Published in final edited form as:

Curr Opin Syst Biol. 2018 April ; 8: 90–96. doi:10.1016/j.coisb.2017.12.009.

Systematic and synthetic approaches to rewire regulatory networks

Jimin Park[1,2] and **Harris H. Wang**[1,3]

[1]Department of Systems Biology, Columbia University Medical Center, New York, USA

[2]Integrated Program in Cellular, Molecular and Biomedical Studies, Columbia University Medical Center, New York, USA

[3]Department of Pathology and Cell Biology, Columbia University Medical Center, New York, USA

Abstract

Microbial gene regulatory networks are composed of cis- and trans-components that in concert act to control essential and adaptive cellular functions. Regulatory components and interactions evolve to adopt new configurations through mutations and network rewiring events, resulting in novel phenotypes that may benefit the cell. Advances in high-throughput DNA synthesis and sequencing have enabled the development of new tools and approaches to better characterize and perturb various elements of regulatory networks. Here, we highlight key recent approaches to systematically dissect the sequence space of cis-regulatory elements and trans-regulators as well as their inter-connections. These efforts yield fundamental insights into the architecture, robustness, and dynamics of gene regulation and provide models and design principles for building synthetic regulatory networks for a variety of practical applications.

Keywords

Gene regulation; Regulatory network rewiring; Cis-regulatory elements; Trans-regulatory proteins; DNA synthesis

Introduction

Microbes rely on precise regulation of gene expression for a myriad of essential processes during growth and adaptation in changing environments. These patterns of gene expression are generated through coordinated interactions between cis-regulatory DNA elements and trans-regulatory proteins [1]. Cis-regulatory elements are stretches along the genome where regulatory modulation can occur during transcription and translation, often found in upstream and downstream untranslated regions (UTRs) flanking coding DNA sequences (CDS). These UTR regions, such as promoters, repressor binding sites, ribosome binding

Corresponding author: Wang, Harris H, Department of Systems Biology, Columbia University Medical Center, New York, USA. (hw2429@columbia.edu).

Conflicts of interest

None.

sites (RBSs), and terminators, contain specific sequences (i.e. single or multipartite sequence motifs) that precisely recruit their corresponding trans-regulatory proteins (e.g. transcription factors, sigma factors) to positively or negatively adjust the expression of associated genes [2]. Each regulatory element and its protein regulators form a regulatory unit that when coupled to other regulatory units make up network motifs, such as feedforward or feedback loops, to facilitate various regulatory functions including signal detection, amplification, propagation, and processing [3]. The constellations of all regulatory units form the global gene regulatory network of the cell, which exhibits a “scale-free” hierarchy with a power-law distribution of regulatory connections [4,5]. Gene regulatory networks possess many features that appear to be conserved across diverse domains of life, ranging from biophysical properties of regulatory factors to global network architectures [6].

While gene regulatory networks must maintain robust performance on a cellular timescale, they are also highly adaptable on an evolutionary timescale [7–9]. The rewiring of regulatory networks through addition or subtraction of connections can produce a variety of adaptive and novel phenotypes [10,11]. Mechanistically, most rewiring events occur either through mutations in a cis-element that alter its binding specificity for a trans-regulatory protein or through a duplicated regulatory protein that subsequently diverges to acquire new network connections and functions [2,12]. Horizontal transfer of regulatory sequences between related species has recently been implicated to also play an important role in regulatory rewiring, suggesting that sharing of regulatory network architectures and motifs may lead to evolutionary advantages during lateral gene transfer events [13].

Even though thousands of sequenced microbial genomes have enabled comparative analysis of regulatory sequences and proteins [14], methods to predict gene expression, network architecture, and regulatory dynamics still remain challenging. Nonetheless, *in vitro* approaches have made key advances towards unraveling essential components of gene regulation. For example, protein binding microarrays have been used to determine the sequence specificities of regulatory proteins [15–17]. Systematic evolution of ligands by exponential enrichment (SELEX) is a foundational method for understanding and modeling protein-DNA interactions [18]. Pairing such approaches with high-throughput sequencing have improved the measurement of regulatory protein affinity to defined DNA targets [19]. A gamut of computational tools (e.g. MEME [20], Bioprospector [21], FIRE [22]) have been developed to extract and identify many new regulatory elements and sequence motifs [23].

