University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Jay F. Storz Publications

Papers in the Biological Sciences

2011

Experimental Approaches to Evaluate the Contributions of Candidate Protein-Coding Mutations to Phenotypic Evolution

Jay F. Storz
University of Nebraska - Lincoln, jstorz2@unl.edu

Anthony J. Zera
University of Nebraska - Lincoln, azera1@unl.edu

Experimental Approaches to Evaluate the Contributions of Candidate Protein-Coding Mutations to Phenotypic Evolution

Jay F. Storz and Anthony J. Zera

University of Nebraska–Lincoln

Abstract

Identifying mechanisms of molecular adaptation can provide important insights into the process of phenotypic evolution, but it can be exceedingly difficult to quantify the phenotypic effects of specific mutational changes. To verify the adaptive significance of genetically based changes in protein function, it is necessary to document functional differences between the products of derived and wild-type alleles and to demonstrate that such differences impinge on higher-level physiological processes (and ultimately, fitness). In the case of metabolic enzymes, this requires documenting in vivo differences in reaction rate that give rise to differences in flux through the pathway in which the enzymes function. These measured differences in pathway flux should then give rise to differences in cellular or systemic physiology that affect fitness-related variation in whole-organism performance. Efforts to establish these causal connections between genotype, phenotype, and fitness require experiments that carefully control for environmental variation and background genetic variation. Here, we discuss experimental approaches to evaluate the contributions of amino-acid mutations to adaptive phenotypic change. We discuss conceptual and methodological issues associated with in vitro and in vivo studies of protein function, and the evolutionary insights that can be gleaned from such studies. We also discuss the importance of isolating the effects of individual mutations to distinguish between positively selected substitutions that directly contribute to improvements in protein function versus positively selected, compensatory substitutions that mitigate negative pleiotropic effects of antecedent changes.

Keywords: adaptation, biochemical adaptation, compensatory evolution, enzyme kinetics, enzyme polymorphism, metabolic adaptation, molecular adaptation

1. Introduction

To assess whether a given protein-coding mutation has contributed to adaptive phenotypic change, it is ultimately necessary to document and explain the mechanistic basis of fitness variation among alternative single-lo-

cus genotypes. This is a tall order because adaptive changes in many phenotypic traits may be attributable to fitness differentials that fall well below the resolving power of "real-time" observational studies. As stated by Gillespie (1), "Selection coefficients for single amino acid substitutions as small as 10^{-4} to 10^{-3} are large enough to dominate genetic drift, yet are refractory to direct experimental investigations. In other words, most of protein evolution could be due to strong natural selection, yet we have no experimental protocol capable of measuring the selective differences." However, fitness differences among alternative genotypes ultimately stem from allelic differences in the concentration and/or biochemical properties of the encoded protein. It will often be possible to measure these allelic differences in protein concentration or protein function even in cases where the net effects on fitness lie below the threshold of experimental detection (2). At the most proximal level of trait variation, in vitro tests of protein function can provide mechanistic insights into the causal link between genotype and biochemical phenotype. After establishing this link – and after confirming that experimental results obtained in vitro are also manifest in vivo – the next challenge is to assess whether the observed change in biochemical phenotype impinges on whole-organism performance in a way that affects fitness.

Identifying the mechanisms by which allelic protein variants contribute to adaptation is a difficult task that requires multiple types of information, at multiple levels of biological organization. For example, adaptive changes in enzyme function can involve modifications of any one of several enzymatic properties such as catalytic efficiency, concentration, stability, and sensitivity to metabolic regulators, and each aspect can be altered in a number of different ways (Figure 1). Moreover, changes in enzyme function do not necessarily affect higher-level physiological phenotypes (e.g., pathway flux).
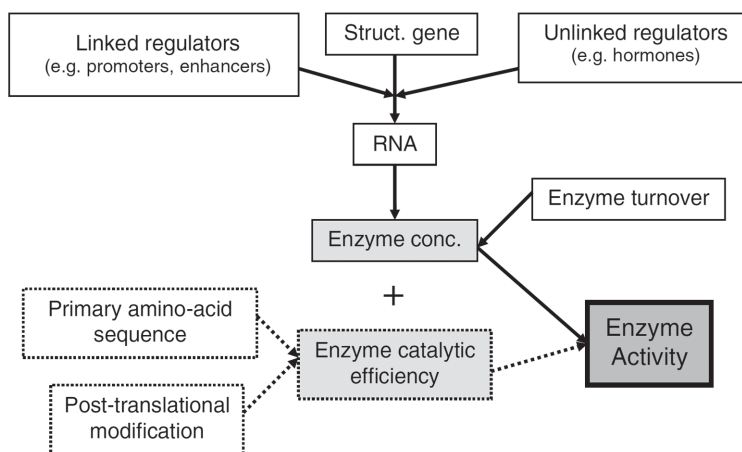

Figure 1. Various factors affecting enzyme activity via influence on enzyme concentration or catalytic efficiency. Modified from Reference 45.

Population-genetic studies of protein polymorphism have traditionally focused on kinetic and thermodynamic properties of metabolic enzymes – parameters that had been identified as key aspects of enzyme adaptation by comparative biochemists (3) and enzymologists (4, 5). This field of study eventually expanded to include investigations into regulatory changes in enzyme concentration as well as enzyme function (6, 7). The importance of experimentally confirming the physiological consequences of protein polymorphism (8–13) came into especially sharp focus with the development of Metabolic Control Analysis, which predicts that change in the catalytic activity of an individual enzyme can often have negligible effects on steady-state pathway flux (14, 15). More recently, advances in molecular biology have permitted more detailed investigations into the regulation of protein-coding genes (e.g., measurement of transcript abundance) and the relationship between protein structure and function.

1.1. The Importance of Elucidating Mechanism in Studies of Adaptation

Once a given protein-coding gene has been implicated in adaptive phenotypic change, why is it important to identify the biochemical/ physiological mechanisms by which the encoded protein exerts its effects on organismal performance? Beyond the obvious point that such information enriches our basic understanding of the biology of adaptation, insights into the mechanistic basis of genetic adaptation can also shed light on why particular genes or particular types of mutations make disproportionate contributions to phenotypic evolution. For example, an understanding of how variation in enzyme activity maps onto fitness can potentially explain why regulatory and structural mutations differ in their relative contributions to adaptive change in flux-based phenotypes, and may also explain why the two classes of mutation have different fixation probabilities at different time points in multistep evolutionary pathways. The stepwise acquisition of cefotaxime resistance in *Escherichia coli* involves a combination of coding and noncoding substitutions in the β-lactamase gene, and the initial steps are consistently attributable to amino-acid substitutions that enhance the catalytic efficiency of the enzyme (16, 17). Regulatory mutants are typically incorporated later in the evolutionary pathway because an increased expression of the enzyme only becomes advantageous once its catalytic efficiency has been optimized by the antecedent structural changes (17). In contrast to the evolution of antibiotic resistance, in the evolution of flux-limited metabolic pathways, mutations that increase gene expression may often be fixed early because an increased enzyme concentration can have a large effect on fitness (18).

