA sequencing instruments coupled with high-throughput capabilities are available to many research labs. On average, bacterial genome sequencing can now be completed within hours or days, at the cost of $25 per Mb assembled sequence (Didelot et al., 2012; Loman et al., 2012). The analysis of the raw sequencing data is largely aided by bioinformatics, which integrates techniques from different disciplines, such as computer science and mathematics, to interpret biological data. Bioinformatics not only helps in genome annotation, but also provides insights on the corresponding protein functions and homologies between species. Proteogenomics, for example, unifies genomic data and protein identification techniques, allowing for new gene discovery and accurate gene annotation (Armengaud, 2010). Proteogenomics plays an important role

in systems biology by providing a detailed picture of cell systems. For example, Banfield et al. (2005) implemented proteogenomics in the characterization of bacterial communities living deep in mine tunnels: specifically those that produce chemoautotrophic biofilms.

At the synthetic biology level, the ability to edit genetic sequences is the basis for manipulating any synthetic system, which creates an obvious need for such editing. Stimulated by this need, DNA synthesis methods have been developing fast over the past few years. Many methods have been developed in the pursuit of efficient, high fidelity, and low cost DNA synthesis techniques. For example, Tian et al. (2004) used photo-programmable microfluidic chips for multiplex gene synthesis, coupled with a hybridization-based method for error correction. The current cost for commercial gene synthesis is $0.28/bp or even lower (Genscript, Inc.), rendering genetic manipulation easier than ever. While chemical synthesis of the whole microbial genome has been demonstrated for *Mycoplasma mycoides*, at the current stage of development, complete chemical synthesis of microbial genome could be complicated and costly (Gibson et al., 2010). An alternative approach to construct large pieces of DNA or metabolic pathways is to assemble multiple existing DNA fragments. For example, the well-established Gibson DNA assembly approach utilized the activity of 5′ exonucleases to generate single-stranded DNA overhangs, which can then be sealed and ligated with high accuracy and efficiency using a DNA polymerase and a ligase (Gibson et al., 2009). This method allows scar-less assembly of multiple DNA fragments in one-pot. Other DNA assembly methods, such as CPEC (circular polymerase extension cloning; Quan and Tian, 2009) and Golden Gate (Engler et al., 2008), have their own advantages and are suitable for certain purposes (**Table 1**).

By the traditional homologous recombination, genomic DNA can be inserted, deleted, and mutated at one site each time (Datsenko and Wanner, 2000). Recent research advances have led to modification of many targeted locations in the chromosome simultaneously. Wang et al. (2009a) demonstrated multiplex automated genome engineering (MAGE), a method based on oligo-mediated allelic replacement, as an efficient tool to introduce mutations to the microbial chromosome, generating rich genome diversity with tuned properties. This technique was applied to optimize multiple genetic components in the 1-deoxy-D-xylulose-5-phosphate (DXP) biosynthesis pathway, and achieved a more than fivefold increase in lycopene production within 3 days. Furthermore, in order to improve the efficiency of MAGE, the conjugative assembly of genome engineering (CAGE) method was developed, which permits large-scale genome modification by assembling modified genome parts. Combining these two techniques, researchers successfully demonstrated the ability to replace all the TAG stop codons with TAA in *E. coli* (Isaacs et al., 2011). The engineered strain has a free TAG codon that could be used in areas such as the incorporation of unnatural amino acids.

Even with these powerful synthetic biology tools available, functional construction of synthetic systems is still hindered by the hidden and complex regulation in the host cell. This fact requires that genetic components be well characterized and

mutually orthogonal. Researchers recently developed a systematic process for refactoring the nitrogen fixation gene cluster, in which they changed the codons of essential coding sequences and put them under the control of synthetic parts, separated by synthetic spacer sequences (Temme et al., 2012). By removing the native regulations, the underlying interactions are simplified, which facilitates the engineering of the nitrogen fixation system in a non-nitrogen-fixation host. All the above-mentioned synthetic biology approaches provide systems biologists new tools to understand gene functions and complex genetic regulatory networks.

TRANSCRIPTOMICS

Involved in both transcription and translation, RNA molecules serve as the link between genes and proteins. While systematic analysis of transcript profiles reveals gene expression patterns, synthetic regulation of these transcript levels can alter protein concentrations. A systematic understanding of transcriptomics is essential for designing synthetic regulatory systems.

The information obtained from genetics can be more precisely understood if we take a step further to the transcriptional level. By quantifying the expression level of related genes under different conditions, an RNA microarray was developed to facilitate the interpretation of the genome function and regulation patterns. This method is high-throughput and inexpensive, but limitations do exist, including the requirement of genome sequence information and errors caused by cross-hybridization (Wang et al., 2009b). Another technique, RNA-Seq, overcame these limitations by enabling the direct sequencing of RNA transcripts. It allows precise detection and quantification of transcripts to a single-base resolution, and can be applied to species with unknown genome sequences. More importantly, it provides a powerful tool to understand the transcriptomic dynamics by monitoring gene expression levels (Mortazavi et al., 2008; Wilhelm et al., 2008). Although challenges regarding library construction and bioinformatics data analysis remain, this method revolutionizes the way scientists analyze transcriptome data (Wang et al., 2009b).

