Bioengineered Bugs 1:6, 378-384; November/December 2010; © 2010 Landes Bioscience

Engineering scalable biological systems

Timothy K. Lu[1,2]

[1]Synthetic Biology Group; Research Lab of Electronics; Department of Electrical Engineering and Computer Science; Massachusetts Institute of Technology; and [2]Broad Institute of MIT and Harvard; Cambridge, MA USA

Key words: synthetic biology, biological circuits, engineered organisms, modelling, high-throughput design, regulatory issues, biological probes, biological modulators

Abbreviations: PCR, polymerase chain reaction; DNA, deoxyribonucleic acid; RNA, ribonucleic acid; Mbp, megabasepairs; Kbp, kilo-basepairs; qRT-PCR, quantitative reverse-transcriptase PCR; FRET, fluorescence resonance energy transfer; IPTG, isopropyl β-D-1-thiogalactopyranoside; SELEX, systematic evolution of ligands by exponential enrichment; RNAi, RNA interference; GMOs, genetically modified organisms

Submitted: 06/24/10

Revised: 07/19/10

Accepted: 07/20/10

Previously published online: www.landesbioscience.com/journals/biobugs/article/13086

DOI: 10.4161/bbug.1.6.13086

Correspondence to: Timothy K. Lu; Email: timlu@mit.edu

Commentary to: Lu TK, Khalil AS, Collins JJ. Next-generation synthetic gene networks. Nature Biotechnology 2009; 27:1139–50. PMID: 20010597; DOI: 10.1038/nbt.1591.

Synthetic biology is focused on engineering biological organisms to study natural systems and to provide new solutions for pressing medical, industrial and environmental problems. At the core of engineered organisms are synthetic biological circuits that execute the tasks of sensing inputs, processing logic and performing output functions. In the last decade, significant progress has been made in developing basic designs for a wide range of biological circuits in bacteria, yeast and mammalian systems. However, significant challenges in the construction, probing, modulation and debugging of synthetic biological systems must be addressed in order to achieve scalable higher-complexity biological circuits. Furthermore, concomitant efforts to evaluate the safety and biocontainment of engineered organisms and address public and regulatory concerns will be necessary to ensure that technological advances are translated into real-world solutions.

In the last century, scientists have made giant strides in identifying and studying biological parts such as proteins and nucleic acids,[1-5] understanding regulatory networks,[6] and constructing engineered organisms using the ever-advancing tools of genetic engineering.[7] In the last decade, synthetic biologists have leveraged the power of modern molecular biology using frameworks translated from traditional disciplines such as electrical engineering, computer science, mechanical engineering and chemical engineering to create a wide range of synthetic biological circuits, including switches,[8-15] oscillators,[16-18] digital logic gates,[19-23] filters,[24-26] modular and interoperable memory devices,[27] counters,[27] sensors,[28,29] and protein scaffolds.[30] Using these circuits, biological engineers have created synthetic organisms that can be used for bioremediation, biosensing, computation, bioenergy and medical therapeutics (reviewed in ref. 31–33). Despite these advances, the realization of synthetic-biology-based applications will require future breakthroughs in our ability to create sufficiently complex and reliable biological systems. Here, I will discuss current limitations and potential solutions for the construction, probing, modulation and debugging of scalable biological systems as well as hurdles for the deployment of engineered organisms from bacteria to mammalian cells which adds to the discussion of next-generation synthetic gene networks in reference 31 (**Fig. 1**).

Physical Construction of Scalable Biological Systems

Construction of early synthetic circuits largely relied on restriction enzymes and polymerase chain reaction (PCR)-based techniques to assemble existing genetic components. These methods do not scale well with increasing complexity due to a lack of sufficient unique restriction sites and the need to have physical DNA templates from which to amplify genetic parts. Standards for library construction and the assembly of parts libraries[34] have been integral in circumventing this dependency on templates and restriction sites. However, since these parts must be devoid of restriction sites used in the defined standards and should ideally be optimized for use in one's organism of choice,[35] the use of whole-gene DNA synthesis is on

