

Adaptive Landscapes in the Age of Synthetic Biology

Xiao Yi^{*1} and Antony M. Dean^{*1,2}

¹BioTechnology Institute, University of Minnesota, St. Paul, MN

²Department of Ecology, Evolution, and Behavior, University of Minnesota, St. Paul, MN

*Corresponding authors: E-mails: xiaoyi0786@gmail.com; deanx024@umn.edu.

Associate editor: Claus Wilke

Abstract

For nearly a century adaptive landscapes have provided overviews of the evolutionary process and yet they remain metaphors. We redefine adaptive landscapes in terms of biological processes rather than descriptive phenomenology. We focus on the underlying mechanisms that generate emergent properties such as epistasis, dominance, trade-offs and adaptive peaks. We illustrate the utility of landscapes in predicting the course of adaptation and the distribution of fitness effects. We abandon aged arguments concerning landscape ruggedness in favor of empirically determining landscape architecture. In so doing, we transform the landscape metaphor into a scientific framework within which causal hypotheses can be tested.

Key words: adaptive landscape, genotype–phenotype gap, epistasis, genotype by environment interaction, distribution of fitness effects, pleiotropy.

Introduction

Ever since their inception, adaptive landscapes have proved compelling metaphors in evolution. Fitness is depicted as height with phenotypes or genotypes arrayed across a geographic grid. Less fit genotypes are replaced by fitter ones as populations creep up slopes, ever onward and upward. Adaptation is analogous to climbing a hill and stops only when a summit is reached.

Adaptive landscapes were first introduced in the early 1930s by Ronald Fisher and Sewall Wright (Box 1). Fisher (1930) conceived an abstract geometrical model to explain why the inheritance of metrical traits is governed by many alleles, each with a small statistically additive effect. Wright (1932) suggested nonadditive interactions (epistasis) would produce rugged landscapes. Each was the basis for a competing vision of the adaptive process. Together they provided a framework for thinking about adaptation that would be further developed and embellished over many decades (Simpson 1944; Lande 1976; Kauffman and Levin 1987; Gavrilets 2004). Today, the concept has found application in protein engineering (Romero and Arnold 2009) and studies of catalytic RNAs (e.g., Pressman et al. 2015).

At the time, little was known of biochemistry, molecular, cell and systems biology. Fisher and Wright therefore framed their landscapes in terms of what was known, and what was known was genetics. They took additivity, dominance, epistasis, and pleiotropy as the elements of their landscapes and applied mathematics to predict evolutionary outcomes.

There are limitations to using phenomenological terms like epistasis and pleiotropy as a basis for theory. A phenomenon might be illusory, just as the sky can be seen but not touched. Seemingly independent phenomena may share a common origin, just as waves and particles are the twinned

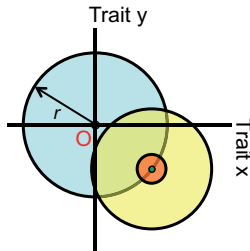
expressions of light. Changes in a phenomenon cannot be predicted in the absence of causative mechanisms, just as the bending of light in a gravitational field cannot be predicted by Newtonian mechanics. Like Newtonian mechanics, population genetics is a theory of consequences, not of causes. It can predict the kinetics of evolution across a landscape, but it cannot address the origins of the landscape's architecture.

The debate between Fisher and Wright over the origins of dominance (Provine 1986), a key descriptor of landscape architecture, is an exemplar of the futility of addressing causes with a theory of consequences. Fisher (1928) argued that dominance itself was an adaptation and invoked a special class of genes, the modifiers of dominance, to explain its evolution. Wright (1929, 1934) and Haldane (1930, 1939) argued that dominance arose as a passive consequence of selection for physiological buffering. The same argument continues today in remote corners (Billiard and Castric 2011; Huber et al. 2018) even though a compelling metabolic theory of dominance was provided decades ago (Kacser and Burns 1981).

That a molecular basis of dominance took decades to be embraced by evolutionists epitomizes a schism entrenched in academe, between studies of ultimate and proximate causes (Mayr 1961; Tinbergen 1963). Evolutionary biologists focus on ultimate causes (the “why?”), often using field studies and framing results in a historical context. Many question the relevance of results from laboratory studies on proximate mechanisms (the “how”) to natural settings, believing that functionalism merely adds detail, and possibly laboratory artifact, to the big picture painted by evolutionary biology (Morange 2011).

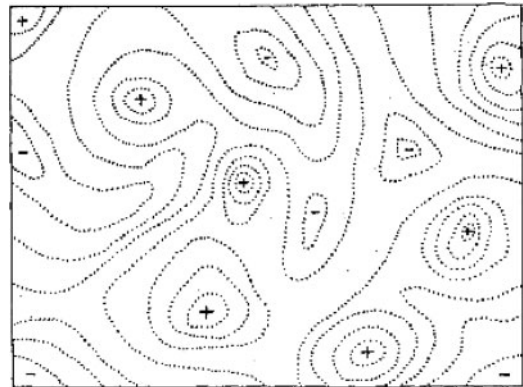
Box 1.

Fisher's (1930) Geometrical Model Consider two metrical traits, x and y , graphed on a plane (right) with their fitness optimum at the origin, O . Place a population (green dot) on the circumference of a circle, radius r , from the origin. Mutations inside the circle are beneficial; those outside are deleterious. Now draw two circles, centered on the population, that differ in size to represent mutations of small and large effect. Assume mutations are distributed uniformly on the circles. A mutation of small effect has an $\sim 50\%$ chance of being beneficial (approximately half the circumference of the small circle lies within r of the optimum). A mutation of large effect has a lower probability of being beneficial (less than half the circumference of the large circle lies within r of the optimum). The difference is not very impressive in two dimensions, but when expanded to many traits on an n -dimensional hypersphere the likelihood that a mutation of large effect is beneficial diminishes rapidly (more dimensions differentially increase the opportunities for deleterious effects). This simple model—amended by [Kimura \(1983\)](#) and then [Orr \(1998\)](#)—provides justification for the observation that metrical traits are governed mostly by many alleles of small, statistically additive, effect.



Wright's (1932) Lumpy Landscape Wright's work on guinea pig coat colors showed that mutations could

affect phenotypes in highly nonadditive ways. This suggested to him that fitness might also be a rugged function of genotype. His difficulty was in visualizing the high dimensionality of genotype space on a two dimensional page. So he settled on a metaphor. A rugged landscape (right) of adaptive peaks (+) separated by maladaptive valleys (−) with the axes representing heuristic orderings of the numerous genotypes. He left the axes unlabeled—there is no way to collapse an n -dimensional array of genotypes onto two dimensions. Later [Lewontin and White \(1960\)](#), analyzing chromosomal inversions in populations of the grasshopper *Moraba scurra*, would redefining Wright's landscape as a plot of mean population fitness against genotype frequencies.



to understand their architecture. Epistasis, dominance, pleiotropy, constraints, and genotype by environment interactions are seen to emerge as natural consequences of underlying proximate mechanisms. They are useful descriptors of landscapes, but they lack causality. Our approach replaces the current paradigm of interpreting observations in light of heuristic theory with direct experimental tests of causal mechanisms. Our framework provides a foundation for novel prediction and discovery.

Landscape Basics

What do all adaptive landscapes have in common? Genotypes, phenotypes, and fitnesses. A fourth element, the environment, enters implicitly when translating genotypes into phenotypes and phenotypes into fitnesses. The relationships between these elements define a landscape (fig. 1). [Fisher's \(1930\)](#) geometrical model captures the idea that the biology flows from genotype to phenotype (the infinitesimal model is his genotype–phenotype map) to fitness (his geometrical model). [Wright's \(1932\)](#) rugged landscape omits phenotypes to focus on the genotype–fitness map in the abstract.

Many attempts to investigate adaptive landscapes confound the landscape itself with the kinetics of the

This attitude has perpetuated for decades the same tired arguments of Fisher and Wright within the same tired intellectual framework, one that owes little to progress elsewhere in the life sciences. The causes of dominance were not found in evolutionary theory as Fisher, Wright, and Haldane believed ([Fisher 1928](#); [Wright 1929, 1934](#); [Haldane 1930, 1939](#)), but rather in the proximate mechanisms of functional biology ([Kacser and Burns 1973, 1981](#)). [Dobzhansky \(1964, 1973\)](#) was wrong. Nothing in evolution makes sense except in the light of biology.

We advocate using functional biology to study adaptive landscapes. In the seven examples discussed (table 1), we show that how knowledge of biochemistry and physiology enriches and extends the current framework. These landscapes are presented in only as much detail as is essential

Table 1. Seven Empirical Landscapes.

		Landscape						
		TEM-1	Opsins	<i>lac</i> Operon	IMDH	Methanol Metabolism	Lac Repressor	Chemotaxis–Growth
Biology Environment Approach	Resistance Constant Bottom up	Vision Constant Bottom up	Competition Constant Bottom up	Biosynthesis Constant Bottom up	Competition Constant Top down	Gene regulation Seasonal Bottom up and top down	Behavior Seasonal Bottom up and top down	
G-P Map	—	Predicted	Observed	Fitted	Observed	Observed	Observed	
P-W Map	—	Not measured	Predicted	Fitted	Fitted	Fitted	Predicted	
G-W Map	Observed	Not measured	Observed	Fitted	Observed	Observed	Observed	
Adaptive paths Uses	Delineated	Delineated	—	Delineated	Delineated	Delineated	Delineated	
Hypotheses confirmed	Local peak, epistasis	Mechanistic G-P map	Mechanistic P-W map	Complete landscape	Climbing to a peak	Climbing a new peak in a new landscape	Trade-off generates a peak in P-W map	
Discoveries	Few adaptive paths	How color vision evolves	Limits to adaptation	Ubiquitous constraint	Single peak in a linear flux	Few mutations to reverse regulation	Evolution of a new behavior	
Predictions	—	Evolutionary physics of vision	U-shaped DFE	—	—	New peak	—	
References	Weinreich et al. (2006)	Collette et al. (2018)	Dean (1989)	Lunzer et al. (2005)	Chou et al. (2014)	Poelwijk et al. (2011)	Yi and Dean (2016)	

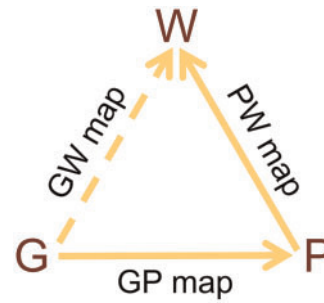


Fig. 1. Three elements of an adaptive landscape: G, genotypes; P, phenotypes; and W, fitnesses. A fourth element, the environment (E, not shown), enters implicitly in transforming genotypes into phenotypes (the GP map) and phenotypes into fitnesses (the PW map).

evolutionary process (commonly a result of conflating Wright’s shifting balance hypothesis with the landscape that underpins it). We follow the logic of Fisher and Wright in keeping the two separated. The population genetic processes of mutation, recombination, migration, mating, drift, and selection are not part of the landscape—they determine how populations move across the landscape—and we shall not discuss them further.

Our focus is on static landscapes. We recognize that many landscapes are dynamic. As evolving populations modify their environments, for example, by consuming resources and producing wastes, so the genotype–phenotype and phenotype–fitness maps can change. The resulting ecoevolutionary feedbacks not only generate frequency-dependent selective effects (Lunzer et al. 2002; Pelletier et al. 2009) but can also create new ecological niches (Rosenzweig et al. 1994; Rainey and Travisano 1998). We omit these fascinating topics to focus on making the simplest adaptive landscapes manifest.

Seven Empirical Landscapes

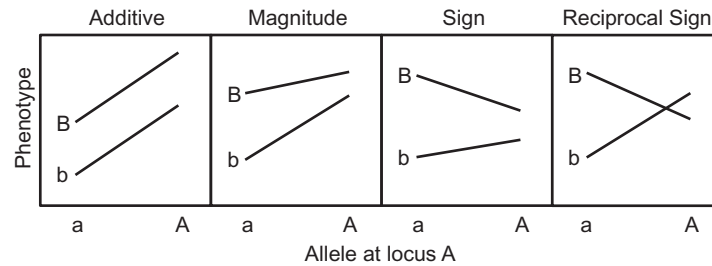
TEM-1: A Genotype–Fitness Map

Resistance to the antibiotic cefotaxime is conferred by five mutations in the plasmid-borne TEM-1 β -lactamase. Weinreich et al. (2006) constructed all $2^5 = 32$ genotypes, transformed them into *Escherichia coli*, and determined the minimum inhibitory concentrations (MICs) needed to prevent growth. The mutational network is characterized by the presence of both magnitude and sign epistasis (Box 2) with only 18 of the $5! = 120$ possible pathways leading to the single adaptive peak (fig. 2). These results are generally held to confirm Wright’s concept of a rugged adaptive landscape.

