Predicting Evolution Using Regulatory Architecture

Nghe, Philippe; de Vos, Marjon G.J.; Kingma, Enzo; Kogenaru, Manjunatha; Poelwijk, Frank J.; Laan, Liedewij; Tans, Sander J.

DOI
10.1146/annurev-biophys-070317-032939

Publication date
2020

Document Version
Final published version

Published in
Annual Review of Biophysics

Annual Review of Biophysics

Predicting Evolution Using Regulatory Architecture

Philippe Nghe,[1,*] Marjon G.J. de Vos,[2,*] Enzo Kingma,[3,*] Manjunatha Kogenaru,[4] Frank J. Poelwijk,[5] Liedewij Laan,[3] and Sander J. Tans[3,6]

[1]Laboratoire de Biochimie, UMR CBI 8231, ESPCI Paris, PSL Research University, 75005 Paris, France

[2]University of Groningen, GELIFES, 9747 AG Groningen, The Netherlands

[3]Bionanoscience Department, Delft University of Technology, 2629HZ Delft, The Netherlands

[4]Department of Life Sciences, Imperial College London, London SW7 2AZ, United Kingdom

[5]cBio Center, Department of Data Sciences, Dana-Farber Cancer Institute, Boston, Massachusetts 02215, USA

[6]AMOLF, 1098 XG Amsterdam, The Netherlands; email: tans@amolf.nl

Annu. Rev. Biophys. 2020. 49:181–97

First published as a Review in Advance on February 4, 2020

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

https://doi.org/10.1146/annurev-biophys-070317-032939

*These authors contributed equally to this article.

Keywords

epistasis, regulation networks, evolutionary constraint, gene regulation, pleiotropy, prediction

Abstract

The limits of evolution have long fascinated biologists. However, the causes of evolutionary constraint have remained elusive due to a poor mechanistic understanding of studied phenotypes. Recently, a range of innovative approaches have leveraged mechanistic information on regulatory networks and cellular biology. These methods combine systems biology models with population and single-cell quantification and with new genetic tools, and they have been applied to a range of complex cellular functions and engineered networks. In this article, we review these developments, which are revealing the mechanistic causes of epistasis at different levels of biological organization—in molecular recognition, within a single regulatory network, and between different networks—providing first indications of predictable features of evolutionary constraint.

Contents

1. INTRODUCTION

Elucidating the range of possibilities and limitations of evolutionary adaptation has been one of the most evocative and complex problems in biology (6, 78). Evolution is sometimes strikingly rapid but can also display long-term stagnation for reasons that often remain obscure. Resolving this conundrum is central to understanding natural and laboratory evolution and important to harnessing evolutionary optimization in protein and cellular engineering applications. At a conceptual level, a wide variety of causes have been invoked to explain why new functions may fail to evolve: Those functions may be impossible biochemically (5) or physically (18); a sufficiently strong selection may be absent (22); or, alternatively, an organism's genetic makeup may hamper evolution (14, 38). Indeed, functional improvements that require multiple genetic changes are difficult to acquire by fixing each mutation one by one. The first empirical studies of such genetic interdependencies were enabled by systematic genetic reconstruction of evolutionary intermediates and laboratory evolution (8, 37, 39, 54, 55, 75). Still, we have only begun to scratch the surface of this multifaceted issue.

It has proven useful to break down the problem of genetic interdependence into a few elementary types of pairwise genetic interactions (14, 38). Briefly, reciprocal sign epistasis refers to cases in which two independent disadvantageous mutations are simultaneously required for an improved phenotype or fitness (56). It is reciprocal because both mutations influence each other's effect, and it is described using the term "sign" because the fitness effects of the mutations switch between negative and positive. It is this type of interaction that constrains adaptive evolution the most, since it implies that both mutations must be fixed simultaneously in a selective sweep—a topic that is discussed further in Section 2. Evolution is less constrained when only one of the two mutations switches its effect between negative and positive, which is referred to as regular sign epistasis (75). In this case, some evolutionary pathways are inaccessible to adaptive evolution, but others remain possible. Finally, one can distinguish between two cases that do not restrict adaptive pathways: magnitude epistasis, where the occurrence of one mutation alters the magnitude of the fitness effect of another mutation, and no epistasis, where mutational effects on fitness are additive (54).

One of the major challenges in current evolutionary research is to go beyond description and toward prediction. In this context, the notion of epistasis is useful because it provides the capacity to classify and quantify evolutionary constraints. However, it only provides a part of the picture, since it does not address the underlying molecular mechanisms. Recently, a series of studies exploited knowledge about the architecture of regulatory networks to begin filling this void. The rationale for these studies is that this type of knowledge provides a mechanistic basis for notions such as constraint and epistasis, which by themselves are mechanism independent. One can determine how the quantifiable network properties, like topology, expression levels of constituent genes, or molecular affinities, affect phenotype and fitness. Moreover, mathematical modeling may

be used to extensively explore the range of possible phenotypes, thus opening up the possibility to predict constraints and epistasis.

This early stage of exploration of constraints in network evolution is characterized by a combination of experimental and theoretical innovations and has focused on elementary questions. For instance, do regulatory trade-offs limit evolution, or can they also accelerate it? Are downstream regulatory elements constrained differently than upstream elements? Does pleiotropy within networks frustrate or facilitate their evolution? In this review, these efforts are organized into three sections, each considering a different level of biological organization, ranging from single molecular interactions to highly interconnected networks.

2. EPISTASIS IN REGULATORY INTERACTIONS

Intermolecular binding is at the heart of all regulatory networks, whether it occurs between membrane-associated effector proteins and kinases to transduce signals (51) or between transcription factors and their DNA binding sites to regulate gene expression (4, 52). In this section, we consider studies of how the physical binding of macromolecules impacts epistasis in a regulatory system, how environmental changes modulate epistasis, and the relationship of these factors to the predictability of epistasis and evolutionary constraint.

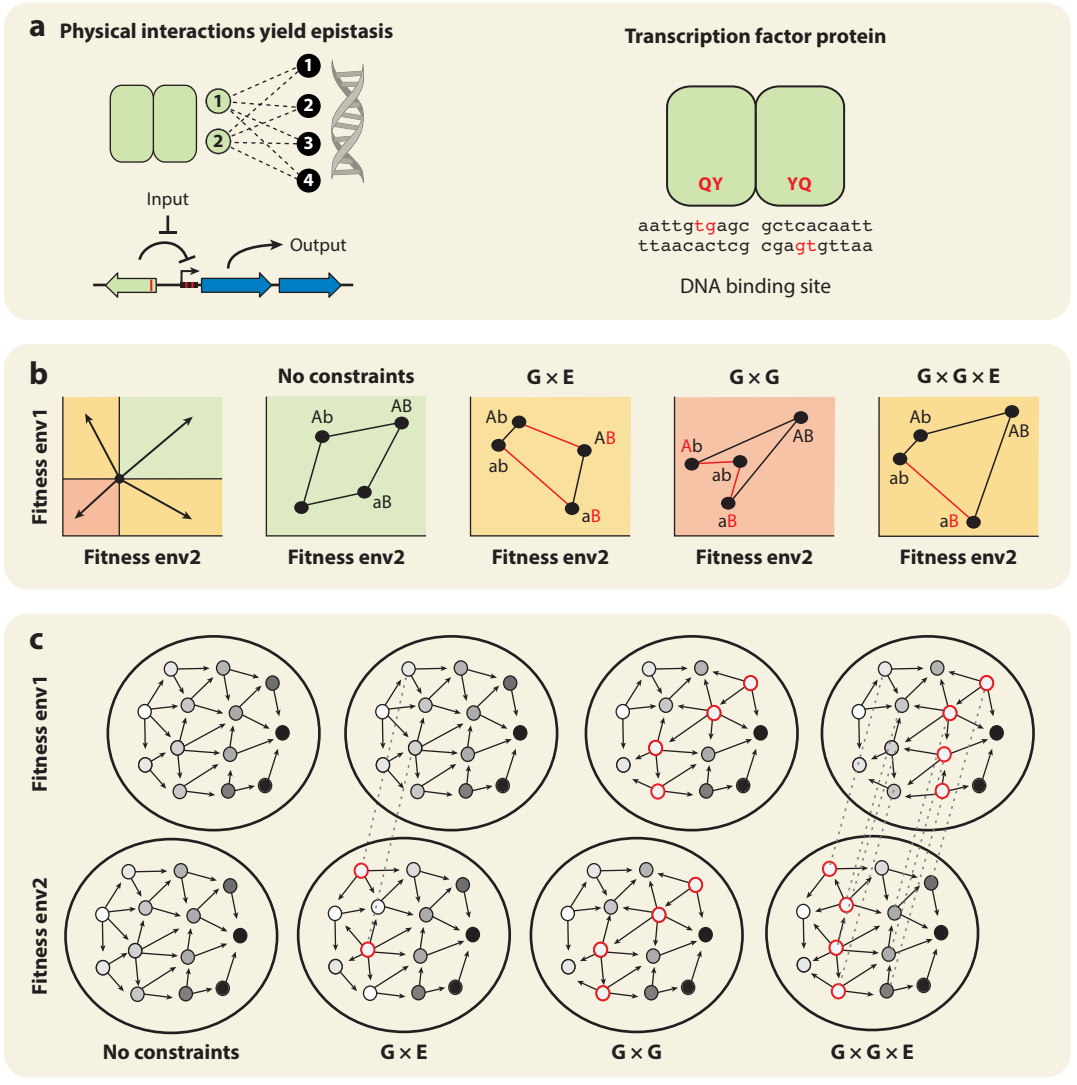
To function, transcription factors must bind their own cognate DNA binding site while avoiding others. It has been proposed that such specific molecular recognition represents an architectural feature that gives rise to epistasis (53). The rationale is that changing such lock–key systems requires modifications in both key (transcription factor) and lock (binding site), as changing only one of them yields nonmatching combinations (**Figure 1*a***). Since scenarios where two (or more) genetic changes occur in the same (selection) period are rare, such an architecture could prohibit mutational trajectories to improved phenotypes under positive selection.

