Annual Review of Biophysics

Understanding Biological Regulation Through Synthetic Biology

Caleb J. Bashor[1] and James J. Collins[1,2,3,4]

[1]Institute for Medical Engineering and Science, Department of Biological Engineering, and Synthetic Biology Center, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA; email: cbashor@mit.edu, jimjc@mit.edu

[2]Harvard–MIT Program in Health Sciences and Technology, Cambridge, Massachusetts 02139, USA

[3]Broad Institute of MIT and Harvard, Cambridge, Massachusetts 02142, USA

[4]Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, Massachusetts 02115, USA

Annu. Rev. Biophys. 2018. 47:399–423

First published as a Review in Advance on March 16, 2018

The *Annual Review of Biophysics* is online at biophys.annualreviews.org

https://doi.org/10.1146/annurev-biophys-070816-033903

Keywords

synthetic biology, regulatory network, synthetic gene circuit, engineering cycle, motif, refactoring

Abstract

Engineering synthetic gene regulatory circuits proceeds through iterative cycles of design, building, and testing. Initial circuit designs must rely on often-incomplete models of regulation established by fields of reductive inquiry—biochemistry and molecular and systems biology. As differences in designed and experimentally observed circuit behavior are inevitably encountered, investigated, and resolved, each turn of the engineering cycle can force a resynthesis in understanding of natural network function. Here, we outline research that uses the process of gene circuit engineering to advance biological discovery. Synthetic gene circuit engineering research has not only refined our understanding of cellular regulation but furnished biologists with a toolkit that can be directed at natural systems to exact precision manipulation of network structure. As we discuss, using circuit engineering to predictively reorganize, rewire, and reconstruct cellular regulation serves as the ultimate means of testing and understanding how cellular phenotype emerges from systems-level network function.

Contents

1. INTRODUCTION: THE DUAL PURPOSE OF SYNTHETIC GENE CIRCUIT ENGINEERING

Synthetic biology is a field of research that aims to establish engineering rules for the forward synthesis of cellular function (43). It was launched in the early 2000s with the publication of two artificial gene regulatory circuits: a genetic toggle switch (39) and a gene expression oscillator (termed repressilator) (28). While the field has diversified in recent years (60, 85), heavy focus continues to be placed on regulatory circuit design and construction as a means to establish user-defined control of cellular function (12). The term circuit is used generically but can be defined as a cellular regulatory program with specific molecular inputs, an intermediate information-processing step, and measurable output. While gene circuit engineering can vary in scope from tinkering with small, synthetic test-bed networks to the construction of complex multi-component circuitry, the process of reaching a desired target behavior is often mired by unpredictable, nonlinear interactions both between circuit components and with the encompassing cellular milieu. As a result, gene circuit engineering is typically iterative and may involve numerous cycles of design, construction, and testing as the bioengineer attempts to move from an initial, abstract description of a circuit's behavior to its experimentally verified, physical implementation. At each step of this circuit engineering cycle, the bioengineer must continuously engage with the allowances and constraints imposed by the underlying biology of the system. Thus, while realizing designed function is usually the stated goal of a gene circuit engineering project, extracting a deeper understanding of biology is often an unavoidable by-product (30).

Synthetic biology: engineering discipline that uses forward engineering to establish control over cellular behavior

Engineering cycle: a single iteration in the circuit engineering process whereby a synthetic circuit design is constructed, tested, and evaluated for its ability to achieve a desired target behavior

We acknowledge that the motivating vision for synthetic biology has always been centered on creating useful—even transformational—technological applications. Progress in engineering regulatory circuits to address problems in medicine (91), metabolic engineering (71), and materials

science (88) has been well reviewed elsewhere. Here, we instead highlight research that embraces using synthesis as a tool for learning about biology and describe how the process of engineering synthetic gene circuits and network connections can enrich our understanding of the organization, function, and evolutionary design of natural regulatory systems.

We begin by discussing how foundational work engineering gene circuits—mostly in prokaryotes—provided quantitative insight into the biology of gene regulation, establishing the circuit engineering cycle as a process that can both complement and motivate basic research. These early lessons have inspired ongoing work on gene and signaling circuit engineering in mammalian systems. Our growing ability to freely manipulate network connections has enabled hypothesis-driven rewiring of native regulatory pathways. As such, circuit engineering tools can be thought of as precision-designed alleles that can be used for making systematic, targeted genetic perturbations. We highlight how synthetic circuits can be fashioned into sophisticated reporter modules capable of converting obscure biological activity into observable, recordable data. What has emerged is a view of gene circuit engineering as a dynamic and essential partner for systems biology capable of systematically reconfiguring, augmenting, or even replacing native networks in order to understand the adaptive forces shaping their design. In this capacity, using synthetic systems may prove the best and only way to understand determinants for complex phenotypes like cell differentiation, morphology, and fitness. We close by describing a far-reaching vision for advancing synthetic circuit engineering toward the long-term goal of system-wide synthetic control of complex phenotype.

Synthetic gene circuit: a unit that features an identifiable molecular input, an intermediate information processing step, and a molecular output

Regulatory network: a physically interlinked collection of molecules used to transmit information intracellularly

Motif: a recurring pattern of molecular connectivity found within networks

2. SYNTHETIC GENE CIRCUITS AS PHYSICAL MODELS FOR STUDYING REGULATORY FUNCTION

A key principle to arise in the contemporary, postgenomic era is the idea that biological systems are organized according to a scalable, modular hierarchy (46, 63). Systems biology has revealed that cellular regulatory networks are constructed from a common set of molecular building blocks and further organized into recurring patterns of connectivity (69). These network motifs often appear in analogous physiological contexts, implying a modularity of function (46). Thus, it would appear that regulatory network behavior is ultimately an emergent property of the arrangement of constituent motifs (1).

The gene circuit engineering process has arisen as a way to test this hierarchical view of cellular regulation by directly evaluating the physical determinants of motif function (**Figure 1**) (5). A critical first step in this process involves searching for molecular parts, usually by examining native networks to identify molecular components that can be modularized, abstracted, and reconnected (31). Through the process of part identification, a bioengineer implicitly evaluates how regulatory connections may have evolved since modular, tunable circuit components may correspond to evolvable features in natural networks (1). After working out a construction scheme, a motif can be reconstituted by establishing user control over circuit input (e.g., inducing transcription with an exogenously administered small-molecule inducer) and connecting it to a measureable output [e.g., a transcriptional green fluorescent protein (GFP) reporter]. Function can then be directly tested in quantitative fashion, and the design can be iteratively adjusted until the desired behavior is achieved (**Figure 1**).

Circuit engineering is typically guided by mathematical models in which parameters represent properties like reaction rate, species concentration, and interaction affinity (58). Model-based, a priori notions of gene circuit behavior may guide initial designs but are usually adjusted following testing and redesign. Achieving close correspondence between model parameters and circuit components creates a powerful experimental framework where questions regarding motif behavior can

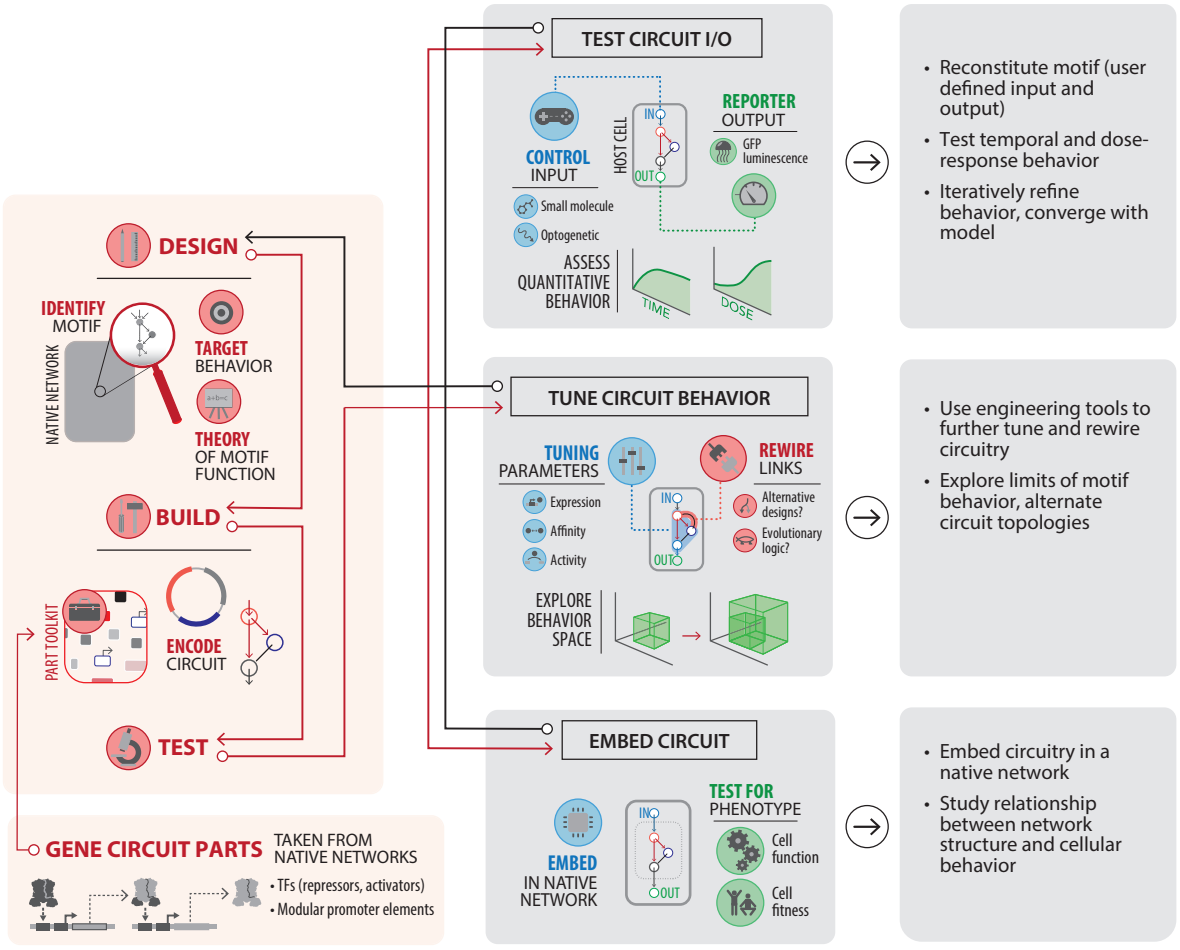


Figure 1

The synthetic gene circuit engineering cycle. Synthetic circuit engineering involves iterative cycles of design, building, and testing, where insight from one cycle is incorporated into the next as a bioengineer attempts to achieve a target behavior. Inception of a circuit engineering project is inherently hypothesis driven; circuits are designed with specific target behaviors in mind, often with natural circuit motifs as inspiration. By attempting to create a synthetic version of the motif, functional modularity is implicitly evaluated. Theory of motif function may inform initial circuit designs, which are built from sets of molecular parts culled from native networks. Initial testing of circuit input/output function may involve connecting to a user-defined input and measurable output, permitting evaluation of both dose- and time-dependent behavior. Convergence of circuit design and model prediction results in a model system for studying regulatory properties of the motif; tuning the circuit through changing the strength of the network, or rewiring with new links, allows systematic assessment of the range of behaviors available to the circuit. Embedding a tunable synthetic circuit within a native network grants insight into the relationship between circuit tunability and phenotype and can be used to investigate relationships between network structure and cell behavior. Through the process of obtaining target behavior, design and testing can feed back (*black lines*) to reshape circuit design, challenging and growing our understanding of regulatory network function. Abbreviations: GFP, green fluorescent protein; I/O, input/output; TF, transcription factor.

be addressed in systematic, quantitative fashion (**Figure 1**): How robust is behavior to parameter tuning? How do molecular properties of the parts shape or constrain behavior? By expanding parameter space with new parts, or rewiring the circuit with new connections, what new behavior regimes can be accessed? Thus, synthetic gene circuits not only serve as physical models for exploring the limits of motif behavior but also permit evaluation of the evolutionary logic of motif

design through comparison with designs not selected by nature. Once the relationship between design and behavior is well understood, a gene circuit can be embedded within a natural network by wiring it to relevant inputs and outputs (**Figure 1**). In this context, the tunable circuit becomes a means to probe the relationship between network structure and phenotype, and different circuit configurations can be tested for their impact on cellular behavior and fitness.