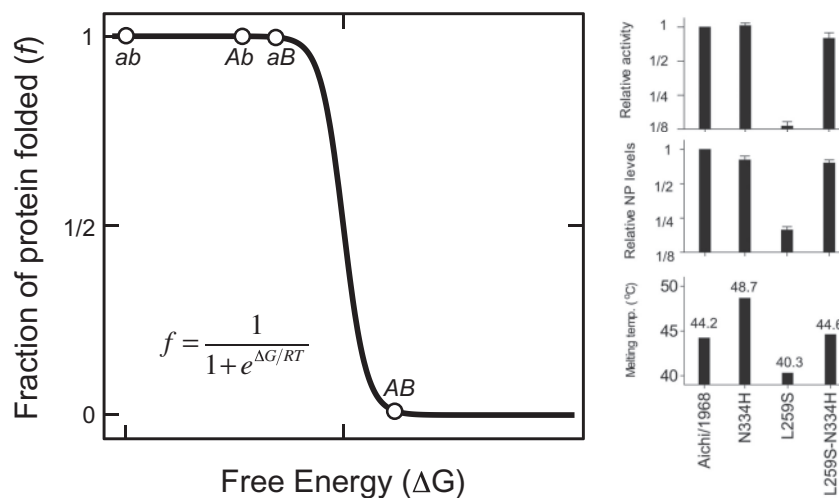
Wright’s vision was far more expansive than the small network explored by Weinreich et al. In this broader context, there might well be many adaptive peaks, or perhaps just the same peak, or perhaps this one peak ceases to be a peak at all as it folds into the shoulder of a still higher peak. The problem is that sign epistasis is necessary but not sufficient to generate isolated adaptive peaks (Crona et al. 2013). Four interacting sites in Streptococcal protein G illustrate the difficulty (Wu et al. 2016). Although reciprocal sign epistasis among mutations at paired sites might seemingly trap protein G on a

Box 2.

Epistasis is a deviation from additivity, the naïve expectation that the phenotypic contribution made by a mutation (a, A) should be independent of the genotype at a second site (b, B). [Weinreich et al. \(2005\)](#) define three types of epistasis below: 1) magnitude, where the effect size differs according to background, 2) sign, where the direction of the effect differs according to background, and 3) reciprocal sign, where the direction of the effects differ at both loci. Only reciprocal sign epistasis produces multi-peaked landscapes ([Poelwijk et al. 2011](#)). However, the converse is not true; the presence of reciprocal sign epistasis does not guarantee multiple peaks exist ([Crona et al. 2013](#)). Identifying multiple peaks in a local landscape is no guarantee that they also exist in the global landscape.



Additivity on one phenotypic scale in the biological hierarchy need not imply additivity on other dependent phenotypic scales. In the two state model of protein folding (below left), the fraction of protein folded (f) is a sigmoidal function of ΔG , the difference in free energy between the folded and unfolded states ([Privalov and Khechinashvili 1974](#)). Two mutations, A and B, each acting additively on the ΔG scale, and each having marginal effects on protein stability, can together completely unfold a protein to obliterate all function. For example, (below right) two replacements in the influenza nucleoprotein Aichi/1968, N334H and L259S, act additively on protein stability and nonadditively at the levels of expression and activity ([Gong et al. 2013](#)).



A perfect fit to data for two mutations (four genotypes) that lie on an arbitrary nonlinear curve (as in the two state model) requires estimating a mean, two additive terms, and a pairwise epistatic (interaction) term. The more mutations, the more terms are needed. For n mutations, 2^n terms are needed to achieve a perfect fit: a mean, n additive terms, and $\sum_{i=2}^n nC_i$ epistatic terms. High-order epistatic terms necessarily exist yet may be too small to estimate. In the unlikely event that we obtain a perfect fit, we still have learnt nothing biological. All we have is a statistical partitioning of variability. Our perfectly fitted model provides no mechanistic understanding of biological cause and effect.

[Weinreich et al. \(2013\)](#) suggest epistasis is a measure of our “surprise.” Yet nonlinearities in biology abound. In light of this fact, we suggest that epistasis quantifies our “ignorance”—whereas our inability to detect the high-order terms quantifies our “technical incompetence.”

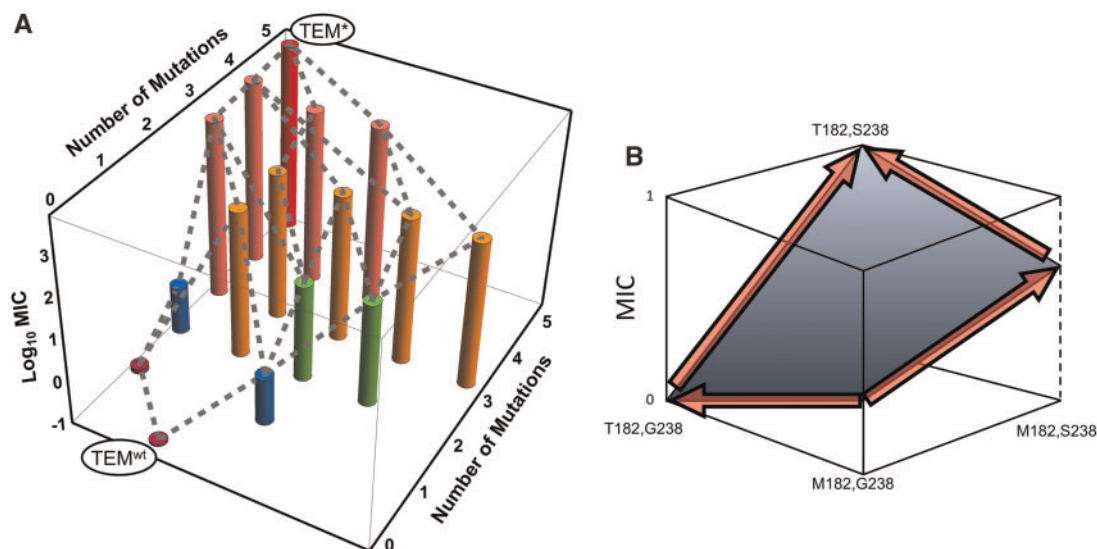


Fig. 2. (A). The genotype network of TEM-1 β -lactamase showing the 18 possible routes of ever increasing fitness (\log_{10} MIC) from wildtype to the cefotaxime resistant TEM*. The 16 genotypes accessible only by mutations that lower fitness are not shown. (B) An example of sign epistasis. Replacement M182T enhances resistance only in the presence of replacement G238S. Evolution from M182, G238 to T182, S238 must proceed counter clockwise through M182, S238.

lower peak, escape to a higher peak is made possible through the temporary acquisition of alternative amino acid replacements at either site or at additional sites.

Absent causal theory one cannot make inferences beyond data that merely associate each genotype with a fitness. Wright's landscapes are uninteresting by themselves because they lack causal mechanisms and so provide few insights and predictions.

Opsins: Bridging the Genotype–Phenotype Gap

That we are unable to predict from primary sequence, the structure, stability, and functional properties of a single protein points to the yawning chasm that is the genotype–phenotype gap. The astronomical number of possible peptide conformations, the difficulty in describing the kinetics of so highly a cooperative process as protein folding (Levinthal 1969), and the need to embed a quantum mechanical description of catalysis within a dynamic active site (Swiderek et al. 2014) suggest that a comprehensive theory bridging the genotype–phenotype gap cannot be envisioned within the foreseeable future.

Nevertheless, studies of visual pigments (a multigene family of transmembrane proteins with covalently bound retinal chromophores that transform light into vision) point to the feasibility of bridging the gap (fig. 3A). Interactions that delocalize electrons into the 11-*cis*-retinal polyene π system and β -ionone ring have the effect of localizing the positive charge at the lysyl Schiff base. This stabilizes the 11-*cis*-retinal ground state. With higher energy photons needed to trigger isomerization, maximum absorbance (λ_{\max}) shifts toward shorter wavelengths. Interactions that delocalize electrons from the polyene π system toward the lysyl Schiff base allow lower-energy photons to trigger isomerization and so λ_{\max} shifts in the opposite direction, toward longer wavelengths. Amino acid replacements produce λ_{\max} shifts by 1) modulating

the interaction between the protonated Schiff base and its counter ion Glu113 (Sakmar et al. 1989; Zhukovsky and Oprian 1989), 2) increasing the planarity of the polyene backbone through steric interactions with the protein, and/or 3) using polar side chains to modify the dipolar environment of the polyene and β -ionone ring (Ernst et al. 2014; Gozem et al. 2017).

Based on phylogenetic analysis and site directed mutation experiments, Yokoyama and coworkers identified the amino acid replacements that control spectral tuning in mammalian medium/long wavelength sensitive (M/LWS) visual pigments (Yokoyama and Radlwimmer 1998, 2001; Yokoyama 2008; Yokoyama et al. 2008). They proposed and refined a “five sites rule” to explain the 50-nm shift in maximum absorbance from red ($\lambda_{\max} = 560$ nm) to green ($\lambda_{\max} = 510$ –530 nm). Identifying these sites enabled hybrid quantum mechanics/molecular mechanics (QM/MM) calculations to explore the chemical basis of spectral tuning (Altun et al. 2008a, 2008b, 2009; Sekharan et al. 2010, 2011, 2012, 2013). This led to a broader “OH-site rule” that accounts for spectral shifts in the visual pigments of species as diverse as monkeys and squid. Inspired by Honig et al. (1976, 1979), Collette et al. (2018) recently explored the role of electrostatics in spectral tuning using a linearized Poisson–Boltzmann/quantum chemical (PBQC) method. They further refined the “OH-site rule” into a “dipole orientation rule” wherein both the position and orientation of a hydroxyl (or other dipolar group interacting with retinal) determine the direction and magnitude of the shift in λ_{\max} .

The molecular basis of ultraviolet (UV) vision has also been established (Tada et al. 2009). QM/MM calculations attribute the shift in λ_{\max} from UV to violet in ancestral SWS1 fish opsins to deleting amino acid Phe86. This mutation rearranges the hydrogen bond network surrounding the retinal (fig. 3C and D), stabilizing the protonated Schiff base over its

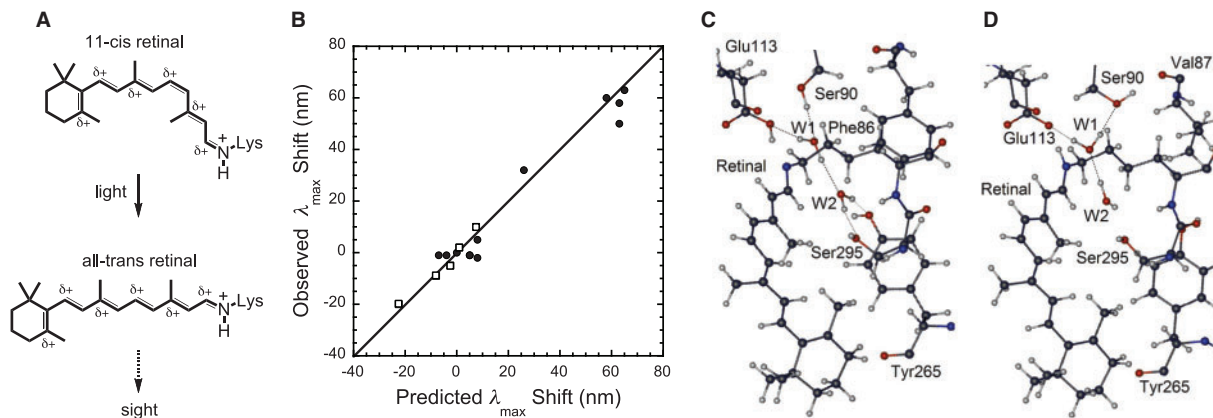


Fig. 3. Spectral tuning in visual pigments is achieved by modulating the relative energies of the ground and excited states. (A) Upon photoexcitation, the retinal chromophore isomerizes from the 11-*cis* ground state to the all-*trans* conformation. (B) QM/MM (Altun et al. 2011) and PBQC/MM (Collette et al. 2018) calculations accurately predict experimentally observed shifts in λ_{\max} : from UV to violet in mutants of ancestral and modern fish SWS1 opsins (dots) and from red to green in mutants of bovine M/LWS opsin (squares). (C and D) QM/MM calculations showed, and mutagenesis experiments confirmed (Tada et al. 2009), that the shift in λ_{\max} from UV to violet in fish SWS1 opsins was achieved by the deleting Phe-86. This rearranged the hydrogen bond network surrounding the retinal and converted the unprotonated Schiff base–linked 11-*cis*-retinal (C) to the protonated form (D). (C) and (D) from Tada et al. (2009).

ancestral unprotonated form. Deleting Phe86 in an earlier ancestor produces only a modest effect on λ_{\max} because, as QM/MM calculations show, the retinal Schiff base in that opsin remains unprotonated.

These hybrid quantum mechanical calculations are remarkably accurate (fig. 3B), yet computationally intensive and not generally practicable. Accurate homology models of extant and ancestral proteins are needed, as is a detailed understanding of the functional chemistry of each protein. Nevertheless, phylogenetic reconstructions combined with protein engineering and QM/MM calculations have bridged the genotype–phenotype gap, revealing the molecular basis of red-green and UV vision and the origins of intramolecular epistasis in ancestral SWS1 fish opsins.

lac Operon: A Phenotype–Fitness Map

The lactose pathway of *E. coli* provides a precise mechanistic biochemical model of Darwinian fitness in which the direction and intensity of natural selection can be predicted solely from knowledge of enzyme kinetics (Dean 1989) (fig. 4A–C). Fitness climbs onto a plateau as enzyme activity is increased (fig. 4D). This relationship inevitably produces diminishing returns epistasis; a small increase in the activity of an inefficient enzyme is strongly favored, whereas the same increase in an efficient enzyme is selectively neutral. The flux–fitness plateau represents a limit of adaptation where evolution proceeds in a neutral fashion (Hartl et al. 1985).

A second source of epistasis arises from interactions among enzymes within a pathway. As an enzyme approaches its flux–fitness plateau, it becomes less rate limiting, forcing other steps in the pathway to become more rate limiting (Kacser and Burns 1973, 1981). A once neutral 50% reduction in β -galactosidase activity (fig. 4E) becomes mildly deleterious in a background with increased porin activity (fig. 4F), whereas a once mildly deleterious 50% reduction in permease activity becomes strongly deleterious. Increased activity at

one step potentiates adaptation by exposing polymorphisms at other steps to selection. The neutrality of an enzyme’s polymorphism is conditional on the activities at other steps in the pathway.

