Published in final edited form as:
Wiley Interdiscip Rev Syst Biol Med. 2011 September ; 3(5): 544–561. doi:10.1002/wsbm.136.

Adaptation of cells to new environments

Aaron N. Brooks,
Molecular and Cellular Biology Program, University of Washington, Seattle, WA 98195, Institute for Systems Biology, 1441 N 34th Street, Seattle, WA 98103

Serdar Turkarslan,
Institute for Systems Biology, 1441 N 34th Street, Seattle, WA 98103

Karlyn D. Beer,
Molecular and Cellular Biology Program, University of Washington, Seattle, WA 98195, Institute for Systems Biology, 1441 N 34th Street, Seattle, WA 98103

Fang Yin Lo, and
Molecular and Cellular Biology Program, University of Washington, Seattle, WA 98195, Institute for Systems Biology, 1441 N 34th Street, Seattle, WA 98103

Nitin S. Baliga
Institute for Systems Biology, 1441 N 34th Street, Seattle, WA 98103,
nbaliga@systemsbiology.org

Abstract

The evolutionary success of an organism is a testament to its inherent capacity to keep pace with environmental conditions that change over short and long periods. Mechanisms underlying adaptive processes are being investigated with renewed interest and excitement. This revival is partly fueled by powerful technologies that can probe molecular phenomena at a systems scale. Such studies provide spectacular insight into the mechanisms of adaptation, including rewiring of regulatory networks via natural selection of horizontal gene transfers, gene duplication, deletion, readjustment of kinetic parameters, and myriad other genetic reorganizational events. Here, we will discuss advances in prokaryotic systems biology from the perspective of evolutionary principles that have shaped regulatory networks for dynamic adaptation to environmental change.

Microorganisms experience myriad environmental factors over their evolutionary history, including those that remain essentially constant over long periods (e.g. geological epochs), change slowly (e.g. general increase in annual temperatures), fluctuate periodically (e.g. day-night cycles and seasonal variations), or change frequently and somewhat randomly (e.g. unpredictable nutrient loading). These changes occur over diverse timescales, ranging from the lifetime of an individual cell to multiple generations. Accordingly, microbes have evolved unique strategies to deal with the peculiarities of their environment (1). Characteristic examples include adaptation to extreme environmental niches (2,3), entrainment of phototrophs to day-night cycles (4,5), and the physiological adjustment of *E. coli* as it passes through the human intestine (6). Organisms respond to short term environmental changes by reversibly adjusting their physiology to maximize resource utilization while maintaining structural and genetic integrity by repairing and minimizing damage to cellular infrastructure (7,8) thereby balancing innovation with robustness. Naturally, physiological response networks emerge as products of evolution by natural selection where they can lend reproductive or fitness advantage, particularly when the environmental change is recurrent (9).

Environmental adaptation of biological systems can be considered from three evolutionary perspectives: (i) acclimation of existing cellular machinery to operate optimally in a new

environmental niche; (ii) acquisition of entirely new capabilities through horizontal gene transfer or neofunctionalization of gene duplications and (iii) reorganization of network dynamics to appropriately adjust existing physiological processes to match dynamic environmental changes. The first type of adaptation can arise through two types of events that differ dramatically in duration. Simple mutations can greatly increase fitness over very short time frames (within one or few generations). Prominent examples of short-term adaptive events include resistance to drugs (10,11) and altered nutrient conditions (12). Alternatively, complex mutations in multiple loci may accumulate over very long time frames, such as the evolution of acidic protein surfaces in halophilic archaea (2,13,14). While the initial transfer of adaptive genes by HGT occurs rapidly (15), full integration of laterally transferred component(s) typically occurs over longer time frames (10s of millions of years), where HGT events often require regulatory rewiring to function optimally in the context of existing cellular networks (16). Finally, physiological readjustment occurs both because of genetic and physiological robustness to withstand stress that accumulates over many generations and latent genetic variance that is revealed after environmental perturbation (17).

This review focuses on the evolution of adaptive mechanisms for acclimation to recurrent yet transient environmental changes. When transient changes are recurrent they select for genetic traits that confer fitness by improving the ability of an organism to rapidly and reversibly adjust physiology to match current environmental conditions. These traits manifest at varying hierarchies of genetic information processing, from receptors for sensing environmental factors, to signal relay, transcriptional, post-transcriptional, translational and post-translational control mechanisms, and also at the metabolic level through modulation of enzyme function (affinities, kinetics etc.). Such adaptive changes occur over intermediate time frames (upwards of 100s of generations in *E. coli* K-12 MG1655; (18)) and, surprisingly, they arise repeatedly and in some cases with some regularity in distinct lineages (19) Fitness, or the number of surviving offspring after one generation (20), is a complex property that emerges from the integration of changes at all these levels. A holistic systems approach, therefore, is necessary to fully appreciate how these varied mechanisms work together when an organism adapts to a new environment.

For the purpose of our discussion, we define environmental conditions to include both abiotic physical variables (such as light and temperature) and biotic components (such as other co-inhabiting organisms). Additionally, we restrict our analysis to asexual prokaryotic systems (including both bacteria and archaea) in which the dynamics and mechanisms of genetic evolution lack the pervasive variation that recombination by sexual reproduction promotes. Sexual populations also may not experience signatures of selection prevalent in asexual populations, such as ‘classic’ sweeps associated with unconditionally advantageous mutations (21). The reader should note, however, that the physiology of archaea and bacteria diverge substantially, with archaea sharing startling similarity to eukaryotes. Infrequently, we may explicitly refer to findings in eukaryotes that reflect mechanisms that likely also occur in prokaryotic systems.

This review will bridge the conceptual gap between *adaptation*, which by definition requires heritable genetic change (20), and *physiological readjustment*, which is a product of adaptation that equips organisms to attune their physiology to dynamic changes in their environments. We will suggest how systems-level methodologies and insights can be applied to better understand the strategies living systems employ to withstand and in some cases take advantage of change in their environments.

The fields of microbiology, molecular evolution, and systems biology are expansive – no single review can possibly cover all adaptive mechanisms and scenarios that may influence

the evolution of natural microbial populations. Instead, we will highlight major themes and new insights in microbial evolution while demonstrating how the principles of systems biology can be leveraged to develop a more comprehensive, integrative understanding of cellular adaptation to new environments. Throughout our analysis we will point the reader to other outstanding reviews that complement our discussion.

TYPES OF ADAPTATIONS AND ASSOCIATED MECHANISMS

What is adaptation?

When we say that an organism has *adapted* to its habitat, we imply that it has evolved molecular mechanisms that allow it to grow optimally under the spatiotemporally varying physicochemical conditions of its environment. Evolution, however, is an unfinished process. Organisms chase fitness optima in constantly changing environments. Subtle fitness differences between individuals (due to genomic plasticity or metabolic flexibility) and phylogenetic complexity (i.e. numbers and diversity of species within a community) can lead to the diversification of species or the extinction of less fit genotypes over time (22). Although many adaptive mutations are lost by random chance (as a result of genetic drift), mutations that confer significant selective advantage have a greater propensity to become fixed within the population, especially in large populations (23). If selection imposed by the environment is particularly strong, fitness-enhancing genotypes will rapidly rise in frequency in the population, often carrying associated, possibly detrimental, genes along with them (i.e. selective sweep) (24) and interfering with one another via clonal interference (25).

Adaptation to linked environmental changes: general stress response or anticipatory behavior?

Conditions within natural environments can change continually over long periods, periodically, or transiently and unpredictably. In the context of evolution, these environmental changes occur simultaneously albeit on different timescales and exert selective pressure to enrich genotypes well matched to particular ecosystems. Not surprisingly, the repertoire of analogous solutions that characterizes success in a given environment can be similar across diverse organisms occupying similar habitats, suggesting convergent evolution of adaptive solutions that were discovered independently in divergent lineages (26,27). While a significant fraction of these responses are condition-specific, most organisms have also evolved robust generalized mechanisms to deal with shared aspects of stress resulting from diverse kinds of environmental changes. In yeast, for example, a set of ~900 genes responds similarly to a diverse array of environmental stresses, sharing common regulatory themes mediated by Yap1p, Msn2p, and Msn4p (28). This generalized stress response typically includes activation of heat shock proteins, phage shock proteins, and oxidative stress response proteins (29,30), although there are notable examples in *Candida albicans* and *Halobacterium salinarum sp. NRC-1* where the central role of general stress response has been called into question (31,32). The alternate perspective offered by these studies is that changes in environmental variables are physically linked and do not occur in isolation. Elevated temperatures, for example, induce a number of associated changes in environmental conditions, including decrease in oxygen solubility (33). Theoretical and experimental work in diverse species suggests that organisms can learn to take advantage of this natural co-variation between environmental parameters (e.g. temperature and oxygen), thereby displaying "anticipatory behavior" (18,33,34). *B. subtilis*, for example, retains short-term and long-term memories to inform sporulation dynamics and degradative enzyme synthesis (34).

Coupled environmental variables have important evolutionary implications that can be assessed by assaying the fitness consequences upon artificially decoupling such associations in the lab. Relative to laboratory populations evolved in an inverted environment, wild-type *E. coli* has a fitness disadvantage when naturally linked parameters, such as temperature and oxygen, are decoupled artificially. Likewise, cyanobacteria mutants with non-functional circadian clocks are less fit and out-competed by their wild-type counterparts (35). The enhanced fitness conferred by diurnal entrainment can be attributed to anticipation of predictable, associated stresses such as damaging UV radiation (4). Adaptation to stress may also prepare cells to better respond to future stresses. *Vibrio parahaemolyticus*, for example, has a greater tolerance to acidity and temperature stresses following growth in media containing 3% NaCl versus 1% NaCl concentrations (36). On the other hand, adaptation to a subset of environmental factors could come at a cost of sacrificing tolerance to others. Propionate adaptation of *Salmonella enteritidis*, for instance, leads to enhanced resistance to stresses experienced inside the host but overall decrease in infectivity (37). For all these reasons, we need to appreciate the high degree of connectedness within environmental networks to interpret causes and functional consequences of complex biological responses.