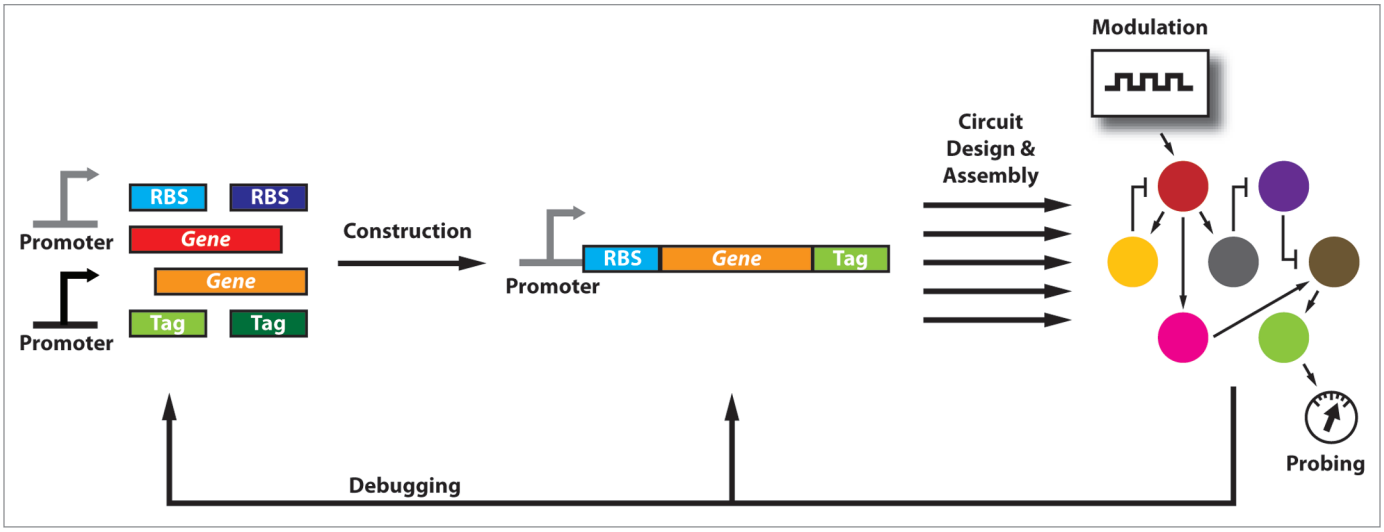

Figure 1. A basic design cycle for synthetic biology includes creating well-characterized parts (e.g., regulatory elements, genes, proteins, RNAs), constructing synthetic devices and modules and designing and assembling higher-order networks. All steps of this cycle are aided by modelling, probes and modulators to analyze circuit performance. Debugging is an iterative process based on parts optimization, fine-tuning regulatory components, modelling and changing circuit architecture.

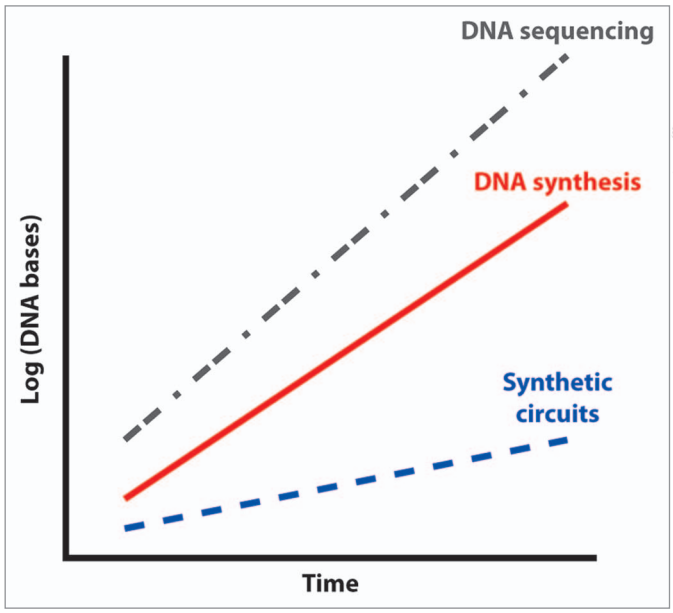

Figure 2. DNA sequencing and synthesis technologies are advancing at exponential rates, outpacing the ability of synthetic biologists to construct useful and scalable biological circuits.[36] These trends are similar to Moore's law for integrated circuits[72] and suggest that there is substantial room for growth in the field of synthetic circuits.