The very same asymptotic relationship between enzyme activity and metabolic flux that generates diminishing returns epistasis also provides a basis for dominance in diploid species (Kacser and Burns 1981). The 50% reductions in activities in figure 4E might just as well represent the activities in the heterozygotes of null mutants, with complete dominance at β -galactosidase, partial dominance at the permease and codominance at the porins. Just as adaptation at one step in a pathway can expose a previously neutral polymorphism to selection at another step, so swapping alleles at one step can modulate the dominance relationships among alleles at other steps. In a wildtype background, permease heterozygotes show partial dominance, whereas in a background where porin activity is increased 100-fold they show codominance (fig. 4F).

β -Isopropylmalate Dehydrogenase: An Ancient Adaptive Landscape

Coenzyme use by β -isopropylmalate dehydrogenase (IMDH), an enzyme in the leucine biosynthetic pathway, is determined by six active site amino acids. Lunzer et al. (2005) constructed 256 *E. coli* IMDH mutants, including transitional amino acids no longer extant, determined their activities with the coenzyme substrates NAD⁺ and NADP⁺, and estimated their fitnesses in competition for glucose as a limiting resource.

Amino acid replacements act additively with respect NAD⁺ and NADP⁺ activities. As in the *lac* operon, the concave dependence of fitness on enzyme activities generates diminishing returns epistasis (fig. 5A). An evident trade-off in activity—no enzyme uses both NAD⁺ and NADP⁺ efficiently—jams the enzymes up against the perimeter of the landscape. This arrangement generates reciprocal sign

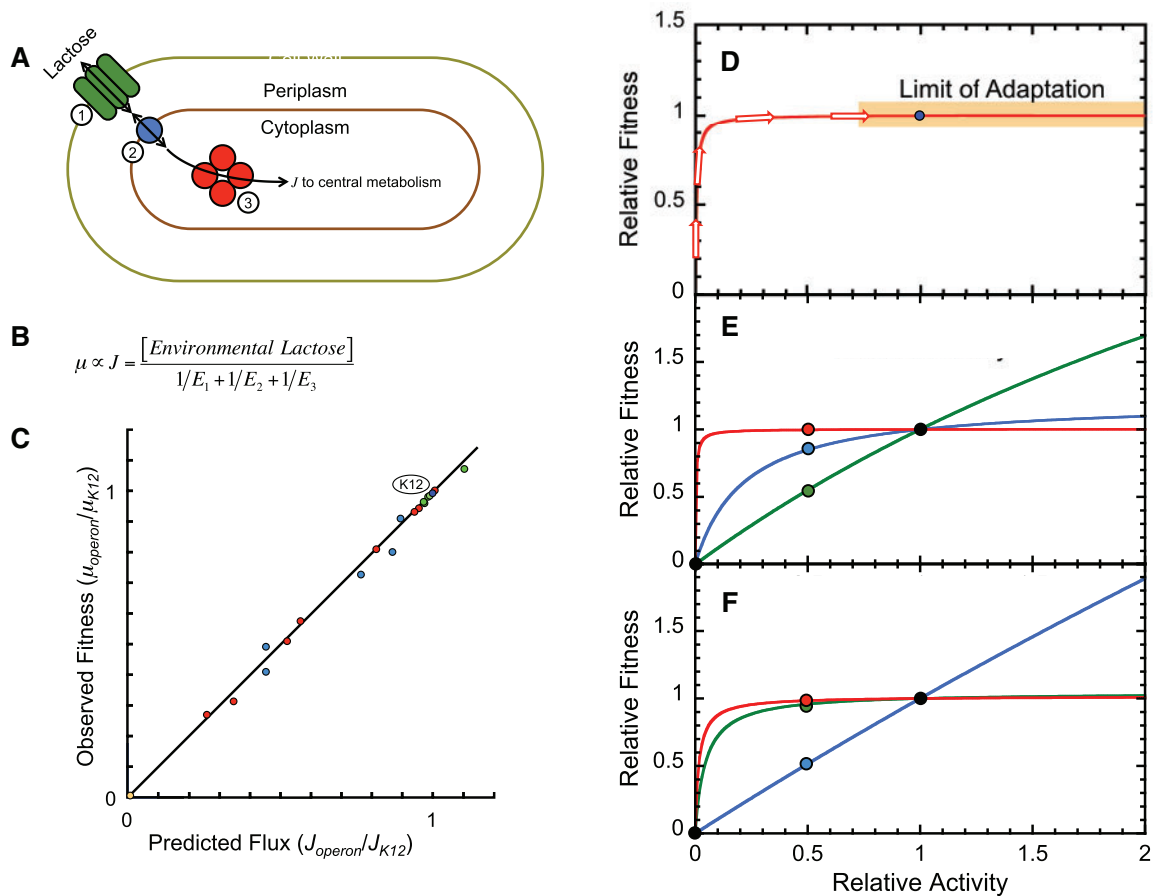


Fig. 4. (A) The lactose pathway of *Escherichia coli* consists of three steps: 1) passive diffusion of lactose through porin pores (green) into the periplasm, 2) active transport of lactose by the *lacY* encoded permease (blue) into the cytoplasm, and 3) irreversible hydrolysis by the *lacZ*-encoded β -galactosidase (red). (B) Starvation in chemostats ensures that the growth rate, μ , is proportional to the flux of lactose, J , into central metabolism. Flux is analogous to current in Ohm's law of resistance, $I = V/R$: $I \equiv J$ is current, $V \equiv [\text{Environmental Lactose}]$ is the potential and $R = \sum 1/C_i \equiv \sum 1/E_i$ is resistance. The conductance of each component (C_i) is analogous to enzyme activity ($E_i \propto [A_i] k_{\text{cat},i}/K_{m,i}$, where A_i is the concentration of active enzyme and $k_{\text{cat},i}$ and $K_{m,i}$ are the Michaelis–Menten parameters). Hence, relative growth rate (relative fitness) equals relative flux ($\mu_{\text{operon}}/\mu_{K12} = J_{\text{operon}}/J_{K12}$). (C) This mechanistic biochemical model (the straight line) accurately predicts relative fitness: dark blue is the *E. coli* K12 operon with β -galactosidase mutants (red), permease mutants (blue), operons from natural isolates (green), and *lac*[−] mutant at the origin (yellow). (D) Strong directional selection drives β -galactosidase activity onto a fitness plateau, a limit of adaptation where evolution is governed by neutral processes. (E) Not all steps in a pathway can lie in a limit of adaptation (Kacser and Burns 1981; Hartl et al. 1985); mutants with half wildtype activity might be selectively neutral (or nearly so) at β -galactosidase (red line), yet mildly deleterious at the permease (blue line) and strongly selected against at the porin step (green line). (F) A 100-fold increase in activity brings the porins close to their limit of adaptation. Necessarily, the β -galactosidase and permease become more rate limiting and so selection against their mutants intensifies, providing an example of intergenic epistasis.

epistasis in fitness; many amino acid replacements that decrease fitness of the NADP⁺-dependent RKYVYR mutant increase the fitnesses of mutants near the NAD⁺-dependent wildtype. Despite the presence of sign epistasis, the NAD⁺-dependent wildtype is accessible from anywhere on the landscape. Access from the NADP⁺-dependent RKYVYR mutant is made possible by three mutations of large functional effect that skirt the maladaptive funnel near the origin. In any landscape, the ruggedness of the phenotype–fitness map must be scaled to the size of the mutational effects in the genotype–phenotype map.

Why is using NADP⁺ less optimal than using NAD⁺? The positively charged nicotinamide ring of either coenzyme lies “above” the γ -isopropyl moiety of the bound substrate (fig. 5B). On reduction to NADPH or NADH, the ring loses

its charge and forms a tight hydrophobic interaction (blue arrow) with the γ -isopropyl moiety of the product “beneath.” Inhibition by abundant intracellular NADPH is severe; <1% of the RKYVYR mutant is available for catalysis. Inhibition by scarce intracellular NADH is weak; >80% of the wildtype enzyme is available for catalysis. Hence, NADP⁺ use is maladaptive (Miller et al. 2006). This landscape is sufficient to explain why all IMDHs use NAD⁺ rather than NADP⁺ and shows that at least one adaptive landscape has remained remarkably stable throughout the entire history of life.

Methanol Metabolism: A Top-Down Approach

The above adaptive landscapes explore relationships among genotype, phenotype, and fitness using predefined genetic variation chosen with prior knowledge. This bottom-up

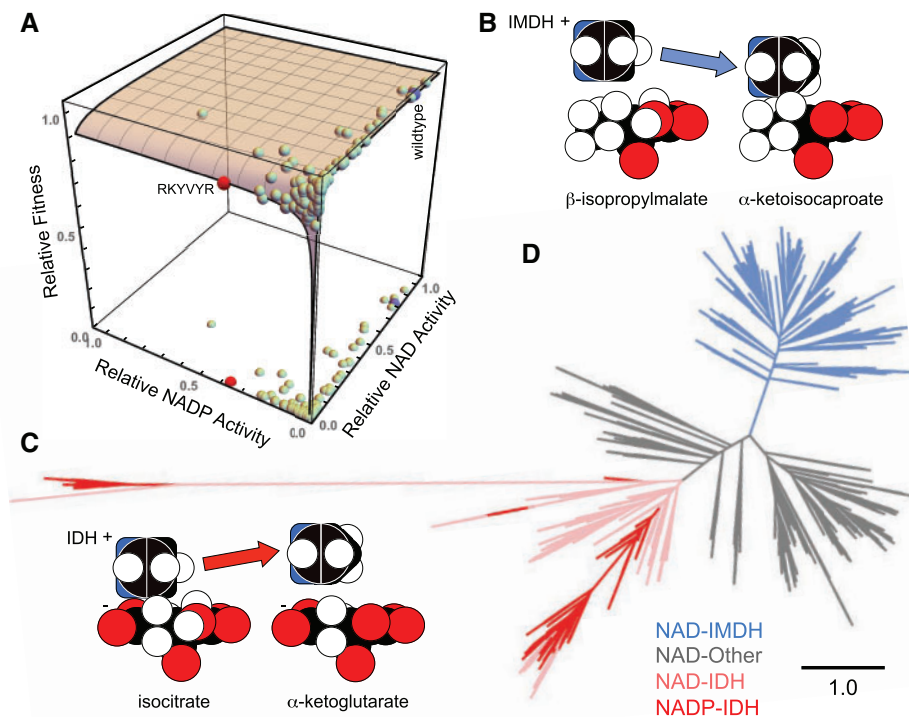


Fig. 5. The adaptive landscape controlling coenzyme use by IMDH (Lunzer et al. 2005). (A) The NAD⁺-dependent wildtype (blue ball) lies on a high fitness plateau (right), whereas the NADP⁺-dependent RKYVYR mutant (red ball) lies on a lower-fitness plateau (left). A trade-off in activity leaves the interior largely devoid of mutants. (B) Structural biology (Gonçalves et al. 2012; Palló et al. 2014) shows the nicotinamide ring of the coenzyme above the γ -isopropyl moiety of the bound substrate/product. NADH and NADPH are potent inhibitors of IMDH because the reduced nicotinamide ring binds the γ -isopropyl moiety tightly (Dean and Dvorak 1995; Miller et al. 2006). (C) NADH and NADPH are weak inhibitors of the related IDH because the reduced nicotinamide rings have no affinity for the negatively charged γ -carboxylate of the isocitrate substrate (Dean and Koshland 1993). (D) A maximum likelihood phylogeny of the IDH-IMDH family of enzymes reveals that all IMDHs use NAD, whereas the related IDHs have evolved NADP use several times.

approach is suitable for testing hypotheses about the underlying architecture of adaptive landscapes, for example, the structural and metabolic origins of epistasis and its impact on adaptation. As powerful as this approach is, it does restrict the scope of discovery to the predefined genetic variation.

Experimental evolution follows populations as they freely adapt to novel environments. This top-down approach allows for relatively unrestricted and open-ended explorations of landscapes. For example, Marx and coworkers (Chou et al. 2014) followed adaptation by *Methylobacterium extorquens* in which methanol catabolism by the native tetrahydromethanopterin-dependent pathway had been replaced by the unrelated glutathione-dependent pathway from *Paracoccus denitrificans*. Combining beneficial mutations that had emerged during laboratory adaptation, all of which reduced expression of the enzymes S-hydroxymethyl GSH dehydrogenase (FlhA) and S-formyl-GSH hydrolase (FghA), allowed the construction of a phenotype–fitness map (fig. 6).

Like the adaptive plateaus seen in the *lac* and IMDH landscapes, increases in fitness decline with further increases in enzyme activity. Unlike those landscapes, increases in enzyme activity are coupled to pleiotropic costs associated with increases in expression. The

resulting peak (positioned where costs and benefits are equiposed) generates both diminishing returns epistasis and sign epistasis. Another feature of this landscape is that zero fitness is offset from the FlhA origin; a minimum FlhA activity is needed to prevent build up of formaldehyde, a toxic metabolite in the pathway.

One might expect that multiple correlated mutational changes in both enzymes would be needed to reach the peak, and that many adaptive walks would be characterized by a series of phenotypic reversals whenever the optimum was overshoot. Yet a single beneficial mutation is sufficient to reach the optimum from the phenotypically distant ancestor. In this adaptive landscape, a variety of genotypes converge to the same phenotypic solution.