LONG-TERM ADAPTATION

Some environmental conditions seldom change. We define long-term adaptation as the response to environmental conditions that remain relatively constant throughout the lifetime of a species. For organisms, this represents the most fundamental level of adaptation. Over many thousands to millions of years or longer, the internal architecture of the cell changes to accurately match the general features of the habitat it occupies. Typically, the features that reflect long-term environmental adaptation are deeply ingrained in the structure of the genome, regulatory schemas, or even the molecules of the cell themselves. These molecular artifacts reflect the organism's uphill struggle towards increased fitness in an environment that changes (being directly modified by biotic factors and even the organism itself) yet remains mostly constant over long timescales. Complete reversion of these features would be difficult, if not impossible, for an organism to achieve over short time scales, suggesting that cells are poorly adapted to large variations in these environmental parameters. *Halobacterium salinarum* NRC-1, for example, thrives in high salt conditions. While *H. salinarum* can withstand lower salt concentrations (as low as 2.5M), anything below this concentration is lethal to cells. Adaptation to high salt conditions requires a number of significant structural changes in the cell, including global alterations at the level of DNA, RNA and protein composition (38). In high salt, water activity is low; this has profound consequences for enzymatic activity and the structural integrity of the cell membrane and genome. Adaptation to high-salt conditions has required the evolution of a highly-acidic proteome, high genomic GC content, and increased intracellular concentration of potassium cation.(39). Other factors including gene redundancy have been reported to enhance survival in fluctuating salt conditions. *Salinibacter ruber*, for instance, possesses two or more copies of several essential genes. It has been proposed that slight differences in the amino acid composition of two versions of the same protein (ecoparalogs) might allow *Salinibacter* to survive broader fluctuations in salt concentration (40). These strategies for thriving in hypersaline environments are generally utilized by diverse halophillic archaea. Halophilic bacteria and eukaryotes, on the other hand, have independently evolved alternate mechanisms such as overproduction of organic osmolytes (sugar and amino acid-derivatives) to live in high salt environments (39).

Adaptation to temperature is another well-studied example of molecular response to long-term environmental pressures. While all microbial species are adapted to some range of permissive temperatures, interesting mechanistic examples of thermal adaption have been reported in both extreme warm and cold environments. Psychrophillic organisms occupy

extremely cold ecosystems (permanent temperatures below 5C) located deep in the ocean, and in polar and alpine regions. The habitats colonized by psychrophiles are of particular interest because they constitute more than three-quarters of the Earth's surface (41). Similarly, hyperthermophiles (optimal growth temperature >80C) occupy high temperature habitats. These organisms are so well adapted to their environments that, under some conditions, their doubling times approach that of *E. coli* grown at 37C (41). In both cases, the microorganisms that occupy these habitats have evolved enzymes and metabolic strategies that allow them to survive and proliferate at temperatures that would be restrictive or lethal to mesophilic organisms. To withstand these extreme temperatures, microbial species have evolved molecular mechanisms that regulate membrane fluidity and conformational flexibility of proteins, and produce thermo-stable protein variants, all of which are significantly challenged at extreme temperatures (42,43). While the specific mechanisms employed to combat extreme temperatures vary across species, common mechanisms, such as increased protein stability through ion pairs, hydrogen bonding, hydrophobic interactions, disulfide bridges, packing, and intersubunit interactions at high temperatures and increased protein flexibility and membrane fluidity and exopolysaccharide production at low temperatures, are common aspects of cellular adaption to extreme environments (44). Like the halophilic organisms described earlier, thermophiles and psychrophiles occupy environments that impose a number of additional harsh constraints on living systems (including metal ion concentrations, nutrient limitation, and sometimes increased pressure), making them useful models to understand adaptation to stressful conditions.

SHORT-TERM ADAPTATION

Environmental variables can fluctuate in unanticipated ways. Such variations are often stressful and, depending on severity, might drive selection of genotypes that are better suited to readjust physiology to manage and mitigate the consequences of stress. Genotypes can be selected on the basis of either genetic or non-genetic components. Asymmetric cell division, in which mother and daughter cells receive disproportionate numbers of molecules, alters the dynamic behavior of cells in response to environmental change and may lead to the selection of genotypes that do not necessarily contain heritable allelic alterations, but rather exhibit a temporary, dynamic state compatible with the altered environment. Sporulation in *B. subtilis*, for instance, couples asymmetric morphological changes with differential gene expression between the forespore and mother cell (Reviewed in 45), leading to divergent cell fate. Due to asymmetric cell division, daughter cells may receive more or fewer ribosomes, transcription factors, or other cellular components, each of which might contribute to their success (or failure) in new environments. Phase variation (or phenotype switching) is another mechanism (non-genetic, though heritable alteration) by which microbial populations leverage stochastic variations in cell components to respond to uncertainty in environmental fluctuations (46). Genetic alterations, on the other hand, can occur anywhere within the cellular hierarchy; advantageous mutations may be located within proteins, affecting the stability or kinetic parameters of the protein, or within regulatory sequences, affecting when, where, and how much of a biological product is made. A common feature of these alterations, however, is that they tend to be simple, i.e. they are the product of one or a few adaptive mutations that spontaneously occur in response to the new environment or preexisting neutral or buffered mutations whose consequence is revealed by perturbation. In this sense, these adaptations are flexible, being easily gained and subsequently lost by genetic drift. The inherent plasticity of these mutations makes them especially important for physiological adaptation, as slight changes can drastically alter the dynamic response of an organism to stress. Short-term adaptation is tightly coupled to the longer-term adaptive mechanisms previously described. Temporary changes in long-term environmental trends may elicit genetic alterations that are short-lived, only becoming fixed

within the population and canalized into regulatory programs if the stress that has elicited the advantageous mutation surpasses a temporal threshold, becoming a regular feature of the environment. Numerous studies have investigated mechanisms of adaptation to altered growth temperature (47,48), nutrient composition (49,50) and population structure (51). Universally, these studies find that fitness to a new environment increases rapidly over the first several thousand generations (52,53). Surprisingly, adaptive mutations discovered in the lab typically occur in one or a few genes, reflecting the earliest events in the adaptive process. Only four mutations in *E. coli*, for example, are responsible for gaining growth advantage in stationary phase: one in the stationary-phase sigma factor , one in the leucine-responsive regulatory protein, two genomic rearrangements of IS5 transposon insertion sequences, and a mutation in the sgaC gene (51).

Short-term, stressful environmental perturbations may also specifically induce increased rates of mutation, potentially facilitating adaptive evolution by rapid exploration of a broader genotypic space. In response to nutrient starvation, for example, the DNA damage and cell-cycle checkpoint control response (SOS response pathway) is induced in *E. coli*. Following induction of this pathway, cells experience higher rates of mutation due to inhibition of mismatch repair and recombinational break repair, and induction of a mutator DNA polymerase (54). In addition, some genetic loci exhibit a disproportionatly high frequency of mutation during hypermutation. These mutational ‘hot’ spots may reflect regions of the genome that are more readily modified and thereby *adaptive.* Taken together, these observations suggest that organisms might possess the capacity to introduce a bias in the distribution of mutational variation along its chromosome(s).

ADAPTATION THROUGH REWIRING OF GENE REGULATORY NETWORKS

Systems-level coordination of cellular functions is accomplished by gene regulatory networks (GRNs) (55) The general form and features of biological networks are illustrated in Figure 1. Central to all GRNs is the interaction of TFs and their cognate DNA binding sites. Remarkably, the origin of DNA-binding domains (DBDs) in all present day TFs can be traced to a few ancestral classes, such as winged-helix and zinc ribbon domains (56). This raises important questions regarding the evolutionary process(es) that led to the diversification of these few DBDs to create a vast array of distinct DNA binding specificities. It could be argued that the common origins of DBDs within TFs and their division into related classes (protein families) should help characterize GRNs in one organism and suggest projections of that information onto orthologous systems in phylogenetically related species. Functions of even structurally similar transcription factors (TFs), however, can diverge substantially through alterations in regulatory domains of either the transcription factor itself or the cis-regulatory sequences of downstream target genes. The divergence of DNA binding specificities and allosteric domains of two FNR (fumarate and nitrate reduction) family TFs in *E. coli* and *B. subtilis* are case in point. As result of subtle DNA binding differences, FNR in *E. coli* controls 135 genes whereas its counterpart in *B. subtilis* regulates only 8 genes (57). Importantly, this result demonstrates that species-specific coevolution of interacting partners (in this case FNR and the target promoters) prohibits simple projection of TF-binding orthologies across species, even when the genes involved share recent ancestry. On the other hand, this finding underscores the flexibility of GRNs – malleability of regulatory network topology can promote variation in gene expression that acclimates a species to the nuances of its environment (58).

Gene regulatory networks evolve through a number of molecular mechanisms that vary in frequency and magnitude of effect. Figure 2 depicts and Box 1 describes several common mechanisms of GRN evolution and their consequences for network topology. While rewiring of GRNs is an efficient mechanism to acquire new features or functions, it is

important to remember that preexisting network topology constrains the space of viable and visible phenotypic outcomes that can result from alterations to its structure (59), especially over short time scales. Historical contingency guides evolution (60), where organisms cannot liberate themselves from their past or innovate beyond the constraints of their current genetic makeup. A related, unresolved problem is how changes in genotype map to phenotype (the representation problem). From this perspective, organisms balance two opposed evolutionary characteristics: evolvability (change) and robustness (resistance to change). Recent experimental studies, for example, have found organisms to be remarkably robust at the level of gene expression, even when challenged by potentially catastrophic rewiring of regulatory components (61) or variations in gene network dosage (62). By contrast, other (primarily theoretical) studies highlight the tendency towards increased evolvability, i.e. the capacity of organisms to generate diversity, in complex, adaptive systems (Reviewed in 63). We suspect that the two counterpoints are a result of disparities in the time dimensions over which these two properties are assessed. Evolvability, by definition, is a property that manifests over a very long period of time whereas robustness is typically assessed in laboratory experiments which are conducted over time frames that are too short to resolve fractional fitness differences. Indeed, evolvablity itself may be a selectable trait in biological systems. Simulations of protein evolution, for example, suggest that fluctuating environments elicit large-scale genetic changes that correlate with the frequency and severity of the environmental change (64). From an engineering standpoint, measures of the topology of biological networks may suggest regions of the network or specific genes that may be more plastic (i.e. evolvable) compared to other regions of the genome, which may guide rational reengineering.