The dawning of low-cost genomic technologies has accelerated the systematic characterization of regulatory elements at a larger scale. Advances in high-throughput DNA synthesis, when combined with deep sequencing and cellular phenotyping, offer important new opportunities to bridge key knowledge gaps by enabling systematic and quantitative dissection of complex regulatory interactions between thousands of cis-regulatory elements and trans-regulatory proteins simultaneously [24,25]. Here, we discuss and highlight recent studies that leverage these new synthetic and systems approaches to analyze microbial regulatory networks to de-convolute the biological complexity that allows even the simplest microbe to exhibit sophisticated, robust, and yet adaptable phenotypes.

Systematic exploration of the sequence-function space of cis-regulatory elements

Since microbial cis-regulatory elements contain only short sequence motifs that modulate gene expression [2], they are often difficult to annotate directly from genomes and to predict their regulatory function [14]. Furthermore, cis-elements exhibit high sequence divergence as well as altered activity in different genomic and local context that make computational analyses challenging. While transcriptomic studies can help to better identify and delineate regulatory units in individual strains [26], such approaches are laborious to scale across many species and are confined to only sequences already present in the genome, which leaves the regulatory plasticity of a strain poorly characterized. Recent advances in oligonucleotide library synthesis (OLS) on DNA microarrays [27] have enabled an unprecedented degree of control in regulatory sequences synthesized and a larger scale of experiments performed and quantitative data generated. In these OLS-based approaches, regulatory sequences are mined from genomes or designed *in silico*, synthesized on made-to-order DNA microarrays, fused to reporter genes, and measured through pooled *in vitro* or *in vivo* assays via next-generation sequencing (NGS) and high-throughput cellular phenotyping (Figure 1a). A variety of OLS-based improvements and variations have been developed over the past several years. Barcoding is a popular strategy to efficiently link the designed sequence with a unique molecular tag that can be easily matched by NGS [28]. To overcome length limitations on oligo synthesis, approaches using pairwise multiplex assembly of oligonucleotides have been employed [29]. Library uniformity (i.e. abundance of individual members), errors that propagate through assays (e.g. DNA synthesis, cloning, or sequencing), and other technical challenges have been addressed to improve the performance of the overall approach and robustness of the results [30–32].

To understand the impact of individual mutations on cis-regulatory element function, a pioneering study by Patwardhan et al. used OLS-based methods to systematically explore the sequence-activity of T3, T7, and SP6 bacteriophage promoters by generating and testing a saturation mutagenesis library of 12,000 variants, including some with multiple base-pair mutations [28]. Following a pooled *in vitro* bacterial transcription reaction, the activity of each variant's DNA and RNA abundance was determined by DNA-and RNA-seq. The high-throughput dataset delineated core regulatory regions or "footprints" of each promoter that are functionally important for transcription, which when mutated resulted in reduced expression. Furthermore, variants containing two distinct mutations often exhibited positive or negative epistasis (~30% of cases), which highlighted the nonlinear interacting determinants of gene expression. To better functionally annotate sequence motifs in the yeast genome, Sharon et al. utilized a similar OLS-based scanning mutagenesis library to generate and characterize the activity of native and mutant UTRs [24]. In a live-cell reporter assay using fluorescence-assisted cell sorting and sequencing (FACS-seq), which relies on sorting library members into defined bins based on the expression levels of a reporter and sequencing each bin to quantitatively determine the expression level of each variant, the authors detected changes in the protein expression level from UTR variants when specific cis-elements were disrupted. The study highlighted that altering key regulatory element parameters such as multiplicity, orientation, and context can significantly affect gene

expression. This OLS-based approach significantly extended beyond previous studies [33] that relied on limited random assemblies of regulatory element fragments to access more complex and defined sequences with single-nucleotide precision.

The activity of cis-regulatory elements can often exhibit context-dependency that necessitates systematic characterizations to unravel their layered complexity. Kosuri et al. attempted to determine the composability and performance of known bacterial cis-elements in synthetic circuits, specifically the degree to which their activity is conserved in different sequence contexts. By generating a combinatorial library of 114 promoters and 111 RBSs and placing them upstream of a fluorescence reporter, the authors could quantify transcription and translation activity levels using RNA-seq, DNA-seq, and FACS-seq. The results revealed frequent context-dependency in these regulatory elements at both transcription and translation levels depending on the combination of promoter and RBS pairing [30]. Furthermore, a significant cross-interaction between transcription and translation was observed: high levels of translation initiation from strong RBSs extended mRNA half-life, while strong secondary structures in the 5′ UTR of mRNA transcripts lowered the translation efficiency. In a complementary set of studies, Mutalik et al. explored ways to better delineate regulatory architectures and to minimize context-dependencies of gene expression [31,32]. Using the experimental data from characterized regulatory libraries, the authors developed a bi-cistronic regulatory architecture that functionally decoupled two cis-elements to minimize context-dependent variations that arise from irregular 5′ UTR and CDS junctions. Mechanistically, the translation of upstream CDS disrupts the mRNA secondary structure at the bi-cistronic junction, enabling translation efficiency of the downstream CDS to be dynamically and independently modulated by another RBS. By eliminating this large source of deviation in gene expression, the studies reported experimentally measured expression levels that matched the predicted values at a success rate of 93%, which is a 87% reduction in mispredicting gene expression in comparison to previous state-of-art in the field [34].