Work by Chou et al. (2014) illustrates the great strength of top-down approaches: Genotype–phenotype and phenotype–fitness maps can be generated, likely adaptive paths delineated, new phenomena discovered and fresh material obtained for studies of mechanistic origins.

lac Repressor: Information Processing on a Phenotype–Fitness Map

Survival depends on sensing and responding to the environment. Hence, organisms process information. To explore the

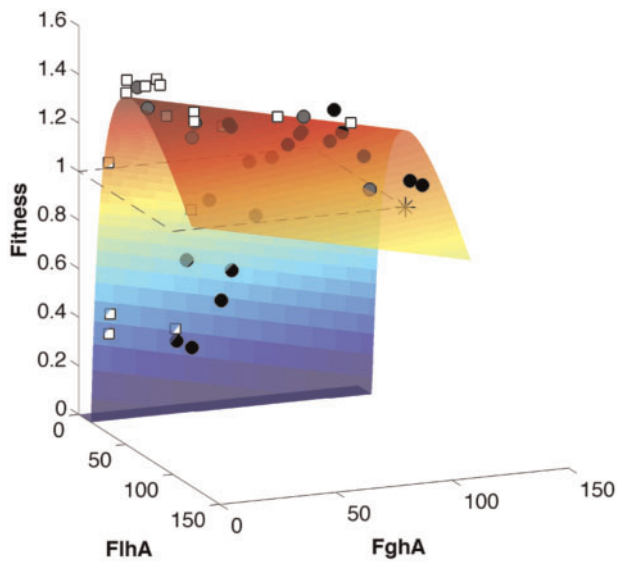


FIG. 6. The adaptive landscape for methanol catabolism by the *Paracoccus denitrificans* glutathione-dependent pathway placed in *Methylobacterium extorquens* reveals a single adaptive peak that could not have been predicted a priori. The surface represents the fit to a model in which fitness is proportional to the methanol flux minus the costs associated with protein expression and with the buildup of formaldehyde, a toxic metabolite. The peak is reached either by a single mutation or by a combination of mutations. Both diminishing returns epistasis and sign epistasis are present. Ancestor (asterisk), single mutants (gray circles), mutational combinations (white squares), and inducible expression from plasmids (black circles). From Chou et al. (2014).

evolution of information processing, Poelwijk et al. (2011) engineered a synthetic operon, using the *E. coli lac* regulatory system to drive the expression of *sacB* (which confers sensitivity to sucrose) and *cmR* (which confers resistance to chloramphenicol) (fig. 7A). This enabled selection for high expression (chloramphenicol medium) or low expression (sucrose medium) in the presence or absence of an environmental cue, the artificial inducer IPTG.

Population growth rates in presence of either chloramphenicol or sucrose were determined across a range of IPTG-induced expression levels (fig. 7B). These were used to predict average doublings per hour in a cyclical environment that alternated between sucrose and chloramphenicol at various combinations of expression levels (fig. 7C). *lacI*⁺ fitness is predicted to be maximized on full induction by IPTG when chloramphenicol is present and complete repression in the absence of IPTG when sucrose is present (green dot). The presence of this adaptive peak was confirmed by the recovery of high fitness *lacI* mutants with the wildtype phenotype (low basal expression levels with high induction by IPTG) from a library of randomly mutated *lacI* repressors that had been propagated in the cyclical environment.

Flipping the environmental cue, IPTG, so that wildtype *lacI*⁺ expression is induced in the presence of sucrose and repressed in the presence of chloramphenicol is predicted to force the wildtype *lacI*⁺ into the maladaptive valley (red dot). After one round of random mutation and selection, *lacI* mutants of similarly improved fitness were isolated across a wide range of expression levels (light ellipse). None responded

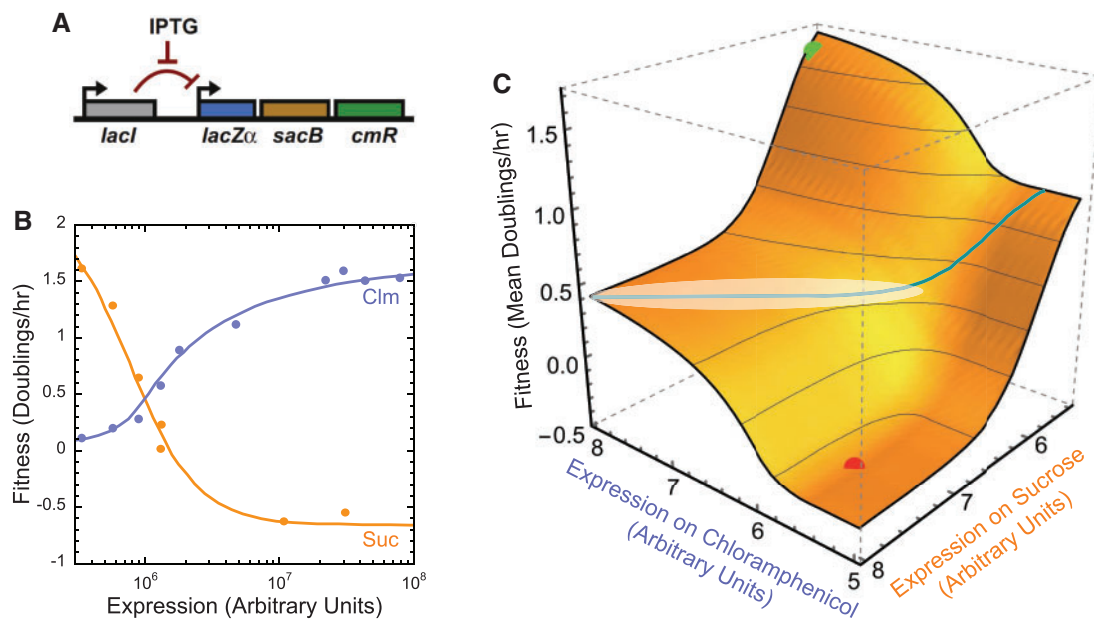


FIG. 7. Evolution of information processing in a synthetic operon. (A) The wildtype regulatory system of the *Escherichia coli lac* operon was used to control expression of *sacB*, which confers sensitivity to sucrose (Suc), and *cmR*, which confers resistance to chloramphenicol (CIm). (B) Induction of the operon by IPTG modulates sensitivity to sucrose (orange) and resistance to chloramphenicol (blue). (C) The phenotype–fitness map of the operon when alternating between two media, one with sucrose and the other with chloramphenicol. Operons expressed only in the presence of chloramphenicol occupy the adaptive peak (green dot). Operons whose expression is insensitive to the environmental cue lie on the blue line according to their level of constitutive expression. Operons expressed only in the presence of sucrose occupy the maladaptive valley (red dot). The light ellipse depicts the region from which deregulated mutants were first isolated.

to IPTG as an environmental cue. Hence, the first response to selection was to deregulate operon expression by obliterating sensing.

Following two more rounds of random mutation and selection, *lacI* mutants were isolated that once again had high levels of expression in the presence of chloramphenicol and low levels of expression in the presence of sucrose. On returning to the adaptive peak, these *lacI* mutants had changed their mode of information processing, from induction by IPTG to corepression with IPTG.

Additional experiments showed that the *lacI* mutants bound the *lac* promoter only in the presence of IPTG. This meant the allosteric response to IPTG had been inverted. In addition to the critical S97P, which lies at the dimer interface key to the IPTG-induced allosteric transition in wildtype *lacI*, any of a number of other amino acid replacements can contribute to repression.

A second means by which *lacI* mutants can switch from induction to repression is to increase the affinity for DNA in general while retaining the same allosteric response to IPTG (Pfahl 1976; Miller and Schmeissner 1979). These mutants bind DNA sufficiently tightly to slow diffusion along the double helix to the point that, in growing cells, they fail to find *lac* promoters before the next round of replication. Adding IPTG weakens their overall affinity for DNA and so increases their rate of diffusion, allowing the mutants to once again find and preferentially bind the *lac* operator, causing repression. This second mechanism involves no change in the allosteric response to IPTG and can be achieved by one of several single amino acid replacements. These experiments, together with those of Poelwijk et al. (2011), show that phenotypic parallelism need not imply mechanistic parallelism, and that the same mechanism need not imply genetic parallelism.

Chemotaxis–Growth: New Behavior Enables Escape from an Adaptive Peak

In a twist on the standard serial transfer protocol of experimental evolution (Yi and Dean 2016), *E. coli* cells that swam into a capillary tube loaded with chemoattractants were used to inoculate fresh medium twice daily (fig. 8A). This protocol defines a phenotype–fitness map, with the growth rate differential on one axis and the motility differential on the other axis (fig. 8B). The phenotype–fitness map has no adaptive peak. Cells must partition the limiting carbon source between growth and chemotaxis (both are energetically expensive). The ensuing trade-off cuts across the contours to produce a peak in the map.

Five replicate populations rapidly adapted to this cyclical environment, first reaching and then moving along the trade-off front (established empirically) toward the adaptive peak where they remained trapped for several weeks. Isolates from these populations were characterized by increased motility at the expense of growth rate, indicating that a premium existed on gaining access to fresh medium through chemotaxis.

After a few more weeks, several populations escaped from the peak, breaking through the Pareto front. Fitness improved through higher growth rates, even as efficient chemotaxis was retained in seeming violation of the trade-off. The contradiction

was resolved when it was shown that, compared with the ancestor, isolates reduced their motility during exponential growth only to increase it on close approach to carrying capacity (fig. 8C). Hence, a new physiological program had been implemented in which phenotypes were matched to each phase of the cyclical environment. Thus, did an evolved behavior mitigate the deleterious consequences of a hard-wired trade-off.

This dramatic change in behavior is largely attributable to a single mutation in a transcription factor specific to chemotaxis and motility (fig. 8D). The Arg220Trp amino acid replacement in FliA eliminates an ionic hydrogen bond to the backbone phosphate of DNA, weakening expression from client promoters, reducing motility during exponential growth, and increasing the proportion of the population that is motile near carrying capacity. A single point mutation at the hub of an existing complex gene network had reprogrammed cellular physiology to produce a new behavior that optimized organismal phenotypes to current environmental demands.

Discussion

These seven functional studies show how proximate mechanisms shape landscape architecture. The emphasis is on delineating chains of causality, from genotype to phenotype to fitness. In this new vision, the old genetic descriptors of additivity, epistasis, dominance, and pleiotropy are assigned no causal roles; they are simply labile epiphenomena. From metabolic architecture to dominance, from active sites to phylogenetics, studies of the proximate mechanisms that underpin adaptive landscapes have enriched our understanding of evolution far beyond what has been achieved previously. After almost 100 years, the adaptive landscape has ceased being a metaphor and emerged as a scientific framework of testable theories.

Adaptive Landscape Uses

The seven landscapes differ in their biology, architecture, and uses (table 1). Weinreich et al.'s (2006) study of TEM-1 β -lactamase is the classic example of a Wrightian hypercube, a local genotypic network shorn of biological causality that illustrates how epistasis can restrict the number of adaptive walks to a local adaptive peak. Quantum mechanical calculations of visual pigments provide a means to predict phenotype from genotype including the molecular basis of intramolecular epistasis (Altun et al. 2008a, 2008b, 2009; Tada et al. 2009; Sekharan et al. 2010, 2011, 2012, 2013; Collette et al. 2018). Experiments with the *lac* operon (Dean 1989) test a mechanistic prediction of fitness from metabolic theory. Studies of IMDH provide a complete fitted adaptive landscape (Lunzer et al. 2005) that informs phylogenetic patterns across the entire tree of life (fig. 5C and D). Chou et al.'s (2014) study of methanol metabolism exemplifies a top-down approach to constructing adaptive landscapes. Experiments by Poelwijk et al. (2011) and Yi and Dean (2016) explore adaptive landscapes in variable environments and illustrate the power of combining bottom-up and top-down approaches.