EVOLUTION OF MICROBIAL COMMUNITIES

Part of the complexity of natural environments is a product of interspecies dynamics. Members of ecological communities influence each other by altering their natural environment, competing for common resources, cooperating with one another to solve common challenges, complementing each other's nutritional needs and preying on one another. Interspecies interactions directly influence natural selection by modifying the selective criteria imposed by the environment. The evolution of a single species is tightly coupled to the co-inhabitants of its environment. Importantly, the evolution of a species in isolation may proceed differently when that same species evolves in the presence of other organisms. Competition between *S. pneumoniae* and *H. infuenzae* for colonization of mucosal surfaces, for example, drives *S. pneumoniae* to develop opsonophagocytosis-resistance, which is associated with invasive nasopharynx diseases. In the absence of *H. influenzae,* this resistance mutation bears a fitness cost; yet, in competition with *H. influenzae,* which actively stimulates neutrophil-mediated opsonophagocytosis, the resistance mutation becomes advantageous (65). Besides adjusting to varied selective criteria that are the consequence of interspecies dynamics, cells in natural environments have access to vast pools of genetic material with which they can diversify and innovate. The prevalence of horizontal genetic diversity within microbial populations was severely underestimated until recently (Reviewed in 66). Faced with stressful conditions that are difficult to address with existing molecular parts, an organism may look outward to find genes that increase its likelihood of survival. Many microorganisms maintain conserved pathways through which they become competent to uptake DNA from the environment. In *B. subtilis*, the activation of competence is mediated by noise in expression of the regulator ComK (67,68). Given stressful conditions, *B. subtilis* dedicates approximately ten percent of its population to survey the environment for new survival strategies. Such mechanisms can dramatically increase the rate of evolution in asexual populations. Another important property of microbes that has only recently been appreciated is genomic plasticity, which refers to variation in genome architectures across individuals of a natural population. It is

becoming increasingly evident that microbial populations are rife with genomic variability, bringing into question the definition of a microbial species itself (69,70). Together, genomic plasticity and metabolic flexibility (the ability of organisms to reversibly adjust biochemical capability by regulating gene expression and enzyme activities) increases the adaptive capability of a microbial community.

SYSTEMS APPROACHES TO STUDY MECHANISMS OF ADAPTATION AND THEIR EVOLUTION

Adaptational events produce molecular signatures at every level of the cellular hierarchy. Some are deeply ingrained in the molecular structures of the cell, whereas others are simpler and therefore rapidly modified. A combination of comparative genomics and systems biology is ideally suited to infer the molecular mechanisms of adaptation from diverse data types collected across taxa. Substantial progress has been made toward this goal. Here we consider advances that have contributed to our understanding of common adaptive mechanisms and highlight emerging technologies and methodologies that will deepen our exploration of evolution in microbial populations.

Comparative Genomics

Since the completion of the first bacterial genome-sequencing project in 1995 (71), the number of completely sequenced organisms has increased rapidly. While initial studies confined themselves to a narrow spectrum of phylogenetic diversity, newer sequencing efforts have leveraged known 16s rRNA associations to suggest additional organisms and lineages whose gene sets may be under sampled (72). Beyond sequencing the genome of a single species, recent metagenomic studies attempt to capture the genetic diversity present in complex environmental samples, characterizing novel genes from a number of microbial species that are uncultured under standard laboratory conditions (73,74). Whole genome sequences from organisms spanning diverse evolutionary domains have enabled comparative analysis between closely (75) and distantly related species and lineages (76,77). Orthology between genes in separate lineages can suggest similarity in function (78), participation in common pathways (79), and shared regulatory motifs (80), although the precise mechanisms behind co-regulation of similar genes is often markedly different between species (81). A challenge of comparative approaches, however, is to separate adaptive events from *exaptive* events; that is, to isolate those events that are either random (i.e. the result of genetic drift) or structural consequences of other acquired traits (82). Exaptations can be prevalent even between closely related species. For instance, small changes in gene regulatory circuits have resulted in substantial diversification of closely related species. In the fungi *Kluyveromyces lactis* and *Candida albicans*, the combinatorial circuits regulated by Mcm1 have diverged significantly (83) through gain or loss of Mcm1 binding sites and combinatorial associations with other transcriptional cofactors. After ~300 million years of divergence between *S. cerevisiae* and these two fungal species, only 15% of the direct Mcm1-target gene interactions remain intact. Determining which of the myriad genetic changes results in increased fitness in new environments is a challenge for laboratory experimental evolution studies. In the subsequent sections, we consider experimental methodologies to characterize molecular mechanisms that drive cellular adaptation to new environments.

Laboratory evolution (directed selection)

Experimental evolution studies have changed our understanding of short-term adaptation in microbial populations. A common strategy employed in these studies is to enrich, over several generations, mutants that are better suited to propagate in a gradually changing environment. In such experiments adaptation to new conditions (improved fitness relative to

the ancestral genotype) emerges quickly, within a few to hundreds of generations depending on the type and severity of stress. Perhaps most well-known and celebrated are Richard Lenski's long-term evolution studies in *E. coli*. In an early study, Travisano and Lenski propagated 12 replicate *E. coli* populations for 2,000 generations in a glucose-limited environment (84). To assay for increased fitness, the authors competed these evolved strains against their ancestor in 11 novel, single-nutrient environments. In cases where the uptake mechanism of the novel nutrient is similar to glucose, the evolved strains exhibited similar levels of increased fitness; in response to nutrients with uptake mechanisms different than glucose, however, the strains behaved more unpredictably, suggesting that each strain achieved adaptation to glucose limitation by an independent mechanism. In follow up to this work, Blount et al. reported the adaptation of *E. coli* to growth in minimal glucose supplemented with citrate (85). Wild-type *E. coli* cannot utilize citrate as a carbon source under oxic conditions. The adaptive ability of *E. coli* to utilize citrate manifested after 33,127 generations and over twenty years of growth under selective conditions. Adaptation of this novel functionality likely required at least three genetic changes. While the exact location and type of mutations are still unknown, the authors suggest that the adapted *E. coli* strains, which have the ability to metabolize citrate but lack the ability to transport it into the cell, may have activated a cryptic citrate transporter. This study is of particular importance both because it demonstrates the emergence of novel functionality during the course of a laboratory experiment and suggests that mutational events have historical contingency. Without at least two "potentiating" mutations in the genetic background, citrate metabolism evolved infrequently. Most recently, Barrick et al. compared the rates of genomic evolution and adaptation in *E. coli*. They confirm the long-standing observation that adaptation slows considerably after several thousand generations (86) and demonstrate that the rate of genomic mutation remains relatively constant for as many as 20,000 generations (53). Surprisingly, most mutations observed in this experiment were beneficial. Taken as a whole, this body of work illuminates important evolutionary mechanisms and raises important questions regarding the reproducibility of adaptation, duration in the periods of rapid evolution followed by stasis, and the role of chance events like mutation and drift in adaptive evolution. Later on we will discuss how complex genetic interactions also play an important role in defining the constraints of adaptive evolution (52).

Studies in *Pseudomonas fluorescens* SBW25 have also elegantly demonstrated the evolution of novel phenotypes under laboratory conditions (87). Depending on conditions of growth (shaken or static), *Psuedomonas* genotypes retain close resemblance to their ancestor or they diversify into a range of sub-types that can occupy unique environmental niches. Remarkably, when subjected to a fluctuating environment that alternately favors one of two phenotypes, *P. fluorescens* evolves bet-hedging mechanisms that permit stochastic switching between the two phenotypes. Similar to Lenski's finding, Beaumont et al. report a limited number of genotypic differences (nine in this case) between the evolved strain and its ancestor. Surprisingly, the final requisite mutation can be attributed to a single non-synonymous mutation in the large subunit of carbamoylphosphate synthetase (CarB), a central enzyme in the pyrimidine and arginine biosynthetic pathways. Similar to Lenski's finding, the final mutation required an accumulation of previous, potentiating mutations, suggesting that complex epistatic interactions between genotypes drives the evolution of novel phenotypes. The methodology in this experiment is of particular interest for future experimental evolution studies, as the authors achieve this new phenotype rapidly by imposing bottlenecks on the population structure at each "selection" event.

A consistent conclusion from these laboratory studies is that significant improvements in fitness are gained through simple mutations in few to single genes. Functional mutations typically reside within the coding region of genes and presumably alter the kinetics or substrate specificity of proteins, although recent studies suggest that simple mutations in

some non-coding elements, such as riboswitches, can also confer selective advantage (88). Such findings, however, are inconsistent with comparative genomic studies, which suggest that regulatory rewiring of cis-regulatory elements is a primary source of early diversification between species. Theoretical insights similarly suggest that modularity, which is the result of regulatory programs, is critical for the robustness and adaptability of living systems (89). Many of the adaptive features that allow an organism to robustly respond to changes in its environment are encoded at the level of gene regulation -- how genes are turned on and off, where and when they are expressed, and what controls their expression (forming the modularity observed in living systems). This raises interesting questions regarding failure of laboratory studies to enrich mutants with improved fitness that results from alterations in GRNs.