RBS: ribosome binding site

TF: transcription factor

2.1. Insights into Prokaryotic Gene Regulation Through Iterative Circuit Engineering

The engineering of small, synthetic gene circuits in prokaryotes—an area of both foundational and ongoing research—is particularly illustrative of how iterative cycles of design and construction can yield new biological insight (15). Early progress in this area was facilitated by a detailed understanding of the functional modularity of native gene networks in *Escherichia coli* (84). This provided the field with a ready-made collection of promoters, transcriptional repressors, activators, and copy-number-controlled plasmids in which tunability, composability, and modularity were prevalidated (**Figure 2**) (67), enabling facile construction, tuning, and modeling. Design of target behavior for both the toggle and repressilator circuits—bistability and periodic oscillation, for the toggle and repressilator, respectively—was model driven and inspired by natural behaviors: Bistability regulated by mutual inhibition was long understood to underpin the lysogenic cycle in bacteriophage (84), while a ring oscillator architecture is a well-known feature of circadian clocks in photosynthetic bacteria (26). Initial circuit designs functioned poorly and had to be iteratively adjusted to realize target behaviors by changing the ribosome binding site (RBS) sequence to balance transcription factor (TF) translation for opposing arms of the toggle and decreasing TF lifetime by appending degradation tags for the repressilator. Together, these studies served as proof of concept that nonlinear behavior could be reconstituted with molecular components not naturally associated with one another and that an area of productive behavior space could be pinpointed by rationally tuning component features within a synthetic design framework.

One phenomenon underscored by both the toggle and repressilator was the role of molecular noise in limiting circuit performance, as apparent cell-to-cell variations in constituent molecular species resulted in nonuniform circuit behavior and led to eventual decay of steady-state behavior. Circuit engineering thus presented both the motivation and experimental means to investigate the molecular origins of gene expression noise—a topic given mostly theoretical treatment up until that point (68). One early inquiry into noise in bacterial gene expression utilized a simple genomically integrated, LacI-regulated GFP reporter in *E. coli* (74). By differentially tuning transcriptional and translational rate (by respectively derepressing LacI with a small-molecule inducer and tuning RBS strength), the authors determined that variability in protein production was the main source of noise. Subsequent work performed in yeast used a similar approach but identified transcript initiation as the primary source of noise (9). The use of single-cell fluorescence reporters to separate noise into intrinsic (promoter-to-promoter fluctuations) and extrinsic (global, cell-to-cell fluctuation) components (29) laid the groundwork for using synthetic circuits to study how noise is propagated within gene networks (81) and resolve contributions to noise in gene regulation (90).

The idea that feedback and feed-forward circuitry could play an information processing role in regulatory networks was first hypothesized by Jacob & Monod (54) in 1961. Subsequent to this insight, theoretical concepts emerged regarding the role of positive and negative autoregulation in gene expression stability (23) that could be directly tested only with the advent of synthetic gene circuit engineering. Becskei & Serrano (7) constructed a simple negative autoregulatory circuit that demonstrated reduced stochastic fluctuations when single-cell expression was measured,

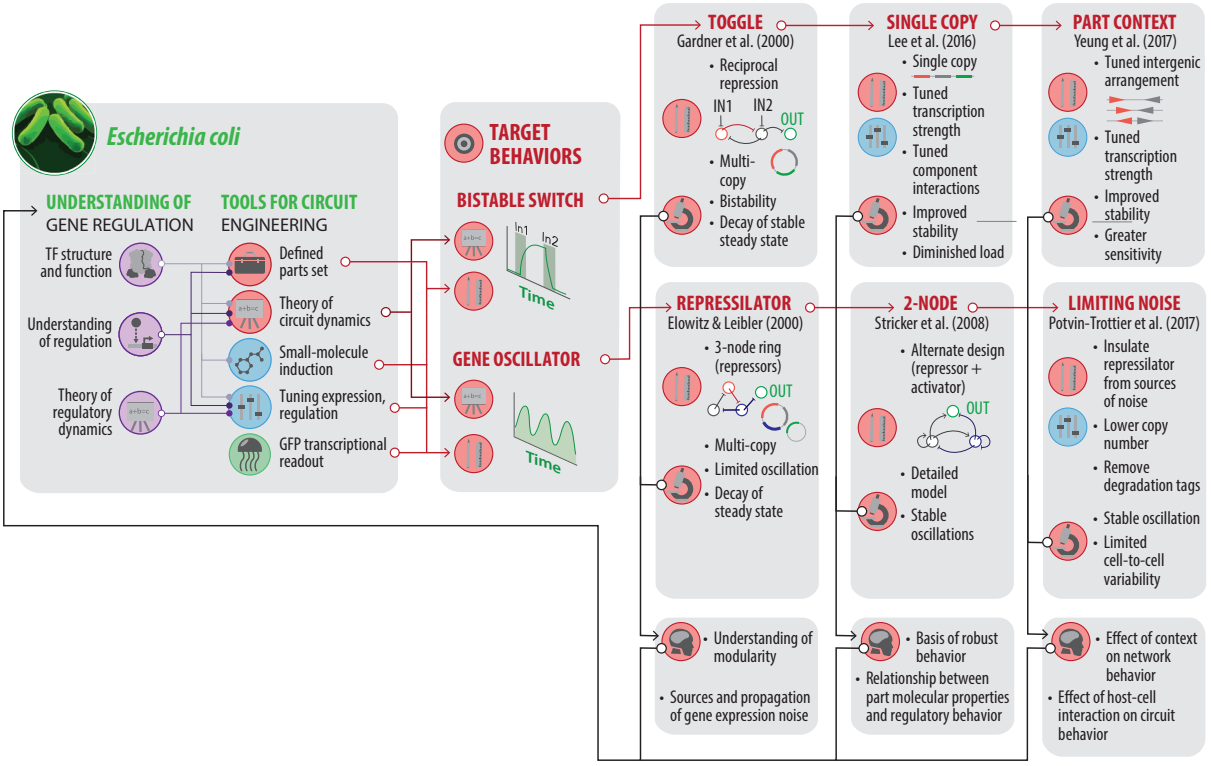


Figure 2

Learning through iterative engineering of model prokaryotic gene circuits. Successive efforts to construct a bistable genetic toggle switch and oscillator circuits have resulted in progressive refinement of gene circuit engineering, leading to concomitant insight into design principles of prokaryotic gene regulation. For both circuits, initial designs drew heavily from historical research into gene regulation in *Escherichia coli*. Deep understanding of the molecular basis of transcription factor (TF)-mediated promoter regulation and theory of gene expression dynamics enabled gathering of parts that could be used to design circuitry embodying simple nonlinear behavior: inducible bistability and stable, periodic oscillations. Inaugural designs (toggle and repressilator) both successfully captured target behaviors by part tuning but failed to demonstrate robust, long-lived steady-state behavior. Success (and shortcomings) of these initial designs motivated inquiry into diverse topics, including sources of gene expression noise and the role of feedback in network dynamics. Next-generation designs used more precise approaches; utilization of more varied parts and updated modeling approaches improved behavioral robustness. These advances included engineering a single-copy toggle switch and, in the case of the oscillator, utilization of an alternative circuit design. More recent design iterations have highlighted the role of circuit context and host–circuit interactions in regulatory network behavior.

while positive autoregulatory feedback was used as a way to achieve population bistability (6, 52). Cumulatively, understanding of the basis for both noise propagation and feedback control yielded models capable of a priori prediction of simple circuit behavior (44).

A number of recent studies have utilized the toggle and the oscillator as test-bed circuits for continued exploration of network design principles. These updates illustrate an underappreciated feature of the field: Beyond design cycles *within* a project, insight can be developed *between* successive projects, as circuit designs are refined over many years (**Figure 2**). Recently, engineering of a single-copy, genomically integrated version of the toggle switch was reported (65). Using both experiment and modeling, the authors demonstrate that the single-copy toggle, which functions at lower component expression levels, is more stable and less prone to mutation—a refinement that highlights design features potentially important for epigenetic state maintenance in

natural systems. Another recent study used toggle circuit engineering to demonstrate how promoter position and orientation are important network design features (107).

Publication of a robust oscillator (97) utilized a design containing opposing positive and negative feedback loops (**Figure 2**). As predicted by theory (47), enhanced stability and tunability of this design compared to the repressilator were borne out by experiment. Subsequent work used engineering design principles to program oscillatory behavior at the population (within a biofilm) (22) and community (between cocultured strains) scales (21). In a more recent study, the repressilator was revisited by simplifying the original circuit design to limit sources of noise; by removing degradation tags and limiting plasmid copy number, stable oscillations could be achieved (82). Toggle and repressilator updates also highlight an initially overlooked aspect of circuit engineering: that interactions between circuit and host cell network components can impinge on circuit function in ways that initial designs often cannot anticipate. Additionally, components from different circuits can interact with each other through shared host machinery. Hasty and colleagues (83) treated this as a feature, rather than a bug, demonstrating that the period and phase of oscillating circuits can be coupled through shared protein degradation machinery.

2.2. Decomposing Mammalian Gene Regulation with Synthetic Circuits

Transcriptional regulation in eukaryotic cells takes place on multiple organizational scales, with contributions from transcription factor binding dynamics, chromatin regulation, and three-dimensional (3D) genome organization all conspiring to regulate promoter activity. Inability to disentangle the integration of these different regulatory modes precludes a systems-level understanding of complex processes like context-dependent cell state maintenance (108) and cell differentiation during tissue development (33). As a handful of recent studies have shown, it may be possible to use engineered circuits to decompose overlapping regulatory contributions.

As with prokaryotes, network structure in eukaryotes is an outgrowth of promoter architecture. In both cases, *trans*-regulation is accomplished by TF binding, but eukaryotic promoters feature multivalent, often cooperative binding of TFs to upstream enhancer sequences—a feature that underlies greater network density and regulatory complexity. Recently, a circuit engineering scheme inspired by eukaryotic promoter design was reported in yeast (61). It featured a set of synthetic zinc-finger transcription factor (ZF-TF) activators engineered to bind synthetic, orthogonal enhancer sequences upstream of a generic core promoter. Transcriptional output could be tuned by adjusting either ZF-TF affinity or the number of enhancers. Appending protein–protein interaction domains to tandem-bound ZF-TFs led to cooperative binding, enabling programming of Boolean-like logic behaviors. Since it simulates ubiquitous features of eukaryotic TF-activator regulation, this platform should scale to metazoan systems, enabling the engineering of multinode networks that approach natural-like sophistication. Furthermore, by leveraging cooperative binding in the upstream promoter complex, programming TF binding assemblies capable of performing complex computations should be possible (75).