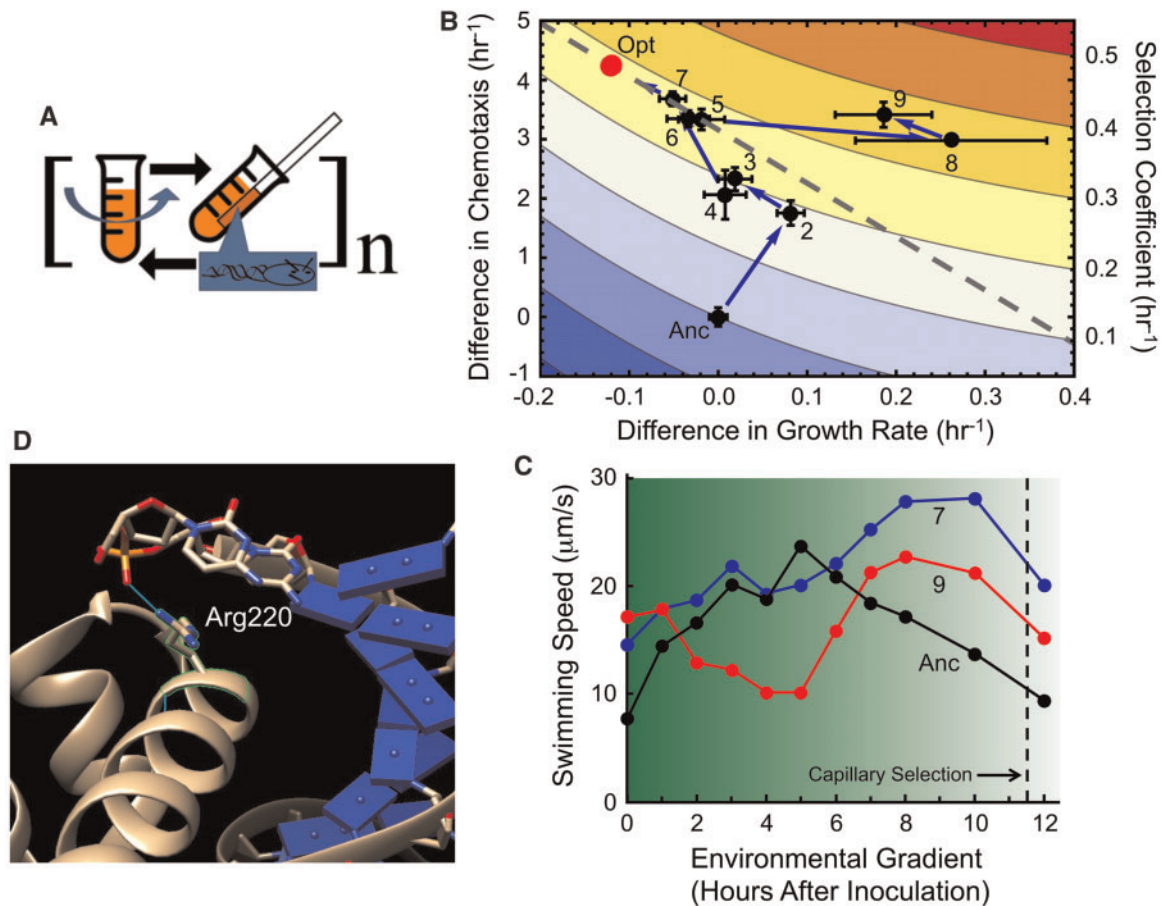


Fig. 8. Evolution of phenotypic plasticity to escape from an adaptive peak. (A) The selection regime where cells must chemotax into a capillary of fresh medium to be transferred to fresh medium. (B) Fitness is a function of growth rate and the ability to chemotax. Logistic growth produces curved contours in the phenotype–fitness map (exponential growth would produce linear contours). The dashed line represents the empirically determined position of the trade-off between chemotaxis and growth rate. The peak (red dot) is broad because the trade-off lies parallel to the contours. Each point is the mean of six weekly isolates from an evolving population. Ancestor (Anc), optimum (Opt), and week sampled (numbers). The large standard standard errors at week 8 are a consequence of transient polymorphisms. (C) The ancestor (Anc) swims fastest during exponential growth when resources are abundant. Week 7 isolates increase the overall swimming speed. Week 9 isolates reduce swimming speed during midlog phase while maintain increased swimming speed during transfer. (D) Amino acid replacement Arg220Trp removes a positive charge and hydrogen bond between transcription factor FliA and the promoter DNA. From Yi and Dean (2016).

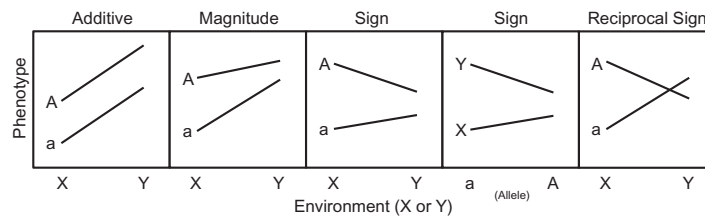
Adaptive landscapes have myriad uses: revealing principles, testing hypotheses, discovering new phenomena, and explaining biological patterns. For example, epistasis turns out to be something of a mirage. Amino acid replacements may act additively on one scale (e.g., protein stability), yet epistatically on another (e.g., metabolic flux). So, do the amino acids interact or do they not? Epistasis is, like a blue sky, an illusion even though it can be quantified. Dominance and diminishing returns epistasis are the twinned expressions of underlying metabolic architecture. To discuss the one and ignore the other makes no sense. Moreover, they are both spandrels (sensu Gould and Lewontin 1979) that evolve in response to selection, even though they themselves are never selected (Kacser and Burns 1981; Hartl et al. 1985). Adaptive landscapes have inspired new experiments. High-throughput approaches have extended Weinreich et al.'s (2006) TEM-1 β -lactamase genotype–fitness map to explore the phenotype–fitness map and the impact

of genotype by environment ($G \times E$) interactions by varying both the type and concentration of antibiotics (Stiffler et al. 2015). Work on *lac* (Dean 1995) identified two fundamental causes of $G \times E$ (Box 3). That $G \times E$ interactions combine with ecoevolutionary feedbacks to produce negative frequency-dependent selection during competition for mixtures of galactosides has been confirmed experimentally and the region where the polymorphism is protected identified (Lunzer et al. 2002). The de novo evolution of specialists within this region has been repeatedly observed (Dykhuizen and Dean 2004), and the mutations associated with each specialization identified (Zhong et al. 2004, 2009). The physiological cause of the trade-off and the presence of two adaptive peaks have yet to be confirmed.

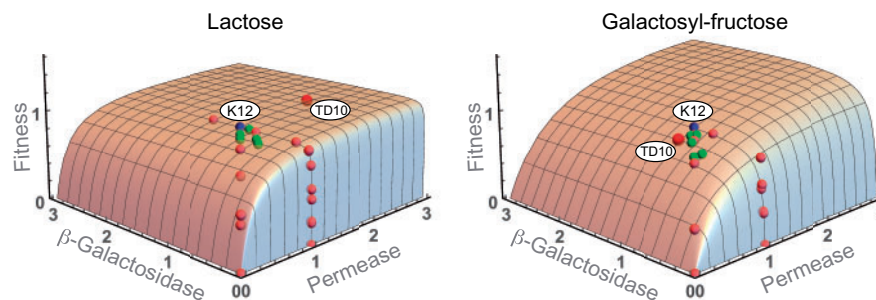
Adaptive landscapes explain the characteristic U-shaped distributions of the fitness effects of new mutations (DFEs). This has important implications for the evolutionary fate of populations (Eyre-Walker and Keightley 2007), from the rate

Box 3.

Environments are constantly changing and so it is of interest to know how changes in the environment affect both phenotypes and fitness. The standard approach is to plot phenotypic values or fitnesses against an environmental treatment (right). Our classification follows the example set by Weirauch et al. (2013) for epistasis. We classify $G \times E$ interactions as additive (no interaction, environment Y contributes equally to genotypes A and a), magnitude $G \times E$ (environment Y contributes more to a than to A). We identify two kinds of sign $G \times E$: either environment Y increases a and decreases A, or substituting A has opposite effects in environments X and Y. With reciprocal sign, $G \times E$ in environment Y increases a and decreases A, whereas substituting A causes an increase in environment X and a decrease in environment Y. As with epistasis, such plots expose the presence of $G \times E$ but reveal nothing of its causes.



Changes in environments can modify landscapes in two ways: 1) by modifying the phenotype–fitness map and 2) by modifying the genotype–phenotype map. Both are evident in the *lac* operon (Dean 1995), where laboratory mutants of the permease and β -galactosidase (pink spheres) were used to define the landscapes (below).



The landscape, almost flat around strain K12 (blue sphere at 1, 1, 1) during competition for lactose, is far steeper during competition for galactosyl-fructose where the permease and β -galactosidase are more rate limiting to metabolic flux. Operons from natural isolates (green spheres), once mildly deleterious on lactose become strongly deleterious on galactosyl-fructose (magnitude $G \times E$), even though their enzyme activities (relative to K12) are unchanged (mostly). Modifications of this phenotype–fitness map are not expected to change the direction of selection at either step in the pathway because fitness is a monotonic function of enzyme activities.

However modifications to the genotype–phenotype map can change the direction of selection (sign $G \times E$) by changing the rank order of relative enzyme activities. For example, the permease of strain TD10 (red sphere) is more active than the permease of strain K12 on lactose, but less active on galactosyl-fructose. Hence, selection favors strain TD10 on lactose and strain K12 on galactosyl-fructose (sign $G \times E$).

$G \times E$ interactions are useful in the design of experimental controls. In many landscape studies, alleles are placed in a common genetic background for estimating fitnesses. To be certain that any selection observed is caused by the alleles of interest, and not by mutations that spontaneously arise in the genetic backgrounds during strain construction, control experiments are conducted in an environment where the alleles are not expected to contribute to fitness. Selection at *lac* disappears during competition for glucose. Selection at TEM-1 β -lactamase disappears in the absence of antibiotic. Without these controls any epistasis detected in a Wrightian genotype–fitness map might plausibly be assigned to mutations elsewhere in the genetic background.

of neutral evolution to the likelihood of an adaptive response (Silander et al. 2007). The asymptotic dependency of flux on enzyme activity (fig. 4D) predicts that mutations obliterating

enzyme function will have low fitness, whereas those with even residual activity have near wildtype fitness leaving few mutants of intermediate fitness. The resulting U-shaped DFE

was first established (Dean et al. 1989) in a high-resolution, low-throughput screen of amino acid replacements in *E. coli* β -galactosidase (fig. 9A).

Adaptive landscapes explain why the fraction of deleterious mutations and the precise distribution of fitness effects vary from protein to protein: ubiquitin (Roscoe et al. 2013; Mavor et al. 2016), Hsp90 (Bank et al. 2015), TEM-1 β -lactamase (Jacquier et al. 2013; Firnberg et al. 2014; Stiffler et al. 2015), an amide hydrolase (Wrenbeck et al. 2017), and 5'ProFAR isomerase (HisA) (Lundin et al. 2018). Each protein's DFE depends not only on its sensitivity to mutation but also on its position along the function–fitness curve. For example, higher concentrations of ampicillin move TEM-1 β -lactamase down its fitness curve exposing more mutations to purifying selection (Stiffler et al. 2015) (fig. 9B and C). Introducing M182T stabilizes wildtype TEM-1 β -lactamase, buffering fitness against the destabilizing effects of random amino acid replacements (Jacquier et al. 2013). DFEs are to be understood as the products of combining genotype–phenotype maps and phenotype–fitness maps. Without this framework, DFEs merely form a series of disconnected anecdotal observations.

Adaptive landscapes explain patterns of variation in natural populations. That most inborn errors of metabolism are recessive (Kacser and Burns 1981) and that most segregating polymorphisms are nearly neutral (Bustamante et al. 2005; Castellano et al. 2018) and are consistent with DFEs produced when the fraction of folded protein is a sigmoidal function of $\Delta\Delta G$ (Box 2; Bershtein et al. 2017; Echave and Wilke 2017; CanaLe et al. 2018) and fitness is a concave function of enzyme activity (Hartl et al. 1985). Such adaptive landscapes buttress Ohta's nearly neutral model of evolution (Ohta 1973, 1977, 1992; Akashi et al. 2012).

Adaptive landscapes provide insights into the broad patterns of functional change and stasis in molecular phylogenies. We have already seen that NADP⁺ use by IMDH is deleterious because the reduced hydrophobic nicotinamide ring of abundant cellular NADPH forms a tight hydrophobic interaction with the γ -isopropyl moiety of the bound substrate/product (fig. 5B, blue arrow). The related isocitrate dehydrogenase (IDH) suffers no such inhibition because the reduced hydrophobic nicotinamide ring has little affinity for the negatively charged γ -carboxylate of its substrate/product (fig. 5C, red arrow). Inhibition by intracellular NADPH is sufficiently weak that IDHs have been free to evolve NADP⁺ use (Zhu et al. 2005) and have done so on multiple occasions (fig. 5D). The pattern of functional evolution and constraint across the tree of life finds its explanation in the structure–function relationships of the active sites that underpin the architectures of the respective adaptive landscapes.

Adaptive landscapes have practical applications. Analysis of TEM-1 β -lactamase landscapes across 15 antibiotics point to the possibility of retarding the evolution of resistance by deploying cyclical treatment paths that select for reversions to the starting state (Mira et al. 2015a). Caution is warranted however. A larger study revealed that although optima vary across 30 landscapes, $G \times E$ interactions are numerous, complex, and can mitigate the impact of sign epistasis to gain

access adaptive trajectories to higher optima (Mira et al. 2015b).

Bridging the Genotype–Phenotype Gap

Bridging the genotype–phenotype gap remains a huge challenge. Deletions, nonsense mutations, etc. reliably obliterate functions. Less catastrophic mutations have less predictable phenotypic effects. For example, silent substitutions are commonly assumed to be functionally equivalent and selectively neutral (Kimura 1983), yet exceptions are known (Ikemura 1981; Sharp and Li 1987; Agashe et al. 2013; Bailey et al. 2014). Gain of function mutations cannot be predicted. They must be identified through mutant screens (Arnold 2015), directed evolution studies (Hartl and Hall 1974), or phylogenetic methods coupled with ancestral sequence resurrection and protein engineering (Siddiq et al. 2017).

High-throughput association studies are now fashionable means to explore, empirically and comprehensively, the impact of mutations on protein phenotypes and fitness (Meng et al. 2005; Berger et al. 2006; Maerkl and Quake 2007; Domingo-Calap et al. 2009; Zykovich et al. 2009; Fillion et al. 2010; Fowler et al. 2010; Bank et al. 2015; Christensen et al. 2011; Wong et al. 2011; Kouyos et al. 2012; Gordan et al. 2013; Stormo 2013; Szendro et al. 2013; Weirauch et al. 2013; Olson et al. 2014; Orenstein and Shamir 2014; Thyagarajan and Bloom 2014; Wu et al. 2014, 2016; Zuo and Stormo 2014; Jolma et al. 2015; Levo et al. 2015; Stiffler et al. 2015; Boucher et al. 2016; Chattopadhyay et al. 2016; Li et al. 2016; Mavor et al. 2016; Puchta et al. 2016; Tripathi et al. 2016; Grossman et al. 2017; Sarkisyan et al. 2016; Wrenbeck et al. 2017; Aguilar-Rodríguez et al. 2018; Le et al. 2018; Robert et al. 2018). Rather than direct assays of the phenotypes of interest, many studies instead employ proxy metrics subject to artifact. Results, often inadequately replicated and lacking suitable experimental controls, are extracted using ad hoc computational modeling and reported as “enrichment scores” or some other opaque statistic (Stormo 2013; Weirauch et al. 2013; Le et al. 2018). Yet even careful empiricism, essential to discovery and foundational to coherent theory, lacks predictive power.