One plausible explanation is that the types of selective pressures used in these studies can be generally categorized as simple nutritional stress. If so, selection with repeated patterns of complex, dynamically changing environmental conditions should enrich for regulatory mutants that can reversibly readjust physiology in novel ways. The study that comes closest to this design is the one conducted by Tagkopolous *et al* (18). In this study the authors enriched a population of *E. coli* with improved fitness to artificially decoupled changes in oxygen and temperature in less than 100 generations. Although the authors did not characterize the mechanistic underpinnings of this improved fitness phenotype, one can speculate that alterations to the GRN structure is the only plausible mechanism to reverse the naturally evolved relationship among processes independently attuned to handle temperature and oxygen-related physiologies. Regardless, we should recollect that selection is a very powerful tool in evolution. While selection imposed by altering a single factor might facilitate analysis and interpretation of adaptive mechanisms discovered in the lab, it does not accurately mirror processes in natural environments, where multiple variables change simultaneously. Environmental heterogeneity and complexity are important and ubiquitous factors in the evolution of natural populations. Interactions among genes, mutations, and environmental factors contribute to adaptability, defining the landscape in which organisms evolve. Heterogeneous environments may create rugged fitness landscapes that contain a multitude of local fitness optima, many of which may be explored by a natural population (49). Simulations suggest that varying environments can actually speed up the rate of evolution (90), especially when new environmental goals are modular – sharing features present at early times. Combinatorial and sequential experimental designs, where cells are exposed to varying sequences or combinations of pressures may reveal natural couplings between environmental events that have been learned by cells (91) and in long-term studies may suggest how organisms anticipate and respond to complex environmental changes.

It is also important to keep in mind that regulatory rewiring events, though fast on an evolutionary timescale, may still require over thousands to millions of years, far beyond the scope of a typical laboratory experiment. Approaches that introduce vast amounts of variation into targeted genetic elements of a natural population prior to selection with an appropriate complex perturbation could circumvent this limitation. Such strategies, while utilized in the past(92), have witnessed renewed interest and enthusiasm because it is now possible to generate fully synthetic genes and genomes (93) and comprehensively identify and track mutants in large populations using NextGen sequencing technologies (94).

Model systems to study adaptation

Microorganisms are ideal for studying the molecular basis for physiological adaptation because of the ease with which they can be genetically manipulated and cultured under controlled environments. In addition, they generally have (i) short generation times, (ii) large effective populations, and (iii) small, relatively simple genomes. These properties

allow for rapid, high-throughput surveys of genetic fitness landscapes over relatively short timescales. Microorganisms commonly employed for molecular evolution studies include *E. coli*, *B. subtilis*, and *S. cerevisiae.* All of these organisms are well characterized, with completely sequenced genomes and a wealth of validated functional annotations. Groundbreaking insights into evolution have been made using each of these organisms. Richard Lenski and colleagues demonstrated adaptation to growth on citrate, where the inability of *E. coli* to proliferate on citrate has traditionally been a characteristic hallmark (50). In *B. subtilis*, several groups have studied adaptation in stress response pathways, including sporulation and competence (Reviewed in 95,96). Finally, the yeast community has produced tremendous work linking gene expression to adaptive evolution (97). Halophillic archaea, like *Halobacterium salinarium sp. NRC-1*, are an especially powerful system in which to study the evolution of cellular stress defense mechanisms as they have evolved in constantly fluctuating and extreme environments, including salinity (10 times that of seawater; 2.5–5.0M), light (>150μmol photons/m^{-2}/s^{-1}), oxygen (<5 μM), temperature (30°C–50°C), nutrients and DNA damaging agents such as UV radiation. Like *E. coli* and *B. subtilis*, *H. salinarum* is relatively simple, readily manipulated, and has a smaller completely sequenced genome (~2,400 genes). All of these features make it an appealing system to disentangle the complex adjustments cells undergo in response to variable environmental conditions. In addition, archaea are evolutionarily unique relative to both bacteria and eukarya. Their basal transcriptional and replication machinery, for instance, shares common ancestry with the eukaryotic machinery, whereas their regulatory systems have bacterial character and they retain the capacity for HGT (98–100). Studies in *H. salinarum* are revealing mechanisms by which a GRN acquires complexity through duplication of transcription factors. By expansion of two eukaryotic general transcription factor families (six TBPs and seven TFBs (TFIIB orthologs)) and subsequent changes to the promoters and coding sequences *H. salinarum* has evolved regulatory programs to modulate the expression of large fractions of genes that allow it to rapidly acclimate to changing environmental conditions (33,101,102).

NEW EXPERIMENTAL APPROACHES TO STUDY ADAPTATION

We are witnessing unprecedented advances in technologies for probing biological phenomena. Powerful high-resolution and high-throughput experimental and computational tools are being used to tackle old biological questions and discover new, often-unanticipated avenues for research. These tools have changed ways in which we think about biological systems, generate hypotheses, and execute experiments. When the first complete genome of a microorganism was sequenced fifteen years ago (103), sequencing a several megabase genome in a day was unimaginable. Now, NextGen sequencing technologies continually push the limits of quality, quantity and cost of sequencing runs. As the cost of these technologies decreases, they are becoming more widespread. Together, these tools give investigators unprecedented access to the signatures of evolution; by comparing larger numbers of related and diverse genomes, and as they occur within laboratory experiments. For instance, we can observe within the time frame of a laboratory experiment evolutionary processes that normally occur over thousands to millions of years in natural environments. We can do this by accelerating evolution using automated multiplexed culturing devices to enrich mutants from microbial populations containing large numbers of random or semi-random variation generated using a combination of gene synthesis and mutagenesis technologies. By using NextGen sequencing, whole genome oligonucleotide arrays, metabolic arrays and mass spectrometry, and high-throughput protein quantitation, it will be possible to link phenotypic outcomes to genetic changes and characterize adaptive mechanisms responsible for improved fitness in new environments (104). Such forward-looking approaches are exemplified by an experiment in which Wang *et al* identified one mutant with fivefold increased lycopene production within 3 days from a population of 4.3

billion *E. coli*. The key technologies used in this remarkable proof-of-concept study included multiplexed automated genome engineering (MAGE) to introduce combinatorial variation into the *E. coli* EcHW2 genome (92) and clever bottlenecking and selection approaches. The combination of these technologies accelerated the emergence of multistep phenotypic adaptations that may have otherwise been extremely improbable due to transient decreases in fitness or genetic drift (87). Especially with respect to GRN evolution, the ability to generate vast amounts of genetic diversity rapidly, followed by deep sequencing of mutants with increased fitness will give us a lens into long-term evolutionary processes that would otherwise be inaccessible. Finally, synthetic approaches that directly manipulate pre-existing cellular architecture can rapidly accelerate evolutionary processes and confirm or repudiate hypothesis of adaptive mechanisms suggested by laboratory evolution or comparative genomics studies. Directly manipulating gene regulatory networks is a powerful approach to disentangle complicated evolutionary scenarios discovered by comparative approaches. Isalan *et al*, for example, studied the effects of regulatory circuit rewiring in E. coli by swapping cis-regulatory and downstream DNA binding domains of TF genes across the TF hierarchy (61). They observed that most rewiring events are neutral under standard growth conditions and can even be advantageous under some stress conditions. Although the phenotypes monitored in this experiment may not fully reveal the consequences of GRN rewiring, the authors suggest that higher-order control of GRNs may minimize the impact of transcriptional rewiring and that such cost-free tinkering with gene regulatory circuits may be a common mechanism in the evolution of GRN. It will be interesting to assess how often this is the case in more complex environments, where the effects of misregulation may manifest themselves more dramatically and to validate the molecular phenotypes of GRN rewiring by gene expression profiling.

INFERENCE AND VISUALIZATION OF BIOLOGICAL NETWORKS

We have described several directions for experimental studies to characterize dynamic evolutionary events that accurately reflect evolution in natural environments or introduce changes into organisms that accelerate the process of evolution itself. To leverage the power of these new biological insights we must integrate data into intuitive-yet-accurate models within which investigators can interactively explore functional relationships among genes and generate new hypotheses from their exploration. As system-wide data becomes more available, it is increasingly important to develop computational methods to handle diverse data types, both individually and collectively. Important advances in the area of network inference have been made over the past several decades. From high-resolution kinetic models to abstract Boolean representations, the field of network inference is rapidly growing in sophistication.

Perhaps the biggest advances have been in the inference of GRNs. In the past, studying evolution of GRNs presented a formidable challenge, even for model organisms with relatively small genomes. It took considerable effort to develop molecular and computational tools to map all regulatory connections to construct reliable network models. Even in this regard "well-defined" GRNs were sparse and mostly restricted to a select few model organisms including *E. coli* (105,106), yeast (107) and *B. subtilis* (108,109). Despite the obvious limitations that come with poor taxonomic coverage of such model organisms, comparative analysis of their GRNs did identify three major mechanisms driving GRN evolution (see Figure 2): (i) alterations in TF-target gene interactions, (ii) gene duplication followed by subfunctionalization or neofunctionalization, and (iii) horizontal gene transfer events between species. Since there were few experimentally derived models of GRNs, investigations relied on *in silico* models whose components include synthetic genes that evolve according to rules encoded in a computer simulation (110–112). Although in silico approaches are extremely informative and will continue to yield valuable insights, they

grossly underestimate complexity of naturally evolved GRNs (113). There are now increasingly sophisticated methods for inferring the structure of GRNs from high-throughput systems biology data (114,115). (Figure 3)