The clearest point of divergence between prokaryotic and eukaryotic gene regulation comes from the organization of eukaryotic DNA into chromatin (59), which is associated with epigenetic memory and the maintenance of heritable, self-reinforcing states of transcriptional activation and repression. A diverse array of chromatin regulators (CRs) add, remove, and are recruited through chemical marks that control histone packing and regulate binding and activity of TFs (55). Because of the number of regulatory states that can arise from combinations of modifications, TF occupancies, and spatial configurations, even a relatively simple eukaryotic promoter can adopt a myriad of regulatory states. Our current understanding of this variation comes from extensive genome-wide mapping of the concurrence of histone modification with CR, TF, and histone

occupancy (32), but details of how these configurations are decoded into a transcriptional output remain obscure.

A pair of recent studies took a bottom-up synthetic approach to understanding the relationship between chromatin regulatory state and transcription, engineering simple test-bed circuits where inducibly expressed CRs were recruited to fluorescent transcriptional reporter loci. The first of these studies, performed in yeast, tested the effects of recruitment of a library of ZF-TF-CR fusions for their ability to either activate or repress reporter transcription (59). When recruited in tandem with a transcriptional activator, CRs repress, enhance, or synergize with the activator. Spatial effects of the CRs were tested: While activating CRs were found to act when recruited to promoters, repressive CRs could regulate at a distance.

In a second study, chromatin regulation dynamics were investigated using transient recruitment of different CR repressors to a reporter locus on a human artificial chromosome (8). The authors found that both repression and reactivation (upon small-molecule washout) proceeded in an all-or-none fashion, and the fraction of cells repressed was proportional to recruitment duration. Postwashout time-course measurements revealed behavioral differences between the repressors, leading the authors to propose a simple model explaining the observed dynamics in terms of transitions between one of three states: active, reversibly silenced, and permanently silenced. As the authors demonstrate, some individually recruited repressors can push promoter regulation into distinct states, while others can yield mixed reversible and irreversible, with the ratio dependent on duration of recruitment.

Despite progress, gene circuit engineering strategies commensurate with the sophistication of mammalian transcriptional regulation are still at an early stage of development. To date, most engineered applications have relied on binary regulatory connections, often by superimposing bacterial TFs onto eukaryotic promoters (100). Studies described above suggest a framework more congruent with natural network design could eventually be realized, combining TF activation with chromatin-mediated regulation, integrating the ability to make arbitrary network connections with epigenetic control over expression timescale and chromosome structure.

2.3. Using Synthetic Circuits to Understand Posttranslational Signaling

Cells—especially those of higher-order taxa—process external signals using networks of signaling proteins that connect upstream receptor activation to downstream outputs, including gene expression, metabolic regulation, and cytoskeletal rearrangement. Designing synthetic signaling circuits is challenging, requiring precise accounting of expression levels, interactions, and activities of freely diffusing pathway components (40). As a result, posttranslational circuit engineering has moved at a much slower overall pace than for gene circuits, with modular parts collection and generalizable design principles still under development.

As with gene circuits, a prerequisite activity for signaling circuit engineering is assessing what commonly repeated, potentially engineerable parts, modules, and motifs can be found to constitute native networks. Reversible chemical modifications are the primary currency of information transfer in most signaling networks. Phosphorylation is the most common, but other examples include acetylation, methylation, and ubiquitination (102). Posttranslational modifications carry information to targets by altering their activity, localization, stability, or interaction specificity. Both addition and removal of marks are catalyzed by enzymes arranged in opposing, push–pull fashion (see the sidebar titled The Push–Pull Cycle as the Fundamental Engineering Unit of Posttranslational Signaling Circuits). Signaling proteins typically contain multiple domains that fall into one of two categories: Catalytic domains harbor the activities that make or remove modifications, while interaction domains account for pathway wiring by specifying localization and

THE PUSH–PULL CYCLE AS THE FUNDAMENTAL ENGINEERING UNIT OF POSTTRANSLATIONAL SIGNALING CIRCUITS

In protein signaling pathways, addition and removal of chemical modifications are catalyzed by pairs of antagonistic enzymes. Phosphorylation, the most common modification, is catalyzed in the forward direction by a kinase and by a phosphatase in the reverse direction. Other examples include acetylation and methylation, which play a role in regulating chromatin state. Opposing enzymes, along with their mutual target, constitute a so-called push–pull motif (24). Changes in forward and reverse reaction rates within a push–pull (e.g., resulting from regulatory input that alters either activity) result in a rapid shift in equilibrium between modified and unmodified target. Both the rate and magnitude of this shift, as well as the rate at which the original equilibrium is restored upon input removal, are dependent upon the molecular features of a push–pull, which include component concentration, interaction affinity, and activity. From an engineering perspective, this correspondence between molecular configuration and quantitative behavior makes the push–pull attractive as a modular building block, and we argue that it should be considered the fundamental unit for posttranslational circuit construction. Circuits built from interconnected phosphorylation push–pulls could be configured to receive upstream input from a receptor or be connected to an output by activating transcription, or by inducing cytoskeletal rearrangement or targeted secretion. Synthetic push–pulls could be linked together to form higher-order information processing circuitry; they could be vertically arranged by coupling target modification of an upstream unit to forward or reverse activities of a downstream unit or, if multiple pairs acted on a single target, programming sophisticated logic functions by coupling combinations of phosphorylation events to an output may be possible.

protein–protein interactions (80). The latter include domains that mediate unregulated interactions [e.g., SH3 (Src homology 3), LIM (Lin11, Isl-1, and Mec-3), and PDZ (postsynaptic density protein, *Drosophila* disc large tumor suppressor, and zonula occludens-1 protein)] but also those where binding is regulated by modifications (e.g., SH2, 14–3–3, and WW that specifically recognize phosphopeptides) (80).

A number of pioneering studies have demonstrated that domain recombination can be used as a design principle for creating new connections in signaling networks, in much the same way that modular promoter elements wire gene networks together through TF *trans*-regulation. Through novel combinations of catalytic and interaction domains, protein switches have been engineered to accept novel inputs and perform elementary Boolean computations (25) and even establish synthetic control over cytoskeletal morphology (106). Adapters and scaffolds are proteins composed of interaction domains that regulate signaling by organizing pathway components into complexes. Several studies have shown that pathway input/output relationships can be reprogrammed by engineering these proteins to accommodate new interactions. One study rewired scaffold-regulated osmolarity and mating response pathways in yeast (79). In another, a normally proliferative epidermal growth factor receptor–mediated input was rerouted to a cell death output via an engineered adapter protein (50).

The relative ease with which pathway connectivity can be synthetically altered suggests that domain recombination may have driven diversification of signaling pathway connectivity during evolution. This possibility was reinforced in a study where response dynamics were reshaped using engineered scaffold recruitment to overlay synthetic feedback circuitry onto the yeast mating pathway (4). This was accomplished by coupling pathway induction to the transcription of positive and negative modulators recruited to the Ste5 scaffold through synthetic interactions. By tuning a limited set of parameters (transcription strength, recruitment affinity, and the sign and order

of the effector recruitment), the authors made feedback circuit connections that radically altered pathway behavior. In a follow-up study, synthetic phospho-regulated (fast) positive feedback was combined with positive transcriptional feedback (slow) to engineer temporally sensitive switching behavior; short-duration inputs were amplified by the fast loop and converted into a permanent memory state by the slow one (41).

Continued progress of posttranslational circuit engineering will be driven by development of diagnostic biosensing (94) and cell-based therapeutics (35). By coupling synthetic signaling circuitry to gene expression, secretion, and cell migration, bionengineers could program cells to sense features of their chemical and mechanical environment and respond in a therapeutically relevant manner through movement, secretion, or differentiation. As the repertoire of parts and design practices used for posttranslational circuit engineering matures, two general strategies could be adopted: (*a*) a bottom-up strategy, where engineered protein components are wired together to construct signaling pathways de novo (while an enabling part toolkit has yet to be developed, adopting such a model-pathway approach would grant precise control of molecular composition and network structure, allowing exploration of the molecular determinants underlying signaling behavior); and (*b*) a more top-down approach, where whole pathway modules, with known signal processing attributes, are appropriated via rewiring of inputs and outputs (such a strategy could be used to investigate the degree to which modules retain their signal processing features when placed in a different context). In an early example of this method, a scaffolded human mitogen–activated protein kinase cascade was orthogonally expressed in yeast (73). Adjusting stoichiometry of both kinases and scaffold, and introducing negative regulation, allowed for tuning of pathway transfer function in model-predicted fashion, suggesting that transplanted modules placed under synthetic control are capable of a high degree of behavioral plasticity.

3. SYNTHETIC GENE CIRCUITS AS TOOLS FOR PERTURBING AND MONITORING NATURAL SYSTEMS

Extensive mapping of cellular networks using the tools and methods of systems biology has resulted in a near-complete inventory of cellular contents and an increasingly accurate map of the interactions between them (3). Unfortunately, our ability to identify network connections has dramatically outpaced our understanding of how connectivity gives rise to function. In this section, we discuss how the techniques of network construction and manipulation that have been developed for synthetic circuit engineering can be appropriated as tools for studying native cellular regulation. This includes the ability to wire in new connections, tune existing ones, or integrate entire circuit modules in cellular networks to report on their behavior.

3.1. Synthetic Rewiring of Natural Regulatory Networks

Traditionally, the study of regulatory network function has been conducted through reverse engineering. Understanding of *how* a specific motif or pathway works has often superseded the equally important question of *why* evolution may have chosen a particular design. Using circuit engineering tools to modify or tune network connectivity precisely and in place provides an experimental means to test both the quantitative behavior and phenotypic consequences of evolutionarily plausible, alternative network topologies (**Figure 3*a***). Furthermore, assessing the behavior and fitness of rewired variants can offer clues as to why certain topologies evolved and even what features of the environment may have driven their selection.

