ACS SyntheticBiology

Review

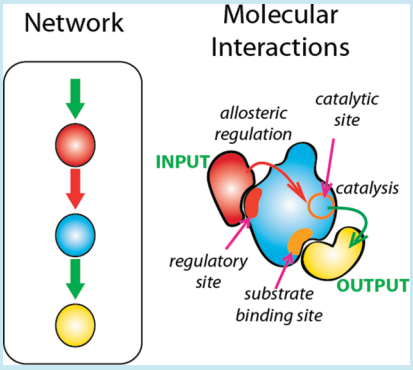
pubs.acs.org/synthbio

Evolutionary Synthetic Biology

Sergio G. Peisajovich*

Department of Cell and Systems Biology, University of Toronto, Toronto, M5S 3G5 Canada

ABSTRACT: Signaling networks process vast amounts of environmental information to generate specific cellular responses. As cellular environments change, signaling networks adapt accordingly. Here, I will discuss how the integration of synthetic biology and directed evolution approaches is shedding light on the molecular mechanisms that guide the evolution of signaling networks. In particular, I will review studies that demonstrate how different types of mutations, from the replacement of individual amino acids to the shuffling of modular domains, lead to markedly different evolutionary trajectories and consequently to diverse network rewiring. Moreover, I will argue that intrinsic evolutionary properties of signaling proteins, such as the robustness of wild type functions, the promiscuous nature of evolutionary intermediates, and the modular decoupling between binding and catalysis, play important roles in the evolution of signaling networks. Finally, I will argue that rapid advances in our ability to synthesize DNA will radically alter how we study signaling network evolution at the genome-wide level.



KEYWORDS: *signaling network, evolution, modular domain, interaction motif, scaffold, synthetic biology, directed evolution*

Through the study of simple synthetic biological systems, synthetic biology is uncovering the design principles of more complex, natural systems. Synthetic biology is also advancing our understanding of how biological systems evolve, analogously to the advancements that directed evolution methods brought about to the study of protein evolution. Directed evolution was originally developed to alter protein function, once it became clear that rational protein engineering methods were limited by our incomplete knowledge of how protein sequences encode function.[1] Protein engineers harnessed the power of evolution, through the creation of large libraries of genetic variants and selections, to alter protein function with relative ease. In addition, protein engineers rapidly realized that directed evolution could be used not only to engineer protein function but also to investigate how proteins evolve in nature.[2] By mimicking evolution in the laboratory, it was possible to uncover evolutionary intermediates,[3,4] compare alternative evolutionary pathways,[4,5] or understand what roles different types of mutations play in evolution.[6]

In this Review, I will discuss how, by incorporating concepts and methodologies originated in directed evolution, synthetic biology is changing our understanding of the evolution of complex biological systems. While the integration of synthetic biology with directed evolution is unraveling aspects of natural evolution that range from the origin of life to alternative biochemistries, I will focus here on studies that shed light on the evolution of regulatory signaling networks.

FACTORS CONTROLLING THE FUNCTION OF REGULATORY SIGNALING NETWORKS

Regulatory networks control the dynamics of most cellular processes. Their function depends on three main factors:[7] (i) network architecture, that is, the particular set of regulatory interactions among all network components; (ii) quantitative parameters that reflect, among others, the concentrations of network components, the affinities of the interactions, and the efficiencies of the catalytic steps; and (iii) noise that results from stochastic fluctuations in network components present at low concentrations. Evolution can alter cellular regulatory networks through mutations affecting any of these factors. Directed evolution studies have addressed extensively how quantitative parameters, such as enzymes' catalytic efficiency and substrate specificity, change as a result of mutations (reviewed in refs 1 and 2). Therefore, here I will only mention aspects that are important to understand the role that changes in enzymatic activity have in the evolution of regulatory networks. Instead, I will focus on evolutionary changes that alter network architecture, as these are well suited to be explored by synthetic biology. Rather than attempting an exhaustive enumeration, I will discuss representative examples and apologize in advance to those whose work, though important, could not be included in this review.

EVOLUTION OF CELLULAR REGULATORY NETWORKS ARCHITECTURE: LESSONS FROM SYNTHETIC BIOLOGY

The architecture of a cellular regulatory network is the set of dynamic interactions between all network components. Proper network functioning depends on the assembly and disassembly of diverse complexes of network components, regulated both in time and space, and in the resulting regulatory changes in the activities of the interacting partners. Evolution could alter the

Received: February 14, 2012
Published: March 29, 2012

   dx.doi.org/10.1021/sb300012g | *ACS Synth. Biol.* 2012, 1, 199–210

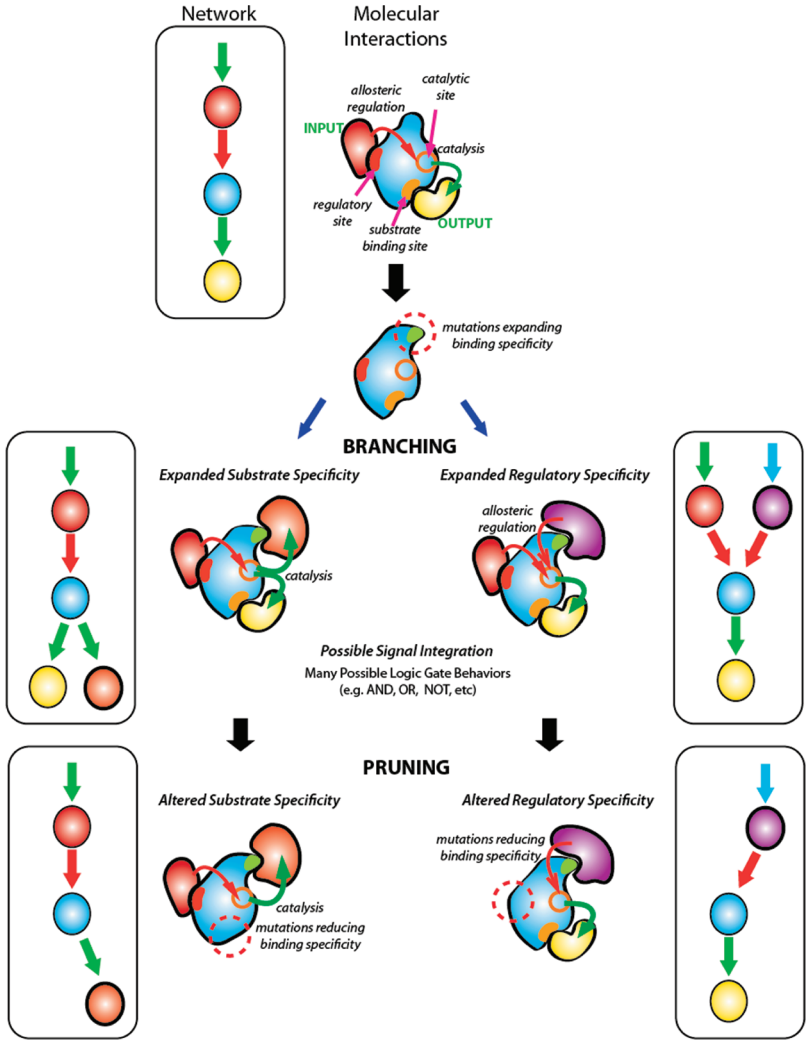


Figure 1. Network rewiring by replacement of specificity-determining residues may proceed through intermediates with expanded specificities. The gain of an interaction partner through amino acid replacements most often does not compromise pre-existing interactions. This effectively expands the range of interactions, adding branches to the pathway. Occasionally, pathway function might depend on precise combinations of binding partners, thus creating novel signal processing capabilities. Eventually, additional mutations might eliminate some interactions, removing pathway branches and leading to a complete pathway rewiring.

function of a regulatory network by changing the identity or the properties of interacting partners, but also the timing and subcellular locations at which interactions occur. Synthetic biology has provided us with examples that demonstrate how changes in any of these properties could alter network function.

1. Evolution of Network Architecture through Rewiring of Protein–Protein Interactions. Protein–protein interactions mediate important signal-processing events in cellular regulatory networks. These interactions bring proteins together in space and time and, more importantly, transfer signaling information from one protein to another by regulating protein function in a stimulus-dependent manner. Some signaling proteins encode regulatory and catalytic activities within a single, highly integrated, structural unit.[8] Others encode regulatory and catalytic activities in separate structural modules, most often different domains or motifs. The input/output relationships encoded in each defined pair of regulatory and catalytic signaling elements determine the architecture of a signaling network. Therefore, mutations that alter relationships between regulatory and catalytic elements are key to understand how signaling networks are rewired by evolution.