These core ideas are testable experimentally, as has been done using the archetypal model system of transcriptional regulation: the *lac* operon in *Escherichia coli* (15). Owing to decades of mutational and physiological study of this regulatory system, it is known that a few key operator base pairs and a few amino acid residues in the binding interface of the *lac* repressor determine binding specificity (34, 60). Consistent with the predictions, these key residues indeed display reciprocal sign epistasis: Mutations in both the DNA binding site and repressor allowed binding improvements, while mutations in either one alone only led to deterioration. Six key sites in the transcription factor and the binding site were mutated. Of the 720 possible mutational pathways going from one specifically binding transcription factor–operator pair to another, none of the trajectories contained only mutations that improved the phenotype. From an adaptive landscape perspective, these data thus indicated local optima separated by a valley (15).

Regulation allows cells to respond to environmental cues. The *lac* repressor, for instance, allows repression of the *lac* operon in the absence of lactose, and expression in its presence, by inducing a conformational change in the transcription factor that lowers its affinity to the DNA binding site. The *lac* repressor should thus be able not only to bind the operator, but also to efficiently release it in the presence of lactose. Analysis of the 720 possible mutational trajectories of the *lac* repressor–operator combinations in this second environment also showed that none of the mutational trajectories allowed continuous improvements. However, alternating between the two environments did open up adaptive trajectories with constant improvements for each mutation. With a computational method that describes the mutational and environmental transitions as a Markov process, the crossing rate from the initial to the final genotype for all trajectories in the landscape, including detours, could be determined. Interestingly, this rate is found to be maximal when the rate of environmental switches compares with the mutation rate (15).

Cross-environmental trade-offs appeared to be responsible for the adaptive accessibility of adaptive trajectories: Sequences that were suboptimal peaks in one environment were transformed into valleys in the other environment, thus allowing escape from a suboptimum (**Figure 1*b*,*c***). In other cases, inaccessible downward slopes were turned into accessible ascending slopes upon environmental change, allowing adaptive trajectories to surf (45) these slopes with positive selective coefficients. Evolutionary constraints can thus be overcome by environment-dependent ratcheting that allows the crossing of otherwise inaccessible regions in sequence space (15, 68).

This highlights the major role that the environment plays in the accessibility of biological functions during selection by modulating genotype–genotype interactions (16) (**Figure 1*b***). This is important for more than just regulatory systems, as constraints due to mutations in nonregulatory but coding sequences can be affected as well by environmental change. A study that focused on environment-dependent fitness effects constructed the genotype space of five mutations in the



(*Caption appears on following page*)

Figure 1 (*Figure appears on preceding page*)

(*a*) Molecular recognition in cellular regulation. The specific binding of a dimeric transcription factor (*light green*) to a DNA binding site, which allows expression control of a downstream gene (*blue*), can be seen as a lock–key interaction. Mutating both binding partners produces new lock–key combinations, while mutating either yields nonfunctional ones, as shown by systematic mutagenesis (46). (*b*) Schematic diagrams illustrating how environmental change can affect the epistatic interactions between two mutations, a-to-A and b-to-B. Mutations are depicted as vectors and denote the change in fitness in environment 1 (env1) and environment 2 (env2). In the case of no genetic constraints, both mutations improve fitness independent of the environment and the genetic background. In the case of G × E interaction, the environment changes the sign of a mutational effect (b-to-B), independently of the genetic background. The mutation b-to-B lowers fitness in env1, and thus is inaccessible by adaptive evolution, but becomes accessible when changing to env2. In the case of G × G interaction, the mutational effects depend on the genetic background, but not on the environment. In this case, both mutational effects change sign depending on the genetic background (but not the environment), and thus correspond to reciprocal sign epistasis. In the case of G × G × E interaction, both the environment and the genetic background determine the mutational effect. In this case, the effect of b-to-B changes sign depending on the genetic background, but only in env1. (*c*) Schematic of an adaptive landscape in two environments. Nodes (*circles*) indicate genotypes, and arrows indicate mutational steps of increasing fitness. Gray scale depicts fitness, with darker tones indicating higher fitness. Note that this schematic depiction does not include the full genetic multidimensionality of an adaptive landscape, e.g., it does not contain all pairwise genetic interactions, which may form epistatic motifs. Without genetic constraints, all mutations are additive, and thus all trajectories from the lowest fitness genotype (*white*) to the optimum (*black*) are accessible. In the case of G × E interaction, mutational effects deviate from additivity in Env2 (*red circles*, detrimental mutants). Yet the landscape can be crossed in Env1 by alternating between Env1 and Env2 (*dotted lines*). In the case of G × G interaction, the mutations are detrimental in both environments, with correlated effects in both environments blocking direct access to the optimum by the creation of an adaptive valley. The landscape cannot be crossed by single-step mutations under positive selection. In the case of G × G × E interaction, in both environments, mutations deviate from additivity as detrimental mutations block direct access to the optimum by the formation of an adaptive valley in each environment. Yet the location of these mutants is different in both environments, and parts of the landscapes are anticorrelated. Therefore, these landscapes can be crossed by single-step mutations under fluctuating selection in Env1 and Env2 (*dotted lines*).

genome of *E. coli* and measured the phenotype of these genotypes in 1,920 environments (21). The fitness effects of the mutations significantly changed in 203 environments. Moreover, by focusing on the adaptive landscapes involving all interactions among these five mutations in the three environments with the most distinct effects, they observed significant changes in the topography of the adaptive landscape; thus, epistatic interactions also differed in the different environments.

Environment-dependent epistasis can also affect the ability of a population to adapt to an environment that is gradually becoming more challenging. In laboratory evolution experiments, the rate of environmental change (which modulates the selective pressure), as well as the chemical nature of the environment (which determines the genotypes that may confer a benefit), determined whether evolving populations could keep up with the imposed environmental change (23, 37). Environmental circumstances can thus alter the sign of a mutational effect, its epistatic interactions with other genetic changes (**Figure 1*b***), and thus the course of evolution (17).