For any synthetic system to function properly, it is key to regulate the expression of the genes involved. Gene expression can be regulated at different stages, among which transcriptional regulation often plays the most important role (Lodish et al., 2000; Romanel et al., 2012). Synthetic control of transcript levels can be achieved by varying the transcription initiation rate, transcript stabilities, or transcription termination frequency. Control of transcription initiation rates by constitutive or inducible promoters with various strengths have been used for many years. This approach still serves as the most effective and robust method for static control of the transcript level. However, sometimes it is advantageous to regulate a gene dynamically according to the environments and the metabolic status of the cell (Holtz and Keasling, 2010). This becomes particularly important when improving the robustness of a synthetic biological system, where many parameters could change as environmental conditions vary. The knowledge of transcription-level regulation from systems biology has inspired synthetic biologists to utilize natural elements to build synthetic regulatory tools. A recent work demonstrated

Table 1 | Systems biology and synthetic biology tools and applications.

	Systems biology		Synthetic biology	
	Tools	**Applications**	**Tools**	**Applications**
Genomics	DNA sequencing	Gene identification, annotation, protein identification	Gibson, CPEC, etc.	Restriction enzyme-free DNA assembly
			Golden Gate	DNA assembly better for library construction
	Bioinformatics		MAGE, CAGE	Genome wide modification
	Proteogenomics		Gene cluster refactoring	Orthogonal genetic components
Transcriptomics	RNA microarray, RNA-Seq	Gene function interpretation, transcriptomic dynamics	Synthetic promoters, ribozymes, aptamers, sRNAs, etc.	Regulate transcript and translation
			RBS calculator	Control translational level
Proteomics	SRM	Protein detection and identification	Modular design of proteins	Build artificial proteins or regulate protein activities
			Post-translational modification	
			Computational protein design	
Metabolomics	GC–MS, LC–MS, NMR, etc.	Identification of novel metabolic pathways, bottleneck steps	Key enzyme overexpression, mutation, and deletion	Optimize metabolic pathways
	Computational tools, FBA, MFA, etc.		Global regulator engineering	
			Synthetic transporter	Control metabolite secretion

the construction of a dynamic sensor-regulator system, in which transcription initiation rate of several heterologous genes are dynamically controlled by the cellular concentration of a key metabolite (Zhang et al., 2012b). In detail, biosensors specific to fatty acyl-CoAs, key intermediates in the biodiesel biosynthetic pathway, were engineered to control both the biosynthesis and the consumption of acyl-CoAs. Similar to natural regulatory systems, the synthetic control tool optimized gene expression levels, preventing the production of unnecessary RNAs and proteins and improving the efficiency of the biodiesel pathway. As a result, biodiesel titers were increased by threefold, reaching 28% of the theoretical yield. Besides the regulation of transcription initiation, mRNA stability can be regulated as well; control of the mRNA degradation rate, folding rate, and ribozyme activity by small molecules could allow for dynamic control of transcript activities in response to metabolites (Beisel et al., 2008; Carothers et al., 2011; Michener et al., 2012). For example, Babiskin and Smolke (2011) constructed an RNA device based on an RNase III enzyme, where they coupled RNA aptamers to the Rnt1p hairpins. Binding to ligand induced a conformational change on the RNA, which inhibited the self-cleavage activity and stabilized the transcript. Furthermore, transcription termination has been regulated by small RNA molecules that interact with mRNAs (Lucks et al., 2011). The naturally occurring transcriptional control mechanism in the *Staphylococcus aureus* plasmid pT181 uses an antisense RNA to induce a conformational change that exposes a transcription termination site.

By producing mutations in the attenuator sequence, researchers were able to create variants that responded to unique antisense RNA molecules. These orthogonal signals were used in tandem to construct logic gates and an RNA-mediated transcriptional cascade.

Once a gene has been transcribed, cellular protein levels can be tuned at the translational level by engineering the ribosome binding sites (RBSs). Salis et al. (2009) developed a mathematical model, called the RBS calculator, to compute the RBS strength. This model considers the energies involved in rRNA–mRNA interaction, mRNA folding, tRNA binding, and the energetic cost of sub-optimal spacing between the RBS and the start codon (Salis, 2011). This computational tool was proved effective in designing RBS sequences to control relative protein levels.

The above-mentioned methods allow the synthetic regulation of a single gene or of multiple genes at either the transcriptional or the translational level. They are particularly powerful when optimizing a specific metabolic pathway. However, to improve the overall behavior of engineered microbes, such as tolerance toward chemicals or stress conditions, it might be useful to regulate gene expression at the whole genome-scale (Zhang et al., 2012a; Woodruff et al., 2013). For example, Alper et al. (2006) employed a global transcription machinery engineering (gTME) approach to improve both ethanol tolerance and production of a yeast strain. In this study, two proteins that regulate the global transcriptome (SPT15, a TATA-binding protein, and TAF25, TATA-binding

protein-associated factor) were subject to random mutagenesis via error-prone PCR. These mutant libraries were introduced to yeast and subject to screening for their abilities to grow in the presence of ethanol. Strains selected from this approach confer both enhanced ethanol tolerance and efficient glucose to ethanol conversion. This study serves a good example for using synthetic biology tools to modify microbes at the systems level.