Mechanistic studies of genetic adaptation can also provide insights into features of adaptive mutations – such as dominance, epistatic interactions, and pleiotropic effects – that exert a strong influence on trajecto-

ries of evolutionary change. A good example is provided by recent studies on the mechanistic basis of adaptive crypsis in lizards from the Chihuahuan desert (19). In several codistributed species of lizards that inhabit the white gypsum dunes of White Sands, New Mexico, the convergent evolution of blanched, substrate-matching coloration has involved different loss-of-function mutations in the melanocortin-1 receptor gene (*Mc1r*). In the little striped whiptail lizard, *Aspidoscelis inornata*, the blanched phenotype is associated with an amino-acid mutation in *Mc1r* that impairs receptor signaling, and in the eastern fence lizard, *Sceloporus undulatus*, the same phenotype is associated with a different amino-acid mutation in *Mc1r* that prevents the receptor from integrating into the melanocyte membrane. Both loss-of-function mutations produce the same blanched phenotype, but the mutation in *A. inornata* is recessive (because the disruption in receptor signaling is compensated by the wild-type allele), whereas the mutation in *S. undulatus* is dominant (because in heterozygotes, cell-surface expression of the wild-type receptor is suppressed as a result of dimerization with the mutant copy). Although both lizard species are presumably subject to similar selection regimes (mediated by visually oriented predators), the differences in dominance of the adaptive mutations are reflected by different spatial patterns of allele frequency variation across ecotonal transitions in substrate color. In these two lizard species, observed differences in the geographic patterning of *Mc1r* allele frequency variation only make sense in light of experimental findings that revealed differences in the penetrance of independently derived, loss-of-function mutations that impair different aspects of receptor function.

As the above examples illustrate, identifying mechanisms of molecular adaptation can provide key insights into the process of phenotypic evolution. Here, we discuss experimental approaches to evaluate the contributions of protein-coding mutations to adaptive phenotypic change. We discuss conceptual and methodological issues associated with in vitro and in vivo studies of protein function, and the evolutionary insights that can be gleaned from such studies. We focus largely on enzymes of intermediary metabolism, as this class of proteins has figured prominently in research efforts to establish causal connections between genotype, phenotype, and fitness in ecologically relevant contexts (9–12, 20). Although we focus primarily on the contributions of amino-acid mutations to adaptive phenotypic evolution, the issues and experimental approaches that we discuss are also applicable to assessments of nonadaptive or maladaptive change.

2. In Vitro Studies of Protein Function: Linking Genotype to Biochemical Phenotype

The main function of an enzyme is to catalyze (increase) the rate of a chemical reaction. Thus, the biochemical phenotypes of interest in evolutionary studies of enzyme polymorphism include catalytic efficiency and enzyme concentration, the two most important factors that influence the rate of an enzyme-catalyzed reaction. Catalytic efficiency, which is an intrinsic kinetic property of an enzyme, can be broadly defined as the rate at which substrate is converted into product per active site of an enzyme molecule under a specified set of conditions. It is important to keep in mind that noncatalytic aspects of enzyme function, such as inhibition or activation by various metabolic regulators, may also contribute to enzymatic adaptation independent of catalytic efficiency (e.g., (21)).

2.1. Enzyme Kinetics

The proper interpretation and measurement of kinetic parameters are best considered in the context of the Michaelis–Menten equation, the basic equation of enzyme kinetics (5, 22, 23). In its simplest form, this equation describes the velocity of an enzyme-catalyzed reaction involving a single substrate that is irreversibly converted into product as a function of two experimentally measurable enzymatic parameters, V_{max} and K_M:

$$v = \frac{V_{max}\ [S]}{K_M + [S]} \tag{1}$$

where V_{max} = maximal velocity, [S] = substrate concentration, K_M = the Michaelis constant, defined as the substrate concentration that results in half-maximal velocity, and v = reaction rate. In addition, $V_{max} = [E] \times k_{cat}$; that is, maximal velocity consists of two components: enzyme concentration ([E]), and the kinetic constant, k_{cat}, or turnover number (i.e., the maximum number of substrate molecules converted to product per active site of an enzyme per unit time). The right-hand side of the above equation can be decomposed into separate components that reflect (a) enzyme concentration [E] and (b) intrinsic kinetic properties of the enzyme (terms on the right-hand side of the equation below):

$$v = [E] \times \frac{k_{cat} \times [S]}{K_M + [S]} \tag{2}$$

The most thorough in vitro kinetic comparisons of allelic enzyme variants (allozymes) have measured and standardized [E]. By controlling for variation in [E], it is then possible to quantify kinetic differences between alternative allozymes. What are the functional meanings of the kinetic parameters in Equa-

tion 2? The answer is a bit complex for K_M, which has often been incorrectly interpreted as a measure of an enzyme's affinity for its substrate. This may or may not be the case depending upon the specific enzyme and substrate under consideration (5, 22–24). In the absence of direct data on substrate binding, K_M is best considered a kinetic component of catalytic efficiency, the other kinetic component being k_{cat}. The relative contribution of k_{cat} and K_M to the catalytic rate is determined by the substrate concentration [S]. When [S] is very low (<0.1 K_M), then [S] in the denominator of Equation 2 can be ignored, and the kinetic part of the equation reduces to $v = (k_{cat}/K_M) \times$ [S]. As [S] increases to intermediate levels (i.e., [S] = K_M), both k_{cat}/K_M and k_{cat} contribute to the rate of catalysis. Because [S] is thought to range from 1.0 to 0.1 K_M in vivo for most enzymes of intermediary metabolism (5), the main kinetic contributor to catalytic efficiency is k_{cat}/K_M alone or some combination of k_{cat}/K_M and k_{cat}. The most thorough kinetic analyses of allozymes have focused on these kinetic constants (25–28).

Many enzymes catalyze complex reactions that cannot be adequately described by Equation 1. For example, dehydrogenase enzymes bind more than one substrate (e.g., a cofactor such as NAD^+ or $NADP^+$, as well as a substrate specific for the particular enzyme). Although the rate equations for such enzymes contain additional terms, simplifying assumptions often have been used to reduce these more complex equations to the form of Equation 1 (e.g., (25)). Many enzymes also catalyze reversible reactions, but again, the analyses have often been simplified by considering the reaction in only one direction.

A more comprehensive and realistic analysis of enzyme catalytic efficiency has been formulated by Albery and Knowles (4), which does not require any simplifying assumptions regarding the relative magnitude of K_M and [S]. Rather than focusing on k_{cat}/K_M alone, estimated values for k_{cat} and K_M are used in the full Michaelis– Mention equation to compare reaction velocities of enzymes across a range of substrate concentrations that occur in vivo. In addition, the approach of Albery and Knowles (4) considers the reaction velocity in both directions, and the primary focus is the net conversion of substrate to product in one direction. This more comprehensive type of analysis was used in a pioneering study of the PGI (phosphoglucose isomerase) polymorphism in *Mytilus edulis* (26, 29).