The level of ruggedness of genotype–phenotype landscapes does not appear to be specific to the *lac* repressor system. A computational analysis based on the in vitro affinity between transcription factors and their binding sites in eukaryotes found that most of these landscapes were relatively rugged (1), i.e., they were neither as rugged as those obtained from randomly shuffled genotypes, nor purely additive. Nonetheless, many of these landscapes were highly navigable, with mutational trajectories in which binding affinity increased at each mutational step. Does this imply a form of evolutionary optimality of the transcription factor–binding site combinations found in nature? Starr et al. (65) found hundreds of alternative transcription factor protein sequences that use diverse binding mechanisms but perform their function at least as well as the transcription factor that has historically evolved. As they noted, this indicates that "the outcome of evolution depends on a serial chain of compounding chance events" (65, p. 409). Thus, they argue that, if evolution had begun from a different ancestral starting point in sequence space, then different genetic and biochemical forms would probably have evolved.

The evolution of regulatory functions can also be constrained by multiple physical interactions. In a regulatory system containing DNA-binding sites for both the RNA polymerase and a transcription factor, both of these proteins compete for binding to an overlapping binding site on the DNA (32). Lagator and colleagues (32) measured the phenotypes of both single and double mutants; based on a thermodynamic model in which the sign and magnitude of the individual mutational effects served as input, they could predict the effects of double mutations. Such models, which take into account key functional parameters, may set the stage for the further prediction of epistasis and, thus, the course of evolution based on the data available for single mutants.

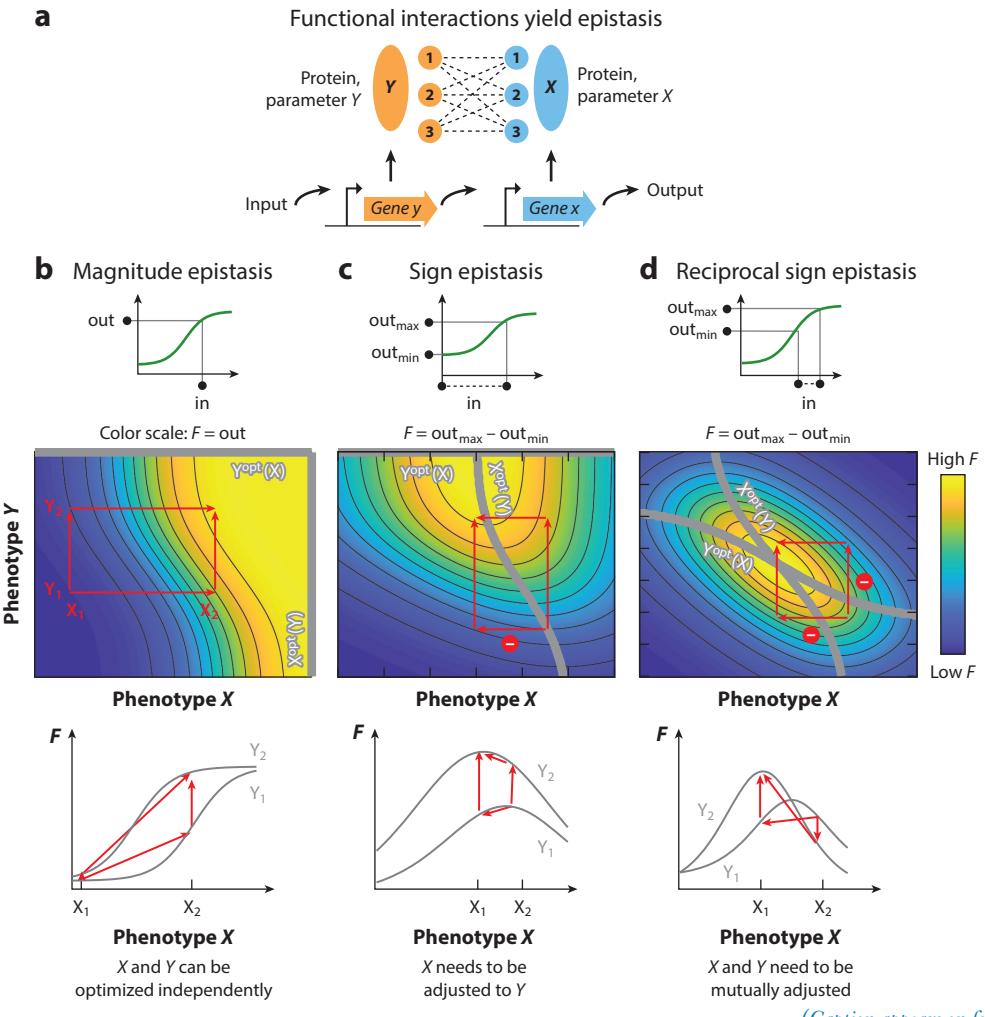
3. EPISTASIS IN REGULATORY PATHWAYS

In this section, we discuss recent studies that have sought to quantify epistasis between genes that act within a regulatory pathway. Unlike the proteins discussed in the previous section, these proteins do not necessarily interact physically, but rather by performing one regulatory function together. We examine the generality and consequences of such functional yet nonphysical interactions and propose interpretations of recent experiments based on phenotype-to-fitness models, also called geometric models. A first interpretation is that network structure can indeed explain how epistasis between two genes arises. However, it is not a direct proxy: One cascade structure can display different types of sign epistasis, depending on other details such as the nature of the variable environment. Second, geometric models can provide a unified framework to interpret both the evolutionary and the more classical phenotypic interpretations of epistasis, such as mutations that can mask the phenotypic effects of other mutations. Finally, one can define general conditions for sign epistasis to arise in any system: It does so when the optimal value of one phenotypic parameter, like the binding constant of a transcription factor, depends on another phenotypic parameter within the network. These findings highlight how functional dependencies within regulatory networks can induce strong constraints in fitness landscapes.

Epistasis was originally used for scenarios in which certain genetic backgrounds masked mutation-induced phenotypic variation (50). Such epistatic interactions between loci across the genome are expected from a purely functional basis, without necessarily implying direct physical interactions between mutated residues, as is, for instance, typically the case for genes within developmental pathways (49). Mechanistic biochemical models indicate that genes in parallel metabolic pathways tend to interact negatively, as the flux catalyzed by a gene can be compensated by flux in a parallel branch, redundantly allowing production of a same final metabolite. However, genes in series within a chain should interact positively, as the removal of any of the catalytic species would strongly reduce the overall metabolic flux (70). This suggests a direct relationship between the network wiring and the observed epistasis (35, 62). However, the sign changes that are key to adaptive fixation were not considered in any of these studies. In parallel, the idea of Fisher geometric models has gained momentum to explain epistasis from biological mechanisms and their function. In this case, generic Gaussian functions describe how fitness depends on a few phenotypic parameters. Although it is heuristic, this assumption can reproduce statistical distributions of epistasis with few parameters (42) and generate a large variety of epistasis distributions (24, 76). This phenotypic view of epistasis and the classical genetic view are starting to be reconciled using geometric models that describe phenotype–fitness relationships in mechanistic terms, using information on the network in question. This approach has so far mainly been applied to small, well-characterized networks (12, 13, 46, 70). While lacking the full cellular context, these studies have revealed general causes of epistasis, and could serve as a basis for more phenomenological long-term evolutionary models (41).

As genes typically represent distinct DNA regions, mutations in one gene logically do not affect the biochemical parameters of another gene, such as its binding constant or enzymatic activity. This genetic modularity has been exploited recently to predict epistasis between genes (45). For instance, in the *lac* operon, a mutation in the transcription factor binding region impacts expression level, and a mutation in the *LacZ* gene affects catalytic rates independently, although they do both participate in the same physiological function. The independence of mutation effects on different genes applies to regulatory cascades, a ubiquitous regulatory motif in cells, where an upstream gene *y* regulates the expression of a downstream gene *x*, which itself regulates an output gene (**Figure 2*a***). Crucially for epistasis, mutational steps that affect the phenotypic parameters X are orthogonal to mutational steps that affect the phenotypic parameters Y within the phenotype space. Thus, either X or Y changes at each mutational step, but not simultaneously.