the rise.[36] Using direct chemical synthesis, circuits can be designed in silico and implemented in DNA with significantly less effort from researchers. As DNA synthesis becomes increasingly economical and efficient, it will become possible to construct complex systems with less reliance on restriction enzymes. For example, DNA synthesis productivity has exceeded 1 Mbp per person per day while Venter and colleagues recently succeeded in synthesizing a 1.08 Mbp genome.[37] However, most synthetic gene circuits to date have not exceeded the 50 Kbp level, indicating that there is a large gap between our ability to read and write DNA and knowing what DNA to write (**Fig. 2**). Just as the decoding of the human genome sequence did not immediately reveal the functions of all human genes, the utility of high-throughput DNA synthesis technology will only gradually become evident as synthetic biologists learn how to create complex systems. For example, future synthesized circuits should be designed with ease of probing, modulating and debugging in mind. These features could be implemented by including validated RNA "handles" that can be easily measured with standard probe sets to determine internal RNA concentrations, gene circuits that allow inducers to modulate synthetic circuit protein levels, and properly situated restriction sites for the rapid cloning of components that need systematic optimization, such as ribosome binding sequences.

Significant advances in well-characterized, interoperable devices are necessary for the construction of higher-order modules that will enable scalable biological systems.[38] The majority of biological circuits have been constructed using a handful of synthetic parts.[31] Furthermore, it is often the case that when new designs for biological parts are developed, only a few instantiations are created and tested, usually in single cellular backgrounds. As a result, there is a need for the systematic

development and characterization of compatible biological parts. Specificity in biological systems largely relies on spatial distribution and chemical interactions. This is in stark contrast to electrical engineering, where specificity is achieved through direct electrical wiring. Thus, strategies for achieving inter-part compatibility include targeting circuits to isolated compartments,[39,40] mutagenesis and directed evolution of existing parts, and using comparative genomics to identify, synthesize, and test homologous proteins or nucleic acids. These efforts may be complicated by unknown global factors (e.g., growth rates, endogenous transcription factors with off-target effects on synthetic circuits, protein-protein interactions, small RNAs) that can confound device testing and render it difficult to use pre-defined parts in a wide range of organisms and environmental conditions without additional alterations and characterization.[41] Therefore, combinatorial methods to test single-component performance, multi-component interactions and biological crosstalk (e.g., cross-activation or cross-repression of transcription, non-specific enzymatic activity, inappropriate triggering of signalling pathways) will be important for parts libraries (**Fig. 3**). These results should be incorporated into mathematical models to aid future model-based design. Indeed, institutions such as BIOFAB are attempting to systematically assemble and characterize libraries of synthetic devices. However, development efforts for certain platforms that are promising for library construction, such as zinc finger proteins and RNA interference, may be slowed by the presence of existing intellectual property.[43]

As an example of combinatorial characterization (**Fig. 3**), suppose one would like to construct multiple interoperable NOR (NOT-OR) gates to constitute a universal logic system. NOR functionality can be built by placing pairwise combinations of unique operator sites for transcriptional repressors within synthetic promoters. To identify orthogonal repressors, one can encode individual transcription factors under inducible control on one set of plasmids and individual cognate operator sites driving expression of a reporter gene on another set of plasmids.