Finally, it is important to understand the limitations associated with measures of V_{max}/K_M. Standard steady-state kinetic analysis provides estimates of K_M and V_{max}, from which k_{cat} is estimated using the relationship: $k_{cat} = V_{max}/[E]$ (see above). However, estimation of k_{cat} requires knowledge of enzyme concentration [E], which in turn requires completely purified enzyme. Since the purification of enzymes to homogeneity is often nontrivial, some studies have focused on allelic differences in V_{max}/K_M in the absence of information on k_{cat}/K_M (30, 31). The problem with this approach is that

allelic differences in V_{max}/K_M do not necessarily reflect intrinsic functional differences between the alternative allozymes. As discussed above, V_{max} is a composite parameter that reflects kinetic properties of the enzyme (k_{cat}) as well as enzyme concentration [E]. Furthermore, genetic variation in [E] can be caused by background genetic variation independent of the structural locus that encodes the allozyme (e.g., *cis*- or *trans*-acting regulatory factors; (32)). Thus, while allelic differences in k_{cat}/K_M are due to intrinsic kinetic differences between the allozymes, the same is not necessarily true for measured differences in V_{max}/K_M. Although measures of V_{max}/K_M are important for understanding the relative catalytic power of allozymes in vivo due to the combined effects of enzyme kinetics and enzyme concentration, it is important to quantify the relative contributions of [E] and k_{cat}/K_M to variation in V_{max}/K_M.

A discussion of experimental methods for estimating enzymatic parameters is beyond the scope of this article (for more detailed treatments, see refs. 22 and 29). A key aspect of in vitro kinetic measurements is that assay conditions should mimic in vivo conditions as much as possible with respect to pH, temperature-pH relations, ionic strength, the presence of various osmolytes, etc. (3, 23, 29). In addition to catalytic aspects of enzyme function, regulatory properties (e.g., inhibition by various metabolites) may often represent key factors in enzyme adaptation. For example, PGI allozymes of the sea anemone *Metridium senile* are differentially inhibited by the pentose-shunt metabolite 6-phosphogluconate (21). This allozyme-dependent inhibition results in differential diversion of carbon through the pentose shunt versus glycolysis, possibly playing a role in allozyme-dependent lipid biosynthesis. Finally, it is important to make biochemical comparisons between enzyme variants of known amino-acid sequence. Allozymes were originally characterized by their electrophoretic mobility, and it is now widely appreciated that electromorphs often contain a substantial amount of amino-acid sequence heterogeneity that could have functional consequences (20, 33). Failure to adequately deal with "cryptic variation" can result in erroneous conclusions regarding adaptive biochemical differences between allozymes and the geographic patterning of allele frequency variation (34, 35). Many classic studies of allozyme polymorphism have involved kinetic characterizations of PGI electromorphs of unknown amino-acid sequence (21, 26, 30, 31), and subsequent DNA-based studies of PGI polymorphism in a number of different animal taxa have revealed substantial levels of cryptic amino-acid variation within the previously characterized electromorphs (20, 33).

A number of studies of enzyme kinetics have been motivated by the results of electrophoretic surveys that revealed striking, locus-specific patterns of clinal variation or correlations with particular environmental variables (9, 20, 36, 37). In particular, latitudinal clines in allele frequency

suggest the hypothesis that adaptive modifications of enzyme function are mediated by factors that are directly or indirectly related to temperature. Of these studies, the best examples involve kinetic analyses in which the products of alternative alleles were purified to homogeneity: ADH and G6PD in *Drosophila melanogaster*, LDH-B in *Fundulus heteroclitus*, and PGI in the blue mussel, *M. edulis* (25–27, 38). In each of these polymorphisms, significant allelic differences in k_{cat}/K_M and its components were identified. These kinetic data provided the motivation for tests of specific physiological hypotheses regarding allozyme-dependent in vivo function (see below). In the case of the ADH and G6PD polymorphisms, functional characterizations were performed on allozymes that were known to differ by a single amino acid. The ADH-F allozyme exhibited a higher k_{cat} for ethanol than the ADH-S allozyme (38), implicating an enhanced ability of the ADH-F allozyme to metabolize alcohol. Relative to G6PD-A, the G6PD-B allozyme exhibited a reduced K_M for glucose-6-phosphate, which is expected to result in a 40% higher k_{cat}/K_M in vivo, and greater flux through the pentose shunt, the pathway in which this enzyme functions. For the *Mytilus* PGI and *Fundulus* LDH polymorphisms, significant temperature-dependent differences in k_{cat}/K_M or overall catalytic efficiency could be interpreted in light of the observed latitudinal clines in allele frequency.

2.2. Enzyme Concentration

As discussed previously, in vivo measures of enzyme activity (e.g., V_{max}/K_M) may reflect differences in enzyme kinetics (k_{cat}/K_M), differences in enzyme concentration ([E]), or both. In addition, regulatory variation in enzyme concentration can result from *cis-* and/ or *trans*-acting factors, which also need to be distinguished (see Chapter 18). Studies of the ADH and G6PD polymorphisms in *D. melanogaster* and the LDH polymorphism in *F. heteroclitus* involved especially thorough examinations of variation in [E] and its underlying causes.

In the case of the G6PD polymorphism in *D. melanogaster*, activity differences between the two main allozymes were not attributable to differences in [E] or k_{cat} (27, 39). In the case of the ADH polymorphism in *D. melanogaster*, the consensus of many studies is that a significantly higher concentration of ADH protein occurs in *Adh^FF* compared with *Adh^SS* genotypes, and this difference accounts for most of the difference in allozyme-specific activity (6). Site-directed mutagenesis studies indicated that the fast/slow amino-acid change responsible for the difference in k_{cat} between the ADH-F and ADH-S allozymes does not affect enzyme concentration (32, 40). Instead, allelic differences in enzyme concentration appear to be largely attributable to a small insertion–deletion polymorphism in the first intron that is in strong linkage disequilibrium with the fast/slow amino-acid polymorphism (41). Thus,

variation in ADH activity among genotypes appears to be attributable to epistatic interaction between a single amino-acid polymorphism and one or more closely linked noncoding polymorphisms. Stam and Laurie (42) suggested that multiple polymorphic sites within the *Adh* gene may be targets of selection. Interestingly, no differences in mRNA transcript abundance or in vivo rates of enzyme degradation were observed between the *Adh^{FF}* and *Adh^{SS}* genotypes (6, 43), so the mechanistic basis of the observed variation in ADH enzyme concentration remains unknown.

In the case of the LDH polymorphism in *Fundulus*, allozyme-associated differences in transcript abundance and enzyme concentration stem from allelic differences at one or more closely linked *cis*-regulatory elements (7, 44). The difference in LDH enzyme concentration makes a much greater contribution to the in vivo rate of catalysis than does the difference in k_{cat}/K_M. Finally, in an analogous study involving crickets (*Gryllus firmus*), the activity of the lipogenic enzyme $NADP^+$- isocitrate dehydrogenase differs dramatically between alternative genotypes that differ in life history and lipid biosynthesis. In this case, in vivo differences in enzyme activity appear to be produced solely by genetic differences in transcript abundance and [E], but not k_{cat}/K_M (45, 46).