The resulting epistasis predictions could be verified experimentally by systematically combining mutations within the different transcription factors that together form a regulatory cascade and quantifying their input–output relationships, as in Reference 46. Based on that study, we



(*Caption appears on following page*)

Figure 2 (*Figure appears on preceding page*)

Predicted sign epistasis in a regulatory cascade. (*a*) A transcriptional cascade, where an input signal, such as an inducer, modulates the expression level of the upstream gene *y*, which itself expresses a transcription factor that regulates a downstream gene *x*, which in turn regulates the level of an output gene. Systematic combinations of mutations in genes *x* and *y* allow testing for the effect of combined changes in the phenotypic parameters respectively called X and Y, as has been explored in Reference 46. For example, X may be the binding constant of regulatory protein x to the promoter of the output gene. (*b*) Scenario yielding magnitude epistasis. (*Top*) The performance or fitness of the system is measured as the output level in response to a single input that is fixed in time, as it would be in a constant environment. (*Middle*) The corresponding phenotype-to-fitness relationship is computed as a function of parameters X and Y, for example, the binding constants of the transcription factors x and y to their target promoters, using a mathematical model (46). Red arrows represent the two ways by which mutations leading from X_1 to X_2 and from Y_1 to Y_2 can be combined to optimize fitness. Importantly, given the independent effects of mutations on X and Y, trajectories are parallel to the axis. Given that, in this case, fitness increases with both X and Y, the maximum values of X as a function of Y and of Y as a function of X, respectively denoted X^{opt} and Y^{opt}, are the straight thick gray lines on the right-hand side and top of the landscape, respectively. (*Bottom*) Given that X^{opt} and Y^{opt} are independent of, respectively, Y and X, the geometric model can only generate magnitude epistasis. (*c*) Scenario yielding regular sign epistasis. (*Top*) Fitness is the total output range $F = \text{out}_{\text{max}} - \text{out}_{\text{min}}$ in response to a variable environment providing wide input variations, with the input going down to zero (signal is absent). (*Middle*) Fitness is optimal for intermediate values X^{opt} of X for fixed Y, and X^{opt} is the thick gray curved line that varies as a function of Y. In the example mutational trajectories (*red arrows*), one path leads to a decreasing step when mutating X first (*circled minus sign*). (*Bottom*) The starting phenotype X_1 is optimal given Y_1; thus, mutating it can only lead to decreased fitness, causing the sign epistasis. (*d*) Scenario yielding reciprocal sign epistasis. (*Top*) The output is evaluated in response to a more restricted range of input signals (the minimum input does not reach zero). (*Middle*) Both X^{opt} and Y^{opt} are curved and can lead to decreasing fitness when mutating X or Y first. (*Bottom*) The mutual dependence of the optima of X and Y on each other's values can generate reciprocal sign epistasis patterns.

discuss how functional relationships between genes produce epistasis and how they can be explained by the shape of the phenotype-fitness functions. We distinguish three major classes of fitness functions.

First, consider a fitness that varies monotonically with phenotypic parameters X and Y, such as the binding affinity of the transcription factor for its operator (**Figure 2*a***). This case arises, for example, when the fitness (or performance) would correspond directly to the output for any fixed input that does not vary in time. We recover the classical notion of phenotypic epistasis, as illustrated in **Figure 2*b***: A mutation with a strong effect on X (in the most extreme case a knock-out) cancels any observable variation of the output that could be caused by changes in Y. Additionally, it is not possible to generate sign epistasis in this scenario. As discussed further below, this is generally the case when the optimal value for X does not depend on Y and vice versa.

A second, qualitatively distinct, scenario is when the optimum of Y does depend on X, while the optimal value of X does not depend on Y (**Figure 2*c***). This situation is found when the fitness corresponds to the dynamic range of the cascade, as quantified by the difference between the minimal and maximal output expression levels, in response to widely varying input signals, as has been shown in Reference 46. The landscape of this fitness function can generate sign epitasis, with the specific property that only mutations in gene *x* can lead to decreasing fitness. This is most easily seen when Y is at its optimum but X is not: Reaching a better X–Y combination cannot be achieved by mutating Y first, as it is already at its maximum in the current X background. However, fitness can increase in a stepwise manner, as X is not yet at its maximum and thus can be improved. In this example, the evolutionary hierarchy reflects the functional hierarchy: The upstream gene *x* must be mutated first because the optimum of the downstream gene *y* must be well tuned to accommodate X expression but not the converse.

The last case is when the optimum of Y depends on X and the optimum of X depends on Y (**Figure 2*d***). This scenario is observed, for example, when the fitness again corresponds to the dynamic range of the output but in response to input signals with variations over a smaller range. By the same geometric reasoning as above, there are starting phenotypes leading to a decrease in fitness when mutating *x* first or when mutating *y* first. This "or" relationship can become an "and"

as exemplified in the mutational trajectory of **Figure 2*d***, where reaching a better combination X–Y in a stepwise manner requires first decreasing the performance of the cascade, independently of the mutated phenotype. This scenario of reciprocal sign epistasis can be understood as a purely functional version of lock–key constraints. It should be noted that geometric models exist where both *x*-related and *y*-related sign epistasis exist, but they do not combine into reciprocal sign epistasis. For example, as shown in Reference 46, slightly tilted Gaussian geometric models do not combine mere sign epistasis into reciprocal sign epistasis, but strongly tilted Gaussian models do.

Hierarchy within a cascade can be applied to interpretation of epistasis in a negative feedback loop in the galactose regulatory system in yeast (48): Mutants of the downstream *GAL80* gene mask mutations in the upstream *GAL3* gene. This corresponds to the geometric model in **Figure 2*b***, where the output corresponds to the intensity of the feedback, and mutations cause strong knock-down effects. Epistatic effects from regulatory structure also appear when integrating signals at the same promoter, as studied by mutating a *lambda* phage promoter repressed by the protein CI (33). In this case, loss of function in polymerase recruitment masks loss of function in CI repression, but CI binding allows tuning expression in the presence of mutated but functional polymerase recruitment, corresponding again to the scheme in **Figure 2*b***. The main point to note is that epistasis is predominantly explained by the regulatory logic, as opposed to the pleiotropy caused by physical interactions (33), indicating a modularity in mutational effects, even within a single regulatory sequence.

Epistasis caused by the functional dependence between the components of a network is also observed in incoherent feedforward network motifs (61), which integrate an environmental signal via two regulatory loops into a single output promoter that defines the phenotype. In this case, the mutational effects were measured in regulatory loops comprising either a double activator or a double repressor cascade. When mutations were introduced into the *cis*-regulatory regions of the networks at each node separately, not all of the possible phenotypic states were accessible. This suggests that the optimal phenotypic parameters of one gene depend on the phenotypic parameters of other genes. However, when mutations were introduced in the *cis*-regulatory regions of the individual nodes simultaneously, this barrier could be overcome through epistasis between the *cis*-regulatory region mutations. Overall, these results support the idea that epistasis originates from the tuning of genetically independent phenotypic parameters with respect to each other.

Overall, we have seen how to use functional dependences to identify causes of sign epistasis: Sign epistasis arises when the optimum of a module needs to be adjusted to the state of another module. Generalization to a large number of phenotypic dimensions is possible (46) and suggests that phenotypes whose phenotypic optima are mutually dependent (independent) generate ruggedness (smoothness) in genotype-to-fitness landscapes. The interplay between phenotypes and their optimality is also crucial during network rewiring, where the accessibility of evolutionary paths requires finely tuned steps to preserve function (64). So far, most mutational scanning studies can be interpreted in large part through loss of function within networks. Revealing sign epistasis systematically will instead require the combination of mutations with mild phenotypic effects in the relative proximity of their optima (24).