PROTEOMICS

Proteins are ubiquitous in biological systems; their complex structure allows them to perform innumerable functions, such as transport, catalysis, signaling, and regulation. Thus a systematic understanding of proteomics, including protein structure, function, concentration, and interactions with other molecules must precede the development of novel synthetic systems.

Experimental protein studies rely heavily on proteomic technologies and instrumentation. Selected reaction monitoring (SRM) is a powerful proteomic technique that can quantitatively detect small numbers of a specific protein. However, SRM can only be used to detect proteins for which assays have been developed. Previously, assay development was an arduous process, limiting the use of SRM (Doerr, 2010). Picotti et al. (2010) devised a high-throughput method for developing SRM assays that allowed them to analyze all the phosphatases and kinases in the proteome of *E. coli*. By synthesizing and analyzing libraries of synthetic peptides, 432 SRM assays were generated in less than 6 h of instrumentation time with an 89% success rate. The ease with which these assays can now be generated will enable a wide variety of new applications, potentially including whole-proteome analysis. For example, Singh et al. (2012) used SRM techniques to optimize flux within the mevalonate pathway in *E. coli*. Malmstroem et al. (2009) applied this technique to determine the average quantity of proteins per cell for 51% of the open reading frames (ORFs), or 83% of the proteome of *Leptospira interrogans*, a human pathogen. They verified their measurements using cryo-electron tomography performed on whole, individual cells, and concluded that the mass spectrometric technique could be quickly and efficiently applied to biological systems.

While systems biology provides information regarding the structure and function of natural proteins, synthetic biology, empowered with such knowledge, can lead to the design of proteins that perform novel functions in synthetic systems. One approach is to design proteins based on modularity (Nash, 2012). Proteins, as well as DNA, RNA, and other small molecules, can often be broken down into modules – discrete parts that perform a specific function. Modules may bind to ligands, transmit information, catalyze a reaction, or accomplish a myriad of other tasks. For example, there are binding domains that mediate protein–peptide interactions by recognizing certain peptide characteristics, such as the SH3 domain that recognizes proline-rich sequences and the PDZ domain that recognizes specific C-terminal sequences (Teyra et al., 2012). Based on the SH3 and PDZ domain, Dueber et al. (2009) constructed synthetic protein scaffolds, which recruit peptide-tagged enzymes for spatial organization of multiple enzymes from a metabolic pathway, preventing the diffusion of metabolic intermediates and improving overall pathway efficiencies. Similar designs have also allowed for the control of signal

transduction pathways through synthetic protein–protein interactions (Good et al., 2011). Furthermore, protein function can also be regulated through post-translational modification. Wang et al. (2010) demonstrated that global protein acetylation allowed for quick responses to changes in environmental conditions, allowing cells to modify their metabolism based on the availability of various carbon sources. Researchers also characterized the role of regulatory enzymes involved in the reversible acetylation process, elucidating a potential global regulatory circuit. Such a system could be engineered for fast regulation of metabolism at protein levels. In addition, protein design using computational approaches has presented more engineering opportunities, providing artificial proteins with novel activities and specific interactions with nucleic acids, small ligands, and other proteins (Mandell and Kortemme, 2009).

METABOLOMICS

Metabolomics focuses on the profile and dynamics of metabolites, revealing the activity of cellular enzymatic reactions as well as metabolic and catabolic pathways. Further, metabolic analyses can be used as diagnostic tools in the study of microbial cell status and environmental conditions. From the perspective of a synthetic biologist, engineering microbial metabolomics has direct links with applications: to degrade toxins (Dziga et al., 2012), herbicides (Sinha et al., 2010), and environmental pollutants (Chien et al., 2010), and to produce chemicals (Curran and Alper, 2012), pharmaceuticals (Paddon et al., 2013), and biofuels (Zhang et al., 2011; Peralta-Yahya et al., 2012).

Metabolite identification and quantification methodologies based on gas chromatography–mass spectrometry (GC–MS), liquid chromatography–mass spectrometry (LC–MS), and nuclear magnetic resonance (NMR) have been developed in recent decades to study metabolite profiles and dynamics (Ludwig and Viant, 2010; Shuman et al., 2011; Bueschl et al., 2013). Additionally, metabolic modeling tools, such as flux balance analysis (FBA; Curran et al., 2012) and metabolic flux analysis (MFA; Tang et al., 2009), were developed. While both FBA and MFA are based on stoichiometric calculation of metabolic reaction rates under pseudo steady state assumptions, MFA uses experimental data instead of targeting biological fitness functions, as does FBA. Further, the combination of metabolic analytical methodologies and modeling tools helps to characterize metabolic networks, to identify novel metabolic pathways and bottleneck steps, and to study the responses of metabolic flux toward genetic modifications or under various environmental conditions.