Beyond 5′ UTR elements, 3′ UTRs encoding transcriptional terminators have been also been studied through OLS-based approaches [35]. Chen et al. systematically mapped the sequence-activity relationship of 582 natural and synthetic bacterial transcriptional terminators in *Escherichia coli*. From the data, the authors developed a biophysical model that incorporated U-/A-tract length, hairpin loop, and stem base sequence information to predict the termination efficiency of a given terminator. These emerging approaches further highlight opportunities to systematically dissect detailed cis-regulatory function via *de novo* DNA synthesis and large-scale functional interrogation.

Network modulation with natural and synthetic trans-regulatory factors

Recent advances in high-throughput genomics have enable large-scale dissection of trans-regulatory modulators of gene expression. Trans-regulatory proteins control gene expression by recognizing and binding to one or more corresponding cis-elements and can be generally divided into two groups: densely connected global regulators that control hundreds of genes (e.g. primary sigma factor), and sparsely connected regulators that control defined regulons (e.g. extracellular function or ECF sigma factors). While regulators can often be annotated

by homology-based analyses [2], accurate prediction of their target cis-regulatory elements remain challenging. Furthermore, it is difficult to *a priori* model how changes to regulatory connections alter the transcriptome and affect cellular phenotype. To better delineate regulatory networks through modulation of transacting regulatory factors, various approaches including regulator mutagenesis, domain-shuffling, and heterologous expression have been explored to establish new regulatory connections by generating alterative trans-regulator to cis-elements assignments (Figure 1b).

Duplication and subsequent functional divergence of regulators play a key role in regulatory network evolution as it minimizes pleotropic effects by enabling one paralog to diverge while maintaining the other's native function [36]. To better mechanistically delineate the evolutionary path of diverging regulatory factors and its associated impact on regulatory networks, studies have employed random or saturation mutagenesis of regulatory proteins themselves. For example, mutations in the DNA-binding domain (DBD) of regulatory protein may alter its binding strength or specificity to cis-regulatory elements, which can lead to alternative or new network connections [2]. Global transcription machinery engineering (gTME) is an approach that utilizes random mutagenesis to generate regulatory protein variants that can yield novel phenotypes [37]. By screening through a library of *E. coli* primary sigma factor RpoD, the approach can isolate mutants with altered phenotypes including higher ethanol and SDS tolerance and increased lycopene production. Transcriptomic analysis of these RpoD variants revealed global changes in transcript levels, including altered expression of multiple genes often associated with stress response (e.g. outer membrane proteins). This approach has been further applied to other microbes, leading to new traits through regulatory network rewiring [38,39]. Additionally, recent advances in directed and saturation mutagenesis methods (e.g. MAGE-seq [40], CREATE [41], and PALS [42]) that leverage high-throughput DNA sequencing and synthesis provide opportunities to more comprehensively assess and precisely engineer the functionalization of new regulatory connections through trans-regulatory protein mutations.

Beyond interrogating single mutations in regulatory proteins, entire domains can be reshuffled and swapped. By comparative analysis, most regulatory proteins and their orthologs appear to organize into defined protein domains that are highly modular, which is perhaps an important evolutionary feature underlying its adaptability. For example, bacterial sigma factors generally encode two distinct DBDs that bind to either −35 or −10 motifs. Through comparative analysis of upstream cis-elements of extracellular function (ECF) sigma factors, Rhodius et al. identified the binding motifs for different ECF sigma factor groups [43]. The authors then generated chimeric sigma factors by combining different −35 and −10 DBDs from these different groups, which remapped sigma factors to new binding motifs. These new regulatory proteins exhibited orthogonal activation of other cis-elements and generated regulatory patterns that are otherwise unavailable to the native network. This type of approach provides an important foundation for future engineering and rewiring of native and synthetic regulatory systems.