3. In Vivo Studies of Protein Function: Linking Genotype to Whole-Organism Physiological Performance

To verify the adaptive significance of an enzyme polymorphism, it is necessary to demonstrate that biochemical differences between the products of alternative alleles give rise to differences in higher-level physiological processes (and ultimately, fitness) (2, 8, 9, 11). It is important that these studies be conducted in vivo under ecologically relevant conditions. There are two reasons why such investigations are especially important. First, the seminal studies of Kacser and Burns (14, 15) and subsequent studies (reviewed in refs. 10, 13, 47, 48) have demonstrated that enzyme activity often exhibits a highly nonlinear (hyperbolic) relationship to pathway flux (Figure 2). Thus, even large changes in enzyme activity often result in only minimal changes in flux. For this reason, the effect of allozyme variation on pathway flux needs to be demonstrated experimentally.

Another reason to carefully consider the physiological context of gene function relates to recent discoveries that enzymes and other types of proteins, such as receptors and transmembrane channels, often have unexpected functions that are distinct from their "traditional" roles, a phenomenon termed "moonlighting" (49, 50). For example, several enzymes involved in intermediary metabolism also function as secreted regulators of various processes such as growth or differentiation (e.g., PGI, cytochrome C). This phenomenon underscores the importance of examining the physiolog-
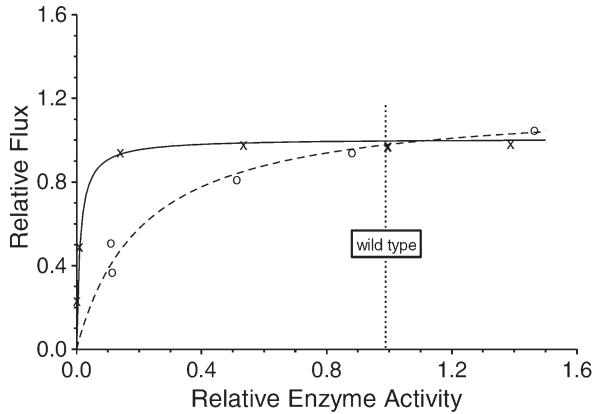
Figure 2. Relationships between activity and pathway flux for components of lactose ca-tabolism in *E. coli. Solid line* = β-galactosidase, *dashed line* = permease. In the case of β-galactosidase, a 50% decrease in activity results in a negligible (0.5%) decrease in path-way flux. In the case of permease, by contrast, the same decrement in enzyme activity re-sults in a 13% decrease in pathway flux. Modified from Reference 10.

ical consequences of protein variation and keeping an open mind with re-spect to the specific aspects of physiology that are chosen for investigation. For example, consider a polymorphic enzyme that is known to function in a particular metabolic pathway. If allelic variation in enzyme function is as-sociated with some fitness-related measure of organismal performance in the absence of any detectable effects on pathway flux, then it is probably worth investigating whether the fitness effects stem from some unantici-pated structural or functional role of the enzyme that is independent of the enzyme's known role in metabolism. These considerations are also relevant to cases where indirect evidence for the locus-specific effects of positive se-lection (based on comparisons of polymorphism and/or divergence at syn-onymous and nonsynonymous nucleotide positions) is not associated with detectable differences in protein function (e.g., (51)). Such cases suggest the possibility that the apparent effects on fitness stem from unexamined pro-tein functions that are only tangentially related to the experimentally as-sayed properties.

3.1. Effects of Allelic Enzyme Variants on Pathway Flux and Higher Level Physiological Processes

What higher-level physiological effects do we need to document to draw firm conclusions about the adaptive significance of a given enzyme poly-morphism? Ideally, we would like to verify that the products of alternative alleles exhibit in vivo differences in reaction rate and that this difference, in

turn, results in a physiologically significant difference in flux through the pathway in which the enzyme functions. Moreover, measured differences in pathway flux should be consistent with the allelic differences in enzyme kinetics that are measured in vitro (9–11). Differences in pathway function should then give rise to differences in cellular or systemic physiology that can ultimately be traced to some fitness-related measure of whole-organism performance.

In an experimental study of the G6PD polymorphism in *D. melanogaster*, Eanes (52) used an ingenious genetic manipulation to unveil allelic differences in enzyme activity between lines that were co-isogenic for 98% of the genome. A number of radio-isotopic studies have demonstrated flux differences through various pathways in metabolism caused by allelic variation in the G6PD and ADH enzymes of *D. melanogaster*. For example, the higher activity of the G6PD-B allozyme is associated with greater pentose shunt flux relative to the lower activity G6PD-A allozyme (39, 53). For ADH in *D. melanogaster*, no flux differences were observed between ADH-F and ADH-S allozymes in the overall conversion of ethanol to triglyceride (energy storage) or CO_2 (energy utilization) in adults (54). However, significant differences were observed in the conversion of ethanol to triglyceride and a variety of other metabolites in juveniles (55, 56). These results suggest that the main fitness effects of the ADH polymorphism are manifest in the juvenile life stage and may result from the enhanced ability of the Adh^{FF} genotype to convert alcohol to triglyceride.

The relationship between enzyme activity and pathway flux has been most thoroughly studied by Dykhuizen, Dean, and colleagues in *E. coli*, focusing on polymorphic genes of lactose metabolism in chemostats (10, 13). In this single-celled organism, flux through the metabolically simple (3 enzymes/proteins) lactose catabolic pathway is proportional to growth (cell division) and fitness under lactose-limiting conditions. Thus, these studies simultaneously characterize the effect of allozyme variation on pathway flux and fitness, using lines that controlled for genetic background. In their landmark studies, Dykhuizen and Dean (10, 13) found that naturally occurring allozymes of β-galactosidase exhibit almost no effect on pathway flux (Figure 2), thus indicating that these allozymes are selectively neutral.

4. The Chain of Causation Linking Genotype, Phenotype, and Fitness

To isolate the phenotypic effects of a specific gene, it is necessary to control for environmental variation as well as background genetic variation. Even at the most proximal level of phenotypic variation, the functional effects of specific proteins can be modulated by changes in the intracellular environment (e.g., temperature, pH, and the presence of cofactor molecules).

In vertebrates, fine-tuned adjustments in blood–oxygen affinity can play a key role in matching cellular oxygen supply and demand. These adjustments may be mediated by changes in the intrinsic oxygen affinity of hemoglobin, changes in the responsiveness of hemoglobin to allosteric cofactors that modulate oxygen affinity (organic phosphates, H^+ and Cl^- ions), and/or changes in the concentrations of these various cofactors within the erythrocyte (57). Allosteric ligands preferentially bind and stabilize the deoxy conformation of the hemoglobin tetramer, thereby decreasing blood–oxygen affinity by shifting the allosteric equilibrium in favor of the low-affinity, deoxy quaternary structure. In hypoxia-tolerant animals that have evolved elevated blood–oxygen affinities, a common mechanism of hemoglobin adaptation involves structural changes that suppress sensitivity to the inhibitory effects of allosteric cofactors (58–60). For example, deer mice (*Peromyscus maniculatus*) that are native to high altitude are characterized by an elevated blood–oxygen affinity that helps safeguard arterial oxygen saturation under hypoxia. Experimental studies of purified hemoglobin revealed that the elevated blood–oxygen affinity of high-altitude mice is largely attributable to structural modifications of hemoglobin that suppress sensitivity to 2,3-diphosphoglycerate (a metabolite of glycolysis) and Cl^- ions (61, 62). The important point is that insights into the mechanistic basis of the adaptive blood phenotype required oxygen-binding experiments on purified hemoglobin. Measuring the oxygen affinity of whole blood can yield physiologically relevant information, but such measures by themselves cannot be related to changes in hemoglobin structure because they are confounded by environmental effects as well as possible genotype–environment interaction effects (Figure 3). This echoes the point made earlier about the importance of disentangling genetic and environmental components of variance when comparing in vivo measures of V_{max}/K_M among enzyme variants.