Guided by knowledge from systems biology, synthetic biology aims to engineer microbial metabolomics, directing the flow of metabolites to desirable pathways. This includes constructing novel biosynthetic pathways for the production of useful molecules and engineering biodegradation routes for environmental applications (Huang et al., 2006). The above-mentioned metabolite analysis methodologies provide diagnostic tools for synthetic pathways, improving their productivity. Various modeling techniques consider metabolic network interactions and predict the optimal genetic modifications for targeted chemical production (Segre et al., 2002; Burgard et al., 2003; Ranganathan et al., 2010; Yang et al., 2011). For example, Ranganathan et al.

(2010) developed an OptForce procedure to identify the potential targets to improve fatty acid production. This method used the flux measurements in a wild type strain, and simulated the optimal flux changes that should be made in the engineered strain for optimal production. The flux changes are then used to identify promising genetic interventions. The consistency between the computational results and the experimental measurements suggested that cell metabolism could be directed through programing design (Ranganathan et al., 2012). Another method developed by Flowers et al. (2013) can systematically identify, instead of one enzyme at each time, multiple target enzymes, whose expression levels could be simultaneously manipulated to obtain the desired phenotype. This method improved the computation efficiency dramatically.

The metabolic profile and regulatory networks obtained from systems biology studies also inspired scientists to develop synthetic control tools at the global level. In a recent study, a global

transcription factor that controls multiple genes involved in fatty acid biosynthesis, degradation, and membrane transport was engineered (Zhang et al., 2012a). Overexpression of this single regulatory protein caused global-scale metabolic changes and was able to increase fatty acid production by fivefold, more significant than the overexpression of many single enzymes in the fatty acid pathway (Zhang et al., 2012a). In addition, the transport of metabolites across the cell membrane can be controlled by pumps identified by systems biology or engineered through synthetic biology methodologies (Dunlop et al., 2011). These pumps are very useful to secrete the product, but not the intermediates, out of the engineered host cell, lowering the stress from chemical accumulation while simplifying downstream processes.

CONCLUSION AND OUTLOOKS

The inherent complexity of genetics presents researchers in systems and synthetic biology with the formidable task of respectively
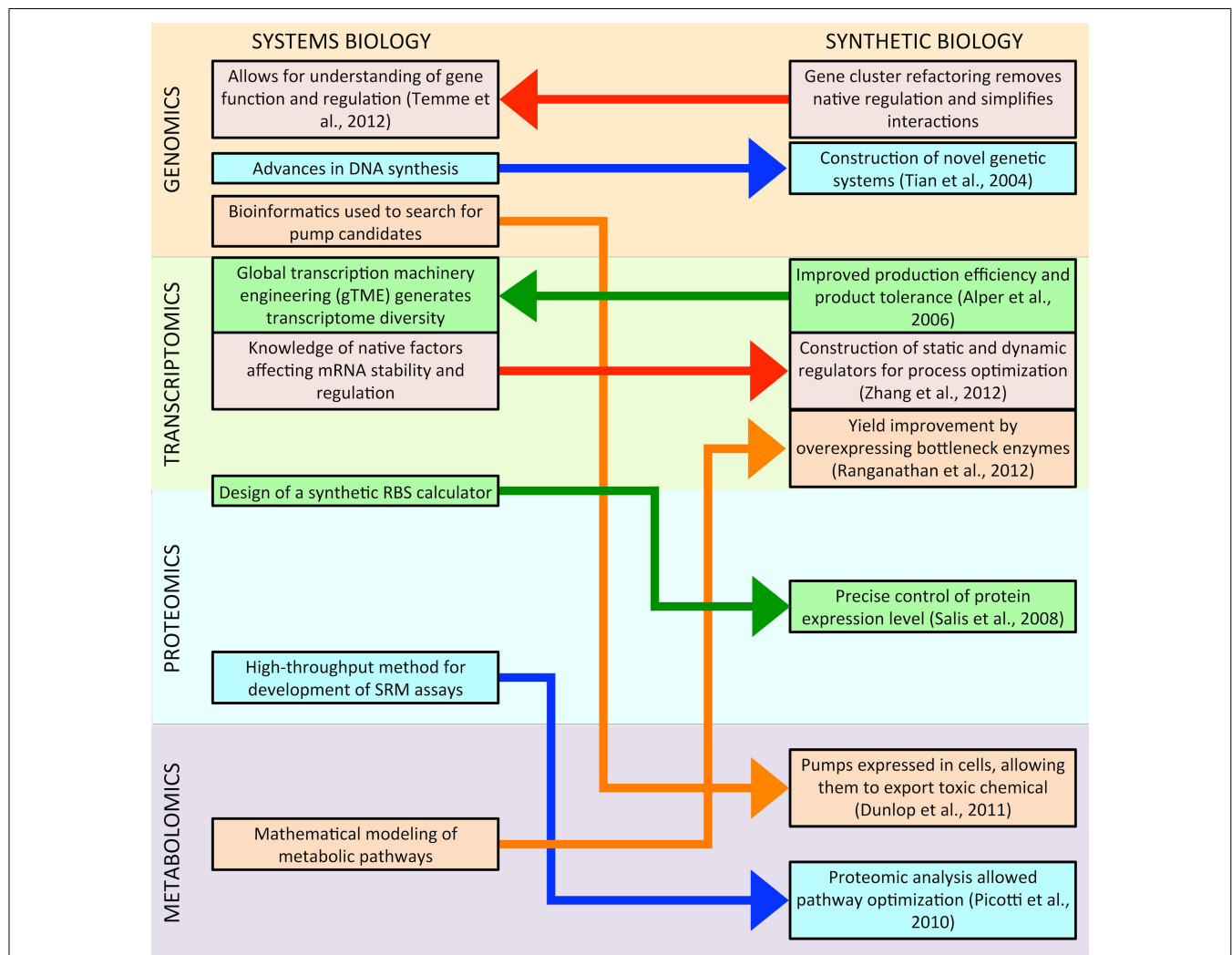
FIGURE 1 | This figure illustrates recent interactions between systems and synthetic biology. In most cases, methodologies developed in systems biology have led to advances in synthetic biology. However, this trend may be changing, as research in synthetic biology has already begun to provide insight to systems biology, and there might be more research at the intersection of the two fields in the near future.