It is important to recognize that key to the success of these inference procedures is the quality, quantity and types of data. Well thought out experiment designs are critical to build models that are predictive and mechanistically accurate. These experiments should appreciate the fundamental properties of biological processes in that they are dynamic, probabilistic and conditional. We will not discuss technical details of network inference algorithms or experiment designs as these topics have been reviewed elsewhere (114,116,117). Through the application of these approaches there is now an ever-increasing number of experimentally verified GRNs (113). Having inferred dynamic models from data, a next step is to make that information useful to investigators. Encoding biological information in a graphical representation has proved valuable for analyzing GRNs and discovering critical topological determinants of information flow. Regulatory influences, protein-protein interaction, and metabolic processes can be represented by a collection of pairwise interactions, where nodes represent regulators, genes, proteins, metabolites or any other biological entity, and edges, which connect nodes, represent arbitrary interactions between the nodes (Figure 1). Mathematically, we refer to this entity as a graph, G = (N,E), where the graph, G, is an ordered pair comprised of nodes, N, and edges, E, which themselves are two element subsets of N. Besides simplifying the representation of biological interactions, graphs have a long, well studied history. Conveniently, many of the mathematical concepts and statistical measures developed for abstract graph structures also apply to biological networks. Analysis of such network depictions of gene regulatory interactions has revealed valuable insight into how cells process information. First, not all regulators are equivalent; few transcription factors (TFs) regulate many genes, whereas the majority of TFs regulate far fewer downstream genes (Reviewed in 118). Second, some patterns of interactions are statistically overrepresented in biological networks -- these overrepresented subgraphs are called motifs (Reviewed in 119). By cataloging frequency of such motifs in *E. coli* Uri Alon and his colleagues discovered overrepresentation of certain kinds of feed-forward and feedback motifs in natural GRNs (120). These regulatory motifs possess interesting dynamic behaviors, which may help cells adjust to changes in their environment. Feed-forward loops, for example, can buffer the effect of noise in highly deterministic biological processes, such as in development (121,122). Last but not least, models of GRNs can be made accessible through visualization software with user-friendly interfaces, such as Cytoscape (123) and BioTapestry (124). By interoperating with other software and databases on the desktop or via the internet using frameworks such as Gaggle (125) and Firegoose (126), scientists can explore the complex networks and disparate data types to formulate hypotheses and drive experimentation. Such systems level investigation has and will continue to revolutionize how we investigate and understand cellular adaptation to a new environment through the evolution of complex GRNs.

Sidebar 1: Evolution of Gene Regulatory Networks (GRNs)

If an organism is challenged by an environmental stress, individuals within the population harboring rewired GRNs that better negotiate the environmental change may enjoy a selective advantage. Here we describe common mechanisms by which GRNs become rewired during evolution, each of which is depicted in Figure 2.

Mutation—Mutational events in transcription factors (TFs) can modify the specificity or affinity of TF DNA binding domains, such as mutations in the base contacting residues of DNA binding proteins that lead to recognition of multiple target DNA sequences (128), or affect protein-protein interaction domains that confer combinatorial specificity to gene

regulatory programs (98). In *Halobacterium*, expansion of the general transcription factors allows for many possible TF-interactions (42), each of which may uniquely control cellular physiology (101). Rewiring may also occur by mutation in downstream target genes. In yeast, the rapid loss of cis-regulatory motifs from multiple genes enables cells to grow rapidly under anaerobic conditions (129).

Gene Duplication—Duplication events are common in microbial genomes. Duplicated genes constitute a genetic toolbox that cells can harness to innovate and expand their phenotypic repertoire (130,131). Nearly half of the regulatory interactions in *E. coli* and yeast appear to evolve through this process (130). Functional divergence of duplicated genes (neofunctionalization and subfunctionalization) can contribute to the development of new cellular functions (132) or specialization in a condition-specific manner (133). Whole genome duplications, though rare, can also contribute to evolution of GRNs(134).

Horizontal gene transfer (HGT)—In prokaryotes, a large proportion of genes have been acquired laterally from different microbial species(66) or even viruses (135). Eukaryotic-derived aminoacyl tRNA synthetases (136), antibiotic-resistance genes (137), and numerous stress-response genes (138) are acquired in diverse lineages through HGT. While entire functional modules may be captured in this way, foreign DNA segments must often be integrated under the control framework of the new host, which can take tens of millions of years (16)

Sidebar 2: Network Inference and Application

High-throughput methodologies generate a wealth of data. Consolidating diverse data types into coherent models of biological interactions and their functional consequences will be critical for understanding systems-level adaptive processes. This presents a challenge for traditional computational tools, which generally represent biological networks as static, uni-dimensional entities. Evolutionary systems biologists, however, are interested in how the structure of multi-dimensional biological networks changes as a result of adaptation and how these alterations in network properties affect dynamic cellular processes and their associated downstream phenotypes. While systems approaches have been successfully applied to a number of biological problems, including abstract boolean network analysis of cell-cycle attractor states (139), flux-balance models of metabolism (Reviewed in 140), and causal biclustering methods for transcriptional regulation (141), to elucidate evolutionary mechanisms future models must (1) integrate across diverse biological processes, (2) represent biological information across multiple time scales with varying degrees of resolution, and (3) develop network representations that track changes in network structures over long (evolutionary) timescales. Such advances will be made by harnessing interdisciplinary approaches, leveraging insights and expertise from a number of disciplines, including mathematics, computation, physics, and engineering. The workflow by which systems-methodologies can be applied to experimental evolution studies is depicted in Figure 3.

Conclusion

We have reviewed known mechanisms by which organisms adapt to new environments and made a case for integrating computational and experimental approaches to study this process and discover new mechanisms. It should be noted that while new technologies have increased the resolution and scale at which we can probe evolutionary processes, they do not override the need for well-designed experiments. Carefully designed selection pressures and culturing technologies will decide the degree to which adaptive mechanisms uncovered in laboratory studies accurately reflect natural processes. The ultimate test of our

understanding of how cells adapt will come from surveying variations within related populations that have evolved naturally to deal with varying environments. This is feasible to do as we now have the capability to determine complete metagenome sequences using NextGen sequencing technologies. It is important that we iteratively refine our models so findings from the laboratory are brought into close apposition with what we are able to observe in a natural environment. Knowing the mechanisms by which organisms evolve will eventually equip us with better engineering principles and strategies to synthetically alter their capabilities. Such rational reengineering has very important implications for how we can address environmental and health related problems in the future (127).

Acknowledgments

This work was supported by grants from NIH (P50GM076547 and 1R01GM077398-01A2), DoE (ENIGMA MAGIC: DE-FG02-07ER64327), and NSF (DBI-0640950) to NSB. We also acknowledge the two reviewers for their constructive critiques that substantially improved some sections of this review.

References

1. Galhardo RS, Hastings PJ, Rosenberg SM. Mutation as a Stress Response and the Regulation of Evolvability. Critical Reviews in Biochemistry and Molecular Biology. 2007; 42(5):399–435. [PubMed: 17917874]
2. Kennedy SP, Ng WV, Salzberg SL, Hood L, DasSarma S. Understanding the adaptation of Halobacterium species NRC-1 to its extreme environment through computational analysis of its genome sequence. Genome Res. 2001; 11(10):1641–1650. [PubMed: 11591641]
3. Kushner, D. Microbial life in extreme environments. London ; New York: Academic Press; 1978.
4. Woelfle MA, Ouyang Y, Phanvijhitsiri K, Johnson CH. The adaptive value of circadian clocks: an experimental assessment in cyanobacteria. Curr Biol. 2004; 14(16):1481–1486. [PubMed: 15324665]
5. Ouyang Y, Andersson CR, Kondo T, Golden SS, Johnson CH. Resonating circadian clocks enhance fitness in cyanobacteria. Proc Natl Acad Sci U S A. 1998; 95(15):8660–8664. [PubMed: 9671734]
6. Mitchell A, Romano GH, Groisman B, Yona A, Dekel E, et al. Adaptive prediction of environmental changes by microorganisms. Nature. 2009; 460(7252):220–224. [PubMed: 19536156]
7. Huisman G, Kolter R. Sensing starvation: a homoserine lactone--dependent signaling pathway in Escherichia coli. Science. 1994; 265(5171):537–539. [PubMed: 7545940]
8. Ram RJ, VerBerkmoes NC, Thelen MP, Tyson GW, Baker BJ, et al. Community Proteomics of a Natural Microbial Biofilm. Science. 2005; 308(5730):1915–1920. [PubMed: 15879173]
9. Hochachka, PW.; Somero, GN. Biochemical adaptation : mechanism and process in physiological evolution. New York: Oxford University Press; 2002.
10. Schrag SJ, Perrot V, Levin BR. Adaptation to the fitness costs of antibiotic resistance in Escherichia coli. Proc Biol Sci. 1997; 264(1386):1287–1291. [PubMed: 9332013]
11. Spain JC, Van Veld PA. Adaptation of natural microbial communities to degradation of xenobiotic compounds: effects of concentration, exposure time, inoculum, and chemical structure. Appl Environ Microbiol. 1983; 45(2):428–435. [PubMed: 16346193]
12. Finkel SE, Kolter R. Evolution of microbial diversity during prolonged starvation. Proc Natl Acad Sci U S A. 1999; 96(7):4023–4027. [PubMed: 10097156]
13. Gribaldo S, Brochier-Armanet C. The origin and evolution of Archaea: a state of the art. Philos Trans R Soc Lond B Biol Sci. 2006; 361(1470):1007–1022. [PubMed: 16754611]
14. Soppa J, Baumann A, Brenneis M, Dambeck M, Hering O, Lange C. Genomics and functional genomics with haloarchaea. Arch Microbiol. 2008; 190(3):197–215. [PubMed: 18493745]
15. Babic A, Lindner A, Vulic M, Stewart E, Radman M. Direct visualization of horizontal gene transfer. Science. 2008; 319(5869):1533–1536. [PubMed: 18339941]
16. Lercher MJ, Pal C. Integration of horizontally transferred genes into regulatory interaction networks takes many million years. Mol Biol Evol. 2008; 25(3):559–567. [PubMed: 18158322]