The use of synthetic rewiring to understand natural network design was superbly illustrated in a series of studies examining the gene circuitry underlying transient switching between competence

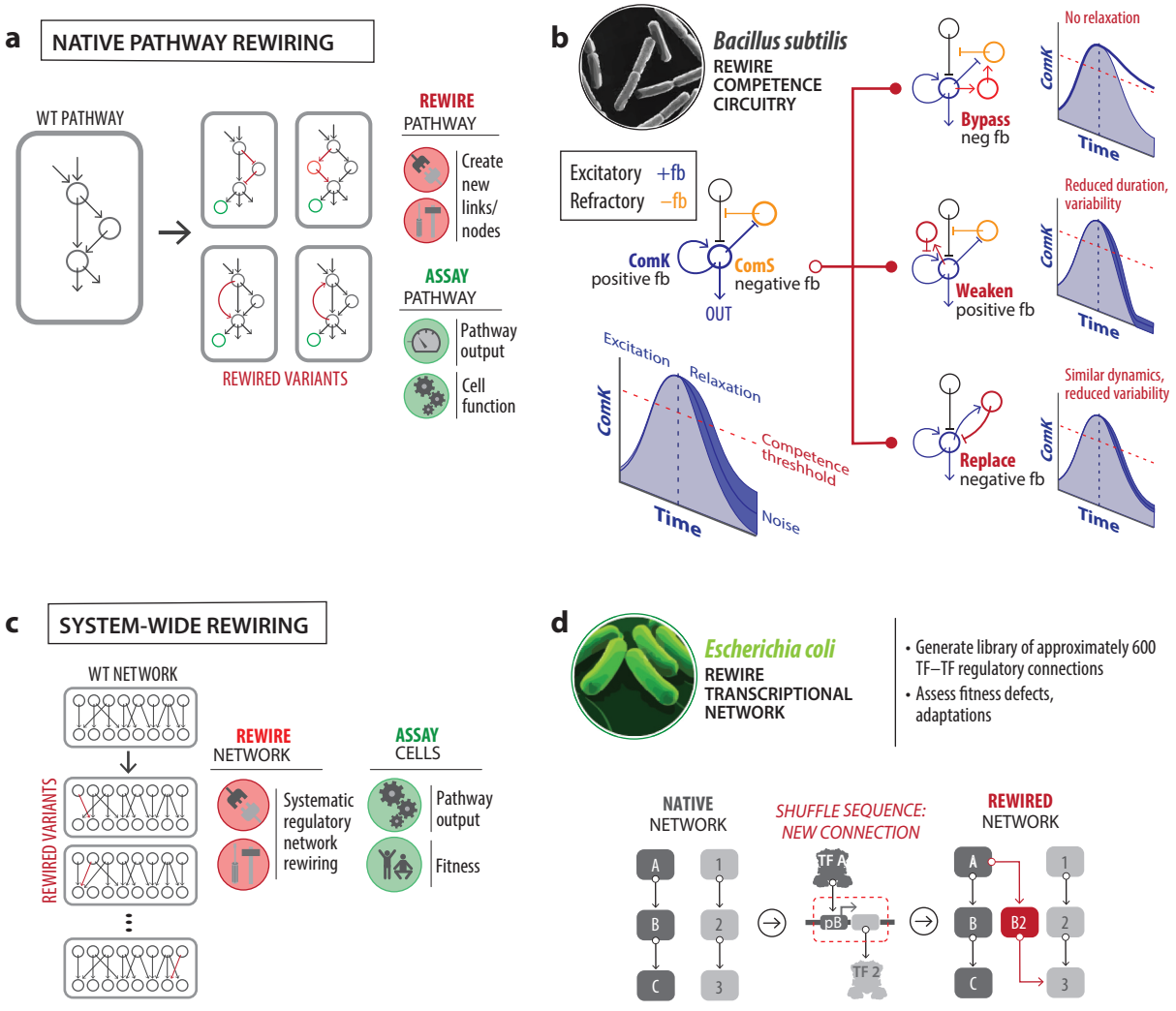


Figure 3

Rewiring native regulatory networks. (*a*) By wiring new linkages, circuit engineering tools can be used to explore the functional consequences of alternate network wiring topologies. (*b*) Rewiring the transcriptional network in *Bacillus subtilis* that mediates competence. Spontaneous, noise-driven elevation in the activity of the master transcriptional regulator ComK results in activation of competence, followed by slow relaxation back to basal activity (sporulation). Dynamics of these excitatory and refractory phases are mediated by positive feedback (*blue*) and negative feedback (*yellow*) circuits, respectively. Various rewiring schemes were used to determine the relationship between circuit architecture and function; rewiring to bypass the negative feedback loop eliminated the refractory phase, whereas adding in a new negative feedback loop to weaken the positive loop resulted in a faster refractory phase while diminishing the variability in the normally noisy refractory phase. Replacing the negative feedback loop with a simpler, protease-mediated loop resulted in similar dynamics but diminished variability. (*c*) Circuit engineering tools can be scaled to phenotypically screen network wiring variants. (*d*) Systematic rewiring of transcriptional regulatory network in *Escherichia coli*. A library of novel network linkages created by shuffling coding sequences of various transcription factors and sigma factors against corresponding promoters. Library members showed few fitness defects and minimal changes in genome-wide transcription. For some members, enhanced fitness was demonstrated under selective conditions, suggesting that the native *E. coli* transcriptional network may be generally robust to spontaneous rewiring that yields specific adaptive advantages. Abbreviations: fb, feedback; pB, promoter B; TF, transcription factor; WT, wild type.

and vegetative growth states in *Bacillus subtilis* (13, 98, 99) (**Figure 3*b***). On the basis of data from time-resolved fluorescence microscopy analysis of single cells, the authors proposed a simple model for switching in which a master regulator, ComK, is activated by an excitatory positive feedback loop, leading to a transient "on" state that slowly relaxes through subsequent action of a refractory negative feedback loop. Fluctuations in gene expression can stochastically trigger activation and also create variability in the distribution of relaxation rates. To verify their model, the authors built several circuit variants: They (*a*) added a positive feedback loop that circumvents the negative feedback loop results in permanent activation (98), (*b*) added a stronger negative feedback loop, which results in a shorter refractory period and reduced population variation in the relaxation time (99), and (*c*) changed the regulatory details of the negative feedback loop (swapping transcriptional repression for protein degradation), reducing variability in competence duration (13). Cumulatively, these results argue that competence switching is determined by the intrinsic timing and noise of the network, and while alternate circuit wirings may be adaptive under particular conditions, noise in the wild type network probably evolved as a bet-hedging strategy for balancing growth with DNA absorption.

Refactoring: for a set of related genes, engineering the replacement native regulation with user-defined, synthetic control

For many years, it was hypothesized that the dynamics of reactivation of latent HIV was driven by positive feedback of the virally encoded transcriptional activator tatA. To investigate this possibility, Weinberger et al. (104) engineered a YFP transcriptional reporter driven by a synthetic tatA autoregulatory circuit. Following genomic integration of the circuit using a lentiviral vector, a combination of fluorescent microscopy and modeling was used to demonstrate that fluctuations between on and off states in the circuit are sufficient to drive a phenotypic bifurcation between latent and active HIV. In a subsequent study, the authors (87) used the same approach to investigate whether activation is triggered by sensing the state of the host cell. By tuning their synthetic latency circuit in a manner that decoupled its dynamics from endogenous cellular machinery, the authors showed transactivation dynamics are, in fact, an intrinsic property of the circuit itself. This surprising result suggests that HIV may have evolved to regulate latency using a bet-hedging strategy, the dynamics of which are agnostic to the physiological state of its host.

Using synthetic rewiring approaches to study natural regulatory networks will become increasingly popular as familiarity of the biological research community with circuit engineering tools grows. Refactoring, an extreme form of rewiring that involves complete removal of native regulatory control over a set of related genes (e.g., a gene cluster or signaling pathway), has been largely application driven but could be used as a tool for decoupling the phenotypic contributions of regulation and expression (see the sidebar titled Refactoring Regulatory Systems to Probe and Perturb Regulation). In recent years, one notable addition to the circuit engineering toolkit is CRISPR (clustered regularly interspaced short palindromic repeats)-Cas9 (CRISPR associated protein 9), which has been used as tool to recruit transcriptional machinery to specific genomic loci (20, 48, 86) and to construct multi-node circuitry (57, 62). In the future, CRISPR-based regulation could be used to wire arbitrary linkages between native loci, permitting systematic exploration of alternate network topologies. Since many complex metazoan phenotypes, like cell differentiation and morphology, are thought to evolve from adaptive variation within a defined molecular and regulatory framework (17), network rewiring could be used to address questions regarding the regulatory basis for phenotypic plasticity and evolutionary optimality.

3.2. Constructing Circuit Engineering Tools for Precision Genetic Screening

To a great extent, confidence in what can be learned from reverse genetic screening—observing phenotype arising from targeted genotypic perturbation—is a matter of precision in both the method used to generate alleles and the resolution of the screening readout. When implementing

REFACTORING REGULATORY SYSTEMS TO PROBE AND PERTURB REGULATION

Refactoring is the process by which a set of related genes are removed from their native regulatory context and placed under synthetic expression control. Refactoring a system establishes the equivalent of a molecular tuning board by eliminating pretranslational regulation of gene expression, recoding constituent genes for optimized translation, and placing them under control of orthogonal promoters of adjustable expression strength. Using gene synthesis combined with combinatorial assembly techniques dramatically augments exploration of regulatory space for a refactored system, allowing for simultaneous tuning of component expression levels and introduction of alternate or mutant alleles. Additionally, wiring user-defined regulatory links specifying input control, intrasystem regulatory connections, or orthogonal output becomes possible. Bacterial gene clusters, which can encode related sense-and-response elements, transcription factors, and biosynthetic enzymes, have been an early target for synthetic refactoring (36). Gene clusters are good targets owing to often opaque regulatory features that result from close proximity and potential coregulation of constituent open reading frames. Recently, the *Nif* operon, which encodes nitrogen-fixing machinery from *Klebsiella oxytoca*, was refactored and expressed orthogonally in *Escherichia coli* (95). After placing all 16 genes under synthetic expression control, the authors tested several hundred variants, finding that their best variant was able to recover most of the nitrogen fixing capacity of the WT (wild type) cluster. Studying mammalian signaling using pathway refactoring to encode a posttranslational network may be possible using a similar approach.

genetic screens, well-constructed controls and careful interpretation of results may be required to weed out false positive or negative hits. Generally, the more precise or targeted the allele, the less likely that data are clouded by pleiotropic effects. Traditionally, genetic screens that utilize gene knockouts, targeted mutations, or overexpression may fail to decouple an allele's phenotype from the secondary effects of genotype creation. Use of RNA interference is one classic solution to this problem—knocking down transcripts decouples expression from the consequences of physically disrupting a gene's expression locus (45). Synthetic circuitry could be used to improve readouts for genetic screens by suppressing noise or to improve resolution of allelic effects by amplifying small changes in output.

Through precise and quantitative control over gene expression, circuits can be used to control the timing and dose of allele presentation. For example, expression of an unstable or toxic allele can be kept caged until the user is ready to assess phenotype. In one recent implementation of this strategy, an orthogonal, inducible protein degradation system was developed for use in bacteria (16). This was accomplished by appropriating Lon protease–specific transfer-messenger RNA degradation tags from *Mesoplasma florum*. Screening of a small library of tags yielded variants that were recognized by *mf*-Lon but not by *E. coli* ClipXP/AP and native Lon. As the authors convincingly demonstrate, induction of *mf*-Lon facilitated rapid and specific degradation of endogenous proteins bearing the tag. Moreover, degradation was tunable, allowing for tightly controlled titration of protein expression levels. A library was composed by appending degradation tags to 238 of 305 essential *E. coli* genes and screened for antibiotic sensitivity, identifying several potential new drug targets.

Network wiring variants can be systematically generated and screened to address important evolutionary questions (**Figure 3*c***). For example, gene duplication and recombination events are thought to give rise to new regulatory connections (72), but whether intermediate alleles provide any adaptive advantages is unknown. Serrano and colleagues (53) addressed this issue by rewiring the *E. coli* transcriptional regulatory network (**Figure 3*d***). To generate new network linkages,

the authors generated an ~600-member plasmid library containing interchanged combinations of promoter and coding sequences selected from genes in which one TF regulates the transcription of another. When introduced into cells, a majority of library members were tolerated; few showed growth defects, while several actually demonstrated increased fitness under certain selective pressures. Thus, the *E. coli* transcriptional network appears robust to individual rewiring events, suggesting recombination through duplication may allow gene networks to explore new, potentially adaptive configurations while incurring little fitness cost.