understanding and manipulating natural genetic systems and their complex control elements. The current available tools and their applications in systems and synthetic biology are summarized in **Table 1**. In the coming years, advancements in genomics will lead to a further decrease in the cost of DNA synthesis, accelerating research. Transcriptomics will experience the development of a wide range of synthetic tools, including control through the modular use of synthetic promoters and RNA elements, e.g., untranslated region (UTRs), RBSs, antisense RNA, and ribozymes. Proteomics will provide a wealth of data in the form of proteome mapping, and metabolomics will incorporate these advances to achieve high yields of desired products.

Figure 1 illustrates the current interactions between systems and synthetic biology; synthetic biology is drawing more tools and knowledge from systems biology than it is reciprocating, but this trend is likely to change, and there might be more research at the intersection of the two fields in the near future. Deeper understanding of systems biology will provide more synthetic biology parts, such as biosensors that can be used for dynamic regulation of synthetic pathways (Zhang and Keasling, 2011). Systems biology knowledge will also make synthetic biology tools more reliable, enabling the precise control of transcription and translation regardless of the under-controlled gene (Mutalik et al., 2013). More powerful systems biology-based computational tools will simplify both the design and the optimization of synthetic metabolic pathways, improving titers and productivities (Colletti et al., 2011). Similarly, simplified genetic systems created in synthetic biology will provide systems biology with insight into the fundamentals of native gene regulation. Powerful synthetic quantification tools will allow the simultaneous collection of omic data at global scales in living cells. Overall, direct communication between systems and synthetic biologists regarding tools and knowledge will hasten progress in genomics, transcriptomics, proteomics, and metabolomics.

ACKNOWLEDGMENTS

REFERENCES

Alper, H., Moxley, J., Nevoigt, E., Fink, G. R., and Stephanopoulos, G. (2006). Engineering yeast transcription machinery for improved ethanol tolerance and production. *Science* 314, 1565–1568. doi: 10.1126/science.1131969

Armengaud, J. (2010). Proteogenomics and systems biology: quest for the ultimate missing parts. *Expert Rev. Proteomics* 7, 65–77. doi: 10.1586/epr.09.104

Babiskin, A. H., and Smolke, C. D. (2011). Synthetic RNA modules for fine-tuning gene expression levels in yeast by modulating RNase III activity. *Nucleic Acids Res.* 39, 8651–8664. doi: 10.1093/nar/gkr445

Banfield, J. F., Verberkmoes, N. C., Hettich, R. L., and Thelen, M. P. (2005). Proteogenomic approaches for the molecular characterization of natural microbial communities. *OMICS* 9, 301–333. doi: 10.1089/omi.2005.9.301

Beisel, C. L., Bayer, T. S., Hoff, K. G., and Smolke, C. D. (2008). Model-guided design of ligand-regulated RNAi for programmable control of gene expression. *Mol. Syst. Biol.* 4, 224. doi: 10.1038/msb2008.62

Brul, S., Mensonides, F. I., Hellingwerf, K. J., and Teixeira De Mattos, M. J. (2008). Microbial systems biology: new frontiers open to predictive microbiology. *Int. J. Food Microbiol.* 128, 16–21. doi: 10.1016/j.ijfoodmicro.2008.04.029

Bueschl, C., Krska, R., Kluger, B., and Schuhmacher, R. (2013). Isotopic labeling-assisted metabolomics using LC–MS. *Anal. Bioanal. Chem.* 405, 27–33. doi: 10.1007/s00216-012-6375-y

Burgard, A. P., Pharkya, P., and Maranas, C. D. (2003). OptKnock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. *Biotechnol. Bioeng.* 84, 647–657. doi: 10.1002/bit.10803

Carothers, J. M., Goler, J. A., Juminaga, D., and Keasling, J. D. (2011). Model-driven engineering of RNA devices to quantitatively program gene expression. *Science* 334, 1716–1719. doi: 10.1126/science.1212209