17. Rutherford S. From genotype to phenotype: buffering mechanisms and the storage of genetic information. Bioessays. 2000; 22(12):1095–1105. [PubMed: 11084625]
18. Tagkopoulos I, Liu YC, Tavazoie S. Predictive behavior within microbial genetic networks. Science. 2008; 320(5881):1313–1317. [PubMed: 18467556]
19. Chaffron S, Rehrauer H, Pernthaler J, von Mering C. A global network of coexisting microbes from environmental and whole-genome sequence data. Genome Res. 2010; 20(7):947–959. [PubMed: 20458099]
20. Barton, NH. Evolution. Cold Spring Harbor, N.Y: Cold Spring Harbor Laboratory Press; 2007.
21. Burke MK, Dunham JP, Shahrestani P, Thornton KR, Rose MR, Long AD. Genome-wide analysis of a long-term evolution experiment with Drosophila. Nature. 2010; 467(7315):587–590. [PubMed: 20844486]
22. Kassen R. Toward a General Theory of Adaptive Radiation. Annals of the New York Academy of Sciences. 2009; 1168(1):3–22. [PubMed: 19566701]
23. Orr HA. Theories of adaptation: what they do and don't say. Genetica. 2005; 123(1):3–13. [PubMed: 15881676]
24. Papadopoulos D, Schneider D, Meier-Eiss J, Arber W, Lenski R, Blot M. Genomic evolution during a 10,000-generation experiment with bacteria. Proceedings of the National Academy of Sciences of the United States of America. 1999; 96(7):3807. [PubMed: 10097119]
25. de Visser JAGM, Rozen DE. Clonal Interference and the Periodic Selection of New Beneficial Mutations in Escherichia coli. Genetics. 2006; 172(4):2093–2100. [PubMed: 16489229]
26. Lauro FM, Chastain RA, Blankenship LE, Yayanos AA, Bartlett DH. The unique 16S rRNA genes of piezophiles reflect both phylogeny and adaptation. Appl Environ Microbiol. 2007; 73(3):838–845. [PubMed: 17158629]
27. Hentschel U, Hopke J, Horn M, Friedrich AB, Wagner M, et al. Molecular evidence for a uniform microbial community in sponges from different oceans. Appl Environ Microbiol. 2002; 68(9): 4431–4440. [PubMed: 12200297]
28. Gasch AP, Spellman PT, Kao CM, Carmel-Harel O, Eisen MB, et al. Genomic expression programs in the response of yeast cells to environmental changes. Mol Biol Cell. 2000; 11(12): 4241–4257. [PubMed: 11102521]
29. Causton HC, Ren B, Koh SS, Harbison CT, Kanin E, et al. Remodeling of yeast genome expression in response to environmental changes. Mol Biol Cell. 2001; 12(2):323–337. [PubMed: 11179418]
30. He G, He B, Racey P, Cui J. Positive Selection of the Bat Interferon Alpha Gene Family. Biochemical Genetics. 2010; 48(9):840–846. [PubMed: 20628805]
31. Enjalbert B, Nantel A, Whiteway M. Stress-induced gene expression in Candida albicans: absence of a general stress response. Mol Biol Cell. 2003; 14(4):1460–1467. [PubMed: 12686601]
32. Kaur A, Van PT, Busch CR, Robinson CK, Pan M, et al. Coordination of frontline defense mechanisms under severe oxidative stress. Mol Syst Biol. 2010; 6:393. [PubMed: 20664639]
33. Whitehead K, Pan M, Masumura K-I, Bonneau R, Baliga NS, Rodriguez-Valera F. Diurnally Entrained Anticipatory Behavior in Archaea. PLoS ONE. 2009; 4(5):e5485. [PubMed: 19424498]
34. Wolf DM, Fontaine-Bodin L, Bischofs I, Price G, Keasling J, Arkin AP. Memory in microbes: quantifying history-dependent behavior in a bacterium. PLoS ONE. 2008; 3(2):e1700. [PubMed: 18324309]
35. Johnson CH, Mori T, Xu Y. A cyanobacterial circadian clockwork. Curr Biol. 2008; 18(17):R816–R825. [PubMed: 18786387]
36. Whitaker WB, Parent MA, Naughton LM, Richards GP, Blumerman SL, Boyd EF. Modulation of Responses of Vibrio parahaemolyticus O3:K6 to pH and Temperature Stresses by Growth at Different Salt Concentrations. Appl Environ Microbiol. 2010; 76(14):4720–4729. [PubMed: 20472729]
37. Calhoun L, Kwon Y. The effect of long-term propionate adaptation on the stress resistance of Salmonella Enteritidis. Journal of Applied Microbiology. 2010; 109(4):1294–1300. [PubMed: 20477896]
38. Dassarma S, Berquist BR, Coker JA, Dassarma P, Müller JA. Post-genomics of the model haloarchaeon Halobacterium sp. NRC-1. Saline Syst. 2006; 2:3. [PubMed: 16542428]

39. Oren A. Microbial life at high salt concentrations: phylogenetic and metabolic diversity. Saline Systems. 2008; 4:2. [PubMed: 18412960]
40. Sanchez-Perez G, Mira A, Nyiro G, Pasic L, Rodriguez-Valera F. Adapting to environmental changes using specialized paralogs. Trends Genet. 2008; 24(4):154–158. [PubMed: 18325625]
41. Feller G, Gerday C. Psychrophilic enzymes: hot topics in cold adaptation. Nat Rev Microbiol. 2003; 1(3):200–208. [PubMed: 15035024]
42. Feller G. Life at low temperatures: is disorder the driving force? Extremophiles. 2007; 11(2):211–216. [PubMed: 17160345]
43. Vieille C, Epting KL, Kelly RM, Zeikus JG. Bivalent cations and amino-acid composition contribute to the thermostability of Bacillus licheniformis xylose isomerase. Eur J Biochem. 2001; 268(23):6291–6301. [PubMed: 11733026]
44. Jaenicke R. Protein stability and molecular adaptation to extreme conditons. European Journal of Biochemistry. 1991; 202(3):715–728. [PubMed: 1765088]
45. Barák I, Wilkinson AJ. Where asymmetry in gene expression originates. Molecular Microbiology. 2005; 57(3):611–620. [PubMed: 16045607]
46. Veening JW, Smits WK, Kuipers OP. Bistability, epigenetics, and bet-hedging in bacteria. Annu Rev Microbiol. 2008; 62:193–210. [PubMed: 18537474]
47. Bennett A, Dao K, Lenski R. Rapid evolution in response to high-temperature selection. Nature. 1990; 346(6279):79–81. [PubMed: 2195353]
48. Bennett AF, Lenski RE. An experimental test of evolutionary trade-offs during temperature adaptation. Proc Natl Acad Sci U S A. 2007; 104 Suppl 1:8649–8654. [PubMed: 17494741]
49. Cooper T, Lenski R. Experimental evolution with E. coli in diverse resource environments. I. Fluctuating environments promote divergence of replicate populations. BMC Evolutionary Biology. 2010; 10(1):11. [PubMed: 20070898]
50. Blount ZD, Borland CZ, Lenski RE. Historical contingency and the evolution of a key innovation in an experimental population of Escherichia coli. Proc Natl Acad Sci USA. 2008; 105(23):7899–7906. [PubMed: 18524956]
51. Zinser ER, Kolter R. Escherichia coli evolution during stationary phase. Res Microbiol. 2004; 155(5):328–336. [PubMed: 15207864]
52. Lenski RE, Travisano M. Dynamics of adaptation and diversification: a 10,000-generation experiment with bacterial populations. Proc Natl Acad Sci U S A. 1994; 91(15):6808–6814. [PubMed: 8041701]
53. Barrick JE, Yu DS, Yoon SH, Jeong H, Oh TK, et al. Genome evolution and adaptation in a long-term experiment with Escherichia coli. Nature. 2009; 461(7268):1243–1247. [PubMed: 19838166]
54. Rosenberg SM. Evolving responsively: adaptive mutation. Nat Rev Genet. 2001; 2(7):504–515. [PubMed: 11433357]
55. McAdams HH, Srinivasan B, Arkin AP. The evolution of genetic regulatory systems in bacteria. Nat Rev Genet. 2004; 5(3):169–178. [PubMed: 14970819]
56. Babu MM, Luscombe NM, Aravind L, Gerstein M, Teichmann SA. Structure and evolution of transcriptional regulatory networks. Curr Opin Struct Biol. 2004; 14(3):283–291. [PubMed: 15193307]
57. Lozada-Chavez I, Angarica VE, Collado-Vides J, Contreras-Moreira B. The role of DNA-binding specificity in the evolution of bacterial regulatory networks. J Mol Biol. 2008; 379(3):627–643. [PubMed: 18466918]
58. López-Maury L, Marguerat S, Bähler J. Tuning gene expression to changing environments: from rapid responses to evolutionary adaptation. Nat Rev Genet. 2008; 9(8):583–593. [PubMed: 18591982]
59. Wagner, A. Robustness and evolvability in living systems. Princeton, N.J: Princeton University Press; 2005.
60. Gould, SJ. Wonderful life : the Burgess Shale and the nature of history. 1st ed.. New York: W.W. Norton; 1989.
61. Isalan M, Lemerle C, Michalodimitrakis K, Horn C, Beltrao P, et al. Evolvability and hierarchy in rewired bacterial gene networks. Nature. 2008; 452(7189):840–845. [PubMed: 18421347]