4. EPISTASIS BETWEEN NETWORKS

In the previous sections, we discuss how epistasis constrains the evolution of regulatory interactions and networks. A focus on the physical (19, 66) or functional (7, 79) features of these regulatory systems allowed prediction of epistasis within a network of interest. Another situation arises when selection drives interactions between networks. An interesting example is a recent experimental

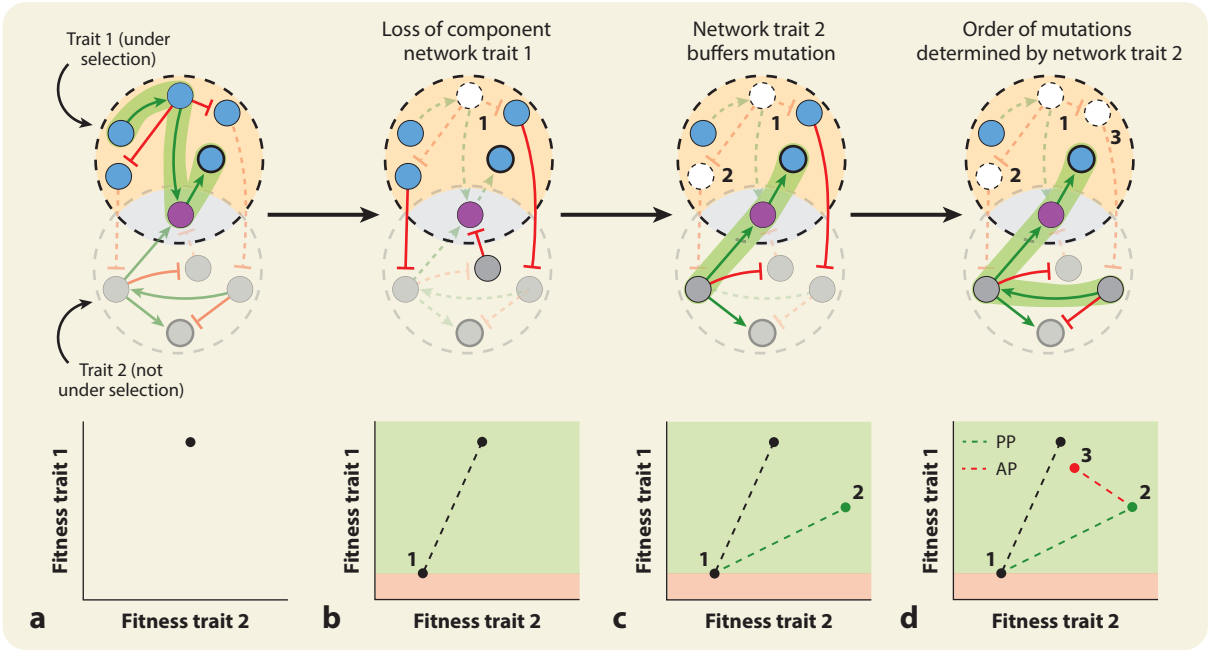
evolution study in *Saccharomyces cerevisiae* (31). In *S. cerevisiae*, formation of a polarized spot of the GTPase Cdc42, a protein that cycles between an active GTP-bound state and an inactive GDP-bound state, is an essential part of the cell cycle. As polarization serves as a paradigm for symmetry breaking, a large amount of experimental and theoretical work has been dedicated to identifying the major components of this network and their interactions (25, 43). The knowledge that this work has generated makes this network an attractive system to study adaptive pathways. Laan et al. (31) applied a strong perturbation to this module by deleting a scaffold protein with a central role in symmetry breaking. The subsequent adaptation to the loss of this component was restricted by sign epistasis, which constrained the order in which other network components were mutated, leading to reproducible mutational pathways across multiple parallel lineages. Surprisingly, this mutational pathway consisted of the inactivation of proteins rather than changes to their biochemical properties, which was confirmed by reconstructing the mutational pathway observed from natural evolution by synthetically deleting the genes. Another puzzling finding was that the epistatic interactions that determined the adaptive pathway were not exclusively among known network components, but also included a protein that was not considered to be part of the polarization network. Thus, despite its well-studied nature, information about the interaction network was insufficient to explain the relevant epistasis. This example highlights a generic challenge when the system increases in size. Many selective pressures, particularly in response to strong perturbations, elicit mutations throughout the genome, which limits predictive approaches.

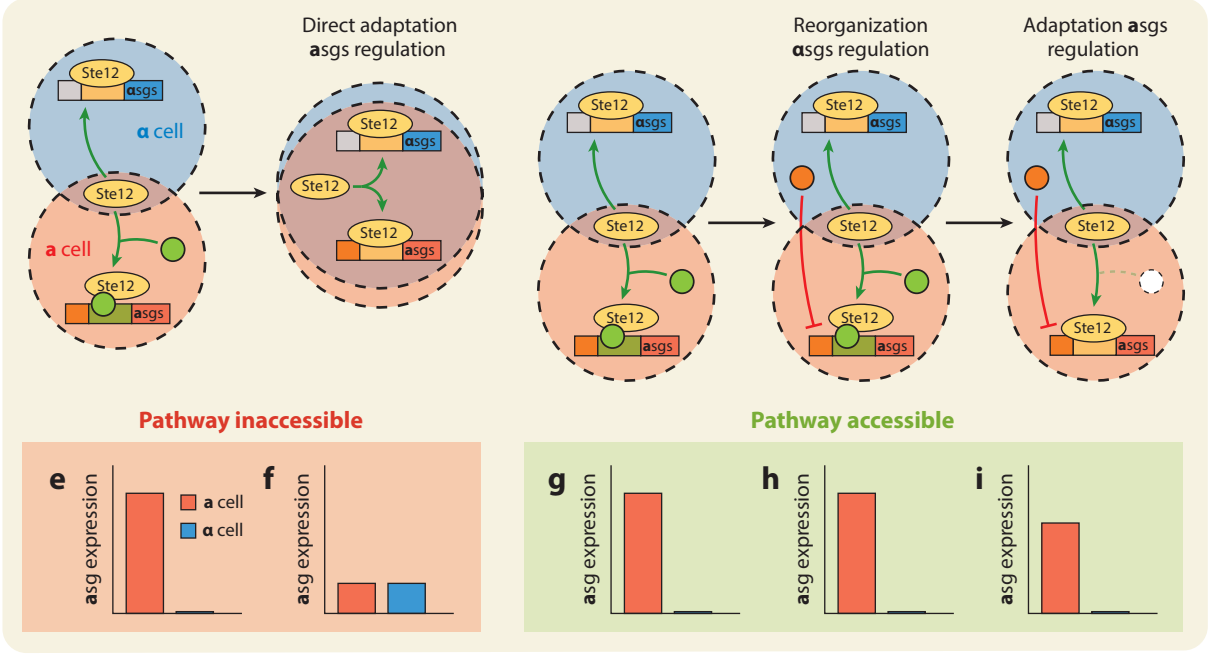
The above discussion indicates that a trait can be restored by co-opting networks with some functional overlap, rather than by reconstructing lost components. How the particular gene deletions that facilitate this process in the polarity network do so remains elusive, but it is known that networks involved in other cellular processes can affect polarity establishment (63, 77). This mode of adaptation appears to not be specific to the example described above: A large-scale study that tracked the adaptive response of *S. cerevisiae* to 187 different single-gene knock-outs showed that compensatory mutations following gene deletions often partially restore affected traits without restoring the original genomic expression pattern (69). Thus, evolution can exploit other networks to rescue perturbed cellular functions (27, 59) by changing the connections between redundant and connected networks, rather than by restoring the original network (**Figure 3*a–d***).

Accessible pathways thus do not always depend only on the topology of the perturbed network, but also on that of compensating networks and on their interconnections. When insight into the relevant network features is incomplete, it is not straightforward to functionally explain the observed epistasis, let alone offer predictions. Another consequence of networks being interconnected is that they typically give rise to pleiotropy: Not only is one trait affected by multiple networks or mutations, but one network or mutation also affects multiple traits. Such pleiotropic properties have long been thought to have important consequences for evolving populations: For example, they are considered to maintain genetic variation (80), produce trade-offs (11, 40) and be an important cause of the persistence of genetic diseases (2, 10).