Chien, M. F., Narita, M., Lin, K. H., Matsui, K., Huang, C. C., and Endo, G. (2010). Organomercurials removal by heterogeneous merB genes harboring bacterial strains. *J. Biosci. Bioeng.* 110, 94–98. doi: 10.1016/j.jbiosc.2010.01.010

Colletti, P. F., Goyal, Y., Varman, A. M., Feng, X., Wu, B., and Tang, Y. J. (2011). Evaluating factors that influence microbial synthesis yields by linear regression with numerical and ordinal variables. *Biotechnol. Bioeng.* 108, 893–901. doi: 10.1002/bit.22996

Curran, K. A., and Alper, H. S. (2012). Expanding the chemical palate of cells by combining systems biology and metabolic engineering. *Metab. Eng.* 14, 289–297. doi: 10.1016/j.ymben.2012.04.006

Curran, K. A., Crook, N. C., and Alper, H. S. (2012). Using flux balance analysis to guide microbial metabolic engineering. *Methods Mol. Biol.* 834, 197–216. doi: 10.1007/978-1-61779-483-4_13

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. doi: 10.1073/pnas.120163297

de Lorenzo, V., and Galperin, M. (2009). Microbial systems biology: bottom up and top down. *FEMS Microbiol. Rev.* 33, 1–2. doi: 10.1111/j.1574-6976.2008.00147.x

Didelot, X., Bowden, R., Wilson, D. J., Peto, T. E. A., and Crook, D. W. (2012). Transforming clinical microbiology with bacterial genome sequencing. *Nat. Rev. Genet.* 13, 601–612. doi: 10.1038/nrg3226

Doerr, A. (2010). Targeted proteomics. *Nat. Methods* 7, 34–34. doi: 10.1038/nmeth.f.284

Dueber, J. E., Wu, G. C., Malmirchegini, G. R., Moon, T. S., Petzold, C. J., Ullal, A. V., et al. (2009). Synthetic protein scaffolds provide modular control over metabolic flux. *Nat. Biotechnol.* 27, 753–759. doi: 10.1038/nbt.1557

Dunlop, M. J., Dossani, Z. Y., Szmidt, H. L., Chu, H. C., Lee, T. S., Keasling, J. D., et al. (2011). Engineering microbial biofuel tolerance and export using efflux pumps. *Mol. Syst. Biol.* 7, 487. doi: 10.1038/msb2011.21

Dziga, D., Wladyka, B., Zielinska, G., Meriluoto, J., and Wasylewski, M. (2012). Heterologous expression and characterisation of microcystinase. *Toxicon* 59, 578–586. doi: 10.1016/j.toxicon.2012.01.001

Engler, C., Kandzia, R., and Marillonnet, S. (2008). A one pot, one step, precision cloning method with high throughput capability. *PLoS ONE* 3:e3647. doi: 10.1371/journal.pone.0003647

Flowers, D., Thompson, R. A., Birdwell, D., Wang, T., and Trinh, C. T. (2013). SMET: systematic multiple enzyme targeting – a method to rationally design optimal strains for target chemical overproduction. *Biotechnol. J.* 8, 605–618. doi: 10.1002/biot.201200233

Gibson, D. G., Glass, J. I., Lartigue, C., Noskov, V. N., Chuang, R. Y., Algire, M. A., et al. (2010). Creation of a bacterial cell controlled by a chemically synthesized genome. *Science* 329, 52–56. doi: 10.1126/science.1190719

Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C. A. III, and Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* 6, 343–345. doi: 10.1038/nmeth.1318

Good, M. C., Zalatan, J. G., and Lim, W. A. (2011). Scaffold proteins: hubs for controlling the flow of cellular information. *Science* 332, 680–686. doi: 10.1126/science.1198701

Gowen, C. M., and Fong, S. S. (2011). Applications of systems biology towards microbial fuel production. *Trends Microbiol.* 19, 516–524. doi: 10.1016/j.tim.2011.07.005

Heinemann, M., and Sauer, U. (2010). Systems biology of microbial metabolism. *Curr. Opin. Microbiol.* 13, 337–343. doi: 10.1016/j.mib.2010.02.005

Holtz, W. J., and Keasling, J. D. (2010). Engineering static and dynamic control of synthetic pathways. *Cell* 140, 19–23. doi: 10.1016/j.cell.2009.12.029

Huang, C. C., Chen, M. W., Hsieh, J. L., Lin, W. H., Chen, P. C.,

and Chien, L. F. (2006). Expression of mercuric reductase from *Bacillus megaterium* MB1 in eukaryotic microalga *Chlorella* sp. DT: an approach for mercury phytoremediation. *Appl. Microbiol. Biotechnol.* 72, 197–205. doi: 10.1007/s00253-005-0250-0

Isaacs, F. J., Carr, P. A., Wang, H. H., Lajoie, M. J., Sterling, B., Kraal, L., et al. (2011). Precise manipulation of chromosomes in vivo enables genome-wide codon replacement. *Science* 333, 348–353. doi: 10.1126/science.1205822

Kohlstedt, M., Becker, J., and Wittmann, C. (2010). Metabolic fluxes and beyond-systems biology understanding and engineering of microbial metabolism. *Appl. Microbiol. Biotechnol.* 88, 1065–1075. doi: 10.1007/s00253-010-2854-2

Lodish, H., Berk, A., and Zipursky, S. L. (2000). *Molecular Cell Biology*. New York: W. H. Freeman.