3.3. Using Synthetic Circuits to Report on Natural Systems

Few tools in biological research have had as much impact as the voltage clamp, which was used to study the physical and molecular basis of electrical conductance across cell membranes. In its original implementation, the device worked by immersing an axon in a saline bath and using an electronic control circuit to maintain (clamp) a constant transmembrane voltage (49). By enabling measurement of the current necessary to sustain the target voltage, the device could record results of changes to membrane potential, enabling decoupling of action potential into the movement of Na^+ and K^+ during respective depolarizing and rectifying phases. We envision a potential role for engineered circuits similar in impact to voltage clamping—by using them to interface with and report on difficult-to-observe events occurring with or experienced by cells.

Synthetic gene circuitry can be used to convert signals of rare or transient molecular species into readable outputs. For example, synthetic riboregulators—messenger RNA (mRNA) secondary structure elements that block RBS access—can be disengaged in the presence of trigger RNA, permitting translation (14, 51). A versatile riboregulator-based circuit engineering platform was recently reported in *E. coli* (42) capable of gating by any RNA sequence and was used to detect endogenous mRNAs. Toehold detection has also been used to construct highly sensitive cell-free, paper-based detection circuits that can be configured to convert the presence of diverse nucleic acid species into a macroscopic readout (76). Potential applications for cell-free diagnostic circuits are numerous, including detection of rare transcripts or proteins in biological or environmental samples (77, 78). In another recent example of circuit-based RNA detection, Benenson and colleagues (105) used intracellular microRNA (miRNA) detection circuitry to report on cell state in mammalian cells (105). By configuring logic gates to detect sets of miRNA species, they engineered a classifier circuit that could identify a cancer cell phenotype and trigger apoptosis.

In one recent example of circuit-mediated interconversion of a molecular signal, circuitry was engineered to detect intracellular DNA sequences (93). The system featured a sensor module composed of two ZF arrays programmed to bind to adjacent target sequences. Tandem sensor binding triggers an intein *trans*-splicing event, linking together associated TF and transcriptional activator domains that can transactivate a GFP reporter module. Sequence sense and response worked effectively in mammalian cells, and the authors demonstrated the ability to couple detection of viral infection to apoptosis. In principle, detection capabilities of the system could be expanded by swapping the ZF arrays for TALEN (transcription activator-like effector nucleases) or CRISPR, enabling versatile real-time detection of the occurrence or movement of rare nucleic acid species (e.g., a latent virus) within a cell population or tissue or as a live readout of specific mutations or chromosomal rearrangements.

One emerging application space is the use of molecular memory circuits to record the occurrence of discrete cellular or environmental events. Recombinase-based switches, where DNA segments flanked by recombination sites are either excised or inverted upon recombinase expression, can be used as digital memory elements (10). Recombinase-based circuitry has been configured to perform simple computations, including Boolean logic (11, 92, 103) and event counting (38).

Analog recording of transient input has been realized using an engineered retron to generate single-stranded DNA that can introduce genomic sequence mutations through strand invasion (34). When distributed across a cell population, such a so-called tape recorder circuit can integrate input duration and intensity based on the proportion of cells that undergo recombination. More recently, a system that records successive events using CRISPR-based mutagenesis coupled to a fluorescence in situ hybridization (FISH) readout was used to track mammalian cell lineage (37).

FISH: fluorescence in situ hybridization

Combining computation with memory could become a powerful tool for understanding the order and timing of cellular events. One recent demonstration utilized recombinase switches to construct state machine circuitry capable of computing the order of any combination of three distinct inputs (89). Synthetic reporter circuits that combine, compute, and record functions with a programmed sense and response could be used for a variety of applications involving extraction of conditional or time-resolved information from complex environments. For example, cells introduced into natural ecosystems could be programmed to detect, amplify, and record rare ecological events. Reporter circuits introduced into cells at early stages of embryonic development could be used to record the order and magnitude of events encountered by different cell lineages during tissue development.

4. FUTURE: USING SYNTHETIC GENE CIRCUITS TO UNDERSTAND THE REGULATORY BASIS OF COMPLEX CELLULAR BEHAVIOR

Circuit engineering offers a formidable set of tools for construction, rewiring, and refactoring regulatory networks. As these techniques continue to advance, we may begin to view synthetic biology as a way to deconstruct the complex biological phenotypes that emerge from regulatory network structure. Below, we discuss several areas where we believe synthetic circuit engineering will make an impact in the years to come.

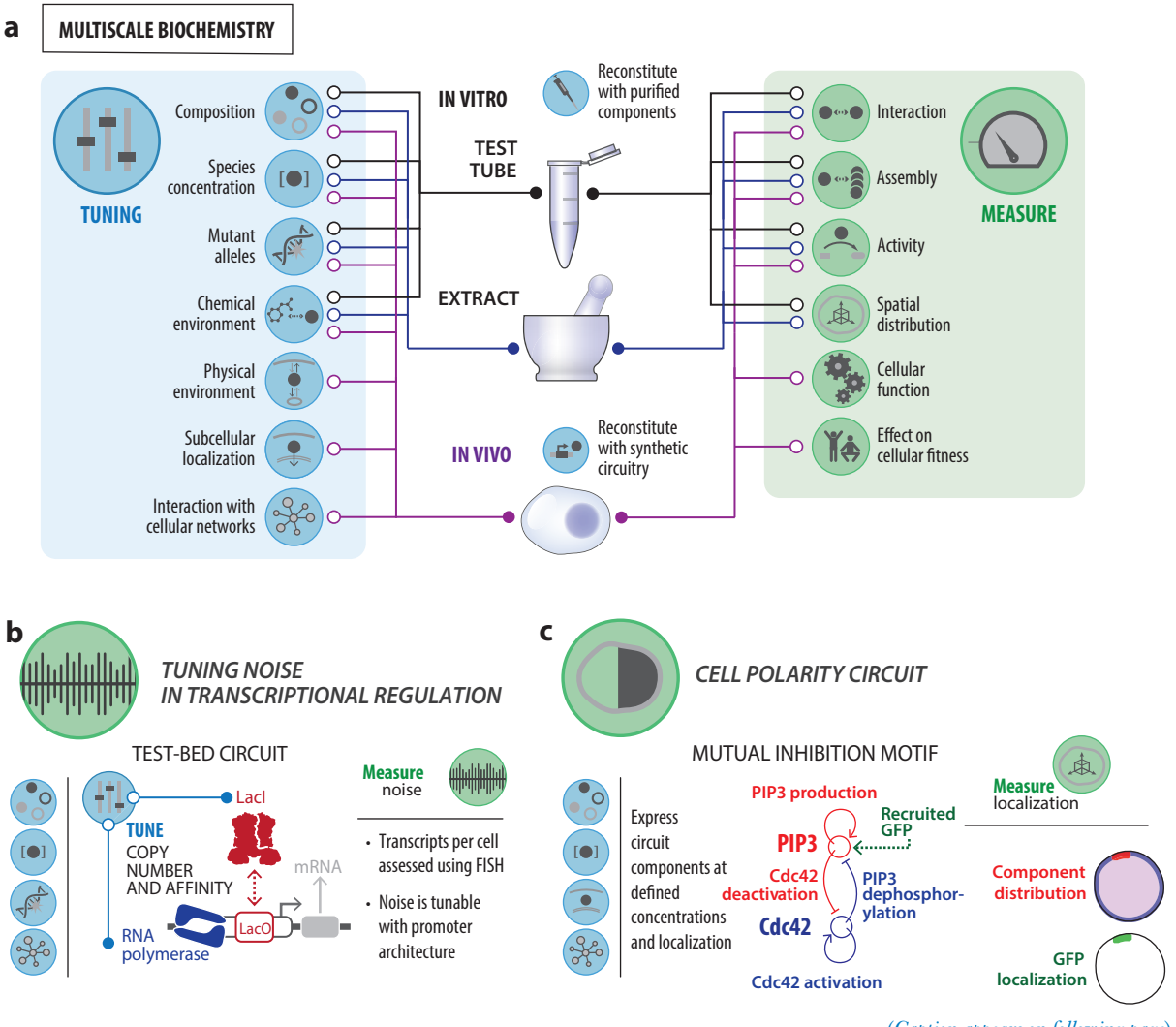
4.1. Synthetic Biology as In Vivo Biochemistry: The Cell as a Test Tube

Much of our understanding of the behavior of biological macromolecules comes from examining them in isolation from their native cellular context. By reconstituting macromolecular species and manipulating their content and concentration in vitro, interaction affinity, complex stoichiometry, and induced changes in activity can be measured (**Figure 4*a***). While using this approach to examine regulatory network components on an individual basis offers mechanistic insight, it yields little understanding of how systems-level behavior emerges from collections of components in vivo. A rich history of reconstituting complex, multispecies systems using cell extracts has served to partially bridge this gap (66). For example, study of transcription and translation apparatuses involving a combination of purification and reconstitution in cell lysates revealed many essential mechanistic details of the central dogma (64).

Traditionally, the goals of in vitro reconstitution have been similar in spirit to synthetic biology: determine functional sufficiency by rebuilding a biochemical system from the bottom up. However, living cells maintain their physical organization, metabolism, and regulatory dynamics far from thermodynamic equilibrium and provide a chemical and physical environment fundamentally unattainable through reconstitution. One possible strategy to overcome this gap is to use circuit engineering tools to reconstitute biochemical systems in a nonnative, in vivo environment, effectively using cells as living test tubes (**Figure 4*a***). In this capacity, synthetic gene circuitry could be used to titrate component expression levels or assemble complexes at a specified rate or in a defined order. Reconstitution using circuits could give the user a degree of control on par with in vitro reconstitution but with the added possibility of directing assembly to a particular

subcellular location, perturbing behavior by changing the intracellular environment or connecting input and output of a biochemical system to endogenous networks.

This approach is exemplified by two recent studies in which synthetic circuitry was used to reconstitute systems. In one study, the authors sought to demonstrate that gene expression noise in *E. coli* can be tuned by repressor–operator binding interactions (56) (**Figure 4*b***). Using FISH to assay mRNA output from a transcriptional reporter, and adjusting repressor expression and affinity, the authors were able to demonstrate that expression noise could be modeled as a function of regulatory configuration. In another study, a minimal biochemical circuit yielding robust cell polarization was reconstituted in yeast (19) (**Figure 4*c***). The circuit motif, which the authors computationally demonstrated could produce a polarized intracellular distribution, featured reciprocally two positive autoregulatory nodes that engage in reciprocally negative regulation. This motif was realized using elements that coupled recruitment to activation of the small GTPase Cdc42 and synthesis of PIP3. Model predictions were largely accurate: cells harboring a



(Caption appears on following page)

Figure 4 (*Figure appears on preceding page*)

Using synthetic circuitry to reconstitute biochemical systems in vivo. (*a*) Gene circuit–based reconstitution can be placed on a continuum with traditional techniques of buffer- and extract-based reconstitution. Control over the composition and concentration of the components of a biochemical system, through the use of in vitro methods, permits quantitative determination of its biophysical characteristics (e.g., interaction affinity, complex stoichiometry, and activity). Extract systems allow in vitro reconstitution of higher complexity systems. While enabling interaction with intracellular components, extracts are unable to fully replicate the physical and regulatory features of an intact cell. In vivo reconstitution uses synthetic circuitry to precisely control the relative stoichiometry, timing, and localization of component expression. Additionally, circuits can be used to integrate biochemical systems into cellular networks to assess effects on cellular function. (*b*) Testing effects of promoter regulatory architecture on transcriptional noise. Using a simple test-bed circuit, the possibility that transcriptional noise can be tuned by altering promoter architecture was tested by counting single-cell transcripts using fluorescence in situ hybridization (FISH). Tuning both expression level and affinity of RNA polymerase and the LacI repressor, researchers tested various promoter configurations against a quantitative model, and the configurations were shown to harbor different characteristic expression noise. (*c*) Synthetic reconstitution of a minimal cell polarization circuit. A self-organizing cell polarization circuit was described in yeast featuring components expressed at defined concentrations and subcellular localizations. A mutual inhibition circuit was created using protein domain fusions; PIP3 production creates localized positive feedback through PH (binds to PIP3) domain–mediated recruitment of PI3 (phosphoinositide 3) kinase, while PH-domain recruitment of a GAP inactivates Cdc42 locally. Active Cdc42 globally inhibits PIP3 production through recruitment of CRIB (Cdc42 and Rac interactive binding) (binds Cdc42) domain–phosphatase fusion. The circuit demonstrated spontaneous formation of individual, long-lived poles as visualized by localized recruitment of a PH-GFP fusion. Abbreviations: Cdc42, cell division control protein 42 homolog; GFP, green fluorescent protein; mRNA, messenger RNA; PH, pleckstrin homology; PIP3, phosphatidylinositol (3,4,5)-trisphosphate.

precisely tuned version of the motif demonstrated individual puncta, indicative of localized circuit activation. Both studies featured circuitry-enabled titration of in vitro–characterized components. More meaningfully, while both systems could have been reconstituted in vitro, using circuitry to reconstitute them in an in vivo context was essential for observing their behavior.