The questions that naturally arise are: How prevalent is pleiotropy in cellular networks? Which adaptive processes cannot be considered without the context of the entire cellular interaction network? The view of universal pleiotropy, where a mutation can potentially affect all selectable traits, is implicit in the original geometric model from Fisher (20) and has been the dominant view for many years (67). Within this view, one may consider whether the number of evolutionary constraints increases as the number of traits increases, seemingly contradicting the emergence and adaptability of complex organisms (47). Due to practical challenges in empirically determining the degree of pleiotropy in organisms, the discussion remained mainly based on theoretical models. However, new molecular biology techniques and interaction network databases have sparked attempts to empirically quantify the number of pleiotropic genes in actual biological systems. For

example, Wang and colleagues (74) used large data sets containing effects of gene deletions on different traits to determine the total pleiotropy on a genome-wide scale for *S. cerevisiae*, *Caenorhabditis elegans* and *Mus musculus*. Comparison to random gene–trait relationships suggested that cellular networks are mostly modular, with only a small percentage of traits (1–9% of the traits included for





(*Caption appears on following page*)

Figure 3 (*Figure appears on preceding page*)

Pleiotropic interactions can facilitate and constrain evolution. (*a–d*). When the networks of two traits share interactions, but only one network is under selection, the network not under selection can compensate for mutations in the other network. (*a*) The initial pathway (*green highlight*) leads to activation of a trait-defining node (*encircled blue node*) through a shared component (*purple node*). (*b*) Deletion of a central component in the network of trait 1 decreases fitness of the trait. (*c*) The network of trait 2 can buffer this deletion by taking over some of the interactions of network 1 through the deletion of other components, (*d*) with the order in which these deletions can take place being constrained by the interactions between the network of trait 1 and the network of trait 2, which can provide a reconstruction of the activation pathway. Only pathways that improve trait 1 are accessible (*green shaded area*); they can have both positive pleiotropic (PP, from mutation 1 to 2) or antagonistic pleiotropic (AP, from mutation 2 to 3) effects on trait 2. (*e–i*) When both traits are under selection, connected networks can lead to additional constraints. (*e*) The two transcription networks regulating mating pathways in yeast both use the transcription factor Ste12, causing them to partially overlap. (*f*) This overlap prevents the evolution of Ste12 binding motifs in **a** cells, as this causes misexpression of **a**-specific genes (**a**sgs) in **α** cells. (*g–i*) To maintain correct regulation of **a**sgs during evolution, the transcription network of **α** cells must first reorganize to obtain an inhibitor for **a**sg expression, which releases the constraint on the evolution of Ste12 binding motifs in **a** cells.

analysis) affected by pleiotropic genes. This result would advocate the view of modular pleiotropy, where sets of traits covary, and only a subset of genes within a network exhibit pleiotropic effects.

However, a common objection is that pleiotropic interactions can remain undetected due to experimental noise and detection limits. The absence of a standardized methodology for extracting statistically significant interactions from experimental data has led to significant variations in the estimates of genes with pleiotropic interactions, with a possible bias for modular pleiotropy (26). The development of new, more sensitive methods for the extraction of significant interactions from databases is required to settle this debate (30, 71).

Although the discussion on the extent of pleiotropy is far from resolved, the notion that pleiotropic interactions exist and impact evolutionary trajectories is widely accepted (73). As discussed above, pleiotropy can lead to epistatic interactions between components of different networks, and both epistasis and pleiotropy make evolution dependent on the genetic background. To understand how pleiotropy affects evolution, it is useful to focus on this interplay between pleiotropic and epistatic interactions. Epistasis is broadly believed to guide evolution by imposing constraints, and from a network perspective, antagonistic pleiotropy (AP) is considered to play a central role in constraining the evolution of networks. The negative correlation between different traits in AP can make it difficult to simultaneously optimize multiple traits and can thus restrict the number of accessible mutational pathways (**Figure 3*d***). The extent of AP in yeast was examined by Qian et al. (58). They tested 4,642 nonessential genes for antagonistic pleiotropy by performing competition experiments of null mutants together with the wild type in different environmental conditions. At least 13.6% of the analyzed genes displayed AP in the considered environments, indicating its importance in restricting mutational pathways. Interestingly, they found signs that antagonistic interactions between networks could be mitigated by changes in *trans*-regulatory molecules that regulate gene expression, rather than in DNA regulatory or coding sequences. The ability of *trans*-regulatory molecules to serve as a source for alleviating constraints is surprising, considering that their evolution is typically regarded to be heavily constrained itself due to their extensive interaction networks (57, 72).

How *trans*-acting elements can resolve AP and alleviate constraints is illustrated by a study of the pheromone response pathways between different yeast species (64). Haploid yeast cells can exist in two different mating types, **a** and **α**. In response to sensing pheromones of the opposite mating type, each of these upregulates the expression of genes required for the pheromone response pathway. Although the upregulation of some of these genes is mating type specific (expression in **a** or **α**), a large portion overlaps in both mating types (expression in **a** and **α**). In both mating types, the upregulation of these genes is induced by the conserved transcription factor Ste12. Despite the conserved function of Ste12, Sorells and colleagues (64) found that different species

maintained different network structures for the upregulation of mating type–specific genes: In *Saccharomyces*, mating type **a**–specific genes (**a**sgs) contain motifs for the direct binding of Ste12, while *Kluyveromyces* and *Candida* required transcriptional coregulators for induction of the same genes. However, evolution of a regulatory network for **a**sgs where Ste12 is recruited by coregulators to one where Ste12 directly binds **a**sg promoters appeared to be inaccessible: Introduction of the Ste12 binding sites of the *Saccharomyces* clade into the *Kluyveromyces* clade resulted in a loss of regulation (**Figure 3*e*,*f***). Instead, the addition of a repressor for **a**sg expression to the regulatory network of mating type **α** cells was required prior to introducing Ste12 binding sites to prevent their (mis)expression in **α** cells (**Figure 3*g–i***).

Note that understanding why the pheromone response pathway in **a** cells cannot directly evolve additional binding sites for **a**sgs requires knowledge of the structure of the pheromone response pathway in **α** cells. Otherwise, the requirement for evolving a repressor for **a**sgs would appear as a hidden parameter in the epistatic landscape, similar to those that appear as hidden parameters for the polarization network of *S. cerevisiae* described above. The way in which interactions between these pheromone response networks constrain evolution shows surprising similarities to sign epistasis between components within a single module: The adaptation in one network becomes beneficial only when the structure of the other network changes (**Figure 3*e–i***). It is tempting to consider the analogy between AP and sign epistasis, but how concepts from epistasis relate to pleiotropy in networks remains to be investigated. Apart from AP, which constrains evolution, interactions that result in a positive correlation in the fitness level of different traits have been found to drive coevolution of traits that are not under selection. For example, Desai and colleagues (28) found that populations of budding yeast adapting to growth at high temperature also improved growth at standard growth temperatures, although the molecular basis of this was not elaborated. This shows how pleiotropic networks and the environment can interact to give surprising evolutionary results. Unraveling these interactions at the molecular level, similar to what has been done for epistatic interactions of the *lac* operon (15), can reveal new concepts that would explain adaptive pathways at both the inter- and the intranetwork levels.

5. CONCLUSIONS

It is evident that evolutionary constraints are inherently interconnected with phenotypes. Indeed, the genetic interactions that underlie constraint are quantified by their impact on phenotypes and fitness (14, 38). As highlighted in this review, epistasis offers a route to predict evolutionary constraint and potential—one of the major goals of evolution research. At the same time, owing to the overwhelming complexity of this phenotypic puzzle, and its many missing pieces, such prediction insights have been difficult to achieve. This humbling reality remains to a large extent, as is also clear from the studies reviewed above. At all levels of biological organization that we address, one encounters unknowns that pose limits to general predictive frameworks. For instance, it is not known if the presence of adaptive valleys is general for molecular interactions other than the well-studied *lac* system, how the geometric landscape prediction method can be applied to naturally occurring pathways, and which redundant networks will be able to compensate for the loss of core cellular functions.

Nonetheless, this new wave of quantitative studies provided the first tools to predict key epistasis features. Within the wide range of studied systems, notable parallels and differences were observed. Specific recognition was found to produce reciprocal sign epistasis, but its limiting effects on adaptation could be mitigated by environmental interactions. The last section showed that interactions between networks can overcome constraint by providing components that are coopted into altered functions. In both cases, additional interactions could alleviate existing constraints,

even as the type of interaction differed. This makes intuitive sense, since those extra interactions (with the environment and with other networks) can be seen as dimensions that are orthogonal to the initial genotypic space, and thus can allow escape from suboptima. Antagonistic pleiotropy played a role in both processes, with mutations having contrasting effects on different phenotypes and networks.