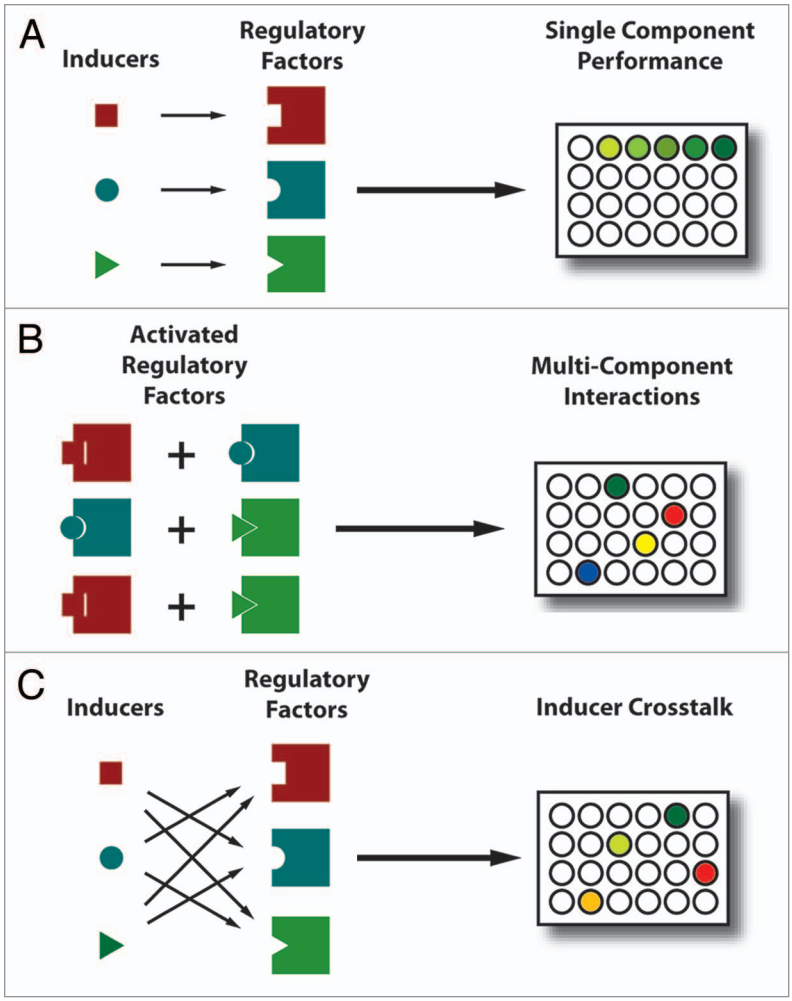

Figure 3. Combinatorial high-throughput methods will be useful in the assembly of well-characterized libraries of synthetic parts and devices. For example, transcriptional regulators and their cognate inducers can be analyzed for (A) single-component performance, (B) interactions between multiple components and (C) inducer crosstalk (e.g., cross-activation and/or cross-inhibition).

Then, all possible combinations of transcription factor plasmids and reporter plasmids can be co-transformed into cells and tested for single-component performance (e.g., when a transcription factor is co-transformed with its cognate operator site) and potential crosstalk interactions (e.g., when a transcription factor is co-transformed with non-cognate operator sites). Standard induction curves can be derived by varying the concentration of transcription factors using the inducible promoters and measuring the resulting output.[42] Based on these results, an optimal set of non-interacting transcription factors and cognate operators can be selected. To create the NOR gates, all possible pairwise combinations of operators can be constructed in synthetic promoters and co-transformed into cells with all pairwise combinations of transcription factors under independent inducible control. Proper NOR gate functionality and crosstalk can then be determined in a high-throughput fashion for all potential gates by varying inducer levels and measuring reporter gene output. In addition to enabling interoperable gate selection, large-scale experiments such as these should yield substantial data for models that can predict the orthogonality of transcription factors and operators for future circuits (e.g., using heuristic or thermodynamically guided algorithms). Moreover, matrices of cross-repression interactions can be constructed and incorporated into