It becomes especially challenging to control for environmental variation and background genetic variation when investigating the effects of specific genes on higher-level physiological processes (as discussed above for studies of pathway flux). In practice, the functional effects of protein-coding genes can be isolated by means of forward genetics (i.e., controlled crosses, including the construction of recombinant inbred lines [RILs] or nearly isogenic lines [NILs]) or reverse genetics (the direct manipulation of gene sequence [e.g., targeted gene replacement] or manipulation of gene expression [e.g., RNAi and gene insertion]). In many cases, isolating the effects of individual genes is not sufficient to identify the specific mechanism of molecular adaptation. As discussed below, it is often necessary to isolate the effects of individual mutations. It is also necessary to measure the functional effects of individual mutations — singly and in different multisite combinations — to address questions about the dynamics of adaptive walks (e.g., What is the distribution of phenotypic effect sizes among adaptive substitutions? What is the role of epistasis in constraining the selective accessibility of mutational pathways to high-fitness genotypes?).
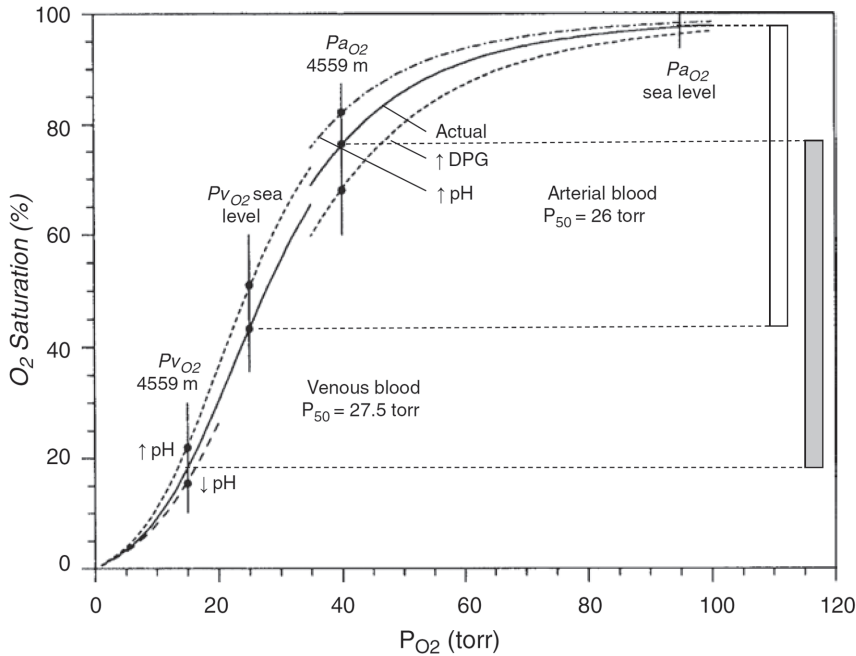
Figure 3. Oxygen-equilibrium curves of human blood and environmentally induced changes in blood–oxygen affinity under hypoxic conditions at high altitude (4,559 m). *Solid curves* represent measured values for arterial blood (P_{50} (the partial pressure of oxygen [PO_2] at which hemoglobin is 50% saturated) = 26 torr, upper section) and venous blood (P_{50} = 27.5 torr). The *broken curves* depict the effects of increased DPG concentration (↑DPG) at unchanged pH, increased pH (↑pH) at unchanged DPG, and the decreased tissue pH (↓pH) that results from metabolic acidosis in the tissues. *Open and shaded vertical columns* denote the oxygen unloaded at sea level (venous PO_2 = 25 torr) and at high altitude (4,559 m; venous PO_2 = 15 torr), respectively. Displacements of the oxygen-equilibrium curve are attributable to environmentally induced changes in DPG and pH in red blood cells. For example, at high altitude, the increased red cell concentration of DPG produces a rightward shift of the curve (i.e., reduced oxygen affinity), and this may be partly counteracted by the effects of respiratory alkalosis (↑pH). Modified from Reference 58.

4.1. Isolating the Effects of Individual Mutations

In cases where specific protein-coding genes are known to contribute to adaptive variation in a particular trait, the products of alternative protein alleles are often distinguished from one another by multiple amino-acid substitutions. To cite just a few examples, in the malaria parasite *Plasmodium falciparum*, derived alleles at the dihydrofolate reductase gene that confer pyrimethamine resistance are distinguished from the wild-type susceptible allele by four amino-acid substitutions (63). In *E. coli*, derived alleles at the β-lactamase gene that confer cefotaxime resistance are distinguished from the wild-type susceptible allele by a combination of four amino-acid substi-

tutions and one noncoding substitution (16). In the sheep blowfly *Lucilia cuprina*, a derived allele at the *E3* esterase gene that confers diazinon resistance is distinguished from the wild type, susceptible allele by five amino-acid substitutions (64). In deer mice, differences in hemoglobin–oxygen affinity between high- and low-altitude populations are attributable to the independent or joint effects of eight amino-acid mutations in the α-chain subunits of the hemoglobin tetramer and four mutations in the β-chain subunits (61, 62, 65, 66). Such cases prompt an obvious question: Are allelic differences in protein function typically attributable to one or two causative mutations, or do the differences in biochemical phenotype depend critically upon the additive or epistatic effects of all mutations in combination? More generally: Are adaptive modifications of protein function typically attributable to a small number of mutations at key positions, or many minor mutations of individually small effect? To answer such questions, it is necessary to isolate the functional effects of individual mutations.

In some cases, the history of recombination in a given population sample may provide a number of naturally occurring sequence variants in which the mutations of interest are essentially randomized against different genetic backgrounds. By judiciously choosing which specimens to use as a source of purified protein for in vitro studies (on the basis of genotypic data), it may be possible to isolate and test a subset of the most important multisite combinations of mutations. In cases where recombination has not provided a sufficient number of naturally occurring sequence variants to make the comparisons of interest, it is necessary to use a protein- engineering approach to isolate the functional effects of individual mutations. This can be accomplished by means of site-directed mutagenesis, followed by in vitro expression and purification of recombinant protein for experimental studies. This experimental approach opens up exciting possibilities for reconstructive inference in studies of protein evolution. For example, it is possible to produce recombinant proteins with primary structures that represent inferred ancestral states (67) or unobserved, intermediate mutational steps in evolutionary pathways (16, 63).