In any conceptual vacuum, perceptions are apt to change as data accumulate. The original empiric additive “three sites rule” for mammalian M/LWS visual pigments (Yokoyama and Yokoyama 1990; Neitz et al. 1991; Merbs and Nathans 1992; Asenjo et al. 1994) eventually transmogrified into an empiric “five sites rule” with five pairwise epistatic interactions (Yokoyama et al. 2008). This more refined empirical fit still lacked predictive power. However, identifying these critical sites was essential before hybrid quantum mechanical/molecular mechanical simulations could be deployed to accurately predict changes in spectral tuning (Altun et al. 2008a, 2008b, 2009; Sekharan et al. 2010, 2011, 2012, 2013; Collette et al. 2018). This example nicely illustrates the marriage of empiricism with mechanism to illumine causality at a most basic level of biological organization—the point mutation.

Our difficulties in bridging the genotype–phenotype gap are further compounded by epistasis and pleiotropy.

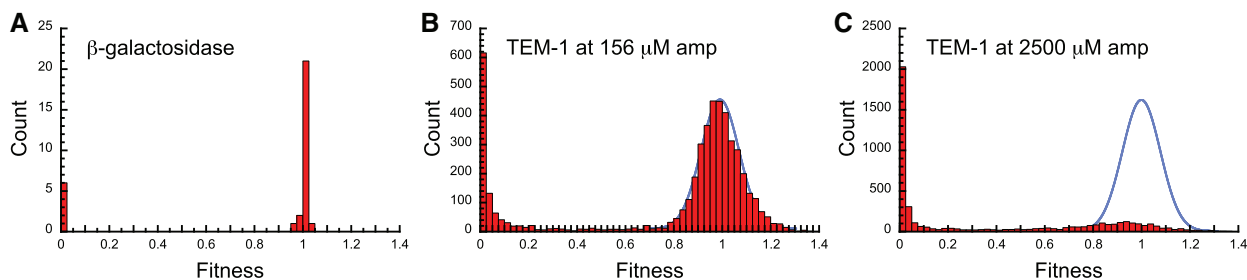


Fig. 9. (A) DFE for amino acid replacements in *Escherichia coli* β -galactosidase with a limit of resolution of $\pm 0.5\%$ /generation (Dean et al. 1988). (B) A similar U-shaped DFE is seen for TEM-1 β -lactamase (Stiffler et al. 2015) in the presence of 156- μ M ampicillin. Near wildtype, the distribution barely differs from the bell curve (blue) in the absence of ampicillin. (C) Increasing the ampicillin concentration effectively pushes the wildtype TEM-1 β -lactamase off its fitness plateau. Now most replacements are strongly deleterious.

Cooperative effects associated with protein stability are obvious sources of epistasis (DePristo et al. 2005; Tokuriki et al. 2008; Tokuriki and Tawfik 2009; Soskine and Tawfik 2010; Jacquier et al. 2013; Melamed et al. 2013; Bank et al. 2015; Olson et al. 2014; Sarkisyan et al. 2016; Echave and Wilke 2017). Many mutations that affect the free energy of folding act additively (or approximately so) and for these the fraction of protein folded can be accurately predicted (Wells 1990; Sandberg and Terwilliger 1991; Gregoret and Sauer 1993; Araya et al. 2012; Melamed et al. 2013). In addition to reduced levels of function, unfolded proteins can impose additional pleiotropic fitness costs (Echave and Wilke 2017). Mutations that affect function directly are often destabilizing (Tokuriki et al. 2008; Tokuriki and Tawfik 2009; Soskine and Tawfik 2010). Indeed, the replacements necessary for a functional change can be so destabilizing that no folded protein is produced. Bloom et al. (2006) showed that prior selection for increased stability of a cytochrome P450 was essential to its subsequent acquisition of novel activities. Detailed structural and biophysical characterizations have elaborated the causes of functional epistasis in a number of proteins but offer few generalizations (Ortlund et al. 2007; Bridgham et al. 2006, 2009; Marciano et al. 2008; Tada et al. 2009; Lunzer et al. 2010; Altun et al. 2011; Kryazhimskiy et al. 2011; Gong et al. 2013; Natarajan et al. 2013; Gong and Bloom 2014; Kaltenbach et al. 2015; Lagator et al. 2017). The problem remains extreme context dependence, which renders each genotype–phenotype map idiosyncratic. Neither theory nor empiricism has yet succeeded in exhaustively characterizing the combined phenotypic impacts of multiple mutations.

Into the Future

This review has focused on just two levels of biological organization, intragenic and metabolic, and how they affect fitness in single celled organisms. Higher levels of biological organization, tissue, organ, development, and behavior are expected to contribute to the landscape architectures of multicellular organisms (Svensson and Calsbeek 2012). These have never been explored in detail owing to difficulties in constructing isogenic strains and estimating genotypic fitnesses in species with complex life histories and long generation times.

Recent attempts to explore the relationships between metrical phenotypes and fitness in higher organisms employ a framework established by Russ Lande (Lande 1976, 1979; Lande and Arnold 1983; Phillips and Arnold 1989; Wood and Brodie 2015). Despite statistical and experimental concerns (Mitchell-Olds and Shaw 1987; Fincke and Hadrys 2001; Kingsolver et al. 2001; Reed and Bryant 2004; Pekkala et al. 2011; Wood and Brodie 2015), Lande’s approach, which uses least squares regression of fitness proxies against other phenotypes, has been used to infer natural selection from morphological data. However, it cannot address the causes of landscape architecture in terms of proximate mechanisms because metrical phenotypes are not amenable to the experimental manipulations needed to delineate specific associations among genotype, phenotype, and fitness.

We anticipate that future studies will be dominated by top-down approaches. Inexpensive genome sequencing combined with bioinformatic analyses can rapidly identify candidate mutations. New techniques in genome editing enable candidate mutations to be isolated and combined in defined genetic backgrounds suitable for fitness studies. GFP and other fluorescent proteins can be introduced as reporters of expression from titratable promoters, allowing the relationships among gene expression, enzyme activity, and fitness to be characterized with unprecedented ease. Hisidine-tagged proteins can be rapidly purified for phenotypic characterization. No longer confined to studies of a few well-defined biochemical systems in model prokaryotes, new studies will explore adaptive landscapes in nonmodel species and in higher eukaryotes. Although genomic analyses offer little more than association studies with no mechanistic insight (Graur et al. 2013; Boyle et al. 2017; Doolittle and Brunet 2017), work on DFEs points to the possibility of coupling comprehensive data sets with mechanistic studies to provide a broader understanding of adaptive landscapes.

Epilog

The age of exploring adaptive landscapes is upon us. As can be seen from the examples described, studying adaptive landscapes is a highly interdisciplinary undertaking involving expertise from many disciplines including physics, chemistry, molecular, structural, cell and systems biology, microbiology, genetics, metabolism, physiology, and behavioral ecology. As

the traditional barriers between disciplines continue to erode, so the study of adaptive landscapes will become increasingly prominent, providing a useful framework to integrate rich diverse and otherwise disparate knowledge of life. Future studies can only deepen and broaden our understanding of the causal basis of evolutionary change.

Acknowledgments

We thank Miriam Barlow, Danielle Drabeck, Dan Dykhuizen, Will Harcombe, Mato Lagator, Ruth Shaw, Mike Travisano, and Shozo Yokoyama for stimulating discussion and constructive criticism.

References

- Agashe D, Martinez-Gomez NC, Drummond AD, Marx CJ. 2013. Good codons, bad transcript: large reductions in gene expression and fitness arising from synonymous mutations in a key enzyme. *Mol Biol Evol.* 30(3):549–560.
- Aguiar-Rodríguez J, Peel L, Stella M, Wagner A, Payne JL. 2018. The architecture of an empirical genotype–phenotype map. *Evolution* 72(6):1242–1260.
- Akashi H, Osada N, Ohta T. 2012. Weak selection and protein evolution. *Genetics* 192(1):15–31.
- Altun A, Morokuma K, Yokoyama S. 2011. H-bond network around retinal regulates the evolution of ultraviolet and violet vision. *ACS Chem Biol.* 6(8):775–780.
- Altun A, Yokoyama S, Morokuma K. 2008a. Spectral tuning in visual pigments: an ONIOM(QM: mM) study on bovine rhodopsin and its mutants. *J Phys Chem B.* 112(22):6814–6827.
- Altun A, Yokoyama S, Morokuma K. 2008b. Mechanism of spectral tuning going from retinal in vacuo to bovine rhodopsin and its mutants: multireference ab initio quantum mechanics/molecular mechanics studies. *J Phys Chem B.* 112(51):16883–16890.
- Altun A, Yokoyama S, Morokuma K. 2009. Color tuning in short wavelength-sensitive human and mouse visual pigments: ab initio quantum mechanics/molecular mechanics studies. *J Phys Chem A.* 113(43):11685–11692.
- Araya CL, Fowler DM, Chen W, Muniez I, Kelly JW, Fields S. 2012. A fundamental protein property, thermodynamic stability, revealed solely from large-scale measurements of protein function. *Proc Natl Acad Sci U S A.* 109(42):16858–16863.
- Arnold FH. 2015. The nature of chemical innovation: new enzymes by evolution. *Q Rev Biophys.* 48(4):404–410.
- Asenjo AB, Rim J, Oprian DD. 1994. Molecular determination of human red/green color discrimination. *Neuron* 12(5):1131–1138.
- Bailey SF, Hinz A, Kassen R. 2014. Adaptive synonymous mutations in an experimentally evolved *Pseudomonas fluorescens* population. *Nature Commun.* 5:4076.
- Bank C, Hietpas RT, Jensen JD, Bolon DN. 2015. A systematic survey of an intragenic epistatic landscape. *Mol Biol Evol.* 32(1):229–238.
- Berger MF, Philippakis AA, Qureshi AM, He FS, Estep PW 3rd, Bulyk ML. 2006. Compact, universal DNA microarrays to comprehensively determine transcription-factor binding site specificities. *Nat Biotechnol.* 24(11):1429–1435.
- Bershtein S, Serohijos AW, Shakhnovich EI. 2017. Bridging the physical scales in evolutionary biology: from protein sequence space to fitness of organisms and populations. *Curr Opin Struct Biol.* 42:31–40.
- Billiard S, Castric V. 2011. Evidence for Fisher's dominance theory: how many 'special cases'? *Trends Genet.* 27(11):441–445.
- Bloom JD, Labthavikul ST, Otey CR, Arnold FH. 2006. Protein stability promotes evolvability. *Proc Natl Acad Sci U S A.* 103(15):5869–5874.
- Boucher JJ, Bolon DNA, Tawfik DS. 2016. Quantifying and understanding the fitness effects of protein mutations: laboratory versus nature. *Protein Sci.* 25(7):1219–1226.
- Boyle EA, Li YI, Pritchard JK. 2017. An expanded view of complex traits: from polygenic to omnigenic. *Cell* 169(7):1177–1186.
- Bridgham JT, Carroll SM, Thornton JW. 2006. Evolution of hormone-receptor complexity by molecular exploitation. *Science* 312(5770):97–101.
- Bridgham JT, Ortlund EA, Thornton JW. 2009. An epistatic ratchet constrains the direction of glucocorticoid receptor evolution. *Nature* 461(7263):515–519.
- Bustamante CD, Fedel-Alon A, Williamson S, Nielsen R, Hubisz MT, Glanowski S, Tanenbaum DM, White TJ, Sninsky JJ, Hernandez RD, et al. 2005. Natural selection on protein-coding genes in the human genome. *Nature* 437(7062):1153–1157.
- Canale AS, Cote-Hammarlof PA, Flynn JM, Bolon DNA. 2018. Evolutionary mechanisms studied through protein fitness landscapes. *Curr Opin Struct Biol.* 48:141–148.
- Castellano D, James J, Eyre-Walker A. 2018. Nearly neutral evolution across the *Drosophila melanogaster* genome. *Mol Biol Evol.* 35:2685–2694.
- Chattopadhyay A, Zandarashvili L, Luu RH, Iwahara J. 2016. Thermodynamic additivity for impacts of base-pair substitutions on association of the Egr-1 zinc-finger protein with DNA. *Biochemistry* 55(47):6467–6474.
- Christensen RG, Gupta A, Zuo Z, Schriefer LA, Wolfe SA, Stormo GD. 2011. A modified bacterial one-hybrid system yields improved quantitative models of transcription factor specificity. *Nucleic Acids Res.* 39(12): e83.
- Chou H-H, Delaney NF, Draghi JA, Marx CJ. 2014. Mapping the fitness landscape of gene expression uncovers the cause of antagonism and sign epistasis between adaptive mutations. *PLoS Genet.* 10(2): e1004149.
- Collette F, Renger T, Müh F, Schmidt am Busch M. 2018. Red/green color tuning of visual rhodopsins: electrostatic theory provides a quantitative explanation. *J Phys Chem B.* 122(18):4828–4837.
- Crona K, Greene D, Barlow M. 2013. The peaks and geometry of fitness landscapes. *J Theor Biol.* 317:1–10.
- Dean AM. 1989. Selection and neutrality in lactose operons of *Escherichia coli*. *Genetics* 123(3):441–454.
- Dean AM. 1995. A molecular investigation of genotype by environment interactions. *Genetics* 139:19–33.
- Dean AM, Dvorak L. 1995. The role of glutamate 87 in the kinetic mechanism of *Thermus thermophilus* isopropylmalate dehydrogenase. *Protein Sci.* 4(10):2156–2167.
- Dean AM, Koshland DE Jr. 1993. The kinetic mechanism of *Escherichia coli* isocitrate dehydrogenase. *Biochemistry* 32(36):9302–9309.
- DePristo MA, Weinreich DM, Hartl DL. 2005. Missense meanderings in sequence space: a biophysical view of protein evolution. *Nat Rev Genet.* 6(9):678–687.
- Dobzhansky T. 1964. Biology, molecular and organismic. *Am Zool.* 4:443–452.
- Dobzhansky T. 1973. Nothing in biology makes sense except in the light of evolution. *Am Biol Teacher* 35(3):125–129.
- Domingo-Calap P, Cuevas JM, Sanjuan R. 2009. The fitness effects of random mutations in single-stranded DNA and RNA bacteriophages. *PLoS Genet.* 5(11): e1000742.
- Doolittle WF, Brunet TDP. 2017. On causal roles and selected effects: our genome is mostly junk. *BMC Biol.* 15(1):116.
- Dykhuizen DE, Dean AM. 2004. Evolution of specialists in an experimental microcosm. *Genetics* 167(4):2015–2026.
- Ernst OP, Lodowski DT, Elstner M, Hegemann P, Brown LS, Kandori H. 2014. Microbial and animal rhodopsins: structures, functions, and molecular mechanisms. *Chem Rev.* 114(1):126–163.
- Echave J, Wilke CO. 2017. Biophysical Models of Protein Evolution: Understanding the Patterns of Evolutionary Sequence Divergence. *Annu. Rev. Biophys.* 46:85–103.
- Eyre-Walker A, Keightley PD. 2007. The distribution of fitness effects of new mutations. *Nat Rev Genet.* 8(8):610–618.