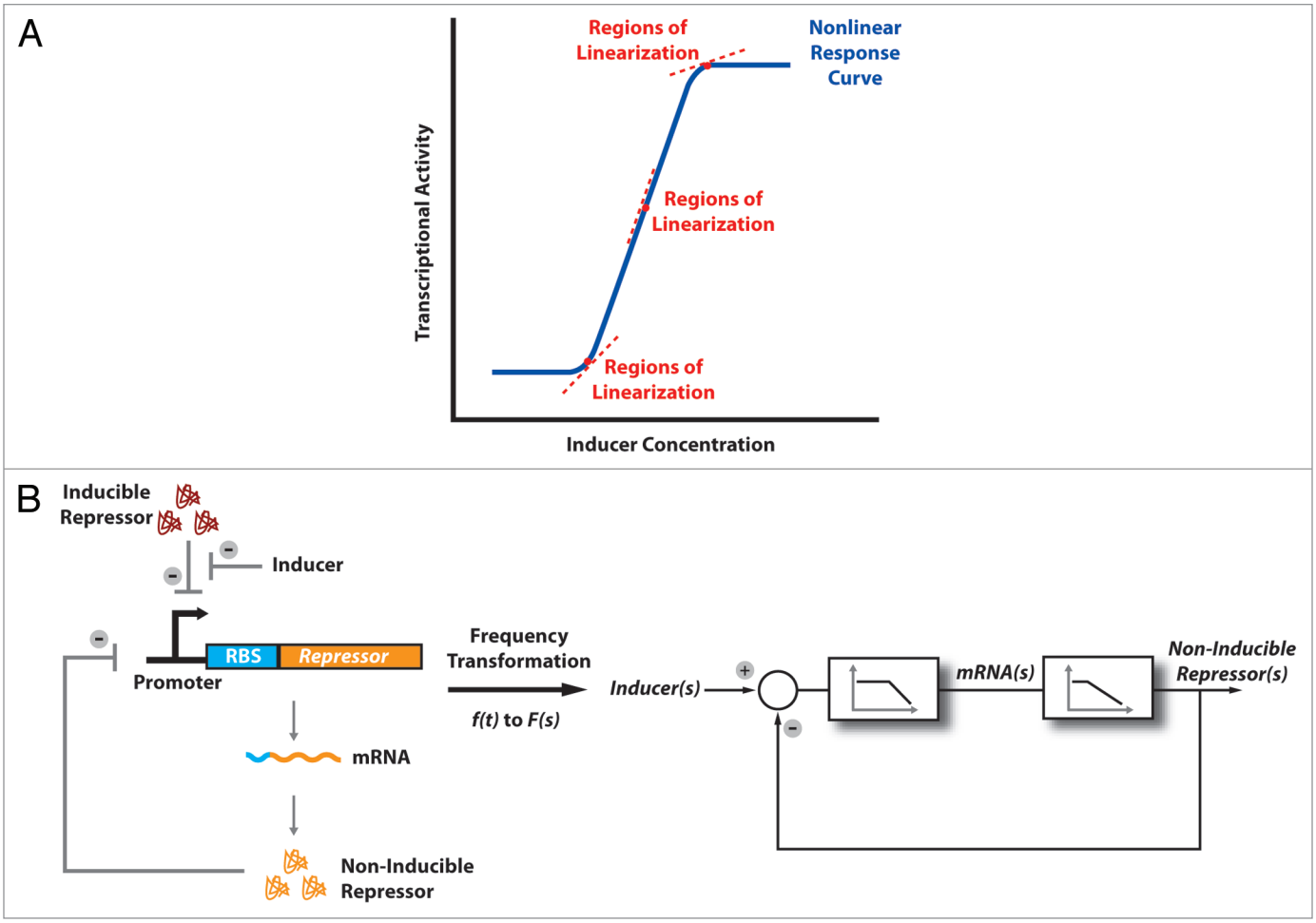

Figure 4. Control theory techniques for modelling synthetic biological circuits. (A) Small-signal linearization of biological components in different regions of operation enables the development of linear models. (B) Linearization can enable frequency-domain analysis, systems modelling using block diagrams and deeper insights into system dynamics. For example, transcription and translation can be understood as low-pass filters and block diagrams can be drawn for simple negative-feedback loops to yield understanding into system interconnections and responses to different input types.[45,46] In the block diagram shown, *s* refers to *j*ω where *j* is $\sqrt{-1}$ and ω is angular frequency.

transcriptional models when cross-interacting transcription factors must be used in other systems.

Model-guided design is crucial for the construction of complicated electrical and mechanical systems. Time-based simulations for electrical and mechanical systems are possible since mathematical models are established and parameters are well known. In contrast, most parameters in biological circuits are unknown and the computational resources required to accurately simulate noise and multiple component interactions are significant. Recent advances in modelling chemical networks, transcription, translation and biological noise using the inherent physics of solid-state electronic devices should enable the construction of large-scale real-time electronic models of synthetic biological systems.[44] Other techniques from control theory such as small-signal linearization and modularization enable tractable modelling and simulations prior to implementation. Biological systems exhibit nonlinearity (e.g., cooperativity) which can be linearized in different regions of operation (**Fig. 4A**). Frequency-domain analysis in linearized systems allows for block modelling and deeper understanding of system dynamics, such as noise, stability, time constants and performance (**Fig. 4B**).[45,46] Small-signal linearization and frequency-domain modelling have not been extensively used for studying and designing synthetic biological circuits even though advances in microfluidics and time-lapse microscopy can now achieve frequency modulation of inputs and long time-scale data collection necessary for frequency-domain analysis.[45,47] Furthermore, microfluidics devices can be coupled with electronic controllers to stabilize and alter the dynamics of synthetic biological circuits similar to electronic controllers that are used to control mechanical systems. The insights that can be gained from linearized block models of complex systems can complement the accuracy of time-domain state-space representation, time-based mathematical simulations and non-linear control theory. To enable successful time-based and frequency-based modelling of biological systems, accurate parameters will need to be derived by high-throughput in vitro

5. Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF, Kerlavage AR, et al. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science 1995; 269:496-512.
6. Johnson AD, Poteete AR, Lauer G, Sauer RT, Ackers GK, Ptashne M. lambda Repressor and cro—components of an efficient molecular switch. Nature 1981; 294:217-23.
7. Itakura K, Hirose T, Crea R, Riggs AD, Heyneker HL, Bolivar F, et al. Expression in *Escherichia coli* of a chemically synthesized gene for the hormone somatostatin. Science 1977; 198:1056-63.
8. Gardner TS, Cantor CR, Collins JJ. Construction of a genetic toggle switch in *Escherichia coli*. Nature 2000; 403:339-42.
9. Kramer BP, Viretta AU, Daoud-El-Baba M, Aubel D, Weber W, Fussenegger M. An engineered epigenetic transgene switch in mammalian cells. Nat Biotechnol 2004; 22:867-70.
10. Isaacs FJ, Hasty J, Cantor CR, Collins JJ. Prediction and measurement of an autoregulatory genetic module. Proc Natl Acad Sci USA 2003; 100:7714-9.
11. Ham TS, Lee SK, Keasling JD, Arkin AP. A tightly regulated inducible expression system utilizing the fim inversion recombination switch. Biotechnol Bioeng 2006; 94:1-4.
12. Ham TS, Lee SK, Keasling JD, Arkin AP. Design and construction of a double inversion recombination switch for heritable sequential genetic memory. PLoS One 2008; 3:2815.
13. Deans TL, Cantor CR, Collins JJ. A tunable genetic switch based on RNAi and repressor proteins for regulating gene expression in mammalian cells. Cell 2007; 130:363-72.
14. Ajo-Franklin CM, Drubin DA, Eskin JA, Gee EP, Landgraf D, Phillips I, et al. Rational design of memory in eukaryotic cells. Genes Dev 2007; 21:2271-6.
15. Dueber JE, Mirsky EA, Lim WA. Engineering synthetic signaling proteins with ultrasensitive input/output control. Nat Biotechnol 2007; 25:660-2.
16. Elowitz MB, Leibler S. A synthetic oscillatory network of transcriptional regulators. Nature 2000; 403:335-8.
17. Stricker J, Cookson S, Bennett MR, Mather WH, Tsimring LS, Hasty J. A fast, robust and tunable synthetic gene oscillator. Nature 2008; 456:516-9.
18. McMillen D, Kopell N, Hasty J, Collins JJ. Synchronizing genetic relaxation oscillators by intercell signaling. Proc Natl Acad Sci USA 2002; 99:679-84.
19. Rinaudo K, Bleris L, Maddamsetti R, Subramanian S, Weiss R, Benenson Y. A universal RNAi-based logic evaluator that operates in mammalian cells. Nat Biotechnol 2007; 25:795-801.
20. Win MN, Smolke CD. Higher-order cellular information processing with synthetic RNA devices. Science 2008; 322:456-60.
21. Anderson JC, Voigt CA, Arkin AP. Environmental signal integration by a modular AND gate. Mol Syst Biol 2007; 3:133.
22. Dueber JE, Yeh BJ, Chak K, Lim WA. Reprogramming control of an allosteric signaling switch through modular recombination. Science 2003; 301:1904-8.
23. Yeh BJ, Rutigliano RJ, Deb A, Bar-Sagi D, Lim WA. Rewiring cellular morphology pathways with synthetic guanine nucleotide exchange factors. Nature 2007; 447:596-600.
24. Hooshangi S, Thiberge S, Weiss R. Ultrasensitivity and noise propagation in a synthetic transcriptional cascade. Proc Natl Acad Sci USA 2005; 102:3581-6.
25. Basu S, Gerchman Y, Collins CH, Arnold FH, Weiss R. A synthetic multicellular system for programmed pattern formation. Nature 2005; 434:1130-4.
26. Sohka T, Heins RA, Phelan RM, Greisler JM, Townsend CA, Ostermeier M. An externally tunable bacterial band-pass filter. Proc Natl Acad Sci USA 2009; 106:10135-40.