Network Architecture Rewiring by Replacement of Specificity-Determining Amino Acids Often Proceeds through Intermediates with Expanded Specificities. In signaling proteins in which regulation and catalysis are encoded within a single structural unit, changes in input/output relationships can evolve by amino acid replacements (or small insertion or deletions) that alter binding sites for regulators or substrates.

The Mitogen-Activated Protein Kinases (MAPK) are a well-studied example of how evolution could rewire signaling networks, by replacing residues governing interactions within a single domain that encodes both regulatory and catalytic activities. Gene duplication and divergence have expanded the MAPK family, from three members in lower eukaryotes to at least 15 members in higher eukaryotes. In the yeast *Saccharomyces cerevisiae*, Fus3 and Hog1 are the MAPK in the mating and high osmolarity signaling pathways, respectively. Ramanathan and co-workers[9] explored how evolution rewires network architecture by altering Fus3 and Hog1 interaction

dx.doi.org/10.1021/sb300012g | *ACS Synth. Biol.* 2012, 1, 199–210

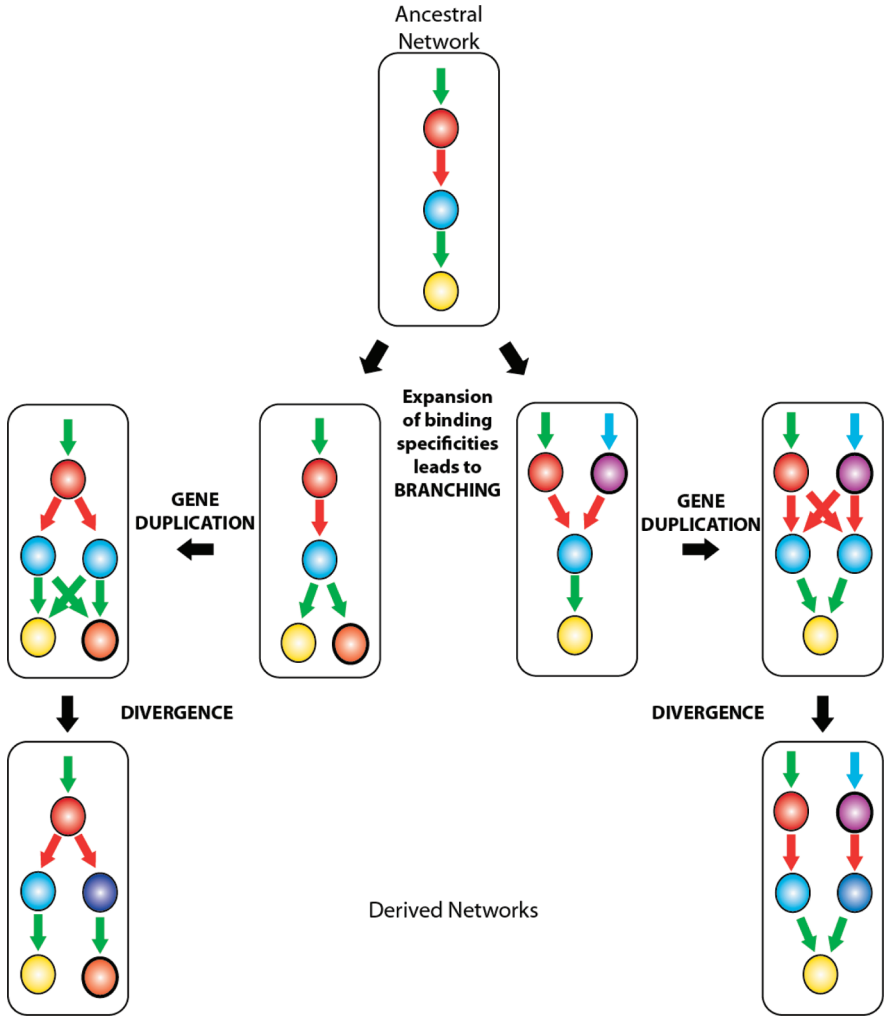


Figure 2. Gene duplication and divergence can lead to the evolution of independent branches. The duplication of pathway components with expanded specificities may initially create redundant branches. The accumulation of additional mutations could diverge interaction specificities, eventually leading to the evolution of independent branches.

specificities. Analysis of MAPK sequence conservation revealed that although Fus3 and Hog1 are structurally similar, they differ in multiple patches of continuous residues that seem to correlate with interaction specificities. To investigate the role that these amino acid patches have in mediating binding specificities, Ramanathan and co-workers created chimeric MAPK by swapping amino acid patches between Fus3 and Hog1. Swapping of only one or a few patches was not enough to switch specificity, leading instead to chimeric kinases with expanded specificities: kinases that could be activated by either mating or high osmolarity stimuli or that could activate both mating and high osmolarity downstream substrates. This indicates that wild type binding specificities are remarkably robust to mutations that add a new interaction partner. Complete switch of binding specificity required the replacement of several patches of residues. Evolution through amino acid replacements is a gradual process. Therefore, network rewiring by mutations of residues that mediate binding interactions is likely to proceed through evolutionary intermediates with expanded specificities, as the intermediates seen by Ramanathan and co-workers (Figure 1). This expansion of the range of binding specificities could effectively add a new branch to the pathway. In some occasions, branching may create a beneficial link between two different cellular processes, providing a selective advantage. Eventually, additional amino acid replacements might complete the switch in binding specificity, pruning the ancestral branch and leaving only the newly created branch. Alternatively, gene duplication might create two copies of the signaling protein encoding multiple interaction specificities (e.g., a hypothetical ancestor of Fus3 and Hog1, in the example studied by Ramanathan and co-workers), each now free to further refine its binding specificity (Figure 2). In this case, each branch could eventually become independent.

Evolutionary trajectories that alter substrate specificity through intermediates with broad specificity have also been described for enzymes involved in metabolism[3] and may represent a general property of proteins in which binding and catalysis are integrated within a structural unit. In the MAPK studied by Ramanathan and co-workers, evolution seems to have selected binding patches that are structurally and functionally decoupled from residues with important roles in stability or folding, as changes in binding specificities can occur without seriously compromising function. This suggests that, even in proteins with regulatory and catalytic functions integrated within a single domain, a certain level of modularity exists.