In nature, genetic materials are often exchanged between bacteria to facilitate rapid adaptation. However, appropriate trans-factors need to exist in the recipient cell to modulate the expression of these horizontally acquired DNA and to generate novel phenotypes. This

process of network rewiring through lateral DNA exchange can be recapitulated in the laboratory, for instance, by cloning a heterologous bacterial sigma factor downstream of a non-native (e.g. constitutive or inducible) promoter on a plasmid that is then transformed into a cell [44,45]. When the primary *Pseudomonas putida* sigma factor RpoD is over-expressed in *E. coli* alongside a genomic library of *P. putida*, certain *E. coli* variants showed a significant increase in tolerance to pinene, a jet fuel precursor, for which *P. putida* possesses natural tolerance [44]. This suggests that transfer of sigma factors from one organism to another may play a role in regulatory adaptation through horizontal transfer. Similarly, when the *Lactobacillus plantarum* sigma factor was introduced to *E. coli*, its transcription machinery could express genes from a diverse set of metagenomic sources that would otherwise remain inactive in the *E. coli* host [45]. Thus, lateral acquisition of trans-regulators could rewire a regulatory network to an alternatively evolved architecture where multiple transcriptional units are re-tuned to yield phenotypes that are found in another organism. While expression of foreign sigma factors may result in broad transcriptomic changes, expression of more specific regulatory factors such as ECFs could modulate only a sub-set of the regulatory network. For instance, the overexpression of a putative pathway-specific ATP-binding regulator of the LuxR family (LAL) protein in *Streptomyces ambofaciens* triggered the activation of an otherwise silent biosynthetic gene cluster (BGC) from which novel macrolides are produced [46]. These results highlight that heterologous expression of regulatory proteins is a powerful approach to study the natural evolution and rewiring of regulatory networks and to generate potentially new phenotypes through a set of simple genetic perturbations.

Rewiring connections between cis-elements and trans-factors across regulatory networks

Cis-elements and trans-regulatory factors work in concert to regulate gene expression across a regulatory network. A key question is how changes in the regulatory connections translate into cellular phenotypes. One approach is to generate libraries of novel cis-element and trans-regulator combinations to mimic network rewiring events, thereby integrating various aspects of regulatory evolution such as cis-element mutations, trans-factor divergence, and horizontal transfer together into a single experimental framework (Figure 1c) [47,48]. In a recent study, two *Salmonella enterica* global regulators, H-NS and StpA, were rewired to take on one another's regulatory input logic by swapping the location of their CDSs [49]. This caused not only altered dynamic expression of the target genes of the two regulators but also another regulator, the stationary phase sigma factor RpoS, and all its regulated genes. Furthermore, the altered transcriptome yielded phenotypic differences such that the rewired strain could outcompete the wild type strain at 37 °C growth but not at 25 °C. This study illustrates how a relatively simple rewiring of two global regulators can significantly alter the transcriptome and confer a distinct phenotype.

A key question in understanding regulatory network architecture and evolution is to delineate its robustness to network rewiring. While changes in protein sequences may be easily predicted and their consequences in the cell directly measurable, the effects of regulatory rewiring on the transcriptome and the cellular phenotype may be more difficult to

unravel. Past studies explored the robustness of regulatory networks by overexpressing or deleting individual regulatory genes [50]. More recently, Isalan et al. generated a library of *E. coli* cis-elements and trans-regulators that re-assigned or added new regulatory connections between regulators [47]. The study tested 600 different network alterations and found that 95% of the new architectures yielded viable cells with 85% having little or no growth defects. This result was especially interesting as many of the regulators were global regulatory proteins (e.g. sigma factors) that controlled hundreds of genes. Some rewired variants exhibited higher fitness than wild type in different selective conditions including heat shock, nutrient starvation, and extended serial passaging, highlighting that many single rewiring events can yield directly beneficial phenotypes.

To further understand the underlying mechanism of the observed cellular adaptations, transcriptomes of 100 selected rewired variants were analyzed and clustered into one of 20 unique profiles, each representing a discrete cellular state [48]. The largest cluster of 25 unique rewired variants were associated with changes in expression of the ribosomal machinery, with ~125 upregulated ribosomal, ATP synthase, and tRNA genes. Interestingly, five different rewired regulators with different cis-elements yielded the same transcriptome profile. These variants affected the same set of master regulators including OmpC, none of which were directly rewired in the study. Similar results were seen for rewired variants that elicited changes in gene expression of flagellar genes driven by the master regulator FilA. Overexpression of OmpC and FilA recapitulated the altered transcriptomes directly, suggesting that rewired networks are tapping into various sets of preconditioned biological responses. These synthetic regulatory rewiring approaches provide opportunities to further generate fundamental insights into how regulatory networks produce novel adaptive responses via defined and tractable evolutionary paths.