62. Acar M, Pando BF, Arnold FH, Elowitz MB, van Oudenaarden A. A General Mechanism for Network-Dosage Compensation in Gene Circuits. Science. 2010; 329(5999):1656–1660. [PubMed: 20929850]
63. Wagner GP, Altenberg L. Perspective: Complex Adaptations and the Evolution of Evolvability. Evolution. 1996; 50(3):967–976.
64. Earl DJ, Deem MW, Chandler D. Evolvability Is a Selectable Trait. Proceedings of the National Academy of Sciences of the United States of America. 2004; 101(32):11531–11536. [PubMed: 15289608]
65. Lysenko ES, Lijek RS, Brown SP, Weiser JN. Within-host competition drives selection for the capsule virulence determinant of Streptococcus pneumoniae. Curr Biol. 2010; 20(13):1222–1226. [PubMed: 20619820]
66. Koonin EV, Makarova KS, Aravind L. Horizontal gene transfer in prokaryotes: quantification and classification. Annu Rev Microbiol. 2001; 55:709–742. [PubMed: 11544372]
67. Schultz D, Wolynes PG, Ben Jacob E, Onuchic JN. Deciding fate in adverse times: sporulation and competence in Bacillus subtilis. Proc Natl Acad Sci U S A. 2009; 106(50):21027–21034. [PubMed: 19995980]
68. Maamar H, Raj A, Dubnau D. Noise in Gene Expression Determines Cell Fate in Bacillus subtilis. Science. 2007; 317(5837):526–529. [PubMed: 17569828]
69. Smith JM, Smith NH, O'Rourke M, Spratt BG. How clonal are bacteria? Proceedings of the National Academy of Sciences of the United States of America. 1993; 90(10):4384–4388. [PubMed: 8506277]
70. Achtman M, Wagner M. Microbial diversity and the genetic nature of microbial species. Nature Reviews Microbiology. 2008; 6(6):431–440.
71. Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF, et al. Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. Science. 1995; 269(5223):496–512. [PubMed: 7542800]
72. Wu Y, Zhang X, Yu J, Ouyang Q. Identification of a topological characteristic responsible for the biological robustness of regulatory networks. PLOS Computational Biology. 2009; 5(7):e1000442. [PubMed: 19629157]
73. Handelsman J. Metagenomics: application of genomics to uncultured microorganisms. Microbiol Mol Biol Rev. 2004; 68(4):669–685. [PubMed: 15590779]
74. Streit WR, Schmitz RA. Metagenomics--the key to the uncultured microbes. Curr Opin Microbiol. 2004; 7(5):492–498. [PubMed: 15451504]
75. Edwards JL, Brown EJ, Uk-Nham S, Cannon JG, Blake MS, Apicella MA. A co-operative interaction between Neisseria gonorrhoeae and complement receptor 3 mediates infection of primary cervical epithelial cells. Cell Microbiol. 2002; 4(9):571–584. [PubMed: 12390350]
76. Koonin EV, Wolf YI, Karev GP. The structure of the protein universe and genome evolution. Nature. 2002; 420(6912):218–223. [PubMed: 12432406]
77. Makarova KS, Grishin NV, Koonin EV. The HicAB cassette, a putative novel, RNA-targeting toxin-antitoxin system in archaea and bacteria. Bioinformatics. 2006; 22(21):2581–2584. [PubMed: 16895922]
78. Tatusov RL, Galperin MY, Natale DA, Koonin EV. The COG database: a tool for genome-scale analysis of protein functions and evolution. Nucleic Acids Res. 2000; 28(1):33–36. [PubMed: 10592175]
79. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 2000; 28(1):27–30. [PubMed: 10592173]
80. Kellis M, Patterson N, Endrizzi M, Birren B, Lander ES. Sequencing and comparison of yeast species to identify genes and regulatory elements. Nature. 2003; 423(6937):241–254. [PubMed: 12748633]
81. Wapinski I, Pfiffner J, French C, Socha A, Thompson DA, Regev A. Gene duplication and the evolution of ribosomal protein gene regulation in yeast. Proceedings of the National Academy of Sciences. 2010; 107(12):5505–5510.

82. Gould SJ, Lewontin RC. The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. Proc R Soc Lond B Biol Sci. 1979; 205(1161):581–598. [PubMed: 42062]
83. Tuch BB, Galgoczy DJ, Hernday AD, Li H, Johnson AD. The evolution of combinatorial gene regulation in fungi. PLoS Biol. 2008; 6(2):352–364.
84. Travisano M, Lenski RE. Long-term experimental evolution in Escherichia coli. IV. Targets of selection and the specificity of adaptation. Genetics. 1996; 143(1):15–26. [PubMed: 8722758]
85. Blount ZD, Borland CZ, Lenski RE. Historical contingency and the evolution of a key innovation in an experimental population of Escherichia coli. Proc Natl Acad Sci U S A. 2008; 105(23):7899–7906. [PubMed: 18524956]
86. Lenski RE. Quantifying fitness and gene stability in microorganisms. Biotechnology. 1991; 15:173–192. [PubMed: 2009380]
87. Beaumont HJ, Gallie J, Kost C, Ferguson GC, Rainey PB. Experimental evolution of bet hedging. Nature. 2009; 462(7269):90–93. [PubMed: 19890329]
88. Tremblay PL, Summers ZM, Glaven RH, Nevin KP, Zengler K, et al. A c-type cytochrome and a transcriptional regulator responsible for enhanced extracellular electron transfer in Geobacter sulfurreducens revealed by adaptive evolution. Environmental Microbiology. 2010:1–11.
89. Kitano H. Biological robustness. Nat Rev Genet. 2004; 5(11):826–837. [PubMed: 15520792]
90. Kashtan N, Noor E, Alon U. Varying environments can speed up evolution. Proc Natl Acad Sci USA. 2007; 104(34):13711–13716. [PubMed: 17698964]
91. Baliga NS. Systems biology. The scale of prediction. Science. 2008; 320(5881):1297–1298. [PubMed: 18535232]
92. Wang HH, Isaacs FJ, Carr PA, Sun ZZ, Xu G, et al. Programming cells by multiplex genome engineering and accelerated evolution. Nature. 2009; 460(7257):894–898. [PubMed: 19633652]
93. Gibson DG, Glass JI, Lartigue C, Noskov VN, Chuang R-Y, et al. Creation of a Bacterial Cell Controlled by a Chemically Synthesized Genome. Science. 2010; 329(5987):52–56. [PubMed: 20488990]
94. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, et al. Genome sequencing in microfabricated high-density picolitre reactors. Nature. 2005; 437(7057):376–380. [PubMed: 16056220]
95. Hamoen L, Venema G, Kuipers O. Controlling competence in Bacillus subtilis: shared use of regulators. Microbiology. 2003; 149(1):9–17. [PubMed: 12576575]
96. Errington J. Regulation of endospore formation in Bacillus subtilis. Nature Reviews Microbiology. 2003; 1(2):117–126.
97. Ferea TL, Botstein D, Brown PO, Rosenzweig RF. Systematic changes in gene expression patterns following adaptive evolution in yeast. Proc Natl Acad Sci U S A. 1999; 96(17):9721–9726. [PubMed: 10449761]
98. Baliga NS, Goo YA, Ng WV, Hood L, Daniels CJ, DasSarma S. Is gene expression in Halobacterium NRC-1 regulated by multiple TBP and TFB transcription factors? Molecular Microbiology. 2000; 36(5):1184–1185. [PubMed: 10844702]
99. Geiduschek EP, Ouhammouch M. MicroReview: Archaeal transcription and its regulators. Molecular Microbiology. 2005; 56(6):1397–1407. [PubMed: 15916593]
100. Edgell DR, Doolittle WF. Archaea and the Origin(s) of DNA Replication Proteins. Cell. 1997; 89(7):995–998. [PubMed: 9215620]
101. Facciotti MT, Reiss DJ, Pan M, Kaur A, Vuthoori M, et al. General transcription factor specified global gene regulation in archaea. Proc Natl Acad Sci U S A. 2007; 104(11):4630–4635. [PubMed: 17360575]
102. Schmid AK, Reiss DJ, Pan M, Koide T, Baliga NS. A single transcription factor regulates evolutionarily diverse but functionally linked metabolic pathways in response to nutrient availability. Mol Syst Biol. 2009; 5(282):1–15.
103. Fleischmann R, Adams M, White O, Clayton R, Kirkness E, et al. Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. Science. 1995; 269(5223):496. [PubMed: 7542800]

