Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Systems Biology

Systematic and synthetic approaches to rewire regulatory networks

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

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.

Addresses

[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

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

Current Opinion in Systems Biology 2018, ■:1–7

This review comes from a themed issue on **Regulatory and metabolic networks (2018)**

Edited by **Bas Teusink** and **Uwe Sauer**

For a complete overview see the Issue and the Editorial

Available online xxx

https://doi.org/10.1016/j.coisb.2017.12.009

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 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

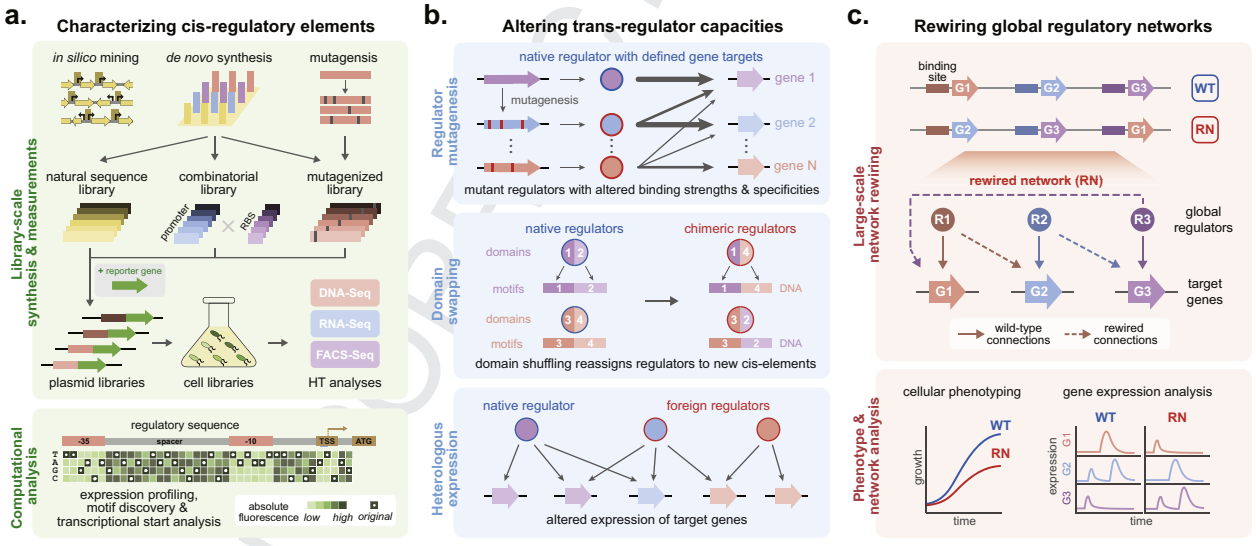
[15–17]. Systematic evolution of ligands by exponential enrichment (SELEX) is a foundational method for understanding and modeling protein-DNA interactions [18]. Pairing such approaches with high-throughput sequencing have improved the measurement of regulatory protein affinity to defined DNA targets [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

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.