In vivo reconstitution, as this approach may come to be known, could be used to study systems with complexity too unwieldy for in vitro work or behaviors that manifest only in an intracellular setting (e.g., modules that exert physical force on cellular shape or 3D organization). By leveraging the rich history of biochemical research derived from buffer and extract systems, this approach could be used to compare in vitro and in vivo properties of reconstituted systems. One intriguing possibility lies in the development of parts frameworks amenable to testing in both settings. Candidate systems might consist of well-behaved, easy-to-purify proteins that could be readily studied in isolation, interfaced with extract systems, and also deployed in vivo using a circuit engineering scheme. This would allow for decoupling of component biophysical properties from the effects of both the intracellular chemical environment and the physical organization of intact cells.

4.2. Evolving Circuits to Explore Adaptive Fitness

Despite circuit engineering advances, minimizing interactions with endogenous host cell systems remains an outstanding challenge. As discussed above, circuits often impose a fitness cost on the host, either through depletion of resources or by interference with native network function. As a result, loss-of-function variants that ameliorate cost may spontaneously arise and rapidly sweep through a cell population. To avoid this, circuit engineering will need to identify and incorporate design features that maximize stability, especially for applications requiring engineered cells to retain circuit function over many generations. One study investigating the molecular determinants of fitness burden in *E. coli* tested a small library of circuits by simultaneously measuring fitness cost and resource depletion using a capacity monitor circuit (18). By examining the relationship between diminished fitness and capacity for different circuit designs, the authors implicated the depletion of free ribosomes as a primary source of fitness burden.

NGS: next-generation sequencing

Further investigation into long-term circuit stability may involve using experimental evolution to monitor cultures of circuit-bearing cells for the acquisition of adaptive mutations. Specific circuit failure modes could be uncovered through the observation of how circuits break upon mutation. By the same token, adaptive mutations that relieve fitness burden in the face of intact circuit function could be identified. Indeed, it may be possible to produce selection schemes that stipulate circuit function as a selective pressure, thereby forcing the host cell to adapt. To achieve this, automated continuous growth platforms designed to maximize throughput while allowing for automated manipulation and live readout of culture conditions (e.g., growth rate, GFP fluorescence) could be employed to simultaneously track culture fitness and circuit function (101), while next-generation sequencing (NGS) could be used to examine the resulting mutational landscape in both the circuit and host genome.

Far from viewing it as merely a factor complicating engineering, we see evolutionary instability of synthetic regulatory circuits as an opportunity to investigate the relationship between network structure and cell fitness. Using experimental evolution, synthetic circuit–host interactions can be used to understand how cells incorporate adaptive regulatory network connections under diverse selective pressures. The consequences of new circuitry or rewired linkages could be studied by examining fitness tradeoffs between the cost of regulatory reconfiguration and fitness advantages conferred by newly acquired phenotypes. Evolution experiments could be used to identify environmental conditions under which rewiring is rejected or accepted, while observing adaptive or compensatory mutations that arise to maximize fitness. One long-term goal might be creating regulatory systems that use novel molecular currencies (e.g., new transcription factors or posttranslational modifications) and operate orthogonally to natural ones. Experimental evolution could be used to assess to what extent integrated parallel programs would interact with native regulatory machinery.

4.3. Scaling Up Circuit Engineering to Accelerate Discovery

Synthetic circuit construction has been traditionally conducted at relatively low throughput, essentially operating at the pace and scale of an artisanal craft (**Figure 5**). Recent work has focused on accelerating the engineering cycle to move beyond testing only a handful of circuits at each turn (15). DNA assembly techniques have enabled rapid, combinatorial construction of circuit variant libraries (27), while bioinformatics-driven parts-mining efforts have expanded parts availability (96). Both these developments have synergized with rapidly decreasing gene synthesis costs, allowing circuit engineering projects to more rapidly move from design to construction. In conjunction with updated modeling techniques that use context-based part compatibility and assembly rules (2), these developments have set the stage for exploration of larger, more complex circuitry, and the field appears poised to enter a more industrial phase where automated, predictive design becomes regular practice (70) (**Figure 5**).

In recent years, NGS technology has radically enhanced our ability to obtain genome-scale information describing cellular content and organization. Application of NGS to synthetic circuit engineering remains largely unexplored but could dramatically enhance engineering cycle throughput. By combining NGS with high-throughput single-cell analysis, barcoded circuit libraries could be synthesized in batch, subjected to screening or selection, and then sequenced. Creating a diversity-based engineering workflow that combines NGS with automation—both for circuit synthesis and for testing—would generate tremendous quantities of data on circuit behavior, part performance, and context rules in each turn of the engineering cycle and permit comprehensive mapping of the relationship between circuit function and design (**Figure 5**).

Transforming gene circuit engineering into a data-driven, postindustrial process would dramatically broaden the scope of all research areas discussed in this review. For example, expanding

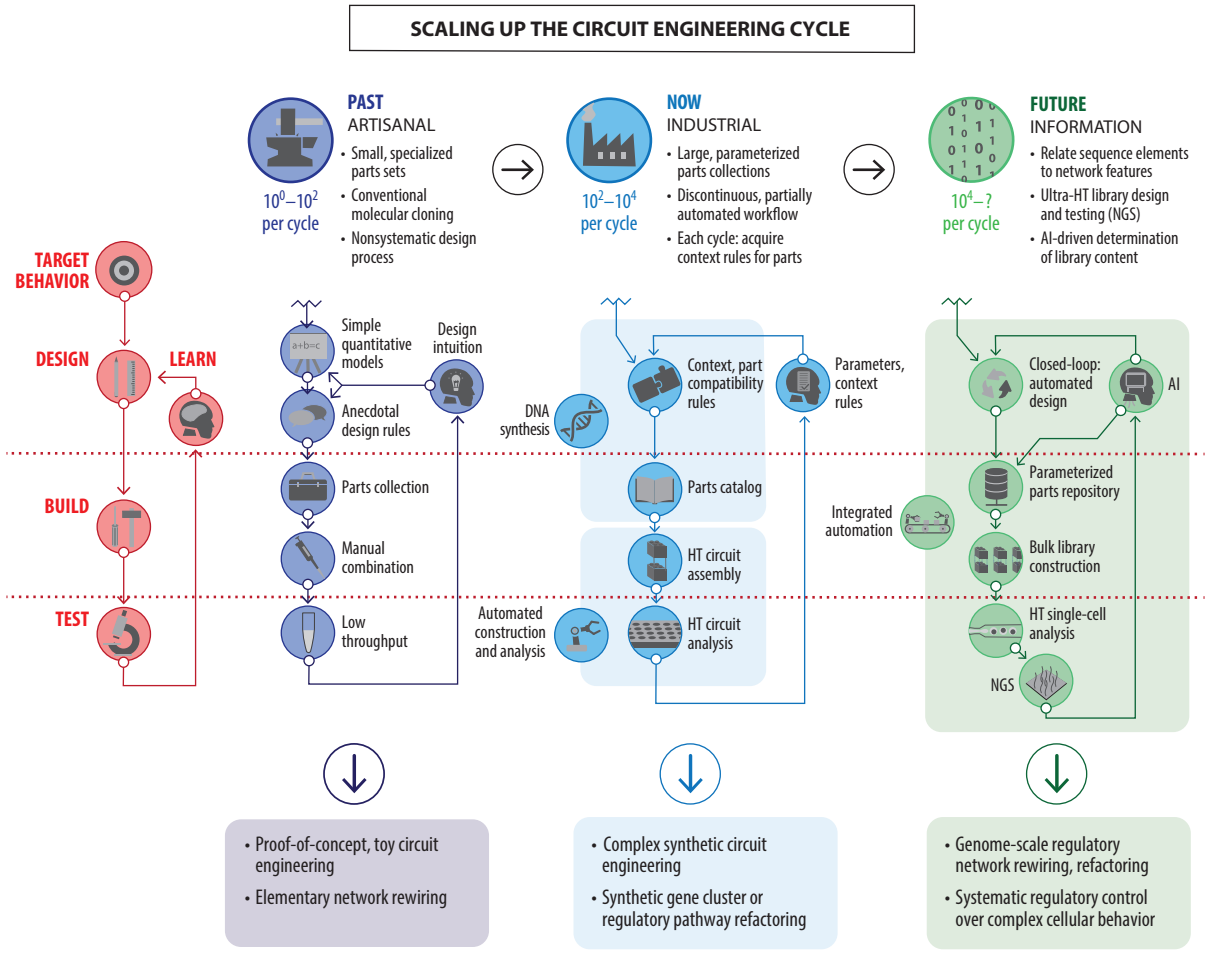


Figure 5

Evolution of the synthetic gene circuit engineering cycle. The scale and scope of inquiry into regulatory network design and function will depend on continued progress in circuit engineering techniques. The field began as an artisanal craft, relying on limited part sets, traditional techniques of molecular cloning, and low-throughput building and testing. While critical for proof-of-concept engineering of simple circuits, insight gained during each cycle was largely intuitive, as nonstandardization of parts and techniques made design rules difficult to generalize. Engineering approaches are beginning to transition into a more industrial mode in which the acquisition of larger collections of better-characterized parts permit more ambitious designs. This work is bolstered by low-cost DNA synthesis and development of circuit assembly techniques, both of which enable automation of building and testing phases. Higher-throughput, more systematic data acquisition in each cycle affords detailed, parameter-oriented profiles of individual parts. This, in turn, enables the engineer to pinpoint more robust, precisely tuned circuit behavior within a multidimensional design space. The throughput and precision of this workflow could be used to address the complexity presented by larger networks where part compatibility and circuit–host interaction are at issue (e.g., rewiring large regulatory modules or refactoring gene clusters). The field awaits transition to a postindustrial, information-based mode, where fully integrated automation of design, build, and test phases is driven by artificial intelligence (AI) and machine learning algorithms. By leveraging large-scale library synthesis and next-generation sequencing (NGS) as a means to scale input and output, large circuit design libraries can be tested at each turn, gathering data that establish rules-based relationships between sequence-level features and synthetic network behavior. Comprehensive, closed-loop characterization of circuit design space could be realized when data collected at each iteration are used to both refine models and suggest subsequent design spaces for testing. This mode of engineering could be used to probe design spaces of the highest complexity, including genome-scale refactoring and perturbation of regulatory networks. Additional abbreviation: HT, high-throughput.

diversity used for circuit construction could redefine our understanding of the relationship between sequence features and regulatory behavior. Massive diversification of part sets and circuit topologies would permit testing a design space that approaches the mutational landscape available to natural systems during evolution. Up until this point, circuit modeling has described essential features of circuit behavior in terms of part-specific parameters. For a diverse circuit library, since precise mechanistic relationships between sequence variants and circuit behavior are difficult to discern, building rules-based models to augment traditional approaches may be necessary. Machine learning algorithms could be used to train a model on a discrete area of design space, then predict a more expansive space, and then be iteratively refined using subsequent rounds of high-throughput design and testing. Upon model convergence, interesting or unusual areas of the space could be examined to understand the molecular basis for model-delineated rules, possibly leading to discovery of previously unknown regulatory phenomena.