- Filion CJ, van Bommel JG, Braunschweig U, Talhout W, Kind J, Ward LD, Brugman W, de Castro IJ, Kerkhoven RM, Bussemaker HJ, et al. 2010. Systematic protein location mapping reveals five principal chromatin types in *Drosophila* cells. *Cell* 143(2):212–224.
- Fincke OM, Hadrys H. 2001. Unpredictable offspring survivorship in the damselfly, *Megalopterus coeruleus*, shapes parental behavior, constrains sexual selection, and challenges traditional fitness estimates. *Evolution* 55(4):762–772.
- Firnberg E, Labonte JW, Gray JJ, Ostermeier M. 2014. A comprehensive, high-resolution map of a gene's fitness landscape. *Mol Biol Evol* 31(6):1581–1592.
- Fisher RA. 1928. The possible modification of the response of the wild type to recurrent mutations. *Am Nat* 62(679):115–126.
- Fisher RA. 1930. The genetical theory of natural selection. Oxford: Oxford University Press.
- Fowler DM, Araya CL, Fleishman SJ, Kellogg EH, Stephany JJ, Baker D, Fields S. 2010. High-resolution mapping of protein sequence-function relationships. *Nat Methods* 7(9):741–746.
- Gavrilets S. 2004. Fitness landscapes and the origin of species. Princeton (NJ): Princeton University Press.
- Gonçalves S, Miller SP, Carrondo MA, Dean AM, Matias PM. 2012. Induced fit and the catalytic mechanism of isocitrate dehydrogenase. *Biochemistry* 51:7089–7115.
- Gong LI, Suchard MA, Bloom JD. 2013. Stability-mediated epistasis constrains the evolution of an influenza protein. *eLife* 2:e00631.
- Gong LI, Bloom JD. 2014. Epistatically interacting substitutions are enriched during adaptive protein evolution. *PLoS Genet* 10(5):e1004328.
- Gordan R, Shen N, Dror I, Zhou T, Horton J, Rohs R, Bulyk ML. 2013. Genomic regions flanking E-box binding sites influence DNA binding specificity of bHLH transcription factors through DNA shape. *Cell Rep* 3(4):1093–1104.
- Gould SJ, Lewontin RC. 1979. The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. *Proc Roy Soc Ser B* 205(1161):581–598.
- Gozem S, Luk HL, Schapiro I, Olivucci M. 2017. Theory and simulation of the ultrafast double-bond isomerization of biological chromophores. *Chem Rev* 117(22):13502–13565.
- Graur D, Zheng Y, Price N, Azevedo RBR, Zufall RA, Elhaik E. 2013. On the immortality of television sets: “function” in the human genome according to the evolution-free gospel of ENCODE. *Genome Biol Evol* 5(3):578–590.
- Gregoret LM, Sauer RT. 1993. Additivity of mutant effects assessed by binomial mutagenesis. *Proc Natl Acad Sci U S A* 90(9):4246–4250.
- Grossman SR, Zhang X, Wang L, Engreitz J, Melnikov A, Rogov P, Tewhey R, Isakova A, Deplancke B, Bernstein BE, et al. 2017. Systematic dissection of genomic features determining transcription factor binding and enhancer function. *Proc Natl Acad Sci U S A* 114(7):E1291–E1300.
- Haldane JBS. 1930. A note on Fisher's theory of the origin of dominance, and on a correlation between dominance and linkage. *Am Nat* 64(690):87–90.
- Haldane JBS. 1939. The theory of the evolution of dominance. *J Genet* 37(2):365–374.
- Hartl DL, Dykhuizen DE, Dean AM. 1985. The limits of adaptation: the evolution of selective neutrality. *Genetics* 111(3):655–674.
- Hartl DL, Hall BG. 1974. Second naturally occurring beta-galactosidase in *E. coli*. *Nature* 248(5444):152–153.
- Honig B, Dinur U, Nakanishi K, Balogh-Nair V, Gawinowicz MA, Arnaboldi M, Motto MG. 1979. An external point-charge model for wavelength regulation in visual pigments. *J Am Chem Soc* 101(23):7084–7086.
- Honig B, Greenberg AD, Dinur U, Ebrely TG. 1976. Visual-pigment spectra: implications of the protonation of the retinal Schiff base. *Biochemistry* 15(21):4593–4599.
- Huber CD, Durvasula A, Hancock AM, Lohmueller KE. 2018. Gene expression drives the evolution of dominance. *Nat Commun* 9(1):2750.
- Ikemura T. 1981. Correlation between the abundance of *Escherichia coli* transfer RNAs and the occurrence of the respective codons in its protein genes: a proposal for a synonymous codon choice that is optimal for the *E. coli* translational system. *J Mol Biol* 151(3):389–409.
- Jacquier H, Birgy A, Le Nagard H, Mechulam Y, Schmitt E, Glodt J, Bercot B, Petit E, Poulain J, Barnaud G, et al. 2013. Capturing the mutational landscape of the beta-lactamase TEM-1. *Proc Natl Acad Sci U S A* 110(32):13067–13072.
- Jolma A, Yin Y, Nitta KR, Dave K, Popov A, Taipale M, Enge M, Kivioja T, Morgunova E, Taipale J. 2015. DNA-dependent formation of transcription factor pairs alters their binding specificity. *Nature* 527(7578):384–388.
- Kacser H, Burns JA. 1973. The control of flux. *Symp Soc Exp Biol* 32:65–104.
- Kacser H, Burns JA. 1981. The molecular basis of dominance. *Genetics* 97(3–4):639–666.
- Kaltenbach M, Jackson CJ, Campbell EC, Hollfelder F, Tokuriki N. 2015. Reverse evolution leads to genotypic incompatibility despite functional and active site convergence. *eLife* 4:e06492.
- Kauffman S, Levin S. 1987. Towards a general theory of adaptive walks on rugged landscapes. *J Theor Biol* 128(1):11–45.
- Kimura M. 1983. The neutral theory of molecular evolution. Cambridge: Cambridge University Press.
- Kingsolver JG, Hoekstra HE, Hoekstra JM, Berrigan D, Vignieri SN, Hill CE, Hoang A, Gibert P, Beerli P. 2001. The strength of phenotypic selection in natural populations. *Am Nat* 157(3):245–261.
- Kouyos RD, Leventhal GE, Hinkley T, Haddad M, Whitcomb JM, Petropoulos CJ, Bonhoeffer S. 2012. Exploring the complexity of the HIV-1 fitness landscape. *PLoS Genet* 8(3): e1002551.
- Kryazhimskiy S, Dushoff J, Bazykin GA, Plotkin JB. 2011. Prevalence of epistasis in the evolution of influenza A surface proteins. *PLoS Genet* 7(2): e1001301.
- Lagator M, Paixa T, Barton NH, Bollback JP, Guet CC. 2017. On the mechanistic nature of epistasis in a canonical cis-regulatory element. *eLife* 6:e25192.
- Lande R. 1976. Natural selection and random genetic drift in phenotypic evolution. *Evolution* 30(2):314–334.
- Lande R. 1979. Quantitative genetic analysis of multivariate evolution, applied to brain: body size allometry. *Evolution* 33(1 Part 2):402–416.
- Lande R, Arnold SJ. 1983. The measurement of selection on correlated characters. *Evolution* 37(6):1210–1226.
- Le DD, Shimko TC, Aditham AK, Keys AM, Longwell SA, Orenstein Y, Fordyce PM. 2018. Comprehensive, high-resolution binding energy landscapes reveal context dependencies of transcription factor binding. *Proc Natl Acad Sci U S A* 115(16): E3702–E3711.
- Levinthal C. 1969. How to fold graciously. Mossbauer spectroscopy in biological systems: proceedings of a meeting held at Allerton House, Monticello, Illinois. *Univ Ill Bull* 67:22–24.
- Levo M, Zalckvar E, Sharon E, Dantas Machado AC, Kalma Y, Lotampompan M, Weinberger A, Yakhini Z, Rohs R, Segal E. 2015. Unraveling determinants of transcription factor binding outside the core binding site. *Genome Res* 25(7):1018–1029.
- Lewontin RC, White MJD. 1960. Interaction between inversion polymorphism of two chromosome pairs in the grasshopper, *Moraba scurra*. *Evolution* 14:116–129.
- Li C, Qian W, Maclean CJ, Zhang J. 2016. The fitness landscape of a tRNA gene. *Science* 352(6287):837–840.
- Lundin E, Tang P-C, Guy L, Näsvall J, Andersson DI. 2018. Experimental determination and prediction of the fitness effects of random point mutations in the biosynthetic enzyme HisA. *Mol Biol Evol* 35(3):704–718.
- Lunzer M, Golding GB, Dean AM. 2010. Pervasive cryptic epistasis in molecular evolution. *PLoS Genet* 6(10):e1001162.
- Lunzer M, Miller SP, Felsheim R, Dean AM. 2005. The biochemical architecture of an ancient adaptive landscape. *Science* 310(5747):499–501.
- Lunzer M, Natarajan A, Dykhuizen DE, Dean AM. 2002. Enzyme kinetics, substitutable resources and competition: from biochemistry to frequency-dependent selection in *lac*. *Genetics* 162(1):485–499.