dx.doi.org/10.1021/sb300012g | *ACS Synth. Biol.* 2012, 1, 199–210

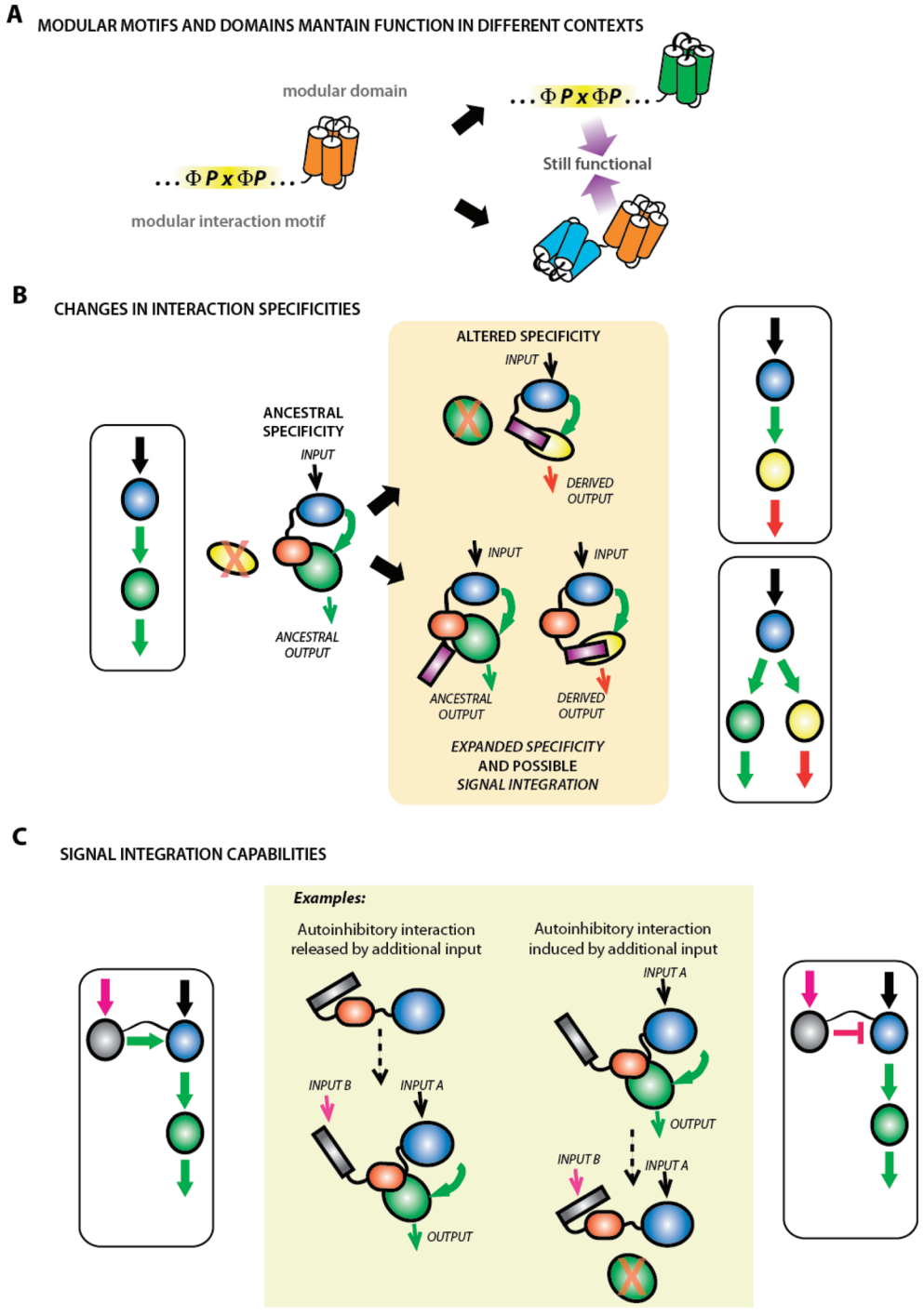


Figure 3. Network rewiring through shuffling of modular motifs or domains. (A) Signaling proteins often contain highly modular short interaction motifs and/or domains that perform functions in a context-independent manner. Modular motifs and domains can still perform their functions when shuffled. (B) Motif or domain shuffling can lead to the gain or loss of interaction specificities in a single mutational step. For example, motif or domain replacement could rewire a pre-existing input into a novel output. Alternatively, the addition, rather than the replacement, of a motif or domain could expand the range of interactions, adding branches to the pathway. In some cases, signal propagation might depend on particular combinations of functional partners, leading to novel signal integration capabilities. (C) Examples of signal integration: left, A AND B logic gate; right, A NOT B logic gate.

Two-component signaling systems are another well-studied example that illustrates how replacement of residues mediating binding interactions could lead to the evolutionary rewiring of signaling networks through intermediates with expanded interaction specificities. Two-component systems are the predominant signaling system in bacteria and are also important in plants. They consist of a sensor histidine kinase (HK) and a response regulator (RR).[10] Activation of the HK leads to its autophosphorylation. The HK then transfers the phorphoryl group to a cognate RR. Once activated, RRs can alter a wide range of cellular processes, often through changes in transcription.[10] Multiple HK-RR pairs are present in a cell. Specific interactions between each pair of HK and RR are an important factor ensuring that specificity is maintained, so that

dx.doi.org/10.1021/sb300012g | ACS Synth. Biol. 2012, 1, 199–210

when one HK is activated, only the proper RR is phosphorylated.[12] Michael Laub and co-workers[11,13] used a computational approach to identify co-varying residues on HK and RR that could be responsible for interaction specificity. In particular, they analyzed ∼1300 HK-RR pairs to identify residues in each HK-RR pair that co-vary. Then, they replaced co-varying residues believed to be responsible for specificity in *E. coli* EnvZ HK with those of *E. coli* RstB HK and found that replacement of only one or two residues expanded the range of phosphotransfer specificity, to include the original EnvZ RR, as well as RstA (RstB's RR). Subsequent replacement of a third residue completely switched the specificity to that of RstA, again indicating that interaction specificities can evolve gradually, through intermediate states of broad specificity that add branches to the network. Moreover, the observation that a few residues suffice to confer binding specificity suggests that evolution could alter specificities through a small number of mutational steps.

Short Linear Docking Motifs Provide Fast Routes for Network Architecture Rewiring. Evolutionary rewiring through changes in binding specificity, when regulatory and catalytic activities are integrated in a single domain, is possible as long as residues responsible for binding are decoupled from those that mediate catalysis, folding, and stability. Replacements of coupled residues could impact folding or catalysis, seriously compromising function. Therefore, one could imagine that only a few of all possible amino acid replacements would be tolerated, somehow limiting the evolutionary potential, or evolvability, of this type of rewiring mechanism. Protein−protein interactions, though, can be mediated not only by residues located in structured domains but also by short linear motifs, usually located within flexible unstructured regions, capable of interacting with specialized binding domains.[14] Examples of short motifs that interact with peptide-binding domains are polyproline motifs that bind SH3 or WW domains, C-terminal motifs that bind PDZ domains, or phosphotyrosine-containing motifs that bind SH2 domains, to name a few. Three major factors make evolutionary network rewiring through changes in short linear motifs less constrained. First, because flexible protein regions are free of structural and catalytic constrains, more amino acid substitutions are likely to be tolerated. Second, linear docking motifs are usually short (e.g., a few amino acids long); therefore gain or loss of docking motifs can occur by a single or a few amino acid changes. Third, linear motifs usually bind to modular domains that are, themselves, decoupled from catalytic domains and are therefore also more likely to accept new interacting partners easily. When a motif is added to a protein (either by amino acid replacement or by shuffling of pre-existing motifs and domains), the resulting protein acquires the ability to interact with a new partner, adding a new connection to the network. Because signaling proteins are often modular, addition of a motif could be sufficient, in some cases, for the new interaction to result in a functional change in the network (Figure 3).

Saito and co-workers[15] analyzed the role of short interacting motifs in directing the flow of information in yeast MAPK-mediated signaling networks. Pbs2 is the MAPK Kinase (MAP2K) in the yeast high osmolarity pathway. Pbs2 is activated by three different MAP2K Kinases (MAP3K): Ssk2, Ssk22, and Ste11. Saito and co-workers first identified a short docking motif present in the N-terminus of Pbs2 that mediates the interaction with Ssk2/22. Then, they demonstrated that the short N-terminal docking motif and the MAP3K/MAP2K phosphorylation step are highly modular. For that, they added Pbs2 docking motif to the N-terminus of Ste7, the MAP2K of the yeast-mating pathway, and found that the resulting chimeric Ste7 variant could be activated by Ssk2/22, effectively rewiring high osmolarity stimulus to mating output. The chimeric Ste7 variant was still able to signal mating input to its natural substrate, the mating MAPK Fus3. Thus, addition of the N-terminal Pbs2 motif to Ste7 did not compromise the function of a second docking motif present in Ste7 and necessary for the interaction between Ste7 and Fus3.[15] This suggests that signaling proteins could accumulate multiple docking motifs, consequently expanding their interaction capabilities. In a complementary study, Won et al.[16] replaced Ste7 kinase domain for that of other yeast MAP2K, in particular Mkk1, Mkk2, and Pbs2, and determined whether the resulting chimeras could functionally substitute Ste7. Because the interaction of Ste7 with its substrate Fus3 occurs when both Ste7 and Fus3 are bound to the scaffold protein Ste5 and the interaction between Ste7 and Ste5 is mediated by Ste7 kinase domain, they also covalently attached the chimeric kinases to the scaffold. In that context, chimeric Mkk2 and Pbs2 could phosphorylate Fus3 in a mating stimulus-dependent manner, indicating that the identities of the chimeric proteins were governed by Ste7 interaction motif, rather than by the catalytic domains. From these studies, we can conclude that interaction motifs and kinase domains are remarkably modular, and gain/loss of interaction motifs can easily override kinase identities in establishing new enzyme−substrate pairs and the resulting novel network connections. These studies also highlight the fact that some signaling kinases are rather promiscuous, willing to accept novel substrates as long as they are brought to close proximity, without the need for specific changes in active site residues.

Thus far, we have focused on examples in which interaction motifs were either swapped or, if added, were added far apart from other pre-existing motifs, therefore not perturbing pre-existing functions. However, mutational events may introduce a new interaction motif in close proximity to, or even partially overlapping, a pre-existing motif. In this case, it may be possible that the addition of one motif could alter the function of the pre-existing motif, sometimes resulting in novel signal-processing capabilities. Sallee et al.[17] rationally designed peptides containing partially overlapping motifs that mediate interactions with either Crk SH3 domain or, when phosphorylated, with a 14−3−3ζ domain, and created mutually exclusive binding patterns that resulted in switch-like signal processing capabilities. For example, incubation of a phosphorylated designed peptide with Crk SH3 resulted in a phosphopeptide-Crk SH3 complex. Subsequent addition of the 14−3−3ζ domain displaced Crk SH3, indicating that overlapping motifs can act as "OR" switches. Thus, in addition to widening the range of specificities, network architecture rewiring through the evolution of interaction motifs can create novel signal-processing functions.