Loman, N. J., Constantinidou, C., Chan, J. Z. M., Halachev, M., Sergeant, M., Penn, C. W., et al. (2012). High-throughput bacterial genome sequencing: an embarrassment of choice, a world of opportunity. *Nat. Rev. Microbiol.* 10, 599–606. doi: 10.1038/nrmicro2850

Lucks, J. B., Qi, L., Mutalik, V. K., Wang, D., and Arkin, A. P. (2011). Versatile RNA-sensing transcriptional regulators for engineering genetic networks. *Proc. Natl. Acad. Sci. U.S.A.* 108, 8617–8622. doi: 10.1073/pnas.1015741108

Ludwig, C., and Viant, M. R. (2010). Two-dimensional J-resolved NMR spectroscopy: review of a key methodology in the metabolomics toolbox. *Phytochem. Anal.* 21, 22–32. doi: 10.1002/pca.1186

Malmstroem, J., Beck, M., Schmidt, A., Lange, V., Deutsch, E. W., and Aebersold, R. (2009). Proteome-wide cellular protein concentrations of the human pathogen *Leptospira interrogans*. *Nature* 460, 762–765. doi: 10.1038/nature08184

Mandell, D. J., and Kortemme, T. (2009). Computer-aided design of functional protein interactions. *Nat. Chem. Biol.* 5, 797–807. doi: 10.1038/nchembio.251

Michener, J. K., Thodey, K., Liang, J. C., and Smolke, C. D. (2012). Applications of genetically-encoded biosensors for the construction and control of biosynthetic pathways. *Metab. Eng.* 14, 212–222. doi: 10.1016/j.ymben.2011.09.004

Mortazavi, A., Williams, B. A., Mccue, K., Schaeffer, L., and Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* 5, 621–628. doi: 10.1038/nmeth.1226

Mutalik, V. K., Guimaraes, J. C., Cambray, G., Lam, C., Christoffersen, M. J., Mai, Q. A., et al. (2013). Precise and reliable gene expression via standard transcription and translation initiation elements. *Nat. Methods* 10, 354–360. doi: 10.1038/nmeth.2404

Nash, P. D. (2012). Why modules matter. *FEBS Lett.* 586, 2572–2574. doi: 10.1016/j.febslet.2012.04.049

Navid, A., Ghim, C. M., Fenley, A. T., Yoon, S., Lee, S., and Almaas, E. (2009). Systems biology of microbial communities. *Methods Mol. Biol.* 500, 469–494. doi: 10.1007/978-1-59745-525-1_16

Paddon, C. J., Westfall, P. J., Pitera, D. J., Benjamin, K., Fisher, K., Mcphee, D., et al. (2013). High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature* 496, 528–532. doi: 10.1038/nature12051

Peralta-Yahya, P. P., Zhang, F., Del Cardayre, S. B., and Keasling, J. D. (2012). Microbial engineering for the production of advanced biofuels. *Nature* 488, 320–328. doi: 10.1038/nature11478

Picataggio, S. (2009). Potential impact of synthetic biology on the development of microbial systems for the production of renewable fuels and chemicals. *Curr. Opin. Biotechnol.* 20, 325–329. doi: 10.1016/j.copbio.2009.04.003

Picotti, P., Rinner, O., Stallmach, R., Dautel, F., Farrah, T., Domon, B., et al. (2010). High-throughput generation of selected reaction-monitoring assays for proteins and proteomes. *Nat. Methods* 7, 43–46. doi: 10.1038/nmeth.1408

Quan, J. Y., and Tian, J. D. (2009). Circular polymerase extension cloning of complex gene libraries and pathways. *PLoS ONE* 4:e6441. doi: 10.1371/journal.pone.0006441

Ranganathan, S., Suthers, P. F., and Maranas, C. D. (2010). OptForce: an optimization procedure for identifying all genetic manipulations leading to targeted overproductions. *PLoS Comput. Biol.* 6:e1000744. doi: 10.1371/journal.pcbi.1000744

Ranganathan, S., Tee, T. W., Chowdhury, A., Zomorrodi, A. R., Yoon, J. M., Fu, Y., et al. (2012). An integrated computational and experimental study for overproducing fatty acids in *Escherichia coli*. *Metab. Eng.* 14, 687–704. doi: 10.1016/j.ymben.2012.08.008

Romanel, A., Jensen, L. J., Cardelli, L., and Csikasz-Nagy, A. (2012). Transcriptional regulation is a major controller of cell cycle transition dynamics. *PLoS ONE* 7:29716. doi: 10.1371/journal.pone.0029716

Salis, H. M. (2011). The ribosome binding site calculator. *Methods Enzymol.* 498, 19–42. doi: 10.1016/B978-0-12-385120-8.00002-4

Salis, H. M., Mirsky, E. A., and Voigt, C. A. (2009). Automated design of synthetic ribosome binding sites to control protein expression. *Nat. Biotechnol.* 27, 946–950. doi: 10.1038/nbt.1568