- Maerkl SJ, Quake SR. 2007. A systems approach to measuring the binding energy landscapes of transcription factors. *Science* 315(5809):233–237.
- Marciano DC, Pennington JM, Wang X, Wang J, Chen Y, Thomas VL, Shoichet BK, Palzkill T. 2008. Genetic and structural characterization of an L201P global suppressor substitution in TEM-1 beta-lactamase. *J Mol Biol.* 384(1):151–164.
- Mavor D, Barlow K, Thompson S, Barad BA, Bonny AR, Cario CL, Gaskins G, Liu Z, Deming L, Axen SD, et al. 2016. Determination of ubiquitin fitness landscapes under different chemical stresses in a classroom setting. *eLife* 5:e15802.
- Mayr E. 1961. Cause and effect in biology. *Science* 134(3489):1501–1506.
- Melamed D, Young DL, Gamble CE, Miller CR, Fields S. 2013. Deep mutational scanning of an RRM domain of the *Saccharomyces cerevisiae* poly(A)-binding protein. *RNA* 19(11):1537–1551.
- Meng X, Brodsky MH, Wolfe SA. 2005. A bacterial one-hybrid system for determining the DNA-binding specificity of transcription factors. *Nat Biotechnol.* 23(8):988–994.
- Merbs SL, Nathans J. 1992. Absorption spectrum of human cone pigments. *Nature* 356(6368):433–435.
- Miller JH, Schmeissner U. 1979. Genetic studies of the lac repressor. X. Analysis of missense mutations in the *lacI* gene. *J Mol Biol.* 131(2):223–248.
- Miller SP, Lunzer M, Dean AM. 2006. Direct demonstration of an adaptive constraint. *Science* 314(5798):458–461.
- Mira PM, Crona K, Greene D, Meza JC, Sturmfels B, Barlow M. 2015. Rational design of antibiotic treatment plans: a treatment strategy for managing evolution and reversing resistance. *PLoS One* 10(9): e0139387.
- Mira PM, Meza JC, Nandipati A, Barlow M. 2015. Adaptive landscapes of resistance genes change as antibiotic concentrations change. *Mol Biol Evol.* 32(10):2707–2715.
- Mitchell-Olds T, Shaw RG. 1987. Regression analysis of natural selection: statistical inference and biological interpretation. *Evolution* 41(6):1149–1161.
- Morange M. 2011. What will result from the interaction between functional and evolutionary biology? *Stud Hist Philos Biol Biomed Sci.* 42(1):69–74.
- Natarajan C, Inoguchi N, Weber RE, Fago A, Moriyama H, Storz JF. 2013. Epistasis among adaptive mutations in deer mouse hemoglobin. *Science* 340(6138):1324–1327.
- Neitz M, Neitz J, Jacobs GH. 1991. Spectral tuning of pigments underlying red-green color vision. *Science* 252(5008):971–974.
- Ohta T. 1973. Slightly deleterious mutant substitutions in evolution. *Nature* 246(5428):96–98.
- Ohta T. 1977. Extension of the neutral mutation drift hypothesis. In: Kimura M, editor. *Molecular evolution and polymorphism*. Mishima (Japan): National Institute of Genetics. p. 148–167.
- Ohta T. 1992. The nearly neutral theory of molecular evolution. *Annu Rev Ecol Syst.* 23(1):263–286.
- Olson CA, Wu NC, Sun R. 2014. A comprehensive biophysical description of pairwise epistasis throughout an entire protein domain. *Curr Biol.* 24(22):2643–2651.
- Orenstein Y, Shamir R. 2014. A comparative analysis of transcription factor binding models learned from PBM, HT-SELEX and ChIP data. *Nucleic Acids Res.* 42:e6.
- Orr HA. 1998. The population genetics of adaptation: the distribution of factors fixed during adaptive evolution. *Evolution* 52(4):935–949.
- Ortlund EA, Bridgham JT, Redinbo MR, Thornton JW. 2007. Crystal structure of an ancient protein: evolution of a new function by conformational epistasis. *Science* 317(5844):1544–1548.
- Palló A, Oláh J, Gráczner E, Merli A, Závodszy P, Weiss MS, Vas M. 2014. Structure and energetic basis of isopropylmalate dehydrogenase enzyme catalysis. *FEBS J.* 281(22):5063–5076.
- Pekkala N, Kotiaho JS, Puurtinen M. 2011. Laboratory relationships between adult lifetime reproductive success and fitness surrogates in a *Drosophila littoralis* population. *PLoS One* 6(9): e24560.
- Pelletier F, Garant D, Hendry AP. 2009. Eco-evolutionary dynamics. *Philos Trans R Soc B.* 364(1523):1483–1489.
- Pfahl M. 1976. lac repressor-operator interaction. Analysis of the X86 repressor mutant. *J Mol Biol.* 106(3):857–869.
- Phillips PC, Arnold SJ. 1989. Visualizing multivariate selection. *Evolution* 43(6):1209–1222.
- Poelwijk FJ, de Vos MGJ, Tans SJ. 2011. Tradeoffs and optimality in the evolution of gene regulation. *Cell* 146(3):462–470.
- Poelwijk FJ, Tănase-Nicola S, Kiviet DJ, Tans SJ. 2011. Reciprocal sign epistasis is a necessary condition for multi-peaked fitness landscapes. *J Theor Biol.* 272(1):141–144.
- Pressman A, Blanco C, Chen IA. 2015. The RNA world as a model system to study the origin of life. *Curr Biol.* 25(19): R953–R963.
- Privalov PL, Khechinashvili NN. 1974. A thermodynamic approach to the problem of stabilization of globular protein structure: a calorimetric study. *J Mol Biol.* 86(3):665–684.
- Provine WB. 1986. Sewall Wright and evolutionary biology. Chicago (IL): University of Chicago Press.
- Puchta O, Cseke B, Czaja H, Tollervey D, Sanguinetti G, Kudla G. 2016. Network of epistatic interactions within a yeast snoRNA. *Science* 352(6287):840–844.
- Rainey PB, Travisano M. 1998. Adaptive radiation in a heterogeneous environment. *Nature* 394(6688):69–72.
- Reed DH, Bryant EH. 2004. Phenotypic correlations among fitness and its components in a population of the housefly. *J Evol Biol.* 17(4):919–923.
- Robert L, Ollion J, Robert J, Song X, Matic I, Elez M. 2018. Mutation dynamics and fitness effects followed in single cells. *Science* 359(6381):1283–1286.
- Romero PA, Arnold FH. 2009. Exploring protein fitness landscapes by directed evolution. *Nat Rev Mol Cell Biol.* 10(12):866–876.
- Roscoe BP, Thayer KM, Zeldovich KB, Fushman D, Bolon DN. 2013. Analyses of the effects of all ubiquitin point mutants on yeast growth rate. *J Mol Biol.* 425(8):1363–1377.
- Rosenzweig RF, Sharp RR, Treves DS, Adams J. 1994. Microbial evolution in a simple unstructured environment: genetic differentiation in *Escherichia coli*. *Genetics* 137:903–917.
- Sakmar TP, Franke RR, Khorana HG. 1989. Glutamic acid-113 serves as the retinylidene Schiff base counterion in bovine rhodopsin. *Proc Natl Acad Sci U S A.* 86(21):8309–8313.
- Sandberg WS, Terwilliger TC. 1991. Energetics of repacking a protein interior. *Proc Natl Acad Sci U S A.* 88(5):1706–1710.
- Sarkisyan KS, Bolotin DA, Meer MV, Usmanova DR, Mishin AS, Sharonov GV, Ivankov DN, Bozhanova NG, Baranov MS, Soylemez O, et al. 2016. Local fitness landscape of the green fluorescent protein. *Nature* 533(7603):397–401.
- Sekharan S, Altun A, Morokuma K. 2010. Photochemistry of visual pigment in a Gq protein-coupled receptor (GPCR)—insights from structural and spectral tuning studies on squid rhodopsin. *Chem-Eur J.* 16(6):1744–1749.
- Sekharan S, Katayama K, Kandori H, Morokuma K. 2012. Color vision: “OH-Site” rule for seeing red and green. *J Am Chem Soc.* 134(25):10706–10712.
- Sekharan S, Mooney VL, Rivalta I, Kazmi MA, Neitz M, Neitz J, Sakmar TP, Yan ECY, Batista VS. 2013. Spectral tuning of ultraviolet cone pigments: an interhelical lock mechanism. *J Am Chem Soc.* 135(51):19064–19067.
- Sekharan S, Yokoyama S, Morokuma K. 2011. Quantum mechanical/molecular mechanical structure, enantioselectivity, and spectroscopy of hydroxyretinals and insights into the evolution of color vision in small white butterflies. *J Phys Chem B.* 115(51):15380–15388.
- Sharp PM, Li W-H. 1987. The codon adaptation index—a measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Res.* 15(3):1281–1295.
- Siddiq MA, Hochberg GK, Thornton JW. 2017. Evolution of protein specificity: insights from ancestral protein reconstruction. *Curr Opin Struct Biol.* 47:113–122.

- Silander OK, Tenaillon O, Chao L. 2007. Understanding the evolutionary fate of finite populations: the dynamics of mutational effects. *PLoS Biol.* 5(4): e94.
- Simpson GG. 1944. Tempo and mode in evolution. New York: Columbia University Press.
- Soskine M, Tawfik DS. 2010. Mutational effects and the evolution of new protein functions. *Nature Rev Genet.* 11(8):572–582.
- Stiffler MA, Hekstra DR, Ranganathan R. 2015. Evolvability as a function of purifying selection in TEM-1 β -lactamase. *Cell* 160(5):882–892.
- Stormo GD. 2013. Modeling the specificity of protein–DNA interactions. *Quant Biol.* 1(2):115–130.
- Svensson EI, Calsbeek R. 2012. The adaptive landscape in evolutionary biology. Oxford: Oxford University Press.
- Swiderek K, Ruiz-Pernía JJ, Moliner V, Tuñón I. 2014. Heavy enzymes—experimental and computational insights in enzyme dynamics. *Curr Opin Chem Biol.* 21:11–18.
- Szendro IG, Schenk MF, Franke J, Krug J, de Visser JAGM. 2013. Quantitative analyses of empirical fitness landscapes. *J Stat Mech Theory Exp.* 2013: P01005.
- Tada T, Altun A, Yokoyama S. 2009. Evolutionary replacement of UV vision by violet vision in fish. *Proc Natl Acad Sci U S A.* 106(41):17457–17462.
- Thyagarajan B, Bloom JD. 2014. The inherent mutational tolerance and antigenic evolvability of influenza hemagglutinin. *eLife* 3:e03300.
- Tinbergen N. 1963. On aims and methods of ethology. *Z Tierpsychol.* 20(4):410–433.
- Tokuriki N, Stricher F, Serrano L, Tawfik DS. 2008. How protein stability and new functions trade off. *PLoS Comput Biol.* 4(2): e1000002.
- Tokuriki N, Tawfik DS. 2009. Stability effects of mutations and protein evolvability. *Curr Opin Struct Biol.* 19(5):596–604.
- Tripathi A, Gupta K, Khare S, Jain PC, Patel S, Kumar P, Pulianmackal AJ, Aghera N, Varadarajan R. 2016. Molecular determinants of mutant phenotypes inferred from saturation mutagenesis data. *Mol Biol Evol.* 33(11):2960–2975.
- Weinreich DM, Delaney NF, DePristo MA, Hartl DL. 2006. Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science* 312(5770):111–114.
- Weinreich DM, Lan Y, Wylie CS, Heckendorn RB. 2013. Should evolutionary geneticists worry about higher-order epistasis? *Curr Opin Genet Dev.* 23(6):700–707.
- Weinreich DM, Watson RA, Chao L. 2005. Sign epistasis and genetic constraint on evolutionary trajectories. *Evolution* 59(6):1165–1174.
- Weirauch MT, Cote A, Norel R, Annala M, Zhao Y, Riley TR, Saez-Rodriguez J, Cokelaer T, Vedenko A, Talukder S, et al. 2013. Evaluation of methods for modeling transcription factor sequence specificity. *Nat Biotechnol.* 31(2):126–134.
- Wells JA. 1990. Additivity of mutational effects in proteins. *Biochemistry* 29(37):8509–8517.
- Wong D, Teixeira A, Oikonomopoulos S, Humburg P, Lone IN, Saliba D, Siggers T, Bulyk M, Angelov D, Dimitrov S, et al. 2011. Extensive characterization of NF-kappaB binding uncovers non-canonical motifs and advances the interpretation of genetic functional traits. *Genome Biol.* 12(7): R70.
- Wood CW, Brodie ED. 2015. Environmental effects on the structure of the G-matrix. *Evolution* 69(11):2927–2940.
- Wrenbeck EE, Azouz LR, Whitehead TA. 2017. Single-mutation fitness landscapes for an enzyme. *Nat Commun.* 8:15695.
- Wrenbeck EE, Faber MS, Whitehead TA. 2017. Deep sequencing methods for protein engineering and design. *Curr Opin Struct Biol.* 45:36–44.
- Wright S. 1929. Fisher's theory of dominance. *Am Nat.* 63(686):274–279.
- Wright S. 1932. The roles of mutation, cross-breeding and selection in evolution. *Proc. 6th Annu. Cong. Genet.*, 1, 356–366. Reprinted in W. B. Provine (1986). Sewall Wright: evolution: selected papers. Chicago (IL): University of Chicago Press. p. 161–177.
- Wright S. 1934. Physiological and evolutionary theories of dominance. *Am Nat.* 67:24–53.
- Wu NC, Dai L, Olson CA, Lloyd-Smith JO, Sun R. 2016. Adaptation in protein fitness landscapes is facilitated by indirect paths. *eLife* 5:e16965.
- Wu NC, Young AP, Al-Mawsawi LQ, Olson CA, Feng J, Qi H, Chen SH, Lu IH, Lin CY, Chin RG, et al. 2014. High-throughput profiling of influenza A virus hemagglutinin gene at single-nucleotide resolution. *Sci Rep.* 4:4942.
- Yi X, Dean AM. 2016. Phenotypic plasticity as an adaptation to a functional trade-off. *eLife* 5:e19307.
- Yokoyama R, Yokoyama S. 1990. Convergent evolution of the red- and green-like visual pigment genes in fish, *Astyanax fasciatus*, and human. *Proc Natl Acad Sci U S A.* 87(23):9315–9318.
- Yokoyama S. 2008. Evolution of dim-light and color vision pigments. *Annu Rev Genomics Human Genet.* 9:259–282.
- Yokoyama S, Radlwimmer FB. 1998. The “five-sites” rule and the evolution of red and green color vision in mammals. *Mol Biol Evol.* 15(5):560–567.
- Yokoyama S, Radlwimmer FB. 2001. The molecular genetics and evolution of red and green color vision in vertebrates. *Genetics* 158(4):1697–1710.
- Yokoyama S, Yang H, Starmer WT. 2008. Molecular basis of spectral tuning in the red- and green-sensitive (MLWS) pigments in vertebrates. *Genetics* 179(4):2037–2043.
- Zhong S, Khodursky A, Dykhuizen DE, Dean AM. 2004. Evolutionary genomics of ecological specialization. *Proc Natl Acad Sci U S A.* 101(32):11719–11724.
- Zhong S, Miller SP, Dykhuizen DE, Dean AM. 2009. Transcription, translation, and the evolution of specialists and generalists. *Mol Biol Evol.* 26(12):2661–2678.
- Zhu G, Golding GB, Dean AM. 2005. The selective cause of an ancient adaptation. *Science* 307(5713):1279–1282.
- Zhukovsky EA, Oprian DD. 1989. Effect of carboxylic acid side chains on the absorption maximum of visual pigments. *Science* 246(4932):928–930.
- Zuo Z, Stormo GD. 2014. High-resolution specificity from DNA sequencing highlights alternative modes of Lac repressor binding. *Genetics* 198(3):1329–1343.
- Zykovich A, Korf I, Segal DJ. 2009. Bind-n-Seq: high-throughput analysis of in vitro protein–DNA interactions using massively parallel sequencing. *Nucleic Acids Res.* 37(22): e151.