Within an individual cell, there are multiple proteins carrying linear interaction motifs that bind to a given domain type, as well as multiple versions of the domain they interact with. For example, in the yeast *S. cerevisiae*, there are 27 different SH3 domains and a large number of proteins possessing SH3-binding motifs. Motifs−domains interactions could, in principle, range from very specific to very promiscuous. Presumably, evolution might favor one or the other, depending on particular circumstances. For instance, the MAPK Fus3 has a

docking site capable of interacting with docking peptides present in its upstream activator Ste7, the substrate Far1, or the phosphatase Msg5, a Fus3 regulator. In this case, evolution seems to have favored a certain level of promiscuity to enable network function and regulation. In other cases, specificity might be preferred. For instance, Zarrinpar et al.[18] analyzed potential cross-reactions between all 27 yeast SH3 domains and a linear interaction motif from the yeast protein Pbs2, known to interact with an SH3 domain from the transmembrane receptor Sho1. They found that while no other SH3 domain present in yeast interacts with the Pbs2 motif, several non-yeast SH3 domains do, suggesting that negative selection has optimized the function of the yeast SH3-peptide interacting network, preventing promiscuous binding. Recent work by Gorelick et al.[19] began to unravel the molecular basis of negative selection-derived binding specificity. In particular, Gorelick et al.[19] studied the interactions between the yeast Bem1 SH3 and Nbp2 SH3 domains and binding motifs from six yeast proteins. Specifically, they identified a Lysine residue highly conserved in Bem1 SH3 domain that, rather than contributing to increase the binding energy for the cognate interaction motif, prevents nonspecific interactions.

Recombination of Modular Domains Enables Network Rewiring without Promiscuous Intermediates. Individual amino acids can be replaced, inserted or deleted, in a process that could lead to the gradual evolution of novel binding specificities and the concomitant change in network architecture and function. While *de novo* creation of domains could also occur gradually, networks are more often rewired as a result of mutational events, such as recombinations, duplications, or transpositions, that alter the domain composition of pre-existing proteins, in a single step. Because these mutational processes are not gradual, evolutionary trajectories that result from domain shuffling do not have to proceed through intermediate states with broad specificities. Instead, new interaction partners might be added or removed at once (Figure 3).

Two-component systems provide an ideal example to appreciate how different types of mutations (e.g., amino acid substitutions versus domain swapping) rewire networks through different evolutionary trajectories. I have discussed above the studies of Laub and co-workers,[11] who analyzed how interaction specificities between HK and RR pairs evolve. As mentioned, they showed that the gradual replacement of specificity-determining amino acids in EnvZ HK alters the specificity of EnvZ HK, from its cognate EnvZ RR to RstA RR, through an intermediate protein capable of phosphorylating both EnvZ RR and RstA RR. In contrast, replacement of the DHp domain in EnvZ HK, which contains all of the specificity-determining residues, for that of RstB results in a chimeric EnvZ HK that can phosphorylate RstA but has lost the ability to phosphorylate EnvZ RR.[11] Thus, while evolution of substrate specificity by sequential replacement of amino acids in EnvZ HK DHp domain proceeds through an intermediate with dual specificity that adds a new branch to the pathway, swapping of DHp domains switches specificities, adding a new branch and removing the ancestral branch, in a single step.

Domain Recombination Can Evolve Novel Signal Processing Capabilities. As with short linear motifs, the addition, rather than the replacement, of a domain might result in an expanded range of interactions. Depending on multiple factors, such as the particular type of domains, their arrangement, or the length and flexibility of interdomain links, to name a few, different multi-input signal processing capabilities could be created. Dueber et al.[20] explored the evolutionary potential of domain shuffling *in vitro*, by creating a small synthetic library where three interaction domains (GBD, PDZ, and SH3) were shuffled with N-WASP catalytic domain and then measuring stimuli-dependent actin polymerization. By varying the order of the domains and the length of the connecting linkers, they built a library of 34 variants. From those, five were constitutively active, nine were constitutively repressed, and 20 showed stimulus-dependent actin polymerization activity. More importantly, 18 of those 20 could respond to two different stimuli, including two "OR" switches (either of the two stimuli suffices for activation), five "AND" switches (two concurrent stimuli needed for activation), and nine with intermediate behaviors. These results (and others[21]) indicate that multidomain switches are highly modular and demonstrate that evolution could generate novel allosteric input/output relationships, by mutational events that alternate regulatory and catalytic domains. Mutation, however, could also insert domains within domains. While one may think that large intradomain insertions would most likely destroy a pre-existing function, experiments by Ostermeier and co-workers[22] have shown that, in some cases, they can generate allosteric switches that, rather than abolishing a pre-existing function, add to it a new level of regulation. Specifically, they created a synthetic library in which TEM-1 β-lactamase was randomly inserted within *E. coli* Maltose-Binding Protein (MBP). Remarkably, ~800 out of ~20,000 library members encoded enzymes that were able to both bind maltose and hydrolyze ampicillin. Among those, they identified multiple variants in which maltose acted as an allosteric regulator, some of them with maltose-dependent increases in catalytic efficiency of ~80%. These results suggest that intradomain insertions can be tolerated and, more importantly, can establish novel regulatory functions.

Domain Recombination Can Alter Network Function by Changing the Time or the Subcellular Localization of an Interaction. Shuffling of protein domains can alter the function of a signaling pathway not only by creating novel allosteric switches but also in other ways. For example, domain shuffling could change the subcellular localization or the time at which an interaction takes place and, in this manner, could alter the dynamics of network activation. The work of Peisajovich et al.[23] provides us with valuable insights about how changes in subcellular localization or interaction dynamics may alter network function. Peisajovich et al.[23] systematically explored the role that domain shuffling may have in the evolution of a signaling network *in vivo*, using the yeast mating pathway as a model system. In particular, they created a library of all possible 66 domain-recombination variants, with domains derived from 11 proteins in the mating pathway. About 15% of the library variants altered the dynamics of mating pathway activation. Among the different phenotypes obtained, there were mutant pathways able to respond either faster or slower to stimulation, pathways that were active even in the absence of stimulus, and pathways that completely failed to respond. Moreover, a correlation was observed between the dynamics of pathway activation, the maximum pathway output, and the overall efficiency of the mating process, with some mutants able to mate more efficiently than wild type under laboratory conditions. While further experiments are needed to fully understand the mechanisms by which domain recombinations alter mating pathway behavior, fluorescence microscopy studies of some of the chimeric protein variants suggest that changes in

