

# Separating spandrels from phenotypic targets of selection in adaptive molecular evolution

Stevan A. Springer<sup>1</sup>, Michael Manhart<sup>2</sup>, and Alexandre V. Morozov<sup>3\*</sup>

<sup>1</sup>Department of Cellular and Molecular Medicine, University of California San Diego, La Jolla 92109, CA, USA

<sup>2</sup>Department of Chemistry and Chemical Biology, Harvard University, Cambridge 02138, MA, USA

<sup>3</sup>Department of Physics and Astronomy, Rutgers University, Piscataway 08854, NJ, USA

\*Corresponding author: [morozov@physics.rutgers.edu](mailto:morozov@physics.rutgers.edu)

## Abstract

**There are many examples of adaptive molecular evolution in natural populations, but there is no existing method to verify which phenotypic changes were directly targeted by selection. The problem is that correlations between traits make it difficult to distinguish between direct and indirect selection. A phenotype is a direct target of selection when that trait in particular was shaped by selection to better perform a function. An indirect target of selection, also known as an evolutionary spandrel, is a phenotype that changes only because it is correlated with another trait under direct selection. Studies that mutate genes and examine the phenotypic consequences are increasingly common, and these experiments could estimate the mutational accessibility of the phenotypic changes that arise during an instance of adaptive molecular evolution. Under indirect selection, we expect phenotypes to evolve toward states that are more accessible by mutation. Deviation from this null expectation (evolution toward a phenotypic state rarely produced by mutation) would be compelling evidence of adaptation, and could be used to distinguish direct selection from indirect selection on correlated traits. To be practical, this molecular test of adaptation requires phenotypic differences that are caused by changes in a small number of genes. These kinds of genetically simple traits have been observed in many empirical studies of adaptive evolution. Here we describe how to use mutational accessibility to separate spandrels from direct targets of selection and thus verify adaptive hypotheses for phenotypes that evolve by adaptive molecular changes at one or a few genes.**

**Keywords:** Adaptation; Selection; Pleiotropy; Molecular Evolution; Mutational Accessibility

## 1. Inferring the phenotypic target of selection.

### *The problems of pleiotropy, correlated traits, and indirect selection.*

Adaptations are phenotypes shaped by selection to perform a function. But we cannot assume that a trait evolved for the function that we happen to assign it (Williams 1966; Gould and Lewontin 1979; Nielsen 2009). To understand the reason a trait evolved, we must formulate and test adaptive hypotheses – scenarios that specify exactly which phenotypic differences created the fitness differences that drove evolution (Williams 1966; Mayr 1983). Tests of adaptive hypotheses are confounded by two problems. The first is that individual mutations typically affect many traits simultaneously, a phenomenon known as pleiotropy. The second is that mutational effects on different traits can be correlated, and if so, a trait can change toward a particular state even when it has no direct consequence for fitness. Traits can evolve solely because they are correlated with some other trait that is important to fitness. We refer to this apparent selection resulting from correlations between traits as indirect selection, and to traits that evolve in this manner as evolutionary spandrels. To verify an adaptive hypothesis, one must be able to distinguish between a spandrel and a trait that was truly a direct target of selection (Pearson 1903; Lande and Arnold 1983).

If we could empirically determine the spectrum of phenotypes available by mutation, we could make two straightforward predictions: *i*) Traits that are not themselves under direct selection should tend toward phenotypic states that are reached frequently by mutation (Stadler et al. 2001). In other words, in the absence of direct selection, traits should evolve toward phenotypes that are mutationally accessible. If the trait in question is weakly coupled to another trait under direct selection, evolution is expected to move its phenotypes toward more accessible states (although, depending on the strength of correlation with the trait under selection, maximally accessible states may never be reached). *ii*) On the other hand, direct selection pushes traits along evolutionary paths that increase fitness, and can fix beneficial mutations even if they are not easily accessible by mutation alone. Similarly, if a trait persists after a series of mutations, even though most potential mutations change it, then it must be under stabilizing selection. These predictions form the basis of the accessibility test described here – a method that uses mutational accessibility to verify the phenotypic targets of adaptive molecular evolution.