Outlook

Next generation sequencing technologies have enabled an unprecedented capacity to determine and compare genomic sequences to better discover, annotate, and dissect regulatory function. In parallel, high-throughput DNA synthesis has facilitated the systematic exploration and analysis of regulatory cis-elements, trans-factors, and their interaction networks. While metagenomic sequencing efforts have generated a significant amount of raw data, better analysis approaches are still needed to assign function to these sequences and to understand the regulatory architecture governing gene expression in individual cells and across whole communities [51,52]. Furthermore, whole-cell models are needed to integrate these complex cellular process with dynamic regulatory programs [53]. More experimental tools leveraging advances in genome engineering, editing, and regulatory control can provide better platforms to rewire and perturb network components and interactions. Emerging CRISPR-mediated gene activation and repression systems could more precisely modulate gene expression in a highly programmable and multiplexed fashion [54]. As large-scale gene and pathway synthesis improves, synthetic regulatory circuits can take on more sophistication, moving beyond simple architectures such as the repressilator [55] towards bigger and more complex designs [56]. These synthetic circuits can be used to better test many features of natural regulatory networks including adaptability, scalability, and robustness to improve our understanding and design of regulatory networks [57,58].

Useful synthetic networks could be implemented in organisms with significant bio-industrial relevance or even in more complex mammalian systems. Ultimately, altering regulatory networks through high-throughput systematic and synthetic approaches constitute a new set of powerful approaches to understand fundamental principles of regulatory structure and evolution and hold great promise to rapidly generate organisms with new phenotypes.

Acknowledgments

We thank members of the Wang lab for helpful scientific discussions and feedback. H.H.W. acknowledges relevant funding support from the NIH (1DP5OD009172–02, U01GM110714-01A1), NSF (MCB-1453219), Sloan Foundation (FR-2015-65795), DARPA (W911NF-15-2-0065), and ONR (N00014-15-1-2704).

References

Papers of particular interest, published within the period of review, have been highlighted as:

* of special interest

** of outstanding interest

1. Sorrells TR, Johnson AD. Making sense of transcription networks. Cell. 2015; 161:714–723. [PubMed: 25957680]
2. Babu MM, Luscombe NM, Aravind L, Gerstein M, Teichmann SA. Structure and evolution of transcriptional regulatory networks. Curr Opin Struct Biol. 2004; 14:283–291. [PubMed: 15193307]
3. Alon U. Network motifs: theory and experimental approaches. Nat Rev Genet. 2007; 8:450–461. [PubMed: 17510665]
4. Barabasi A-L, Albert R. Emergence of scaling in random networks. Science. 1999; 286:509–512. [PubMed: 10521342]
5. Yu H, Gerstein M. Genomic analysis of the hierarchical structure of regulatory networks. Proc Natl Acad Sci. 2006; 103:14724–14731. [PubMed: 17003135]
6. Boyle AP, Araya CL, Brdlik C, Cayting P, Cheng C, Cheng Y, Gardner K, Hillier LW, Janette J, Jiang L, et al. Comparative analysis of regulatory information and circuits across distant species. Nature. 2014; 512:453–456. [PubMed: 25164757]
7. Draghi JA, Parsons TL, Wagner GP, Plotkin JB. Mutational robustness can facilitate adaptation. Nature. 2010; 463:353–355. [PubMed: 20090752]
8. Habib N, Wapinski I, Margalit H, Regev A, Friedman N. A functional selection model explains evolutionary robustness despite plasticity in regulatory networks. Mol Syst Biol. 2012; 8:1–18.
9. Payne JL, Wagner A. The robustness and evolvability of transcription factor binding sites. Science. 2014; 343:875–877. [PubMed: 24558158]
10. McAdams HH, Srinivasan B, Arkin AP. The evolution of genetic regulatory systems in bacteria. Nat Rev Genet. 2004; 5:169–178. [PubMed: 14970819]
11. Taylor TB, Mulley G, Dills AH, Alsohim AS, McGuffin LJ, Studholme DJ, Silby MW, Brockhurst MA, Johnson LJ, Jackson RW. Evolutionary resurrection of flagellar motility via rewiring of the nitrogen regulation system. Science. 2015; 347:1014–1017. [PubMed: 25722415]
12. Perez JC, Groisman EA. Evolution of transcriptional regulatory circuits in bacteria. Cell. 2009; 138:233–244. [PubMed: 19632175]
13. Oren Y, Smith MB, Johns NI, Kaplan Zeevi M, Biran D, Ron EZ, Corander J, Wang HH, Alm EJ, Pupko T. Transfer of noncoding DNA drives regulatory rewiring in bacteria. Proc Natl Acad Sci. 2014; 111:16112–16117. [PubMed: 25313052]
14*. Thompson D, Regev A, Roy S. Comparative analysis of gene regulatory networks: from network reconstruction to evolution. Annu Rev Cell Dev Biol. 2015; 31:399–428. [PubMed: 26355593]