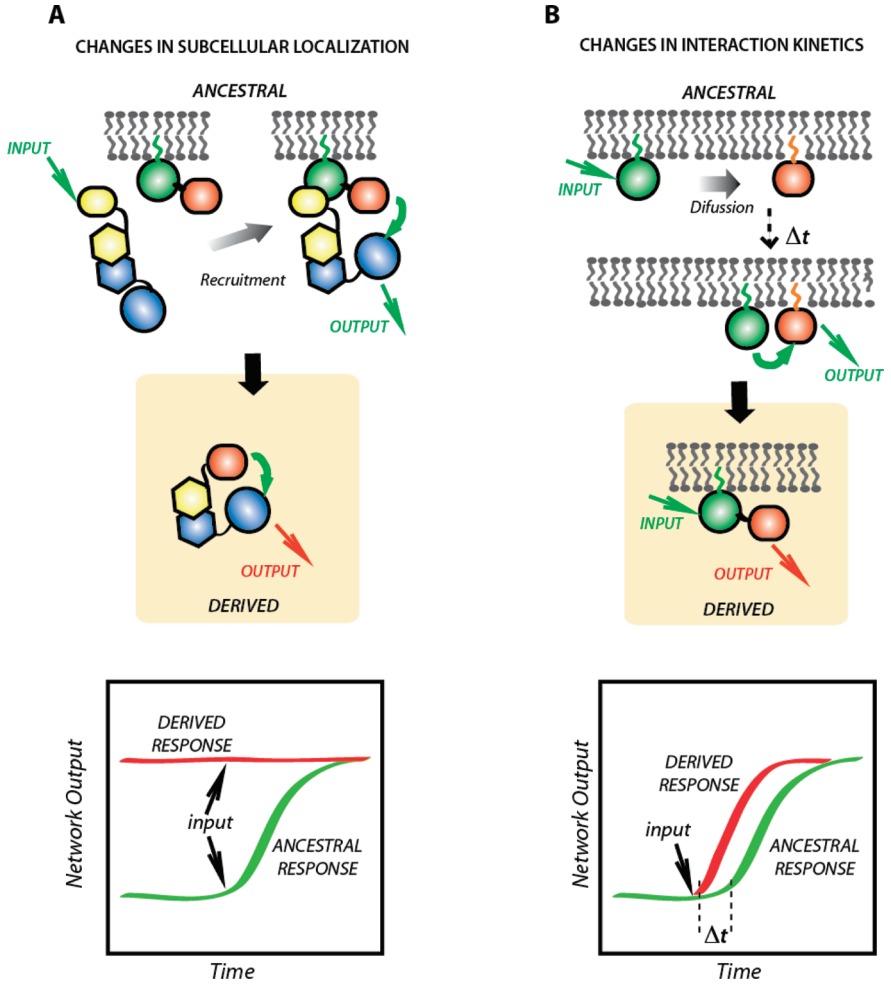


Figure 4. Recombination of modular domains/motifs could alter the place or time at which an interaction takes place. Shuffling of modular domains or motifs could also affect network function by changing interactions' spatial or temporal dynamics. (A) In this example, shuffling creates a direct interaction between an upstream activator and its downstream substrate, bypassing the need for input-dependent mediators. This scenario, analogous to the shuffling of the yeast mating pathway Ste50 SAM domain and Ste20 kinase domain,[23] leads to the constitutive activation of the pathway. (B) In this example, shuffling creates a covalent interaction between domains that previously interacted only transiently, analogous to the shuffling of the yeast mating pathway Ste18 and Cdc42 proteins.[23] By eliminating the diffusion-controlled search of both interaction partners on the membrane surface, the signal propagates faster.

subcellular localization played an important role. For instance, in wild type yeast, activation of the mating pathway leads to the recruitment of the MAP3K Ste11 to the plasma membrane, via interactions with the mating scaffold Ste5. In this way, Ste11 co-localizes with its activator, the PAK kinase Ste20. The interaction between Ste11 and Ste20 is indirect: Ste11 interacts with the adaptor protein Ste50 (through Ste50 N-t SAM domain), which in turn interacts with the membrane-localized small GTPase Cdc42. Ste20 also binds to Cdc42, an interaction mediated by Ste20 N-t RBD domain. Through domain shuffling, Peisajovich et al.[23] created a chimeric protein containing Ste50 N-t SAM domain fused to Ste20 C-t kinase domain. This chimeric protein binds to Ste11 via the N-t SAM domain, relocalizing the site of Ste11 activation away from the plasma membrane to the cytoplasm and resulting in the constitutive activation of Ste11 by Ste20 C-t kinase domain and the subsequent activation of the mating pathway. Domain shuffling can also alter the kinetics of network activation by altering the timing of an interaction. For instance, upon yeast mating pathway activation, the membrane-bound G protein γ-subunit Ste18 (in a complex with the β-subunit Ste4) is released form the G-α subunit, and diffuses on the membrane to bind Cdc42. Recombination of Ste18 and Cdc42 into a single polypeptide leads to a mutant pathway that is activated by mating pheromone twice as fast as wild type.[23] Presumably, in wild type cells diffusion of the Ste4/Ste18 complex on the membrane surface to find Cdc42 delays signal transmission, while preassembly of the Ste18-Cdc42 complex by covalent fusion bypasses diffusion and propagates the signal faster (Figure 4).

Scaffold Proteins Are Highly Evolvable Network Hubs. Signaling scaffolds are proteins capable of binding multiple signaling components, organizing signaling complexes in space and time. In addition, by confining activated signaling proteins and their proper substrates to specific locations, scaffolds contribute to signaling fidelity, preventing spurious cross talks. Scaffolds have been identified in a large number of signaling pathways, including the yeast Ste5 and Pbs2, involved in mating and high osmolarity responses, respectively, or Far1 and Bem1, involved in organizing protein complexes responsible for cell polarity. Scaffolds are key components of signaling networks in higher eukaryotes as well; examples include mammalian KSR,

dx.doi.org/10.1021/sb300012g | *ACS Synth. Biol.* 2012, 1, 199–210

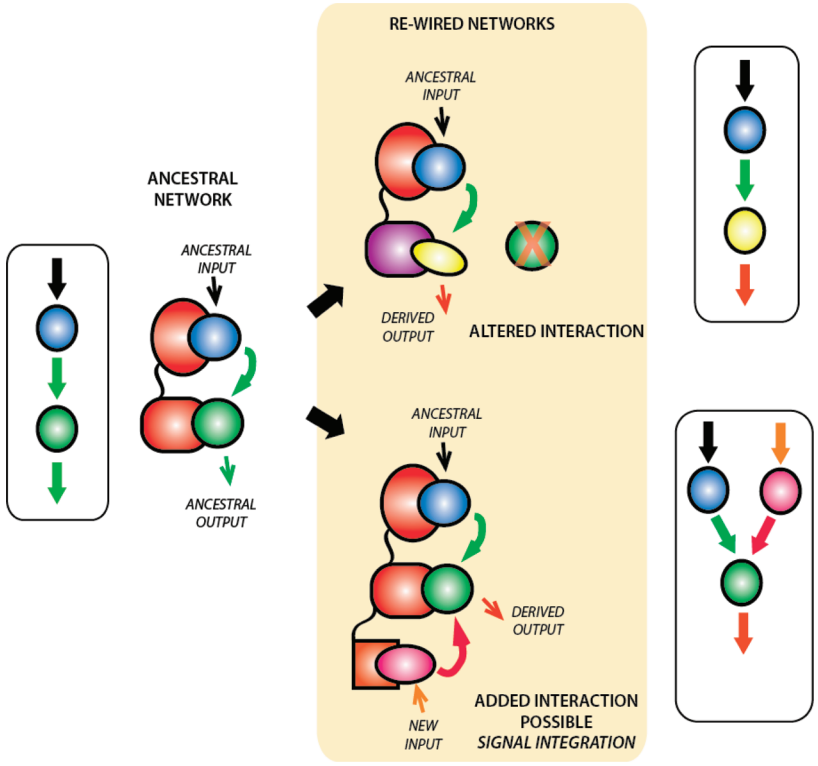


Figure 5. Network rewiring through mutations in scaffolds. Signaling scaffolds are proteins capable of binding multiple signaling components, organizing signaling complexes in space and time. Scaffolds are highly modular signaling hubs and thus have high evolutionary potential. The gain or loss of interaction modules in the scaffold could alter network properties by recruiting (or removing) additional signaling components. For example, switching an interaction module for another could rewire a pre-existing input to a novel output. Alternatively, the addition, rather than the replacement, of an interaction module could add a branch to the network and/or alter signal integration capabilities.

which assembles a MAPK signaling complex that contains Ras, Raf, MEK and ERK, the T-cell signaling scaffolds LAT and SLP-76, or the synaptic PSD-95.[24] While some scaffolds may only contribute to the assembly of a signaling complex, others also contain catalytic domains that are necessary for signaling function. For instance, Pbs2, the scaffold in the yeast high osmolarity pathway, is also the MAP2K in the pathway, whereas Ste5, the scaffold in the yeast mating pathway, catalytically unlocks the MAPK Fus3 to allow its activation by the MAP2K Ste7.[25]

Scaffolds control the assembly of functional signaling complexes through modular interactions with multiple binding domains or motifs. Because of their modularity, scaffolds have an enormous evolutionary potential. From amino acid changes that alter the interaction specificity of a scaffold domain or motif, to the blunt addition or removal of whole domains or motifs, mutations could easily alter the composition of signaling complexes through scaffold remodeling. A comparison between the *S. cerevisiae* scaffolds Ste5 and Far1 provides an example of amino acid replacements that alter scaffold's binding specificities, without gain or loss of domains. Ste5 and Far1 have similar domain compositions. They both contain a protein-interaction RING domain, followed by a phosphoinositide-interaction PH domain, and a protein interaction vWA domain.[26] Still, the precise binding specificities of some (or all) of these interaction domains are different, as they bind to different targets (e.g., Ste5 binds Ste11, Ste7 and Fus3, while Far1 binds Cdc24 and Cdc42, among other proteins). A plausible evolutionary scenario is that Ste5 and Far1 are derived from the duplication of an ancestral gene and the subsequent divergence through amino acid replacements and small insertions and deletions, but without the gain or loss of whole domains. In contrast, *Candida albicans* Ste5 ortholog has lost the C-terminal vWR domain, suggesting that interactions have migrated to other regions in the protein or that *S. cerevisiae* and *C. albicans* Ste5 functions have diverged.[26]

Synthetic biology has provided us with several examples illustrating how evolution could alter the function of a signaling pathway by mutational changes in scaffolds. Park et al.[27] created a chimeric scaffold in yeast that combined Ste5 binding sites for the upstream mating signaling components Ste4 and Ste11, with Pbs2 binding sites for the downstream high osmolarity signaling component Hog1. This chimeric scaffold was able to redirect a mating input into a high osmolarity output. Similarly, Howard et al.[28] created a chimeric protein that fused the EGF receptor interacting domain of Grb2 with the Caspase binding domain of DED. In this way, they were able to redirect a receptor tyrosine kinase-mediated proliferation signal to an apoptotic Caspase pathway. Taken together, these studies suggest that changes in the co-localization of signaling components, as a consequence of mutations in scaffolds, may suffice to rewire network architecture (Figure 5).