5. CONCLUSION

Historically, our understanding of cellular regulation has come from perturbation and measurement working interdependently, leveraging classical tools of genetics and molecular biology to map network connectivity and information flow. In recent years, with the advent of genomics, high-throughput single-cell analysis, and multiscale imaging, our measurement and mapping capabilities have raced far ahead of our ability to make the proportional perturbations needed to understand their function. We foresee the tools of synthetic gene circuit engineering providing the means to close this gap. In this capacity, the role of synthetic biology may be one of final arbiter for models forwarded by systems biology. At present, our ability to rewire and recapitulate natural network design is at different stages for different systems. While constructing complex, natural-like circuitry in prokaryotic gene networks is well developed, the same level of bottom-up control in eukaryotic and metazoan systems remains a work in progress. Combining the ability to make new connections with an accelerated engineering cycle should facilitate rewiring, refactoring, and reconstitution of increasingly complex regulatory systems. To this end, rather than studying discrete circuit function using a single type of network connection (e.g., a circuit composed of entirely transcriptional network connections), focus should turn to obtaining comprehensive multiscale engineering control over entire cellular subsystems in a manner that subsumes native behavior, while granting the ability to systematically test alternate, unnatural designs. Such a synthetic regulatory capture approach could be used to refactor an organism's stress response program to understand the operational design rules underlying fitness or to acquire engineered control over an animal developmental module to investigate the regulatory determinants of morphology. To this end, other branches of synthetic biology, like synthetic genome engineering, could merge with gene circuit engineering to enable large-scale reorganization of regulatory networks at the whole-organism level. One medium-term step in this direction might involve layering regulatory machinery over a minimal bacterial genome to assess fitness advantages gained through increasing degrees of regulatory control. As our ability to reconstruct network function becomes more sophisticated, we anticipate that understanding of cellular regulatory network design will increasingly come from using forward engineering and manipulation, causing meaningful distinctions between synthetic biology and systems biology to gradually disappear.

SUMMARY POINTS

1. Designing, building, and testing synthetic gene circuits can be used to evaluate hypotheses regarding the modular organization and function of regulatory networks.

2. As the most developed area of the field, synthetic gene circuits in prokaryotes exemplify how biological understanding can be acquired through iterative engineering cycles.
3. While part sets and strategies for circuit engineering in metazoan systems are still emerging, test-bed circuits can be used to decompose and understand the modular design principles of both transcriptional and posttranslational regulation.
4. By furnishing biologists with a toolkit for precise manipulation of regulatory network connectivity, gene circuit engineering enables the creation of sophisticated genetic perturbations for investigating systems-level function.
5. Circuit engineering can be used to construct reporter circuits capable of converting obscure molecular species or cellular events into measureable, recordable data.
6. By scaling up circuit engineering approaches, it may be possible to modify or reconstruct regulatory networks in a way that allows us to understand how network structure shapes complex phenotype.

FUTURE ISSUES

1. Emphasis should be placed on developing techniques that accelerate the gene circuit engineering cycle. Automated combinatorial library assembly, high-throughput screening, and next-generation sequencing should be incorporated. A diversity-based workflow could be used to comprehensively map the relationship between network structure and function.
2. An engineering grammar should be devised for bottom-up construction of posttranslational signaling networks—similar to modular building blocks used for transcriptional networks. Signaling motifs reconstituted with such a system could be used as physical models for studying signal transduction.
3. Focus on prokaryotic gene circuit engineering should migrate toward using experimental evolution of synthetic gene networks to study the relationship between regulation and fitness. Experimental tools will need to be developed to assess the relationship between environmental pressure and circuit–host coevolution.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

We would like to thank Pankaj Mehta and Arnaud Gutierrez for useful discussions.

LITERATURE CITED

1. Alon U. 2003. Biological networks: the tinkerer as an engineer. *Science* 301:1866–67
2. Appleton E, Madsen C, Roehner N, Densmore D. 2017. Design automation in synthetic biology. *Cold Spring Harb. Perspect. Biol.* 9:a023978

3. Barrett CL, Kim TY, Kim HU, Palsson BO, Lee SY. 2006. Systems biology as a foundation for genome-scale synthetic biology. *Curr. Opin. Biotechnol.* 17:488–92
4. Bashor CJ, Helman NC, Yan S, Lim WA. 2008. Using engineered scaffold interactions to reshape MAP kinase pathway signaling dynamics. *Science* 319:1539–43
5. Bashor CJ, Horwitz AA, Peisajovich SG, Lim WA. 2010. Rewiring cells: synthetic biology as a tool to interrogate the organizational principles of living systems. *Annu. Rev. Biophys.* 39:515–37
6. Becskei A, Seraphin B, Serrano L. 2001. Positive feedback in eukaryotic gene networks: cell differentiation by graded to binary response conversion. *EMBO J.* 20:2528–35
7. Becskei A, Serrano L. 2000. Engineering stability in gene networks by autoregulation. *Nature* 405:590–93
8. Bintu L, Yong J, Antebi YE, McCue K, Kazuki Y, et al. 2016. Dynamics of epigenetic regulation at the single-cell level. *Science* 351:720–24
9. Blake WJ, Kærn M, Cantor CR, Collins JJ. 2003. Noise in eukaryotic gene expression. *Nature* 422:633–37
10. Bonnet J, Subsoontorn P, Endy D. 2012. Rewritable digital data storage in live cells via engineered control of recombination directionality. *PNAS* 109:8884–89
11. Bonnet J, Yin P, Ortiz ME, Subsoontorn P, Endy D. 2013. Amplifying genetic logic gates. *Science* 340:599–603
12. Brophy JA, Voigt CA. 2014. Principles of genetic circuit design. *Nat. Methods* 11:508–20
13. Çağatay T, Turcotte M, Elowitz MB, Garcia-Ojalvo J, Süel GM. 2009. Architecture-dependent noise discriminates functionally analogous differentiation circuits. *Cell* 139:512–22
14. Callura JM, Dwyer DJ, Isaacs FJ, Cantor CR, Collins JJ. 2010. Tracking, tuning, and terminating microbial physiology using synthetic riboregulators. *PNAS* 107:15898–903
15. Cameron DE, Bashor CJ, Collins JJ. 2014. A brief history of synthetic biology. *Nat. Rev. Microbiol.* 12:381–90
16. Cameron DE, Collins JJ. 2014. Tunable protein degradation in bacteria. *Nat. Biotechnol.* 32:1276–81
17. Carroll SB. 2008. Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* 134:25–36
18. Ceroni F, Algar R, Stan GB, Ellis T. 2015. Quantifying cellular capacity identifies gene expression designs with reduced burden. *Nat. Methods* 12:415–18
19. Chau AH, Walter JM, Gerardin J, Tang C, Lim WA. 2012. Designing synthetic regulatory networks capable of self-organizing cell polarization. *Cell* 151:320–32
20. Chavez A, Scheiman J, Vora S, Pruitt BW, Tuttle M, et al. 2015. Highly efficient Cas9-mediated transcriptional programming. *Nat. Methods* 12:326–28
21. Chen Y, Kim JK, Hirning AJ, Josić K, Bennett MR. 2015. Emergent genetic oscillations in a synthetic microbial consortium. *Science* 349:986–89
22. Danino T, Mondragon-Palomino O, Tsimring L, Hasty J. 2010. A synchronized quorum of genetic clocks. *Nature* 463:326–30
23. Del Vecchio D, Dy AJ, Qian Y. 2016. Control theory meets synthetic biology. *J. R. Soc. Interface* 13:20160380
24. Detwiler PB, Ramanathan S, Sengupta A, Shraiman BI. 2000. Engineering aspects of enzymatic signal transduction: photoreceptors in the retina. *Biophys. J.* 79:2801–17
25. Dueber JE, Yeh BJ, Chak K, Lim WA. 2003. Reprogramming control of an allosteric signaling switch through modular recombination. *Science* 301:1904–8
26. Dunlap JC. 1999. Molecular bases for circadian clocks. *Cell* 96:271–90
27. Ellis T, Adie T, Baldwin GS. 2011. DNA assembly for synthetic biology: from parts to pathways and beyond. *Integr. Biol.* 3:109–18
28. Elowitz MB, Leibler S. 2000. A synthetic oscillatory network of transcriptional regulators. *Nature* 403:335–38
29. Elowitz MB, Levine AJ, Siggia ED, Swain PS. 2002. Stochastic gene expression in a single cell. *Science* 297:1183–86
30. Elowitz MB, Lim WA. 2010. Build life to understand it. *Nature* 468:889–90
31. Endy D. 2005. Foundations for engineering biology. *Nature* 438:449–53
32. Ernst J, Kheradpour P, Mikkelsen TS, Shoresh N, Ward LD, et al. 2011. Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature* 473:43–49