4.2. Epistasis and Compensatory Substitutions: When Adaptive Walks Take Two Steps Forward and One Step Back

Even if an advantageous change in protein function can be attained by means of a single substitution at a key position in the active site, such changes may often require additional compensatory substitutions to offset negative pleiotropic effects of the original change. These compensatory mutations may be neutral or deleterious by themselves—they are only advantageous when preceded by the initial, function-altering substitution. By definition, the fitness effects of compensatory mutations vary as a function

of the genetic background in which they appear. This context-dependence, therefore, represents a special form of epistasis that can influence the selective accessibility of mutational pathways through sequence space (68–70). If compensatory substitutions are pervasive, then in comparisons between the products of alternative alleles—or in comparisons between the products of orthologous genes in different species—it may often be exceedingly difficult to determine which substitutions were responsible for producing the primary adaptive improvements in protein function and which substitutions represent secondary compensatory changes. Experimental studies of protein function involving different multisite combinations of mutations will also be required to identify seemingly "compensated" genotypes that do not actually involve the correction of individually deleterious mutations (71, 72).

A number of recent comparative studies provide indirect evidence for the pervasiveness of compensatory substitutions by identifying cases where a pathogenic amino-acid mutation in one species appears as the wild-type amino acid at the same site in the orthologous proteins of other species. In order for potentially pathogenic mutations to become fixed, compensatory substitutions must have occurred at other sites in the same protein, or possibly even in different, interacting proteins (73–75). The former case seems to be more common, as evidence from a phylogenetically diverse array of organisms indicates that deleterious mutations and their associated compensatory mutations typically occur in the same gene and they also tend to be nonrandomly clustered in the tertiary structure of the encoded protein (76, 77).

The phenomenon of compensatory epistasis highlights the importance of considering structure–function relationships in multiple dimensions. In many cases, trade-offs between protein function and particular biophysical properties can be predicted based on principles of structural biology. Since catalytic activity requires mechanical flexibility, many aspects of protein function, protein stability, and aggregation propensity are intrinsically linked (78, 79). The latter two properties are interrelated because destabilizing mutations lead to a reduced ratio of folded and unfolded molecules, and the increased cellular concentration of unfolded molecules results in an increased rate of aggregation. The important implication is that evolutionary changes in catalytic activity will often have negative pleiotropic effects that require compensatory mutations to restore stability and/or resistance to aggregation. In bacterial pathogens, mutations that confer antibiotic resistance often have negative side effects on stability and aggregation. Thus, the evolution of drug-resistant bacteria often involves numerous compensatory substitutions that offset the destabilizing effects of substitutions at catalytic residues. In *E. coli*, for example, primary substitutions in the β-lactamase gene that increase cefotaximine resistance are typically coupled with stabilizing/ aggregation-inhibiting substitutions at residues remote from the active site (78, 80).

Compensatory epistasis has important implications for the rate and mode of protein evolution, and it also poses interpretive challenges for microevolutionary studies of molecular adaptation. It may be that adaptive modifications of protein function (catalytic efficiency, ligand affinity, etc.) can often be achieved through a small number of substitutions at key residues, but the resultant perturbations of other biophysical properties may require numerous compensatory substitutions at remote sites. In such cases, both types of substitutions may be attributable to positive, directional selection, but substitutions of the latter type are only beneficial on specific genetic backgrounds. This highlights the important point that positive selection does not necessarily produce adaptive improvements in protein function – the conditionally beneficial compensatory substitutions may simply maintain the status quo with respect to structural stability or other biophysical properties. In comparative studies of sequence evolution, bursts of compensatory substitutions could produce an elevated level of protein divergence relative to neutral expectations, but most of the changes contributing to the elevated divergence may not have any direct effects on protein function. This point is seldom appreciated in comparative studies that seek to identify specific codons that have experienced an excess of nonsynonymous substitutions, as evidence for positive selection on coding sequence is often implicitly equated with evidence for adaptive evolution of protein function.

4.3. Future Prospects

Noteworthy successes in establishing causal connections between genotype, phenotype, and fitness come from studies of single-locus protein polymorphisms that have clearly defined effects on whole-organism physiological performance (2, 11, 12, 23, 81, 82). With the advent of genomic, transcriptomic, and proteomic technologies, efforts to dissect the mechanistic basis of fitness variation have expanded to focus more on adaptive changes at the level of whole pathways or networks, rather than functional properties of individual proteins (48, 56). Genomic analyses of sequence variation and expression variation in metabolic pathways can be expected to shed light on long-standing questions in evolutionary biology (e.g., What is the relative role of structural vs. regulatory changes in adaptive evolution?). Such studies can also be expected to push new questions to the forefront (e.g., How does the pathway position of a given protein influence its evolutionary dynamics?). Indeed, population genomic studies at the pathway or network level can be expected to make important contributions to the developing field of evolutionary systems biology. The simultaneous examination of multiple pathway components in a population-genetic framework should also provide important insights into the evolution of complex traits (e.g., floral pigmentation (83, 84); intermediary metabolism in insects (45, 85–90)).

References

1. Gillespie JH (1991) The Causes of Molecular Evolution, Oxford University Press, New York
2. Storz JF, Wheat CW (2010) Integrating evolutionary and functional approaches to infer adaptation at specific loci, Evolution **64**: 2489–2509
3. Hochachka PW, Somero GN (2002) Biochemical Adaptation. Mechanism and Process in Physiological Evolution
4. Albery WJ, Knowles JR (1976) Evolution of enzyme function and the development of catalytic efficiency, Biochemistry **15**: 5631–5640
5. Fersht A (1999) Structure and Mechanism in Protein Science, W. H. Freeman and Co., New York
6. Laurie CC, Stam LF (1988) Quantitative analysis of RNA produced by slow and fast alleles of *Adh* in *Drosophila melanogaster*, Proc Natl Acad Sci USA **85**:5161–5165
7. Crawford DL, Powers DA (1989) Molecular basis of evolutionary adaptation at the lactate dehydrogenase-B locus in the fish *Fundulus heteroclitus*. Proc Natl Acad Sci USA **86**: 9365–9369
8. Clarke B (1975) The contribution of ecological genetics to evolutionary theory: detecting the direct effects of natural selection at particular polymorphic loci. Genetics **79**:101–113
9. Koehn RK, Zera AJ, Hall JG (1983) Enzyme polymorphism and natural selection. In: Nei M, Koehn RK, (eds.) Evolution of genes and proteins, Sinauer, Sunderland, MA
10. Dykhuizen DE, Dean AM (1990) Enzyme activity and fitness: evolution in solution. Trends Ecol Evol **5**:257–262
11. Watt WB, Dean AM (2000) Molecular-functional studies of adaptive genetic variation in prokaryotes and eukaryotes. Ann Rev Genet **34**:593–622
12. Dean AM, Thornton JW (2007) Mechanistic approaches to the study of evolution: the functional synthesis. Nat Rev Genet **8**:675–688
13. Dykhuizen DE, Dean AM (2009) Experimental evolution from the bottom up. In: Garland T, Rose MR (eds.) Experimental evolution: concepts, methods, and applications of selection experiments. Univ Calif Press, Berkeley
14. Kacser H, Burnes JA (1979) Molecular democracy: who shares the controls? Trans Biochem Soc **7**:1149–1160
15. Kacser H, Burnes JA (1981) The molecular basis of dominance. Genetics **97**:639–666
16. Weinreich DM, Delaney NF, DePristo MA et al (2006) Darwinian evolution can follow only very few mutational paths to fitter proteins. Science **312**:111–114
17. Brown KM, DePristo MA, Weinreich DM et al (2009) Temporal constraints on the incorporation of regulatory mutants in evolutionary pathways. Mol Biol Evol **26**:2455–2462
18. Hall BG, Hauer B (1993) Acquisition of new metabolic activities by microbial populations. In: Abelson JN, Simon MI, Zimmer EA et al (eds.) Molecular evolution: producing the biochemical data. Academic Press Inc., San Diego
19. Rosenblum EB, Rompler H, Schoneberg T et al (2009) Molecular and functional basis of phenotypic convergence in white lizards at White Sands. Proc Natl Acad Sci USA **107**: 2113–2117
20. Eanes WF (1999) Analysis of selection on enzyme polymorphisms. Ann Rev Ecol Systemat **30**:301–326
21. Zamer WE, Hoffman RJ (1989) Allozymes of glucose-6-phosphate isomerase differentially modulate pentose-shunt metabolism in the sea anemone *Metridium senile*, Proc Natl Acad Sci USA **86**:2737–2741