In addition to organizing signaling complexes, scaffolds can also affect signaling responses in a quantitative way. For example, membrane recruitment of the mating scaffold Ste5 has been shown to facilitate the propagation of weak signals, effectively transforming a signaling pathway that otherwise would have an ultrasensitive response into a pathway with a graded response.[28] Bashor et al.[29] investigated how evolution could reshape signaling dynamics by altering the function of the

dx.doi.org/10.1021/sb300012g | *ACS Synth. Biol.* 2012, 1, 199–210

yeast mating scaffold. For that, they engineered an additional interaction domain in Ste5 and used that domain to recruit mating pathway modulators to the scaffold. First, they constitutively expressed an additional copy of the negative regulator Msg5 (a phosphatase that reversibly inactivates the MAPK Fus3) or the positive regulator Ste50 (the adaptor protein that connects the MAP3K Ste11 with its upstream activator). When the modulators were recruited to the scaffold via the engineered interaction domain, they altered pathway output: Msg5 reduced pathway output, while Ste50 increased it. In contrast, expression of unrecruited Msg5 or Ste50 had only marginal effects on pathway function. Then, they created positive or negative feedback loops by controlling the expression of the scaffold-recruited pathway modulators with mating-responsive promoters. Finally, they combined scaffold-recruited positive and negative modulators to create a wide range of dynamic signaling behaviors, including pulse generators, accelerators, delays, and ultrasensitive switches. The work of Bashor et al.[29] demonstrates that scaffolds confer substantial versatility to signaling responses and suggests that scaffolds are attractive targets for pathway evolution. Recombination of modular interaction domains or motifs, as well as amino acid replacements that alter or modulate interaction specificities, could change the composition of scaffold-organized signaling complexes and, consequently, evolve novel signaling properties.

2. Evolution of Network Architecture through Rewiring of Transcriptional Connections. Signaling network activation often leads to changes in transcription regulation. Thus, evolution could rewire signaling networks not only by altering protein−protein binding specificities but also by mutational events that change interactions between proteins and DNA. For example, mutations affecting the DNA-binding region of a transcription factor or mutations in promoter DNA could affect transcription regulation. Cis-Regulatory Elements (CREs) and Open Reading Frames (ORFs) are highly modular. Therefore, networks could be rewired by swapping CREs and ORFs, analogously to the swapping of regulatory and catalytic domains in signaling proteins. For brevity, rather than discussing how evolution could alter CREs or DNA-binding transcriptional regulators, I will focus here on studies that illustrate how system-level rewiring could be achieved by shuffling of modular CREs and ORFs.

Combinatorial Recombination of CREs and ORFs Leads to Transcriptional Networks with Diverse Behaviors. Some of the simplest regulatory networks are composed of a set of interconnected transcription regulators (either activators or repressors) and their corresponding CREs. Guet et al.[30] used a combinatorial approach to investigate how evolution could generate diverse patterns of regulated gene expression, simply by altering the connectivity between transcription regulators. In particular, they used the lacI repressor, the tetR repressor, the phage λ cI transcription regulator, and five promoters, two repressed by lacI with different strength, one repressed by tetR, one repressed by λ cI, and one activated by λ cI. Binding of lacI and tetR to their corresponding promoters can be controlled by the small molecules β-D-isopropyl-thiogalactopyranoside (IPTG) and anhydrotetracycline (aTc), respectively. To explore a wide range of evolutionary scenarios, they created a library of all possible combinations of three promoter/transcriptional regulator pairs, such as that expression of each of the three transcriptional regulators in the network was controlled by one of the five promoters. They varied network inputs by altering concentrations of IPTG and aTc and determined network output by measuring the expression of GFP controlled by the λ cI repressible promoter. In this way, they identified clones encoding networks in which the GFP output was controlled by a binary logical function of the two inducers. DNA sequencing of 30 clones with distinct behaviors identified 13 different network architectures, indicating that changes in network connectivity are a major source of phenotypic diversity. In some instances a single change in the connectivity of the network (e.g., by transposition, recombination or gene duplication) would suffice to convert the network operation from one logical function to another. As expected, network architecture is not the only determinant of function: some networks had similar connectivity but performed different functions, suggesting that evolution could alter network function in subtle ways that do not require changes in architecture. This observation also reminds us that it will be difficult to predict network function simply from connectivity. Furthermore, they also found that networks with different architectures could perform similar functions, suggesting that evolutionary history, and not only performance, could dictate which networks are present in natural organisms.

Novel Phenotypes Can Evolve by Rewiring of Master Transcriptional Regulators. Transcriptional networks, though often more complex than those analyzed by Guet et al.,[30] have fundamental roles in many biological processes. In *E. coli*, for instance, nine master regulators control, in direct or indirect ways, expression of about half of the genome. This highly centralized form of control has been proposed to lead to system robustness, as random changes are unlikely to affect a small number of highly connected genes. On the other hand, when changes in any one of the master regulators do occur, they could have cascading effects on a wide range of cellular processes. Isalan et al.[31] explored the functional consequences of transcriptional network rewiring in *E. coli* by analyzing a library of almost 600 duplications of shuffled pairs of promoters and transcription regulators that included seven master regulators, seven σ factors, and eight downstream transcription factors. Rewired network outputs were determined by measuring the expression levels of a GFP reporter controlled by csgD, one of the eight transcription factors included in the library. Analysis of GFP expression levels already led to unexpected observations. First, in some cases GFP levels depended on the identity of the duplicated transcription factor, independently of which promoter controlled its expression. This suggests that some transcription factors achieve expression levels that are somehow promoter-independent and therefore could be robust to evolutionary change. Second, in some instances GFP levels for rewired pairs that were predicted to encode either positive or negative feedback loops were similar, indicating that feedback loop behavior *in vivo* is difficult to predict and suggesting that nonobvious mechanisms of network control could mask the effects of introduced direct feedback elements.

Many of the analyzed networks included a duplicated master regulator or sigma factor under control of a rewired promoter, likely resulting in the altered expression of a large number of genes. To determine possible global effects resulting from these drastic changes in regulation, Isalan et al.[31] measured growth rates as indicators of overall fitness. Surprisingly, under the experimental conditions tested, ~84% of rewired networks had little or no effect on growth, suggesting that the transcriptional

network of *E. coli* might be more robust than expected, even to mutations that change the expression of master regulators. In addition, while rewiring seemed well-tolerated under standard laboratory conditions, the authors were able to select, from the pool of ~600 library clones, rewired networks that were better adapted to serial passages, to continuous incubation at stationary phase, or to heat shock, suggesting that transcriptional rewiring could be, in some cases, sufficient for adaptation. Finally, transcriptome-wide analysis of gene expression for some of the selected networks determined that rewiring perturbed the expression of only ~10% of the genes, indicating that some transcriptional modules are insulated, as the effect of rewiring does not propagate across the whole network.

Parameter Tuning by Base Substitutions Could Be Needed for Proper Function of a Rewired Network. While recombination of CREs and ORFs can alter network architecture, the establishment of novel complex regulatory behaviors may also require adjustments in parameters, such as binding affinities or components concentrations, governing the rewired interactions. Yokobayashi et al.[32] explored how evolution could tune two such parameters, ribosome binding site strength and protein abundance, in a synthetic regulatory network. For that, they first designed a circuit in which a constitutively expressed lacI repressor blocks the expression of a second transcriptional repressor, lambda phage cI. In turn, when expressed, cI blocks the expression of the fluorescent reporter GFP controlled by a lambda phage Pro12 promoter. In principle, addition of IPTG would release repression by lacI, leading to the expression of cI and the subsequent repression of GFP. However, proper functioning of this simple circuit requires that the expression levels of cI match the concentration range of the lambda phage Pro12 promoter. In fact, a first implementation of this design failed, as leaky expression of cI, in the absence of the IPTG inducer, repressed GFP expression. Yokobayashi et al.[32] then used a directed evolution approach to select mutant circuits that would express GFP only in the absence of IPTG. In particular, they created libraries with mutations in the ribosome binding site controlling cI expression or in the cI gene, assuming that changes in the basal levels of cI were necessary for proper circuit function. About 50% of all library mutants expressed GFP in the absence of IPTG. From those, ~5−10% were non-fluorescent in the presence of IPTG. DNA sequencing of the selected circuits revealed that multiple solutions existed. A set of variants presented mutations near the ribosome-binding site or even disrupted the start codon of cI, presumably reducing translation efficiency and, consequently, cI concentration. A second set of mutants acquired a premature STOP codon that eliminated a C-terminal fragment of cI needed for homo-oligomerization. Reduction in oligomerization ability presumably reduces DNA binding affinity, thus attenuating repression of GFP expression. A third set of variants presented base substitutions in the cI coding region (some only synonymous substitutions!), probably decreasing transcription or translation efficiency or reducing the concentration of active protein. Thus, while Yokobayashi et al.[32] did not fully explored the mechanisms at play, it is possible to conclude from their experiments that, even when constrained in the type of mutations available (e.g., libraries were built by error-prone PCR leading mostly to base substitutions), evolution can still find multiple ways to optimize circuit parameters.