15. Mukherjee S, Berger MF, Jona G, Wang XS, Muzzey D, Snyder M, Young RA, Bulyk ML. Rapid analysis of the DNA-binding specificities of transcription factors with DNA microarrays. Nat Genet. 2004; 36:1331–1339. [PubMed: 15543148]

16. Berger MF, Philippakis AA, Qureshi AM, He FS, Estep PW, Bulyk ML. Compact, universal DNA microarrays to comprehensively determine transcription-factor binding site specificities. Nat Biotechnol. 2006; 24:1429–1435. [PubMed: 16998473]

17. Berger MF, Bulyk ML. Universal protein-binding microarrays for the comprehensive characterization of the DNA-binding specificities of transcription factors. Nat Protoc. 2009; 4:393–411. [PubMed: 19265799]

18. Stoltenburg R, Reinemann C, Strehlitz B. SELEX—a (r)evolutionary method to generate high-affinity nucleic acid ligands. Biomol Eng. 2007; 24:381–403. [PubMed: 17627883]

19. Darmostuk M, Rimpelova S, Gbelcova H, Ruml T. Current approaches in SELEX: an update to aptamer selection technology. Biotechnol Adv. 2015; 33:1141–1161. [PubMed: 25708387]

20. Bailey TL, Elkan C. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proceedings Int Conf Intell Syst Mol Biol. 1994; 2:28–36.

21. Liu X, Brutlag DL, Liu JS. BioProspector: discovering conserved DNA motifs in upstream regulatory regions of co-expressed genes. Pac Symp Biocomput. 2001; 138:127–138.

22. Elemento O, Slonim N, Tavazoie S. A universal framework for regulatory element discovery across all genomes and data types. Mol Cell. 2007; 28:337–350. [PubMed: 17964271]

23. Das MK, Dai H-K. A survey of DNA motif finding algorithms. BMC Bioinf. 2007; 8:S21.

24. Sharon E, Kalma Y, Sharp A, Raveh-Sadka T, Levo M, Zeevi D, Keren L, Yakhini Z, Weinberger A, Segal E. Inferring gene regulatory logic from high-throughput measurements of thousands of systematically designed promoters. Nat Biotechnol. 2012; 30:521–530. [PubMed: 22609971]

25. Mogno I, Kwasnieski JC, Cohen BA. Massively parallel synthetic promoter assays reveal the in vivo effects of binding site variants. Genome Res. 2013; 23:1908–1915. [PubMed: 23921661]

26. Faith JJ, Hayete B, Thaden JT, Mogno I, Wierzbowski J, Cottarel G, Kasif S, Collins JJ, Gardner TS. Large-scale mapping and validation of Escherichia coli transcriptional regulation from a compendium of expression profiles. PLoS Biol. 2007; 5:e8. [PubMed: 17214507]

27. Kosuri S, Church GM. Large-scale de novo DNA synthesis: technologies and applications. Nat Methods. 2014; 11:499–507. [PubMed: 24781323]

28**. Patwardhan RP, Lee C, Litvin O, Young DL, Pe'er D, Shendure J. High-resolution analysis of DNA regulatory elements by synthetic saturation mutagenesis. Nat Biotechnol. 2009; 27:1173–1175. [PubMed: 19915551]

29. Klein JC, Lajoie MJ, Schwartz JJ, Strauch E-M, Nelson J, Baker D, Shendure J. Multiplex pairwise assembly of array-derived DNA oligonucleotides. Nucleic Acids Res. 2016; 44:e43–e43. [PubMed: 26553805]

30*. Kosuri S, Goodman DB, Cambray G, Mutalik VK, Gao Y, Arkin AP, Endy D. Church GM: composability of regulatory sequences controlling transcription and translation in Escherichia coli. Proc Natl Acad Sci. 2013; 110:14024–14029. [PubMed: 23924614]

31. Mutalik VK, Guimaraes JC, Cambray G, Mai Q-A, Christoffersen MJ, Martin L, Yu A, Lam C, Rodriguez C, Bennett G, et al. Quantitative estimation of activity and quality for collections of functional genetic elements. Nat Methods. 2013; 10:347–353. [PubMed: 23474467]