The level of prediction was found to depend on the scale of the system, as well as on the selective pressures acting on it. At the smallest scale—that of molecular binding—epistasis was merely implied by the specific nature of the recognition and ultimately depended on the details of the molecular binding interface; the resulting genotype–phenotype map required experimental reconstruction (15). Notably, such molecular details no longer appeared relevant at the intermediate scale of a single pathway, where epistasis instead emerged in a more predictable fashion from the network topology (32, 33, 46). The key changes in binding affinity could be achieved by a wide spectrum of mutations, suggesting that a coarse-grained description suffices when studying evolution at the network scale. This is a promising realization: As larger systems are considered (59), the basis of constraint does not necessarily become less predictable. However, for the largest scale considered in this review—that of networks interacting in core cellular functions—it was shown that predictions are more challenging because such networks are more interconnected with unknown factors (31, 69).

Approaches similar to the ones discussed in this review, which exploit any functional information for the prediction of constraint, can be applied more broadly. For instance, it is intriguing to consider epistatic constraints in RNA molecules (36) and metabolic networks (3), as well as in the evolution of complex traits (9), ecosystems (29), and disease (44).

DISCLOSURE STATEMENT

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

ACKNOWLEDGMENTS

This work was supported by the Netherlands Organization for Scientific Research (NWO). L.L. gratefully acknowledges funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant 758132).

LITERATURE CITED

1. Aguilar-Rodriguez J, Payne JL, Wagner A. 2017. A thousand empirical adaptive landscapes and their navigability. *Nat. Ecol. Evol.* 1(2):45
2. Amell A, Roso-Llorach A, Palomero L, Cuadras D, Galván-Femenía I, et al. 2018. Disease networks identify specific conditions and pleiotropy influencing multimorbidity in the general population. *Sci. Rep.* 8:15970
3. Bachmann H, Molenaar D, Branco Dos Santos F, Teusink B. 2017. Experimental evolution and the adjustment of metabolic strategies in lactic acid bacteria. *FEMS Microbiol. Rev.* 41(Supp_1):S201–19
4. Balleza E, López-Bojorguez LN, Martínez-Antonio A, Resendis-Antonio O, Lozada-Chávez I, et al. 2009. Regulation by transcription factors in bacteria: beyond description. *FEMS Microbiol. Rev.* 33(1):133–51
5. Beardmore RE, Gudelj I, Lipson DA, Hurst LD. 2011. Metabolic trade-offs and the maintenance of the fittest and the flattest. *Nature* 472(7343):342–46
6. Bell MA, Futuyma DJ, Eanes WF, Levinton JS. 2010. *Evolution since Darwin: The First 150 Years*. Sunderland, MA: Sinauer Assoc.

7. Boucher B, Jenna S. 2013. Genetic interaction networks: better understand to better predict. *Front. Genet.* 4:290
8. Bridgham JT, Carroll SM, Thornton JW. 2006. Evolution of hormone-receptor complexity by molecular exploitation. *Science* 312(5770):97–101
9. Carlborg O, Haley CS. 2004. Epistasis: too often neglected in complex trait studies? *Nat. Rev. Genet.* 5(8):618–25
10. Carter AJ, Nguyen AQ. 2011. Antagonistic pleiotropy as a widespread mechanism for the maintenance of polymorphic disease alleles. *BMC Med. Genet.* 12:160
11. Cooper VS, Lenski RE. 2000. The population genetics of ecological specialization in evolving *Escherichia coli* populations. *Nature* 407(6805):736–39
12. Corson F, Siggia ED. 2012. Geometry, epistasis, and developmental patterning. *PNAS* 109(15):5568–75
13. Cotterell J, Sharpe J. 2013. Mechanistic explanations for restricted evolutionary paths that emerge from gene regulatory networks. *PLOS ONE* 8(4):e61178
14. de Visser JA, Krug J. 2014. Empirical fitness landscapes and the predictability of evolution. *Nat. Rev. Genet.* 15(7):480–90
15. de Vos MGJ, Dawid A, Sunderlikova V, Tans SJ. 2015. Breaking evolutionary constraint with a tradeoff ratchet. *PNAS* 112(48):14906–11
16. de Vos MGJ, Poelwijk FJ, Battich N, Ndika JDT, Tans SJ. 2013. Environmental dependence of genetic constraint. *PLOs Genet.* 9(6):e1003580
17. de Vos MGJ, Schoustra SE, de Visser JAGM. 2018. Ecology dictates evolution? About the importance of genetic and ecological constraints in adaptation. *EPL* 122(5):58002
18. Dill KA, Ghosh K, Schmit JD. 2011. Physical limits of cells and proteomes. *PNAS* 108(44):17876–82
19. Diss G, Lehner B. 2018. The genetic landscape of a physical interaction. *eLife* 7:e32472
20. Fisher RAS. 1930. *The Genetical Theory of Natural Selection*. Oxford, UK: Clarendon Press
21. Flynn KM, Cooper TF, Moore FB-G, Cooper VS. 2013. The environment affects epistatic interactions to alter the topology of an empirical fitness landscape. *PLOS Genet.* 9(4):e1003426
22. Gavrilets S, Hastings A. 1993. Maintenance of genetic variability under strong stabilizing selection: a two-locus model. *Genetics* 134(1):377–86
23. Gorter FA, Aarts MGM, Zwaan BJ, de Visser JAGM. 2018. Local fitness landscapes predict yeast evolutionary dynamics in directionally changing environments. *Genetics* 208(1):307–22
24. Gros PA, Le Nagard H, Tenaillon O. 2009. The evolution of epistasis and its links with genetic robustness, complexity and drift in a phenotypic model of adaptation. *Genetics* 182(1):277–93
25. Halatek J, Brauns F, Frey E. 2018. Self-organization principles of intracellular pattern formation. *Philos. Trans. R. Soc. Lond. B* 373(1747):20170107
26. Hill WG, Zhang XS. 2012. Assessing pleiotropy and its evolutionary consequences: Pleiotropy is not necessarily limited, nor need it hinder the evolution of complexity. *Nat. Rev. Genet.* 13(4):296. Author reply. 2012. *Nat. Rev. Genet.* 13(4):296
27. Hottes AK, Freddolino PL, Khare A, Donnell ZN, Liu JC, Tavazoie S. 2013. Bacterial adaptation through loss of function. *PLOS Genet.* 9(7):e1003617
28. Jerison ER, Kryazhimskiy S, Mitchell JK, Bloom JS, Kruglyak L, Desai MM. 2017. Genetic variation in adaptability and pleiotropy in budding yeast. *eLife* 6:e27167
29. Kassen R, Rainey PB. 2004. The ecology and genetics of microbial diversity. *Annu. Rev. Microbiol.* 58:207–31
30. Koch EN, Costanzo M, Deshpande R, Andrews B, Boone C, Myers CL. 2017. Systematic identification of pleiotropic genes from genetic interactions. bioRxiv 112326. **https://doi.org/10.1101/112326**
31. Laan L, Koschwanez JH, Murray AW. 2015. Evolutionary adaptation after crippling cell polarization follows reproducible trajectories. *eLife* 4:e09638
32. Lagator M, Paixão T, Barton NH, Bollback JP, Guet CC. 2017. On the mechanistic nature of epistasis in a canonical *cis*-regulatory element. *eLife* 6:e25192
33. Lagator M, Sarikas S, Acar H, Bollback JP, Guet CC. 2017. Regulatory network structure determines patterns of intermolecular epistasis. *eLife* 6:e28921