SYNTHETIC BIOLOGY APPROACHES TO INVESTIGATE SIGNALING NETWORK EVOLUTION AT THE GENOME-WIDE LEVEL

In the studies discussed above, mutations were introduced in predefined genes or DNA regulatory regions. While this targeted approach facilitates analysis and enables us to study how evolution affects our pathways of choice, natural evolution acts blindly, at the level of the whole genome. The study of regulatory signaling networks will certainly benefit from the adoption of methodologies that enable genome-wide mutagenesis and analysis. In fact, numerous methods have already been developed that can target multiple genes, even any gene in the genome, at once. However, thus far they have been mostly used either to explore how organisms adapt to stress, how metabolism evolves, or for practical applications in metabolic engineering. Methods that follow the evolution of microorganisms in response to laboratory-controlled selection pressures have been in use for many years. From the pioneering work of Luria and Delbruck, showing that adaptive mutations predate selection,[33] or that of Hall,[34] demonstrating that novel enzymatic functions can arise without compromising existing functions, genome-wide experimental evolution has been used to explore fundamental aspects of natural evolution. More recently, the availability of microarrays and whole genome sequencing technologies has added a radically new dimension to the study of genome-wide evolution. It is now possible to identify, with single-base resolution, all (or at least most) mutations, as they occur during the evolutionary process, their frequencies in the population, and their specific contributions to phenotype (for illustrative examples see refs 35, and 36). Experimental microbial evolution, combined with selection pressures based on responses to specific external stimuli, could be used to investigate the evolution of regulatory signaling networks.

Experimental evolution can be based on an organism's natural mutation rate or could be accelerated by the use of mutagens, to increase mutation rates across the whole genome. Alternatively, methods exist to augment mutation rates for defined genes of interest. Church and co-workers[37,38] adapted an oligo-mediated allelic replacement method[39,40] to enable the continuous randomization of multiple genes in parallel in prokaryotes. The approach, named "Multiplex Automated Genome Engineering" (MAGE), has been used for metabolic engineering, as well as for genome-wide codon replacement in *E. coli*. By targeting multiple genes with increased mutations rates at once, MAGE is well suited to investigate how different genes within a network (or in different networks) co-evolve.

Exploring mutational landscapes beyond amino acids substitutions at the genome-wide level is more challenging. Still, several approaches have been developed that allow combinatorial cloning of multiple large inserts, *in vitro* or *in vivo*.[41−43] These methods, combined with site-directed or random genome insertions (e.g., by transposons or recombinases) and genome shuffling,[44,45] are opening the door to studies aimed at understanding the roles that structural variation (e.g., duplications, translocations, recombinations, etc.) has in the evolution of regulatory signaling networks. As our ability to synthesize genomes improves,[46−49] we shall witness a radical revolution in the study of regulatory signaling networks evolution.

FINAL REMARKS

For centuries, we learned about nature by observation. During the twentieth century, we acquired the ability to perturb biological systems and, in that way, were able to better understand biological functions. With the advent of synthetic biology, we are now gaining the ability to create novel biological systems. By combining the concepts and methods of synthetic biology with those of directed evolution, we are advancing our understanding of the evolutionary processes that shape complex biological systems. Furthermore, useful applications are likely to be derived from the integration of synthetic biology and the unique bioengineering abilities of evolution.

AUTHOR INFORMATION

Corresponding Author

*Phone: +1 (416) 978-2143. E-mail: sergio.peisajovich@utoronto.ca.

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Bloom, J. D., and Arnold, F. H. (2009) In the light of directed evolution: pathways of adaptive protein evolution. *Proc. Natl. Acad. Sci. U.S.A. 106* (Suppl 1), 9995−10000.

(2) Peisajovich, S. G., and Tawfik, D. S. (2007) Protein engineers turned evolutionists. *Nat. Methods 4*, 991−994.

(3) Aharoni, A., Gaidukov, L., Khersonsky, O., Mc, Q. G. S., Roodveldt, C., and Tawfik, D. S. (2005) The 'evolvability' of promiscuous protein functions. *Nat. Genet. 37*, 73−76.

(4) Peisajovich, S. G., Rockah, L., and Tawfik, D. S. (2006) Evolution of new protein topologies through multistep gene rearrangements. *Nat. Genet. 38*, 168−174.

(5) Weinreich, D. M., Delaney, N. F., Depristo, M. A., and Hartl, D. L. (2006) Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science 312*, 111−114.

(6) Stemmer, W. P. (1994) Rapid evolution of a protein in vitro by DNA shuffling. *Nature 370*, 389−391.

(7) Suel, G. M., Kulkarni, R. P., Dworkin, J., Garcia-Ojalvo, J., and Elowitz, M. B. (2007) Tunability and noise dependence in differentiation dynamics. *Science 315*, 1716−1719.

(8) Bhattacharyya, R. P., Remenyi, A., Yeh, B. J., and Lim, W. A. (2006) Domains, motifs, and scaffolds: the role of modular interactions in the evolution and wiring of cell signaling circuits. *Annu. Rev. Biochem. 75*, 655−680.

(9) Mody, A., Weiner, J., and Ramanathan, S. (2009) Modularity of MAP kinases allows deformation of their signalling pathways. *Nat. Cell Biol. 11*, 484−491.

(10) Stock, A. M., Robinson, V. L., and Goudreau, P. N. (2000) Two-component signal transduction. *Annu. Rev. Biochem. 69*, 183−215.

(11) Groban, E. S., Clarke, E. J., Salis, H. M., Miller, S. M., and Voigt, C. A. (2009) Kinetic buffering of cross talk between bacterial two-component sensors. *J. Mol. Biol. 390*, 380−393.

(12) Skerker, J. M., Perchuk, B. S., Siryaporn, A., Lubin, E. A., Ashenberg, O., Goulian, M., and Laub, M. T. (2008) Rewiring the specificity of two-component signal transduction systems. *Cell 133*, 1043−1054.

(13) Siryaporn, A., Perchuk, B. S., Laub, M. T., and Goulian, M. (2010) Evolving a robust signal transduction pathway from weak cross-talk. *Mol. Syst. Biol. 6*, 452.

(14) Remenyi, A., Good, M. C., and Lim, W. A. (2006) Docking interactions in protein kinase and phosphatase networks. *Curr. Opin. Struct. Biol. 16*, 676−685.

(15) Tatebayashi, K., Takekawa, M., and Saito, H. (2003) A docking site determining specificity of Pbs2MAPKK for Ssk2/Ssk22 MAPKKKs in the yeast HOG pathway. *EMBO J. 22*, 3624−3634.

(16) Won, A. P., Garbarino, J. E., and Lim, W. A. (2011) Recruitment interactions can override catalytic interactions in determining the functional identity of a protein kinase. *Proc. Natl. Acad. Sci. U.S.A. 108*, 9809−9814.

(17) Sallee, N. A., Yeh, B. J., and Lim, W. A. (2007) Engineering modular protein interaction switches by sequence overlap. *J. Am. Chem. Soc. 129*, 4606−4611.

(18) Zarrinpar, A., Park, S. H., and Lim, W. A. (2003) Optimization of specificity in a cellular protein interaction network by negative selection. *Nature 426*, 676−680.

(19) Gorelik, M., Stanger, K., and Davidson, A. R. (2011) A Conserved residue in the yeast Bem1p SH3 domain maintains the high level of binding specificity required for function. *J. Biol. Chem. 286*, 19470−19477.

(20) Dueber, J. E., Yeh, B. J., Chak, K., and Lim, W. A. (2003) Reprogramming control of an allosteric signaling switch through modular recombination. *Science 301*, 1904−1908.