22. Fromm HJ (1975) Initial rate enzyme kinetics. Springer, Berlin
23. Zera AJ, Koehn RK, Hall JG (1985) Allozymes and biochemical adaptation. In: Kerkut GA, Gilbert LI (eds.) Comprehensive insect physiology: biochemistry and pharmacology. Pergamon, Oxford
24. Greaney GS, Somero GN (1980) Contributions of binding and catalytic rate constants to evolutionary modifications in Km of NADH for muscle type (M4) lactate dehydrogenase, J Comp Physiol B **137**:115–121
25. Place AR, Powers DA (1979) Genetic variation and relative catalytic efficiencies: lactate dehydrogenase- B allozymes of *Fundulus heteroclitus*. Proc Natl Acad Sci USA **76**: 2354–2358
26. Hall JG (1985) Temperature-related kinetic differentiation of glucose-phosphate isomerase alleloenzymes isolated from the blue mussel, *Mytilus edulis* Biochem Genet **23**:705–728
27. Eanes W, Katona L, Longtine M (1990) Comparison of in vitro and in vivo activities associated with G6PD allozyme polymorphisms in *Drosophila melanogaster*. Genetics **125**: 845–853
28. White MW, Mane SD, Richmond RC (1988) Studies of esterase 6 in *Drosophila melanogaster*. XVIII. Biochemical differences between the slow and fast allozymes. Mol Biol Evol **5**:41–62
29. Hall JG, Koehn RK (1983) The evolution of catalytic efficiency and adaptive inference from steady-state kinetic data. Evol Biol **16**:53–96
30. Watt WB (1983) Adaptation at specific loci. II. Demographic and biochemical elements in the maintenance of the *Colias* PGI polymorphism. Genetics **103**:691–724
31. Watt WB, Donohue K, Carter PA (1996) Adaptation at specific loci. VI. Divergence vs. parallelism of polymorphic allozymes in molecular function and fitness component effects among *Colias* species (Lepidoptera; Pieridae). Mol Biol Evol **13**:699–709
32. Chambers GK (1988) The Drosophila alcohol dehydrogenase gene-enzyme system, Advances Genet **25**:39–107
33. Wheat CW, Watt WB, Pollock DD et al (2006) From DNA to fitness differences: sequences and structures of adaptive variants of *Colias* phosphoglucose isomerase (PGI), Mol Biol Evol **23**:499–512
34. Verrelli BC, Eanes WF (2001) The functional impact of *Pgm* amino acid polymorphism on glycogen content in *Drosophila melanogaster*. Genetics **159**:201–210
35. Verrelli BC, Eanes WF (2001) Clinal variation for amino acid polymorphisms at the *Pgm* locus in *Drosophila melanogaster*. Genetics **157**:1649–1663
36. Powers DA, Lauerman T, Crawford D et al (1991) Genetic mechanisms for adapting to a changing environment. Ann Rev Genet **25**:629–659
37. Sezgin E, Duvernell D, Matzkin LM et al (2004) Single-locus latitudinal clines and relationship to temperate adaptation in metabolic genes and derived alleles in *Drosophila melanogaster*. Genetics **168**:923–931
38. Winberg JO, Hovik R, McKinley-McKee JS (1985) The alcohol dehydrogenase alleloenzymes Adh[S] and Adh[F] from the fruitfly *Drosophila melanogaster*: an enzymatic rate assay to determine the active-site concentration. Biochem Genet **23**:205–216
39. Labate J, Eanes WF (1992) Direct measurement of in vivo flux differences between electrophoretic variants of G6PD in *Drosophila melanogaster*. Genetics **132**:783–787
40. Laurie CC, Bridgham J, Choudhary M (1991) Associations between DNA sequence variation and variation in expression of the Adh gene in natural populations of *Drosophila melanogaster*. Genetics **129**:489–499
41. Laurie CC, Stam LF (1994) The effect of an intronic polymorphism on alcohol dehydrogenase expression in *Drosophila melanogaster*. Genetics **138**:379–385
42. Stam LF, Laurie CC (1996) Molecular dissection of a major gene effect on a quantitative trait: the level of alcohol dehydrogenase expression in *Drosophila melanogaster*. Genetics **144**:1559–1564
43. Anderson SM, McDonald JF (1983) Biochemical and molecular analysis of naturally occurring *Adh* variants in *Drosophila melanogaster*. Proc Natl Acad Sci USA **80**:4798–4802

44. Crawford DL, Powers DA (1992) Evolutionary adaptation to different thermal environments via transcriptional regulation. Mol Biol Evol **9**:806–813

45. Zera AJ, Harshman LG (2011) Intermediary metabolism and the biochemical-molecular basis of life history variation and trade-offs in two insect models. In: Flatt T, Heywood A, (eds.) Molecular mechanisms of life-history evolution. Oxford Univ Press, New York

46. Schilder RJ, Zera AJ, Black C et al (2011) The biochemical basis of life history adaptation: molecular and enzymological causes of NADP+-isocitrate dehydrogenase activity differences between morphs of *Gryllus firmus* that differ in lipid biosynthesis and life history. In Press

47. Fell D (1997) Understanding the control of metabolism. Portland Press, London

48. Eanes WF (2011) Molecular population genetics and selection in the glycolytic pathway. J Exp Biol **214**:165–171

49. Jeffery C (1999) Moonlighting proteins. Trends Biochem Sci **24**:8–11

50. Copley S (2003) Enzymes with extra talents: moonlighting functions and catalytic promiscuity. Curr Opin Chem Biol **7**:265–272

51. Runck AM, Weber RE, Fago A et al (2010) Evolutionary and functional properties of a two-locus β-globin polymorphism in Indian house mice. Genetics **184**:1121–1131