33. Erwin DH, Davidson EH. 2009. The evolution of hierarchical gene regulatory networks. *Nat. Rev. Genet.* 10:141–48
34. Farzadfard F, Lu TK. 2014. Genomically encoded analog memory with precise in vivo DNA writing in living cell populations. *Science* 346:1256272
35. Fischbach MA, Bluestone JA, Lim WA. 2013. Cell-based therapeutics: the next pillar of medicine. *Sci. Transl. Med.* 5:179ps7
36. Fischbach MA, Voigt CA. 2010. Prokaryotic gene clusters: a rich toolbox for synthetic biology. *Biotechnol. J.* 5:1277–96
37. Frieda KL, Linton JM, Hormoz S, Choi J, Chow KK, et al. 2017. Synthetic recording and in situ readout of lineage information in single cells. *Nature* 541:107–11
38. Friedland AE, Lu TK, Wang X, Shi D, Church G, Collins JJ. 2009. Synthetic gene networks that count. *Science* 324:1199–202
39. Gardner TS, Cantor CR, Collins JJ. 2000. Construction of a genetic toggle switch in *Escherichia coli*. *Nature* 403:339–42
40. Gordley RM, Bugaj LJ, Lim WA. 2016. Modular engineering of cellular signaling proteins and networks. *Curr. Opin. Struct. Biol.* 39:106–14
41. Gordley RM, Williams RE, Bashor CJ, Toettcher JE, Yan S, Lim WA. 2016. Engineering dynamical control of cell fate switching using synthetic phospho-regulons. *PNAS* 113:13528–33
42. Green AA, Silver PA, Collins JJ, Yin P. 2014. Toehold switches: de-novo-designed regulators of gene expression. *Cell* 159:925–39
43. Guido NJ, Lee P, Wang X, Elston TC, Collins JJ. 2007. A pathway and genetic factors contributing to elevated gene expression noise in stationary phase. *Biophys. J.* 93:L55–57
44. Guido NJ, Wang X, Adalsteinsson D, McMillen D, Hasty J, et al. 2006. A bottom-up approach to gene regulation. *Nature* 439:856–60
45. Hannon GJ, Rossi JJ. 2004. Unlocking the potential of the human genome with RNA interference. *Nature* 431:371–78
46. Hartwell LH, Hopfield JJ, Leibler S, Murray AW. 1999. From molecular to modular cell biology. *Nature* 402:C47–52
47. Hasty J, Dolnik M, Rottschafer V, Collins JJ. 2002. Synthetic gene network for entraining and amplifying cellular oscillations. *Phys. Rev. Lett.* 88:148101
48. Hilton IB, D'Ippolito AM, Vockley CM, Thakore PI, Crawford GE, et al. 2015. Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nat. Biotechnol.* 33:510–17
49. Hodgkin AL, Huxley AF. 1952. A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol.* 117:500–44
50. Howard PL, Chia MC, Del Rizzo S, Liu FF, Pawson T. 2003. Redirecting tyrosine kinase signaling to an apoptotic caspase pathway through chimeric adaptor proteins. *PNAS* 100:11267–72
51. Isaacs FJ, Dwyer DJ, Ding C, Pervouchine DD, Cantor CR, Collins JJ. 2004. Engineered riboregulators enable post-transcriptional control of gene expression. *Nat. Biotechnol.* 22:841–47
52. Isaacs FJ, Hasty J, Cantor CR, Collins JJ. 2003. Prediction and measurement of an autoregulatory genetic module. *PNAS* 100:7714–19
53. Isalan M, Lemerle C, Michalodimitrakis K, Horn C, Beltrao P, et al. 2008. Evolvability and hierarchy in rewired bacterial gene networks. *Nature* 452:840–45
54. Jacob F, Monod J. 1961. Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.* 3:318–56
55. Jenuwein T, Allis CD. 2001. Translating the histone code. *Science* 293:1074–80
56. Jones DL, Brewster RC, Phillips R. 2014. Promoter architecture dictates cell-to-cell variability in gene expression. *Science* 346:1533–36
57. Jusiak B, Cleto S, Perez-Pinera P, Lu TK. 2016. Engineering synthetic gene circuits in living cells with CRISPR technology. *Trends Biotechnol.* 34:535–47
58. Kærn M, Blake WJ, Collins JJ. 2003. The engineering of gene regulatory networks. *Annu. Rev. Biomed. Eng.* 5:179–206

59. Keung AJ, Bashor CJ, Kiriakov S, Collins JJ, Khalil AS. 2014. Using targeted chromatin regulators to engineer combinatorial and spatial transcriptional regulation. *Cell* 158:110–20
60. Khalil AS, Collins JJ. 2010. Synthetic biology: applications come of age. *Nat. Rev. Genet.* 11:367–79
61. Khalil AS, Lu TK, Bashor CJ, Ramirez CL, Pyenson NC, et al. 2012. A synthetic biology framework for programming eukaryotic transcription functions. *Cell* 150:647–58
62. Kiani S, Beal J, Ebrahimkhani MR, Huh J, Hall RN, et al. 2014. CRISPR transcriptional repression devices and layered circuits in mammalian cells. *Nat. Methods* 11:723–26
63. Kirschner M, Gerhart J. 1998. Evolvability. *PNAS* 95:8420–27
64. Kornberg A. 1960. Biologic synthesis of deoxyribonucleic acid. *Science* 131:1503–8
65. Lee JW, Gyorgy A, Cameron DE, Pyenson N, Choi KR, et al. 2016. Creating single-copy genetic circuits. *Mol. Cell* 63:329–36
66. Liu AP, Fletcher DA. 2009. Biology under construction: in vitro reconstitution of cellular function. *Nat. Rev. Mol. Cell Biol.* 10:644–50
67. Lutz R, Bujard H. 1997. Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucleic Acids Res.* 25:1203–10
68. McAdams HH, Arkin A. 1997. Stochastic mechanisms in gene expression. *PNAS* 94:814–19
69. Milo R, Shen-Orr S, Itzkovitz S, Kashtan N, Chklovskii D, Alon U. 2002. Network motifs: simple building blocks of complex networks. *Science* 298:824–27
70. Nielsen AA, Der BS, Shin J, Vaidyanathan P, Paralanov V, et al. 2016. Genetic circuit design automation. *Science* 352:aac7341
71. Nielsen J, Keasling JD. 2016. Engineering cellular metabolism. *Cell* 164:1185–97
72. Nocedal I, Johnson AD. 2015. How transcription networks evolve and produce biological novelty. *Cold Spring Harb. Symp. Quant. Biol.* 80:265–74
73. O'Shaughnessy EC, Palani S, Collins JJ, Sarkar CA. 2011. Tunable signal processing in synthetic MAP kinase cascades. *Cell* 144:119–31
74. Ozbudak EM, Thattai M, Kurtser I, Grossman AD, van Oudenaarden A. 2002. Regulation of noise in the expression of a single gene. *Nat. Genet.* 31:69–73
75. Panne D. 2008. The enhanceosome. *Curr. Opin. Struct. Biol.* 18:236–42
76. Pardee K, Green AA, Ferrante T, Cameron DE, DaleyKeyser A, et al. 2014. Paper-based synthetic gene networks. *Cell* 159:940–54
77. Pardee K, Green AA, Takahashi MK, Braff D, Lambert G, et al. 2016. Rapid, low-cost detection of Zika virus using programmable biomolecular components. *Cell* 165:1255–66
78. Pardee K, Slomovic S, Nguyen PQ, Lee JW, Donghia N, et al. 2016. Portable, on-demand biomolecular manufacturing. *Cell* 167:248–59.e12
79. Park SH, Zarrinpar A, Lim WA. 2003. Rewiring MAP kinase pathways using alternative scaffold assembly mechanisms. *Science* 299:1061–64
80. Pawson T. 1995. Protein modules and signalling networks. *Nature* 373:573–80
81. Pedraza JM, van Oudenaarden A. 2005. Noise propagation in gene networks. *Science* 307:1965–69
82. Potvin-Trottier L, Lord ND, Vinnicombe G, Paulsson J. 2017. Synchronous long-term oscillations in a synthetic gene circuit. *Nature* 538:514–17
83. Prindle A, Selimkhanov J, Li H, Razinkov I, Tsimring LS, Hasty J. 2014. Rapid and tunable post-translational coupling of genetic circuits. *Nature* 508:387–91
84. Ptashne M, Johnson AD, Pabo CO. 1982. A genetic switch in a bacterial virus. *Sci. Am.* 247:128–40
85. Purnick PE, Weiss R. 2009. The second wave of synthetic biology: from modules to systems. *Nat. Rev. Mol. Cell Biol.* 10:410–22
86. Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, et al. 2013. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* 152:1173–83
87. Razooky BS, Pai A, Aull K, Rouzine IM, Weinberger LS. 2015. A hardwired HIV latency program. *Cell* 160:990–1001
88. Rice MK, Ruder WC. 2014. Creating biological nanomaterials using synthetic biology. *Sci. Technol. Adv. Mater.* 15:014401
89. Roquet N, Soleimany AP, Ferris AC, Aaronson S, Lu TK. 2016. Synthetic recombinase-based state machines in living cells. *Science* 353:aad8559

90. Rosenfeld N, Young JW, Alon U, Swain PS, Elowitz MB. 2005. Gene regulation at the single-cell level. *Science* 307:1962–65
91. Ruder WC, Lu T, Collins JJ. 2011. Synthetic biology moving into the clinic. *Science* 333:1248–52
92. Siuti P, Yazbek J, Lu TK. 2013. Synthetic circuits integrating logic and memory in living cells. *Nat. Biotechnol.* 31:448–52
93. Slomovic S, Collins JJ. 2015. DNA sense-and-respond protein modules for mammalian cells. *Nat. Methods* 12:1085–90
94. Slomovic S, Pardee K, Collins JJ. 2015. Synthetic biology devices for in vitro and in vivo diagnostics. *PNAS* 112:14429–35
95. Smanski MJ, Bhatia S, Zhao D, Park Y, Woodruff LBA, et al. 2014. Functional optimization of gene clusters by combinatorial design and assembly. *Nat. Biotechnol.* 32:1241–49
96. Stanton BC, Nielsen AA, Tamsir A, Clancy K, Peterson T, Voigt CA. 2014. Genomic mining of prokaryotic repressors for orthogonal logic gates. *Nat. Chem. Biol.* 10:99–105
97. Stricker J, Cookson S, Bennett MR, Mather WH, Tsimring LS, Hasty J. 2008. A fast, robust and tunable synthetic gene oscillator. *Nature* 456:516–19
98. Süel GM, Garcia-Ojalvo J, Liberman LM, Elowitz MB. 2006. An excitable gene regulatory circuit induces transient cellular differentiation. *Nature* 440:545–50
99. Süel GM, Kulkarni RP, Dworkin J, Garcia-Ojalvo J, Elowitz MB. 2007. Tunability and noise dependence in differentiation dynamics. *Science* 315:1716–19
100. Tigges M, Fussenegger M. 2009. Recent advances in mammalian synthetic biology—design of synthetic transgene control networks. *Curr. Opin. Biotechnol.* 20:449–60
101. Toprak E, Veres A, Yildiz S, Pedraza JM, Chait R, et al. 2013. Building a morbidostat: an automated continuous-culture device for studying bacterial drug resistance under dynamically sustained drug inhibition. *Nat. Protoc.* 8:555–67
102. Ubersax JA, Ferrell JE Jr. 2007. Mechanisms of specificity in protein phosphorylation. *Nat. Rev. Mol. Cell Biol.* 8:530–41
103. Weinberg BH, Pham NTH, Caraballo LD, Lozanoski T, Engel A, et al. 2017. Large-scale design of robust genetic circuits with multiple inputs and outputs for mammalian cells. *Nat. Biotechnol.* 35:453–62
104. Weinberger LS, Burnett JC, Toettcher JE, Arkin AP, Schaffer DV. 2005. Stochastic gene expression in a lentiviral positive-feedback loop: HIV-1 Tat fluctuations drive phenotypic diversity. *Cell* 122:169–82
105. Xie Z, Wroblewska L, Prochazka L, Weiss R, Benenson Y. 2011. Multi-input RNAi-based logic circuit for identification of specific cancer cells. *Science* 333:1307–11
106. Yeh BJ, Rutigliano RJ, Deb A, Bar-Sagi D, Lim WA. 2007. Rewiring cellular morphology pathways with synthetic guanine nucleotide exchange factors. *Nature* 447:596–600
107. Yeung E, Dy AJ, Martin KB, Ng AH, Del Vecchio D, et al. 2017. Biophysical constraints arising from compositional context in synthetic gene networks. *Cell Syst.* 5:11–24.e12
108. Young RA. 2011. Control of the embryonic stem cell state. *Cell* 144:940–54

Annual Review of Biophysics
Volume 47, 2018

Contents

Indexes

Errata

An online log of corrections to *Annual Review of Biophysics* articles may be found at http://www.annualreviews.org/errata/biophys