104. van Opijnen T, Bodi KL, Camilli A. Tn-seq: high-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms. Nat Methods. 2009; 6(10):767–772. [PubMed: 19767758]
105. Salgado H, Santos-Zavaleta A, Gama-Castro S, Peralta-Gil M, Penaloza-Spinola MI, et al. The comprehensive updated regulatory network of Escherichia coli K-12. BMC Bioinformatics. 2006; 7(5)
106. Madan Babu M, Teichmann S. Evolution of transcription factors and the gene regulatory network in Escherichia coli. Nucleic Acids Research. 2003; 31(4):1234–1244. [PubMed: 12582243]
107. Lee TI, Rinaldi NJ, Robert F, Odom DT, Bar-Joseph Z, et al. Transcriptional regulatory networks in Saccharomyces cerevisiae. Science. 2002; 298(5594):799–804. [PubMed: 12399584]
108. Moreno-Campuzano S, Janga SC, Perez-Rueda E. Identification and analysis of DNA-binding transcription factors in Bacillus subtilis and other Firmicutes--a genomic approach. BMC Genomics. 2006; 7:147. [PubMed: 16772031]
109. Makita Y, Nakao M, Ogasawara N, Nakai K. DBTBS: database of transcriptional regulation in Bacillus subtilis and its contribution to comparative genomics. Nucleic Acids Research. 2004; 32(Database Issue):D75. [PubMed: 14681362]
110. Wagner A. Genetic redundancy caused by gene duplications and its evolution in networks of transcriptional regulators. Biol Cybern. 1996; 74(6):557–567. [PubMed: 8672563]
111. Milo R, Itzkovitz S, Kashtan N, Levitt R, Shen-Orr S, et al. Superfamilies of evolved and designed networks. Science. 2004; 303(5663):1538–1542. [PubMed: 15001784]
112. Siegal ML, Promislow DE, Bergman A. Functional and evolutionary inference in gene networks: does topology matter? Genetica. 2007; 129(1):83–103. [PubMed: 16897451]
113. De Smet R, Marchal K. Advantages and limitations of current network inference methods. Nat Rev Micro. 2010; 8(10):717–729.
114. Bonneau R. Learning biological networks: from modules to dynamics. Nat Chem Biol. 2008; 4(11):658–664. [PubMed: 18936750]
115. Gardner TS, di Bernardo D, Lorenz D, Collins JJ. Inferring genetic networks and identifying compound mode of action via expression profiling. Science. 2003; 301(5629):102–105. [PubMed: 12843395]
116. Facciotti M, Bonneau R, Hood L, Baliga N. Systems biology experimental design-considerations for building predictive gene regulatory network models for prokaryotic systems. Current Genomics. 2004; 5(7):527–544.
117. Hayete B, Gardner TS, Collins JJ. Size matters: network inference tackles the genome scale. Mol Syst Biol. 2007; 3:77. [PubMed: 17299414]
118. Barabasi AL, Oltvai ZN. Network biology: understanding the cell's functional organization. Nat Rev Genet. 2004; 5(2):101–113. [PubMed: 14735121]
119. Alon U. Network motifs: theory and experimental approaches. Nat Rev Genet. 2007; 8(6):450–461. [PubMed: 17510665]
120. Milo R, Shen-Orr S, Itzkovitz S, Kashtan N, Chklovskii D, Alon U. Network motifs: simple building blocks of complex networks. Science. 2002; 298(5594):824–827. [PubMed: 12399590]
121. Mangan S, Alon U. Structure and function of the feed-forward loop network motif. Proc Natl Acad Sci USA. 2003; 100(21):11980–11985. [PubMed: 14530388]
122. Levine M, Davidson EH. Gene regulatory networks for development. Proc Natl Acad Sci U S A. 2005; 102(14):4936–4942. [PubMed: 15788537]
123. Shannon P, Markiel A, Ozier O, Baliga N, Wang J, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Research. 2003; 13(11):2498. [PubMed: 14597658]
124. Longabaugh W, Davidson E, Bolouri H. Computational representation of developmental genetic regulatory networks. Developmental biology. 2005; 283(1):1–16. [PubMed: 15907831]
125. Shannon PT, Reiss DJ, Bonneau R, Baliga NS. The Gaggle: an open-source software system for integrating bioinformatics software and data sources. BMC Bioinformatics. 2006; 7:176. [PubMed: 16569235]

126. Bare JC, Shannon PT, Schmid AK, Baliga NS. The Firegoose: two-way integration of diverse data from different bioinformatics web resources with desktop applications. BMC Bioinformatics. 2007; 8:456. [PubMed: 18021453]

127. Koide T, Lee Pang W, Baliga NS. The role of predictive modelling in rationally re-engineering biological systems. Nat Rev Micro. 2009; 7(4):297–305.

128. Luscombe NM, Thornton JM. Protein-DNA interactions: amino acid conservation and the effects of mutations on binding specificity. J Mol Biol. 2002; 320(5):991–1009. [PubMed: 12126620]

129. Ihmels J, Bergmann S, Gerami-Nejad M, Yanai I, McClellan M, et al. Rewiring of the yeast transcriptional network through the evolution of motif usage. Science. 2005; 309(5736):938–940. [PubMed: 16081737]

130. Teichmann SA, Babu MM. Gene regulatory network growth by duplication. Nat Genet. 2004; 36(5):492–496. [PubMed: 15107850]

131. Gevers D, Vandepoele K, Simillon C, Van de Peer Y. Gene duplication and biased functional retention of paralogs in bacterial genomes. Trends Microbiol. 2004; 12(4):148–154. [PubMed: 15116722]

132. Li, W.; Olmstead, R. Molecular evolution. Sunderland, MA: Sinauer Associates; 1997.

133. Wapinski I, Pfiffner J, French C, Socha A, Thompson DA, Regev A. Gene duplication and the evolution of ribosomal protein gene regulation in yeast. Proc Natl Acad Sci U S A. 107(12): 5505–5510. [PubMed: 20212107]

134. Kellis M, Patterson N, Birren B, Berger B, Lander ES. Methods in comparative genomics: genome correspondence, gene identification and regulatory motif discovery. J Comput Biol. 2004; 11(2–3):319–355. [PubMed: 15285895]

135. Rohwer F, Thurber RV. Viruses manipulate the marine environment. Nature. 2009; 459(7244): 207–212. [PubMed: 19444207]

136. Woese CR, Olsen GJ, Ibba M, Soll D. Aminoacyl-tRNA synthetases, the genetic code, and the evolutionary process. Microbiol Mol Biol Rev. 2000; 64(1):202–236. [PubMed: 10704480]

137. Davies J. Origins and evolution of antibiotic resistance. Microbiologia. 1996; 12(1):9–16. [PubMed: 9019139]

138. Nelson K, Clayton R, Gill S, Gwinn M, Dodson R, et al. Evidence for lateral gene transfer between Archaea and bacteria from genome sequence of Thermotoga maritima. Nature. 1999; 399(6734):323–329. [PubMed: 10360571]

139. Li F, Long T, Lu Y, Ouyang Q, Tang C. The yeast cell-cycle network is robustly designed. Proceedings of the National Academy of Sciences of the United States of America. 2004; 101(14):4781. [PubMed: 15037758]

140. Kauffman K, Prakash P, Edwards J. Advances in flux balance analysis. Current opinion in biotechnology. 2003; 14(5):491–496. [PubMed: 14580578]

141. Bonneau R, Facciotti MT, Reiss DJ, Schmid AK, Pan M, et al. A predictive model for transcriptional control of physiology in a free living cell. Cell. 2007; 131(7):1354–1365. [PubMed: 18160043]

142. Barabási A-L, Oltvai ZN. Network biology: understanding the cell's functional organization. Nat Rev Genet. 2004; 5(2):101–113. [PubMed: 14735121]

143. Yu H, Gerstein M. Genomic analysis of the hierarchical structure of regulatory networks. Proc Natl Acad Sci USA. 2006; 103(40):14724–14731. [PubMed: 17003135]

144. Hartwell LH, Hopfield JJ, Leibler S, Murray AW. From molecular to modular cell biology. Nature. 1999; 402 Suppl(6761):C47–C52. [PubMed: 10591225]

145. Alon U. Network motifs: theory and experimental approaches. Nat Rev Genet. 2007; 8(6):450–461. [PubMed: 17510665]

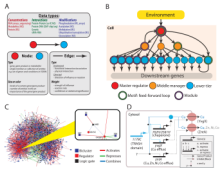

Figure 1. Generation and properties of networks

Networks are a useful formalism to represent, catalogue, and analyze biological information. (A) Networks are generic entities that are used to represent many types of biological interactions. The fundamental units of a network (or graph) are nodes, and edges. Three types of biological information are commonly represented as a network: transcriptional, metabolic, and protein-protein interactions. The types of data used to build the network and the features that define nodes and edges vary by application. (B) Biological networks share common features. Here, we represent a transcriptional network. Three features commonly define network topology and have important dynamic consequences for the behavior of the network: (1) hierarchy: transcriptional networks are close to scale-free in the distribution of regulatory connections and exhibit hierarchical arrangement of connections (142). At the beginning of many biological pathways are few "master regulators" that initiate response to environmental or internal cues. These master regulators propagate information to "middle managers" that have many additional regulatory connections, mostly to "lower-tier" regulators that mediate specific biological functions (143) (2) modularity: biological networks aggregate pathways and functions into modules, which are defined by groups of genes, regulators, and gene products that are somehow interconnected and interdependent (144). Here, we denote a transcriptional "module" by a purple ring surrounding the nodes in the module, though we note that biological modules are often composed of a diversity of part types that need not be coregulated. (3) motifs: interesting dynamic behaviors of gene circuits are often mediated by particular wiring of the parts, which defines a motif. The members of a particularly well-studied motif, the feed-forward loop, are depicted in this illustration by green circles surrounding the nodes (145). (C) Biological networks learned from experimental data, such as the Environmental and Gene Regulatory Influence Network (EGRIN) for *H. salinarum* sp. NRC-1, contain many layers of information that can be mined to aid hypothesis generation (141).

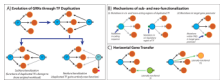

Figure 2. Molecular changes that shape GRN evolution

Several molecular mechanisms mediate topological changes within GRNs. (A) Duplication of transcription factor (TF) and/or target genes (TG). Duplicated copies of a TF or downstream gene initially share the same interactions as its ancestor. Following duplication, however, either copy can subfunctionalize to contain a subset of those ancestral connections or neofunctionalize by gaining new interactions. Sub- and neofunctionalization generally occur via random mutations. (B) Mutations can occur in the coding or cis-regulatory sequences of either TFs or TGs. Mutations in the cis-regulatory sequences of a TG only affect interaction with that particular target, while mutations in coding and cis regions of TF may affect all downstream interactions. (C) Microbial genomes can be extensively modified by horizontal gene transfers. Genomes can horizontally inherit new TF (green circle), TG (yellow circle) or both TF and its target simultaneously (not shown). Transcription factors and target genes are depicted in blue and orange circles respectively. Cis-regulatory regions are denoted by gray boxes attached to the circles.

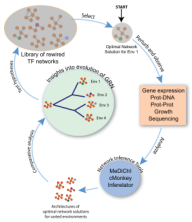

Figure 3. Systems approaches to study GRN evolution

The first step in understanding GRN evolution starts with a naturally evolved GRN that has increased fitness in a given environment. Initially, the architecture of the network is unknown (indicated by gray shading). To delineate the connectivity of the network, we perturb it via genetic or environmental alterations and observe the phenotypic response across multiple data types (orange box). The data is analyzed and integrated using network inference tools to build a comprehensive view of the changes to identify architectures of optimal network solutions (Blue circle). Many iterations of this cycle for different environments yield to library of optimal network solutions which can be compared to derive evolutionary insights (Green circle) and further formulate hypotheses. These hypotheses are tested by construction of rewired GRN by introducing variations at specific components (Gray eclipse). Rewired GRN are selected at different environments and characterized in the next iteration.