(21) Yeh, B. J., Rutigliano, R. J., Deb, A., Bar-Sagi, D., and Lim, W. A. (2007) Rewiring cellular morphology pathways with synthetic guanine nucleotide exchange factors. *Nature 447*, 596−600.

(22) Guntas, G., Mitchell, S. F., and Ostermeier, M. (2004) A molecular switch created by in vitro recombination of nonhomologous genes. *Chem. Biol. 11*, 1483−1487.

(23) Peisajovich, S. G., Garbarino, J. E., Wei, P., and Lim, W. A. (2010) Rapid diversification of cell signaling phenotypes by modular domain recombination. *Science 328*, 368−372.

(24) Good, M. C., Zalatan, J. G., and Lim, W. A. (2011) Scaffold proteins: hubs for controlling the flow of cellular information. *Science 332*, 680−686.

(25) Good, M., Tang, G., Singleton, J., Remenyi, A., and Lim, W. A. (2009) The Ste5 scaffold directs mating signaling by catalytically unlocking the Fus3MAP kinase for activation. *Cell 136*, 1085−1097.

(26) Cote, P., Sulea, T., Dignard, D., Wu, C., and Whiteway, M. (2011) Evolutionary reshaping of fungal mating pathway scaffold proteins. *mBio 2*, e00230−00210.

(27) Park, S. H., Zarrinpar, A., and Lim, W. A. (2003) Rewiring MAP kinase pathways using alternative scaffold assembly mechanisms. *Science 299*, 1061−1064.

(28) Howard, P. L., Chia, M. C., Del Rizzo, S., Liu, F. F., and Pawson, T. (2003) Redirecting tyrosine kinase signaling to an apoptotic caspase pathway through chimeric adaptor proteins. *Proc. Natl. Acad. Sci. U.S.A. 100*, 11267−11272.

(29) Bashor, C. J., Helman, N. C., Yan, S., and Lim, W. A. (2008) Using engineered scaffold interactions to reshape MAP kinase pathway signaling dynamics. *Science 319*, 1539−1543.

(30) Guet, C. C., Elowitz, M. B., Hsing, W., and Leibler, S. (2002) Combinatorial synthesis of genetic networks. *Science 296*, 1466−1470.

(31) Isalan, M., Lemerle, C., Michalodimitrakis, K., Horn, C., Beltrao, P., Raineri, E., Garriga-Canut, M., and Serrano, L. (2008) Evolvability and hierarchy in rewired bacterial gene networks. *Nature 452*, 840−845.

(32) Yokobayashi, Y., Weiss, R., and Arnold, F. H. (2002) Directed evolution of a genetic circuit. *Proc. Natl. Acad. Sci. U.S.A. 99*, 16587−16591.

(33) Luria, S. E., and Delbruck, M. (1943) Mutations of bacteria from virus sensitivity to virus resistance. *Genetics 28*, 491−511.

(34) Hall, B. G. (1978) Experimental evolution of a new enzymatic function. II. Evolution of multiple functions for ebg enzyme in E. coli. *Genetics 89*, 453−465.

(35) Gresham, D., Desai, M. M., Tucker, C. M., Jenq, H. T., Pai, D. A., Ward, A., DeSevo, C. G., Botstein, D., and Dunham, M. J. (2008) The repertoire and dynamics of evolutionary adaptations to controlled nutrient-limited environments in yeast. *PLoS Genet. 4*, e1000303.

(36) Barrick, J. E., Yu, D. S., Yoon, S. H., Jeong, H., Oh, T. K., Schneider, D., Lenski, R. E., and Kim, J. F. (2009) Genome evolution and adaptation in a long-term experiment with Escherichia coli. *Nature 461*, 1243−1247.

(37) Isaacs, F. J., Carr, P. A., Wang, H. H., Lajoie, M. J., Sterling, B., Kraal, L., Tolonen, A. C., Gianoulis, T. A., Goodman, D. B., Reppas, N.

B., Emig, C. J., Bang, D., Hwang, S. J., Jewett, M. C., Jacobson, J. M., and Church, G. M. (2011) Precise manipulation of chromosomes in vivo enables genome-wide codon replacement. *Science 333*, 348−353.
(38) Wang, H. H., Isaacs, F. J., Carr, P. A., Sun, Z. Z., Xu, G., Forest, C. R., and Church, G. M. (2009) Programming cells by multiplex genome engineering and accelerated evolution. *Nature 460*, 894−898.
(39) Thomason, L., Court, D. L., Bubunenko, M., Costantino, N., Wilson, H., Datta, S., and Oppenheim, A. (2007) Recombineering: genetic engineering in bacteria using homologous recombination, in *Current Protocols in Molecular Biology* (Ausubel, F. M., et al., Eds.) Chapter 1, Unit 1 16, Wiley, New York.
(40) Sawitzke, J. A., Thomason, L. C., Costantino, N., Bubunenko, M., Datta, S., and Court, D. L. (2007) Recombineering: in vivo genetic engineering in E. coli, S. enterica, and beyond. *Methods Enzymol. 421*, 171−199.
(41) Benders, G. A., Noskov, V. N., Denisova, E. A., Lartigue, C., Gibson, D. G., Assad-Garcia, N., Chuang, R. Y., Carrera, W., Moodie, M., Algire, M. A., Phan, Q., Alperovich, N., Vashee, S., Merryman, C., Venter, J. C., Smith, H. O., Glass, J. I., and Hutchison, C. A. 3rd. (2010) Cloning whole bacterial genomes in yeast. *Nucleic Acids Res. 38*, 2558−2569.
(42) Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C. A. 3rd, and Smith, H. O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods 6*, 343−345.
(43) Gibson, D. G., Benders, G. A., Axelrod, K. C., Zaveri, J., Algire, M. A., Moodie, M., Montague, M. G., Venter, J. C., Smith, H. O., and Hutchison, C. A. 3rd. (2008) One-step assembly in yeast of 25 overlapping DNA fragments to form a complete synthetic Mycoplasma genitalium genome. *Proc. Natl. Acad. Sci. U.S.A. 105*, 20404−20409.
(44) Patnaik, R., Louie, S., Gavrilovic, V., Perry, K., Stemmer, W. P., Ryan, C. M., and del Cardayre, S. (2002) Genome shuffling of Lactobacillus for improved acid tolerance. *Nat. Biotechnol. 20*, 707−712.
(45) Zhang, Y. X., Perry, K., Vinci, V. A., Powell, K., Stemmer, W. P., and del Cardayre, S. B. (2002) Genome shuffling leads to rapid phenotypic improvement in bacteria. *Nature 415*, 644−646.
(46) Gibson, D. G., Glass, J. I., Lartigue, C., Noskov, V. N., Chuang, R. Y., Algire, M. A., Benders, G. A., Montague, M. G., Ma, L., Moodie, M. M., Merryman, C., Vashee, S., Krishnakumar, R., Assad-Garcia, N., Andrews-Pfannkoch, C., Denisova, E. A., Young, L., Qi, Z. Q., Segall-Shapiro, T. H., Calvey, C. H., Parmar, P. P., Hutchison, C. A. 3rd, Smith, H. O., and Venter, J. C. (2010) Creation of a bacterial cell controlled by a chemically synthesized genome. *Science 329*, 52−56.
(47) Lartigue, C., Vashee, S., Algire, M. A., Chuang, R. Y., Benders, G. A., Ma, L., Noskov, V. N., Denisova, E. A., Gibson, D. G., Assad-Garcia, N., Alperovich, N., Thomas, D. W., Merryman, C., Hutchison, C. A. 3rd, Smith, H. O., Venter, J. C., and Glass, J. I. (2009) Creating bacterial strains from genomes that have been cloned and engineered in yeast. *Science 325*, 1693−1696.
(48) Gibson, D. G., Benders, G. A., Andrews-Pfannkoch, C., Denisova, E. A., Baden-Tillson, H., Zaveri, J., Stockwell, T. B., Brownley, A., Thomas, D. W., Algire, M. A., Merryman, C., Young, L., Noskov, V. N., Glass, J. I., Venter, J. C., Hutchison, C. A. 3rd, and Smith, H. O. (2008) Complete chemical synthesis, assembly, and cloning of a Mycoplasma genitalium genome. *Science 319*, 1215−1220.
(49) Dymond, J. S., Richardson, S. M., Coombes, C. E., Babatz, T., Muller, H., Annaluru, N., Blake, W. J., Schwerzmann, J. W., Dai, J., Lindstrom, D. L., Boeke, A. C., Gottschling, D. E., Chandrasegaran, S., Bader, J. S., and Boeke, J. D. (2011) Synthetic chromosome arms function in yeast and generate phenotypic diversity by design. *Nature 477*, 471−476.