Segre, D., Vitkup, D., and Church, G. M. (2002). Analysis of optimality in natural and perturbed metabolic networks. *Proc. Natl. Acad. Sci. U.S.A.* 99, 15112–15117. doi: 10.1073/pnas.232349399

Shuman, J. L., Cortes, D. F., Armenta, J. M., Pokrzywa, R. M., Mendes, P., and Shulaev, V. (2011). Plant metabolomics by GC–MS and differential analysis. *Methods Mol. Biol.* 678, 229–246. doi: 10.1007/978-1-60761-682-5_17

Singh, P., Batth, T. S., Juminaga, D., Dahl, R. H., Keasling, J. D., Adams, P. D., et al. (2012). Application of targeted proteomics to metabolically engineered *Escherichia coli*. *Proteomics* 12, 1289–1299. doi: 10.1002/pmic.201100482

Sinha, J., Reyes, S. J., and Gallivan, J. P. (2010). Reprogramming bacteria to seek and destroy an herbicide. *Nat. Chem. Biol.* 6, 464–470. doi: 10.1038/nchembio.369

Tang, Y. J., Martin, H. G., Myers, S., Rodriguez, S., Baidoo, E. E., and Keasling, J. D. (2009). Advances in analysis of microbial metabolic fluxes via (13)C isotopic labeling. *Mass Spectrom. Rev.* 28, 362–375. doi: 10.1002/mas.20191

Temme, K., Zhao, D., and Voigt, C. A. (2012). Refactoring the nitrogen fixation gene cluster from *Klebsiella oxytoca*. *Proc. Natl. Acad. Sci. U.S.A.* 109, 7085–7090. doi: 10.1073/pnas.1120788109

Teyra, J., Sidhu, S. S., and Kim, P. M. (2012). Elucidation of the binding preferences of peptide recognition modules: SH3 and PDZ domains. *FEBS Lett.* 586, 2631–2637. doi: 10.1016/j.febslet.2012.05.043

Tian, J. D., Gong, H., Sheng, N. J., Zhou, X. C., Gulari, E., Gao, X. L., et al. (2004). Accurate multiplex gene synthesis from programmable DNA microchips. *Nature* 432, 1050–1054. doi: 10.1038/nature03151

Wang, H. H., Isaacs, F. J., Carr, P. A., Sun, Z. Z., Xu, G., Forest, C. R., et al. (2009a). Programming cells by multiplex genome engineering and accelerated evolution. *Nature* 460, 894–898. doi: 10.1038/nature08187

Wang, Z., Gerstein, M., and Snyder, M. (2009b). RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* 10, 57–63. doi: 10.1038/nrg2484

Wang, Q., Zhang, Y., Yang, C., Xiong, H., Lin, Y., Yao, J., et al. (2010). Acetylation of metabolic enzymes coordinates carbon source utilization and metabolic flux. *Science* 327, 1004–1007. doi: 10.1126/science.1179687

Wilhelm, B. T., Marguerat, S., Watt, S., Schubert, F., Wood, V., Goodhead, I., et al. (2008). Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution. *Nature* 453, 1239–1243. doi: 10.1038/nature07002

Woodruff, L. B., Pandhal, J., Ow, S. Y., Karimpour-Fard, A., Weiss, S. J., Wright, P. C., et al. (2013). Genome-scale identification and characterization of ethanol tolerance genes in *Escherichia coli*. *Metab. Eng.* 15, 124–133. doi: 10.1016/j.ymben.2012.10.007

Yang, L., Cluett, W. R., and Mahadevan, R. (2011). EMILiO: a fast algorithm for genome-scale strain design. *Metab. Eng.* 13, 272–281. doi: 10.1016/j.ymben.2011.03.002

Zaneveld, J. R., Parfrey, L. W., Van Treuren, W., Lozupone, C., Clemente, J. C., Knights, D., et al. (2011). Combined phylogenetic and genomic approaches for the high-throughput study of microbial habitat adaptation. *Trends Microbiol.* 19, 472–482. doi: 10.1016/j.tim.2011.07.006

Zhang, F., and Keasling, J. (2011). Biosensors and their applications in microbial metabolic engineering. *Trends Microbiol.* 19, 323–329. doi: 10.1016/j.tim.2011.05.003

Zhang, F., Ouellet, M., Batth, T. S., Adams, P. D., Petzold, C. J., Mukhopadhyay, A., et al. (2012a). Enhancing fatty acid production by the expression of the regulatory transcription factor FadR. *Metab. Eng.* 14, 653–660. doi: 10.1016/j.ymben.2012.08.009

Zhang, F. Z., Carothers, J. M., and Keasling, J. D. (2012b). Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids. *Nat. Biotechnol.* 30, 354–359. doi: 10.1038/nbt.2149

Zhang, F., Rodriguez, S., and Keasling, J. D. (2011). Metabolic engineering of microbial pathways

for advanced biofuels production. *Curr. Opin. Biotechnol.* 22, 775–783. doi: 10.1016/j.copbio.2011.04.024