32. Mutalik VK, Guimaraes JC, Cambray G, Lam C, Christoffersen MJ, Mai Q-A, Tran AB, Paull M, Keasling JD, Arkin AP, et al. Precise and reliable gene expression via standard transcription and translation initiation elements. Nat Methods. 2013; 10:354–360. [PubMed: 23474465]

33. Gertz J, Siggia ED, Cohen BA. Analysis of combinatorial cis-regulation in synthetic and genomic promoters. Nature. 2009; 457:215–218. [PubMed: 19029883]

34. Salis HM, Mirsky EA, Voigt CA. Automated design of synthetic ribosome binding sites to control protein expression. Nat Biotechnol. 2009; 27:946–950. [PubMed: 19801975]

35. Chen Y-J, Liu P, Nielsen AAK, Brophy JAN, Clancy K, Peterson T, Voigt CA. Characterization of 582 natural and synthetic terminators and quantification of their design constraints. Nat Methods. 2013; 10:659–664. [PubMed: 23727987]

36. Voordeckers K, Pougach K, Verstrepen KJ. How do regulatory networks evolve and expand throughout evolution? Curr Opin Biotechnol. 2015; 34:180–188. [PubMed: 25723843]

37**. Alper H, Stephanopoulos G. Global transcription machinery engineering: a new approach for improving cellular phenotype. Metab Eng. 2007; 9:258–267. [PubMed: 17292651]

38. Alper H, Moxley J, Nevoigt E, Fink GR, Stephanopoulos G. Engineering yeast transcription machinery for improved ethanol tolerance and production. Science. 2006; 314:1565–1568. [PubMed: 17158319]

39. Klein-Marcuschamer D, Stephanopoulos G. Assessing the potential of mutational strategies to elicit new phenotypes in industrial strains. Proc Natl Acad Sci. 2008; 105:2319–2324. [PubMed: 18252819]

40. Kelsic ED, Chung H, Cohen N, Park J, Wang HH, Kishony R. RNA structural determinants of optimal codons revealed by MAGE-seq. Cell Syst. 2016; 3:563–571. [PubMed: 28009265]

41. Garst AD, Bassalo MC, Pines G, Lynch SA, Halweg-Edwards AL, Liu R, Liang L, Wang Z, Zeitoun R, Alexander WG, et al. Genome-wide mapping of mutations at single-nucleotide resolution for protein, metabolic and genome engineering. Nat Biotechnol. 2016; 35:48–55. [PubMed: 27941803]

42. Kitzman JO, Starita LM, Lo RS, Fields S, Shendure J. Massively parallel single-amino-acid mutagenesis. Nat Methods. 2015; 12:203–206. [PubMed: 25559584]

43*. Rhodius VA, Segall-Shapiro TH, Sharon BD, Ghodasara A, Orlova E, Tabakh H, Burkhardt DH, Clancy K, Peterson TC, Gross CA, et al. Design of orthogonal genetic switches based on a crosstalk map of s, anti- s, and promoters. Mol Syst Biol. 2014; 9:702–702.

44*. Tomko TA, Dunlop MJ. Expression of heterologous sigma factor expands the searchable space for biofuel tolerance mechanisms. ACS Synth Biol. 2017; 6:1343–1350. [PubMed: 28319371]

45. Gaida SM, Sandoval NR, Nicolaou SA, Chen Y, Venkataramanan KP, Papoutsakis ET. Expression of heterologous sigma factors enables functional screening of meta-genomic and heterologous genomic libraries. Nat Commun. 2015; 6:7045. [PubMed: 25944046]

46. Laureti L, Song L, Huang S, Corre C, Leblond P, Challis GL, Aigle B. Identification of a bioactive 51-membered macrolide complex by activation of a silent polyketide synthase in Streptomyces ambofaciens. Proc Natl Acad Sci. 2011; 108:6258–6263. [PubMed: 21444795]

47**. Isalan M, Lemerle C, Michalodimitrakis K, Horn C, Beltrao P, Raineri E, Garriga-Canut M, Serrano L. Evolvability and hierarchy in rewired bacterial gene networks. Nature. 2008; 452:840–845. [PubMed: 18421347]

48*. Baumstark R, Hänzelmann S, Tsuru S, Schaerli Y, Francesconi M, Mancuso FM, Castelo R, Isalan M. The propagation of perturbations in rewired bacterial gene networks. Nat Commun. 2015; 6:10105. [PubMed: 26670742]