As an example, consider a protein that binds to a ligand but must be correctly folded to do so. The protein has two relevant traits: folding stability (quantified by the fraction of proteins in the cell that are properly folded), and binding strength (quantified by the fraction of proteins in the cell that are bound to a ligand). These two traits are inevitably correlated because a protein can only bind when properly folded, so that the fraction of folded proteins is always greater than or equal to the bound fraction (Fig. 1). In principle, selection could act directly on one of these traits and not the other – solely to improve either the binding interaction, or the stability of folding. But even if selection directly acts on only one of these traits, indirect selection can drive change in the other trait if the effects of mutations on the two traits are correlated (Manhart and Morozov 2015). For example, even if there is selection only for folding (Fig. 1, blue horizontal arrow), the fraction of bound proteins will tend to increase simply because proteins that bind properly are more abundant among proteins that fold properly. In other words, proteins that bind well become more accessible by mutation as folding improves. Improved binding could thus be a spandrel that evolves in the absence of direct selection.

Molecular tests of phenotypic accessibility are only useful if we can quantify the space of possible phenotypes available by mutation for a particular trait. For quantitative traits, this difficult problem can be avoided by assuming that any small phenotypic change can be achieved by mutation (Maynard Smith 1978; Grafen 1984) (Fig. 2). Measures of the variances and co-variances of quantitative traits can identify the phenotypic targets of single selective events and find Pareto-optimal compromises between different selective regimes (Lande and Arnold 1983; Crespi 1990; Shoval et al. 2012). However, extending these kinds of analyses from quantitative traits to traits with a simple genetic basis (phenotypic differences caused by mutations at a few specific genomic loci) requires information or assumptions about mutational constraints on phenotypic change.

### ***Natural adaptations can have simple genetics.***

A striking empirical finding is that some instances of natural adaptive evolution have a simple genetic basis (Orr and Coyne 1992; Bell 2009; Conte et al. 2012; Martin and Orgogozo 2013; Gallant et al. 2014; Rosenblum and Parent 2014). Natural phenotypic changes sometimes occur by mutations in one or a few

genes (Nachman et al. 2003; Bradshaw and Schemske 2003; Hoekstra et al. 2006; Storz et al. 2007), and the same genes or even the same mutations can evolve in parallel (Wichman et al. 1999; Holder and Bull 2001; Cresko et al. 2004; Colosimo et al. 2005; Zhang 2006; Musset et al. 2007; McDonald et al. 2009; Jiang et al. 2012; Shen et al. 2012; Frankel et al. 2012; Springer et al. 2014; Wessinger and Rausher 2015). Artificial mutants of lab organisms can have phenotypes that resemble related species, sometimes due to the same genes or pathways that cause the natural difference (Koufopanou and Bell 1991; Parichy and Johnson 2001; Shapiro et al. 2004; Owen and Bradshaw 2011). Mimics are known to use the same genes as the organisms they model to achieve similar forms, and even regulatory elements can evolve in parallel (Reed et al. 2011; Gallant et al. 2014). Co-evolving proteins maintain their partners over long time intervals (Clark et al. 2009; Hellberg et al. 2012), and conflicts between the same proteins can arise independently in response to similar ecological interactions (Feldman et al. 2012; Dobler et al. 2012; Zhen et al. 2012). Evolution sometimes uses the same genetic elements repeatedly, so for some adaptations there must be only a few loci whose mutations can create an appropriate phenotypic change (Fig. 2). In such cases, it may be possible to move beyond just identifying the genes that cause phenotypic differences, by quantitatively estimating which phenotypic changes are common and which are rare when these causal loci are experimentally mutated. For natural phenotypic differences with simple genetic causes we can thus empirically evaluate the accessibility of the derived phenotype by mutagenesis.