34. Lehming N, Sartorius J, Kisters-Woike B, von Wilcken-Bergmann B, Müller-Hill B. 1990. Mutant *lac* repressors with new specificities hint at rules for protein–DNA recognition. *EMBO J.* 9(3):615–21
35. Lehner B. 2011. Molecular mechanisms of epistasis within and between genes. *Trends Genet.* 27(8):323–31
36. Li C, Zhang J. 2018. Multi-environment fitness landscapes of a tRNA gene. *Nat. Ecol. Evol.* 2(6):1025–32
37. Lindsey HA, Gallie J, Taylor S, Kerr B. 2013. Evolutionary rescue from extinction is contingent on a lower rate of environmental change. *Nature* 494(7438):463–67
38. Lobkovsky AE, Koonin EV. 2012. Replaying the tape of life: quantification of the predictability of evolution. *Front. Genet.* 3:246
39. Lunzer M, Miller SP, Felsheim R, Dean AM. 2005. The biochemical architecture of an ancient adaptive landscape. *Science* 310(5747):499–501
40. MacLean RC, Bell G, Rainey PB. 2004. The evolution of a pleiotropic fitness tradeoff in *Pseudomonas fluorescens*. *PNAS* 101(21):8072–77
41. Martin G. 2014. Fisher's geometrical model emerges as a property of complex integrated phenotypic networks. *Genetics* 197(1):237–55
42. Martin G, Elena SF, Lenormand T. 2007. Distributions of epistasis in microbes fit predictions from a fitness landscape model. *Nat. Genet.* 39(4):555–60
43. Martin SG. 2015. Spontaneous cell polarization: Feedback control of Cdc42 GTPase breaks cellular symmetry. *Bioessays* 37(11):1193–201
44. Moore JH. 2003. The ubiquitous nature of epistasis in determining susceptibility to common human diseases. *Hum. Hered.* 56(1–3):73–82
45. Mustonen V, Lassig M. 2009. From fitness landscapes to seascapes: non-equilibrium dynamics of selection and adaptation. *Trends Genet.* 25(3):111–19
46. Nghe P, Kogenaru M, Tans SJ. 2018. Sign epistasis caused by hierarchy within signalling cascades. *Nat. Commun.* 9:1451
47. Orr HA. 2000. Adaptation and the cost of complexity. *Evolution* 54(1):13–20
48. Peng W, Liu P, Xue Y, Acar M. 2015. Evolution of gene network activity by tuning the strength of negative-feedback regulation. *Nat. Commun.* 6(1):6226
49. Peter IS, Davidson EH. 2011. Evolution of gene regulatory networks controlling body plan development. *Cell* 144(6):970–85
50. Phillips PC. 2008. Epistasis: the essential role of gene interactions in the structure and evolution of genetic systems. *Nat. Rev. Genet.* 9(11):855–67
51. Podgornaia AI, Laub MT. 2015. Protein evolution: pervasive degeneracy and epistasis in a protein-protein interface. *Science* 347(6222):673–77
52. Poelwijk FJ, de Vos MGJ, Tans SJ. 2011. Tradeoffs and optimality in the evolution of gene regulation. *Cell* 146(3):462–70
53. Poelwijk FJ, Heyning PD, de Vos MGJ, Kiviet DJ, Tans SJ. 2011. Optimality and evolution of transcriptionally regulated gene expression. *BMC Syst. Biol.* 5:128
54. Poelwijk FJ, Kiviet DJ, Weinriech DM, Tans SJ. 2007. Empirical fitness landscapes reveal accessible evolutionary paths. *Nature* 445(7126):383–86
55. Poelwijk FJ, Socolich M, Ranganathan R. 2019. Learning the pattern of epistasis linking genotype and phenotype in a protein. *Nat. Commun.* 10:4213
56. Poelwijk FJ, Tañase-Nicola S, Kiviet DJ, Tans SJ. 2011. Reciprocal sign epistasis is a necessary condition for multi-peaked fitness landscapes. *J. Theor. Biol.* 272(1):141–44
57. Prud'homme B, Gompel N, Carroll SB. 2007. Emerging principles of regulatory evolution. *PNAS* 104(Suppl. 1):8605–12
58. Qian W, Ma D, Xiao C, Wang Z, Zhang J. 2012. The genomic landscape and evolutionary resolution of antagonistic pleiotropy in yeast. *Cell Rep.* 2(5):1399–410
59. Rojas Echenique JI, Kryazhimskiy S, Nguyen Ba AN, Desai MM. 2019. Modular epistasis and the compensatory evolution of gene deletion mutants. *PLOS Genet.* 15(2):e1007958
60. Sartorius J, Lehming N, Kisters B, von Wilcken-Bergmann B, Müller-Hill B. 1989. *lac* repressor mutants with double or triple exchanges in the recognition helix bind specifically to *lac* operator variants with multiple exchanges. *EMBO J.* 8(4):1265–70

61. Schaerli Y, Jiménez A, Duarte JM, Mihajlovic L, Renggli J, et al. 2018. Synthetic circuits reveal how mechanisms of gene regulatory networks constrain evolution. *Mol. Syst. Biol.* 14(9):e8102
62. Segrè D, Deluna A, Church GM, Kishony R. 2005. Modular epistasis in yeast metabolism. *Nat. Genet.* 37(1):77–83
63. Slaughter BD, Unruh JR, Das A, Smith SE, Rubinstein B, Li R. 2013. Non-uniform membrane diffusion enables steady-state cell polarization via vesicular trafficking. *Nat. Commun.* 4:1380
64. Sorrells TR, Booth LN, Tuch BB, Johnson AD. 2015. Intersecting transcription networks constrain gene regulatory evolution. *Nature* 523(7560):361–65
65. Starr TN, Picton LK, Thornton JW. 2017. Alternative evolutionary histories in the sequence space of an ancient protein. *Nature* 549(7672):409–13
66. Starr TN, Thornton JW. 2016. Epistasis in protein evolution. *Protein Sci.* 25(7):1204–18
67. Stearns FW. 2010. One hundred years of pleiotropy: a retrospective. *Genetics* 186(3):767–73
68. Steinberg B, Ostermeier M. 2016. Environmental changes bridge evolutionary valleys. *Sci. Adv.* 2(1):e1500921
69. Szamecz B, Boross G, Kalapis D, Kovács K, Fekete G, et al. 2014. The genomic landscape of compensatory evolution. *PLOS Biol.* 12(8):e1001935
70. Szathmáry E. 1993. Do deleterious mutations act synergistically? Metabolic control theory provides a partial answer. *Genetics* 133(1):127–32
71. Tyler AL, Crawford DC, Pendergrass SA. 2016. The detection and characterization of pleiotropy: discovery, progress, and promise. *Brief Bioinform.* 17(1):13–22
72. Voordeckers K, Pougach K, Verstrepen KJ. 2015. How do regulatory networks evolve and expand throughout evolution? *Curr. Opin. Biotechnol.* 34:180–88
73. Wagner GP, Zhang J. 2011. The pleiotropic structure of the genotype-phenotype map: the evolvability of complex organisms. *Nat. Rev. Genet.* 12(3):204–13
74. Wang Z, Liao BY, Zhang J. 2010. Genomic patterns of pleiotropy and the evolution of complexity. *PNAS* 107(42):18034–39
75. Weinreich DM, Delaney NF, Depristo MA, Hartl DL. 2006. Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science* 312(5770):111–14
76. Weinreich DM, Watson RA, Chao L. 2005. Perspective: sign epistasis and genetic constraint on evolutionary trajectories. *Evolution* 59(6):1165–74
77. Woods B, Lai H, Wu CF, Zyla TR, Savage NS, Lew DJ. 2016. Parallel actin-independent recycling pathways polarize Cdc42 in budding yeast. *Curr. Biol.* 26(16):2114–26
78. Wright S. 1932. The roles of mutation, inbreeding, crossbreeding, and selection in evolution. In *Proceedings of the Sixth International Congress of Genetics*, Vol. 1, ed. DF Jones, pp. 356–66. Austin, TX: Genet. Soc. Am.
79. Yang YF, Cao W, Wu S, Qian W. 2017. Genetic interaction network as an important determinant of gene order in genome evolution. *Mol. Biol. Evol.* 34(12):3254–66
80. Zhang XS, Hill WG. 2005. Genetic variability under mutation selection balance. *Trends Ecol. Evol.* 20(9):468–70

Annual Review of Biophysics
Volume 49, 2020

Contents

Indexes

Errata

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