49. Fitzgerald S, Dillon SC, Chao T-C, Wiencko HL, Hokamp K, Cameron ADS, Dorman CJ. Re-engineering cellular physiology by rewiring high-level global regulatory genes. Sci Rep. 2015; 5:17653. [PubMed: 26631971]

50. Albert, Jeong, Barabasi. Error and attack tolerance of complex networks. Nature. 2000; 406:378–382. [PubMed: 10935628]

51. Simon C, Daniel R. Metagenomic analyses: past and future trends. Appl Environ Microbiol. 2011; 77:1153–1161. [PubMed: 21169428]

52. Quince C, Walker AW, Simpson JT, Loman NJ, Segata N. Shotgun metagenomics, from sampling to analysis. Nat Biotechnol. 2017; 35:833–844. [PubMed: 28898207]

53. Karr JR, Sanghvi JC, Macklin DN, Gutschow MV, Jacobs JM, Bolival B, Assad-Garcia N, Glass JI, Covert MW. A whole-cell computational model predicts phenotype from genotype. Cell. 2012; 150:389–401. [PubMed: 22817898]

54. Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, Lim WA. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell. 2013; 152:1173–1183. [PubMed: 23452860]

55. Elowitz MB, Leibler S. A synthetic oscillatory network of transcriptional regulators. Nature. 2000; 403:335–338. [PubMed: 10659856]

56. Nielsen AAK, Der BS, Shin J, Vaidyanathan P, Paralanov V, Strychalski EA, Ross D, Densmore D, Voigt CA. Genetic circuit design automation. Science. 2016:352.

57. Lim WA, Lee CM, Tang C. Design principles of regulatory networks: searching for the molecular algorithms of the cell. Mol Cell. 2013; 49:202–212. [PubMed: 23352241]

58. Cardinale S, Arkin AP. Contextualizing context for synthetic biology - identifying causes of failure of synthetic biological systems. Biotechnol J. 2012; 7:856–866. [PubMed: 22649052]

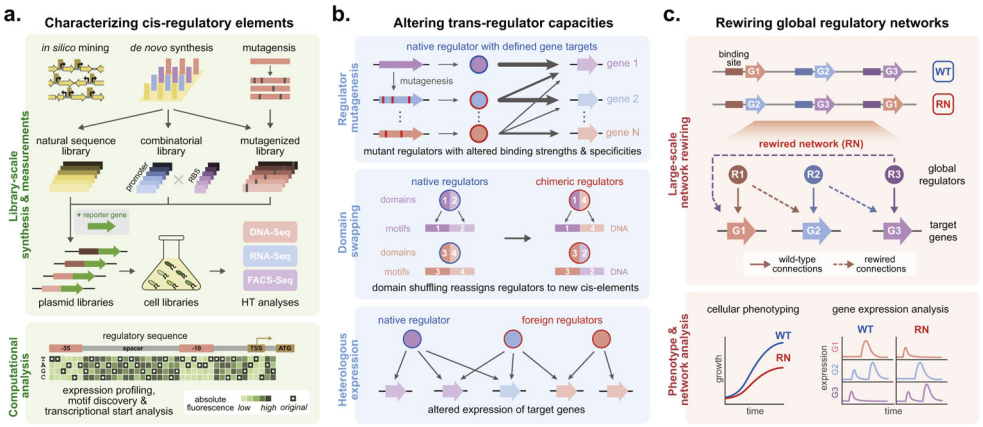

Figure 1. Approaches to systematically dissect regulatory network architectures and processes

(a) Systematic analysis of cis-regulatory elements can be performed through *in silico* library design, *in vitro* DNA synthesis, and *in vivo* characterization by various methods including DNA-seq, RNA-seq, and FACS-seq. Detailed analyses of high-throughput datasets yield new regulatory motifs and transcriptomic information **(b)** Modulating gene expression by altering trans-regulator capacities via deep mutagenesis, domain swapping, and heterologous expression. Mutagenesis of native regulators yields variants with altered cis-element specificity and strength. Domain swapping generates chimeric regulators with altered binding profiles to native regulators. Heterologous expression of foreign regulators can activate cis-elements that are not normal targets of native regulators. **(c)** Rewiring regulatory networks through combinatorial libraries that alternatively assign cis-elements and trans-regulators with new connections to explore new network architectures. Two regulatory networks are shown, wild type (WT) and rewired network (RN). In the rewired network, global regulators (R1–R3) are assigned to different target genes (dashed lines) compared to their wild-type targets (solid lines). Various characterization tools for phenotyping or network analysis can be utilized to assess network architecture, dynamics, and performance.