52. Eanes WF (1984) Viability interactions, *in vivo* activity and the G6PD polymorphism in *Drosophila melanogaster*. Genetics **106**: 95–107

53. Cavener DR, Clegg MT (1981) Evidence for biochemical and physiological differences between enzyme genotypes in *Drosophila melanogaster*. Proc Natl Acad Sci USA **78**: 11666–11670

54. Middleton RJ, Kacser H (1983) Enzyme variation, metabolic flux and fitness: alcohol dehydrogenase in *Drosophila melanogaster*. Genetics **105**:633–650

55. Freriksen A, de Ruiter BLA, Scharloo W et al (1994) *Drosophila* alcohol dehydrogenase polymorphism and carbon-13 fluxes: opportunities for epistasis and natural selection. Genetics **137**:1071–1078

56. Zera AJ (2011) Microevolution of intermediary metabolism: evolutionary genetics meets metabolic biochemistry. J Exp Biol **214**:179–190

57. Weber RE, Fago A (2004) Functional adaptation and its molecular basis in vertebrate hemoglobins, neuroglobins and cytoglobins. Respir Physiol Neurobiol **144**:141–159

58. Weber R E (2007) High-altitude adaptations in vertebrate hemoglobins. Respir Physiol Neurobiol **158**:132–142

59. Storz JF, Moriyama H (2008) Mechanisms of hemoglobin adaptation to high-altitude hypoxia. High Alt Med Biol **9**:148–157

60. Storz JF, Scott GR, Cheviron ZA (2010) Phenotypic plasticity and genetic adaptation to high-altitude hypoxia in vertebrates. J Exp Biol **213**:2565–2574

61. Storz JF, Runck AM, Sabatino SJ et al (2009) Evolutionary and functional insights into the mechanism underlying high-altitude adaptation of deer mouse hemoglobin. Proc Natl Acad Sci USA **106**:14450–14455

62. Storz JF, Runck AM, Moriyama H et al (2010) Genetic differences in hemoglobin function between highland and lowland deer mice. J Exp Biol **213**:2565–2574

63. Lozovsky E R, Chookajorn T, Brown KM et al (2009) Stepwise acquisition of pyrimethamine resistance in the malaria parasite. Proc Natl Acad Sci USA **106**:12025–12030

64. Newcomb RD, Campbell PM, Ollis DL et al (1997) A single amino acid substitution converts a carboxylesterase to an organophosphorus hydrolase and confers insecticide resistance on a blowfly. Proc Natl Acad Sci USA **94**: 7464–7468

65. Storz JF, Sabatino SJ, Hoffmann FG et al (2007) The molecular basis of high-altitude adaptation in deer mice. PloS Genet **3**:e45

66. Storz JF, Kelly JK (2008) Effects of spatially varying selection on nucleotide diversity and linkage disequilibrium: insights from deer mouse globin genes. Genetics **180**:367–379

67. Thornton JW (2004) Resurrecting ancient genes: experimental analysis of extinct molecules. Nat Rev Genet **5**:366–375

68. Weinreich DM, Watson RA, Chao L (2005) Sign epistasis and genetic constraint on evolutionary trajectories. Evolution **59**:1165–1174

69. Poelwijk FJ, Kiviet DJ, Weinreich DM et al (2007) Empirical fitness landscapes reveal accessible evolutionary paths. Nature **445**:383–386

70. Carneiro M, Hartl DL (2005) Adaptive landscapes and protein evolution. Proc Natl Acad Sci USA **107**:1747–1751

71. Haag ES, Molla MN (2005) Compensatory evolution of interacting gene products through multifunctional intermediates. Evolution **59**: 1620–1632

72. Haag ES (2007) Compensatory vs. pseudocompensatory evolution in molecular and developmental interactions. Genetica **129**:45–55

73. Kondrashov AS, Sunyaev S, Kondrashov FA (2002) Dobzhansky-Muller incompatibilities in protein evolution. Proc Natl Acad Sci USA **99**:14878–14883

74. Gao L, Zhang J (2003) Why are some human disease-associated mutations fixed in mice? Trends Genet **19**:678–681

75. Kulathinal RJ, Bettencourt BR, Hartl DL (2004) Compensated deleterious mutations in insect genomes. Science **306**:1553–1554

76. Poon A, Davis BH, Chao L (2005) The coupon collector and the suppressor mutation: estimating the number of compensatory mutations by maximum likelihood. Genetics **170**: 1323–1332

77. Davis BH, Poon AFY, Whitlock MC (2009) Compensatory mutations are repeatable and clustered within proteins. Proc Royal Soc B-Biol Sci **276**:1823–1827

78. Wang XJ, Minasov G, Shoichet BK (2002) Evolution of an antibiotic resistance enzyme constrained by stability and activity trade-offs. J Mol Biol **320**:85–95

79. DePristo MA, Weinreich DM, Hartl DL (2005) Missense meanderings in sequence space: a biophysical view of protein evolution, Nat Rev Genet **6**:678–687

80. Sideraki V, Huang WZ, Palzkill T et al (2001) A secondary drug resistance mutation of TEM-1 beta-lactamase that suppresses misfolding and aggregation. Proc Natl Acad Sci USA **98**:283–288

81. Feder ME, Watt WB (1992) Functional biology of adaptation. In: Berry RJ, Crawford TJ, Hewitt GM (eds.) Genes in ecology. Blackwell Scientific Publications, Oxford, UK

82. Dalziel AC, Rogers SM, Schulte PM (2009) Linking genotypes to phenotypes and fitness: how mechanistic biology can inform molecular ecology. Mol Ecol **18**:4997–5017

83. Rausher MD (2008) Evolutionary transitions in floral color. Int J Plant Sci **169**:7–21

84. Streisfeld MA, Rausher MD (2010) Population genetics, pleiotropy, and the preferential fixation of mutations during adaptive evolution. Evolution **65**:629–642

85. Eanes WF, Merritt TJS, Flowers JM et al (2006) Flux control and excess capacity in the enzymes of glycolysis and their relationship to flight metabolism in *Drosophila melanogaster*. Proc Natl Acad Sci USA **103**:19413–19418

86. Eanes WF, Merritt TJS, Flowers JM et al (2009) Direct evidence that genetic variation in glycerol-3-phosphate and malate dehydrogenase genes (*Gpdh* and *Mdh1*) affects adult ethanol tolerance in *Drosophila melanogaster*. Genetics **181**:607–614

87. Flowers JM, Sezgin E, Kumagai S et al (2007) Adaptive evolution of metabolic pathways in *Drosophila*. Mol Biol Evol **24**:1347–1354

88. Greenberg AJ, Stockwell SR, Clark AG (2008) Evolutionary constraint and adaptation in the metabolic network of *Drosophila*. Mol Biol Evol **25**:2537–2546

89. Merritt TJS., Kuczynski C, Sezgin E et al (2009) Quantifying interactions within the NADP(H) enzyme network in *Drosophila melanogaster*. Genetics **182**:565–574

90. Merritt TJS, Sezgin E, Zhu CT et al (2006) Triglyceride pools, flight and activity variation at the *Gpdh* locus in *Drosophila melanogaster*. Genetics **172**:293–304