27. Friedland AE, Lu TK, Wang X, Shi D, Church G, Collins JJ. Synthetic gene networks that count. Science 2009; 324:1199-202.
28. Win MN, Smolke CD. A modular and extensible RNA-based gene-regulatory platform for engineering cellular function. Proc Natl Acad Sci USA 2007; 104:14283-8.
29. Bayer TS, Smolke CD. Programmable ligand-controlled riboregulators of eukaryotic gene expression. Nat Biotechnol 2005; 23:337-43.
30. Bashor CJ, Helman NC, Yan S, Lim WA. Using engineered scaffold interactions to reshape MAP kinase pathway signaling dynamics. Science 2008; 319:1539-43.
31. Lu TK, Khalil AS, Collins JJ. Next-generation synthetic gene networks. Nat Biotechnol 2009; 27:1139-50.
32. Lim WA. Designing customized cell signalling circuits. Nat Rev Mol Cell Biol 2010; 11:393-403.
33. Khalil AS, Collins JJ. Synthetic biology: applications come of age. Nat Rev Genet 2010; 11:367-79.
34. Smolke CD. Building outside of the box: iGEM and the BioBricks Foundation. Nat Biotechnol 2009; 27:1099-102.
35. Okano H, Kobayashi TJ, Tozaki H, Kimura H. Estimation of the source-by-source effect of autorepression on genetic noise. Biophys J 2008; 95:1063-74.
36. Carlson R. The changing economics of DNA synthesis. Nat Biotechnol 2009; 27:1091-4.
37. Gibson DG, Glass JI, Lartigue C, Noskov VN, Chuang RY, Algire MA, et al. Creation of a bacterial cell controlled by a chemically synthesized genome. Science 2010; 329:52-6.
38. Andrianantoandro E, Basu S, Karig DK, Weiss R. Synthetic biology: new engineering rules for an emerging discipline. Mol Syst Biol 2006; 2:2006 0028.
39. Parsons JB, Frank S, Bhella D, Liang M, Prentice MB, Mulvihill DP, et al. Synthesis of empty bacterial microcompartments, directed organelle protein incorporation, and evidence of filament-associated organelle movement. Mol Cell 2010; 38:305-15.
40. Tanaka S, Sawaya MR, Yeates TO. Structure and mechanisms of a protein-based organelle in *Escherichia coli*. Science 2010; 327:81-4.
41. Klumpp S, Zhang Z, Hwa T. Growth rate-dependent global effects on gene expression in bacteria. Cell 2009; 139:1366-75.
42. Kelly JR, Rubin AJ, Davis JH, Ajo-Franklin CM, Cumbers J, Czar MJ, et al. Measuring the activity of BioBrick promoters using an in vivo reference standard. J Biol Eng 2009; 3:4.
43. Scott CT. The zinc finger nuclease monopoly. Nat Biotechnol 2005; 23:915-8.
44. Mandal S, Sarpeshkar R. Log-domain circuit models of chemical reactions. Circuits and Systems, 2009 ISCAS 2009 IEEE International Symposium on 2009; 2697-700.
45. Cox CD, McCollum JM, Austin DW, Allen MS, Dar RD, Simpson ML. Frequency domain analysis of noise in simple gene circuits. Chaos 2006; 16:026102.
46. Simpson ML, Cox CD, Sayler GS. Frequency domain analysis of noise in autoregulated gene circuits. Proc Natl Acad Sci USA 2003; 100:4551-6.
47. Mettetal JT, Muzzey D, Gomez-Uribe C, van Oudenaarden A. The frequency dependence of osmo-adaptation in *Saccharomyces cerevisiae*. Science 2008; 319:482-4.
48. Tyagi S. Imaging intracellular RNA distribution and dynamics in living cells. Nat Methods 2009; 6:331-8.
49. Andersen JB, Sternberg C, Poulsen LK, Bjorn SP, Givskov M, Molin S. New unstable variants of green fluorescent protein for studies of transient gene expression in bacteria. Appl Environ Microbiol 1998; 64:2240-6.
50. Uttamapinant C, White KA, Baruah H, Thompson S, Fernandez-Suarez M, Puthenveetil S, et al. A fluorophore ligase for site-specific protein labeling inside living cells. Proc Natl Acad Sci USA 2010; 107:10914-9.
51. Malmstrom J, Beck M, Schmidt A, Lange V, Deutsch EW, Aebersold R. Proteome-wide cellular protein concentrations of the human pathogen *Leptospira interrogans*. Nature 2009; 460:762-5.
52. Hamaguchi N, Ellington A, Stanton M. Aptamer beacons for the direct detection of proteins. Anal Biochem 2001; 294:126-31.
53. Lee SK, Chou HH, Pfleger BF, Newman JD, Yoshikuni Y, Keasling JD. Directed evolution of AraC for improved compatibility of arabinose- and lactose-inducible promoters. Appl Environ Microbiol 2007; 73:5711-5.
54. Levskaya A, Weiner OD, Lim WA, Voigt CA. Spatiotemporal control of cell signalling using a light-switchable protein interaction. Nature 2009; 461:997-1001.
55. Boyden ES, Zhang F, Bamberg E, Nagel G, Deisseroth K. Millisecond-timescale, genetically targeted optical control of neural activity. Nat Neurosci 2005; 8:1263-8.
56. Whitehead KA, Langer R, Anderson DG. Knocking down barriers: advances in siRNA delivery. Nat Rev Drug Discov 2009; 8:129-38.
57. Ellis T, Wang X, Collins JJ. Diversity-based, model-guided construction of synthetic gene networks with predicted functions. Nat Biotechnol 2009; 27:465-71.
58. Salis HM, Mirsky EA, Voigt CA. Automated design of synthetic ribosome binding sites to control protein expression. Nat Biotechnol 2009; 27:946-50.
59. Widmaier DM, Tullman-Ercek D, Mirsky EA, Hill R, Govindarajan S, Minshull J, et al. Engineering the Salmonella type III secretion system to export spider silk monomers. Mol Syst Biol 2009; 5:309.
60. Wang HH, Isaacs FJ, Carr PA, Sun ZZ, Xu G, Forest CR, et al. Programming cells by multiplex genome engineering and accelerated evolution. Nature 2009; 460:894-8.
61. Zaslaver A, Bren A, Ronen M, Itzkovitz S, Kikoin I, Shavit S, et al. A comprehensive library of fluorescent transcriptional reporters for *Escherichia coli*. Nat Methods 2006; 3:623-8.
62. Forster AC, Church GM. Towards synthesis of a minimal cell. Mol Syst Biol 2006; 2:45.
63. Parens E, Johnston J, Moses J. Ethics. Do we need "synthetic bioethics"? Science 2008; 321:1449.
64. Bugl H, Danner JP, Molinari RJ, Mulligan JT, Park HO, Reichert B, et al. DNA synthesis and biological security. Nat Biotechnol 2007; 25:627-9.
65. Varela-Rohena A, Carpenito C, Perez EE, Richardson M, Parry RV, Milone M, et al. Genetic engineering of T cells for adoptive immunotherapy. Immunol Res 2008; 42:166-81.
66. Anderson JC, Clarke EJ, Arkin AP, Voigt CA. Environmentally controlled invasion of cancer cells by engineered bacteria. J Mol Biol 2006; 355:619-27.
67. Lu TK, Collins JJ. Dispersing biofilms with engineered enzymatic bacteriophage. Proc Natl Acad Sci USA 2007; 104:11197-202.
68. Lu TK, Collins JJ. Engineered bacteriophage targeting gene networks as adjuvants for antibiotic therapy. Proc Natl Acad Sci USA 2009; 106:4629-34.
69. Kiskinis E, Eggan K. Progress toward the clinical application of patient-specific pluripotent stem cells. J Clin Invest 2010; 120:51-9.
70. Molin S, Boe L, Jensen LB, Kristensen CS, Givskov M, Ramos JL, et al. Suicidal genetic elements and their use in biological containment of bacteria. Annu Rev Microbiol 1993; 47:139-66.
71. Lim GJ, Zare S, Van Dyke M, Atala A. Cell microencapsulation. Adv Exp Med Biol 2010; 670:126-36.
72. Moore GE. Cramming more components onto integrated circuits. Electronics Magazine 1965; 38.