### ***Distinguishing traits.***

There are fundamental limitations on our ability to resolve direct and indirect selection. It will not be possible to distinguish the phenotypic targets of selection from their pleiotropic effects when these components of phenotypic variation are strongly correlated (when for any initial state, the maximum phenotypic change in one trait nearly always corresponds to the maximum phenotypic change in another). But selection is also powerless to make this distinction – in terms of evolutionary response, perfectly-correlated phenotypes can in fact be considered one trait (Lewontin 1978; Stadler et al. 2001; Wagner and Zhang 2011). All tests of adaptation (whether quantitative genetic or molecular) thus rest on the fact that we can only separate direct and indirect selection when mutational effects on the measured traits are not too strongly correlated. In other words, for a set of mutations with the same effect on the trait under selection,

each must have an independent suite of effects on other traits (Lewontin 1978; Stadler et al. 2001).

Fortunately, the extent of correlations between traits can be verified experimentally. For example, directed protein evolution experiments commonly observe weak or no correlations between various biophysical traits: mutations that modify protein-ligand affinity or catalytic ability typically have varying effects on stability, and non-pleiotropic mutations which affect one trait and not another are nearly always available (Bloom and Arnold 2009). Lack of correlation between traits is also observed in natural contexts. For example, mutations of the Agouti locus have a suite of effects on light and dark coloration in mice, but selection can separate one trait from another (Linnen et al. 2013). In this paper, we will assume that when we refer to a trait, we mean the feature being measured, together with all the features that are strongly correlated with it (i.e., with the linear correlation coefficient close to  $\pm 1.0$ ). Therefore, strictly speaking, a trait is a set of phenotypic features with correlations that cannot be broken by mutation.

## **2. Practical exploration of phenotype space.**

For traits with a simple genetic basis, it is possible to recreate the set of mutations that caused a natural phenotypic difference, and explore their phenotypic effects one-by-one (Dean and Thornton 2007; Harms and Thornton 2013). In one of the first examples, Weinreich et al. recreated mutational paths during the evolution of the TEM antibiotic resistance gene (Weinreich et al. 2006). Each path is a different way of ordering the set of mutations that separate the ancestral and derived alleles. The researchers measured each mutation's phenotypic effect on cefotaxime resistance, and estimated its fixation probability by assuming that selection acted only on this measured trait. Exploring phenotype space in this way is ambitious even for the simplest traits. As the number of differences between the ancestral and derived alleles increases, the number of possible intermediate alleles goes up exponentially, and the number of possible mutational paths increases factorially (Weinreich et al. 2005). To estimate accessibility, we need methods that generate vast molecular diversity and measure its phenotypic consequences. Equally importantly, we need evolutionary models that describe tractable representations of phenotype space.

### ***Creating and phenotyping molecular diversity with combinatorial molecular biology.***

Modern molecular biology methods can be used to measure the phenotypic effects of many protein variants, and link each phenotype to its DNA sequence (Scott and Smith 1990). Combinatorial mutagenesis can generate mutational diversity exceeding  $10^{13}$  variants (Weiss et al. 2000; Overstreet et al. 2012). Together, these methods have been used to explore huge spaces of molecular variation, seeking new drug-binding partners, and mapping functional residues in protein-protein interactions. For example, the binding energetics of human growth hormone (hGH) and its receptor (hGHR) have been completely mapped by quantitative saturation mutagenesis: every possible single amino acid mutation of the hGH binding site has been created and tested for receptor affinity (Pál et al. 2006). High-throughput combinatorial methods can recreate ancestral proteins (Zhu et al. 2005; Lunzer et al. 2005; Szendro et al. 2013), find mutations that interconvert the activity of related but functionally-divergent proteins (O'Maille et al. 2008), and explore the function of intermediates between the ancestral and derived state (Weinreich et al. 2006; Bridgham et al. 2009; Field and Matz 2010). Libraries of cis-regulatory regions can measure the effect of regulatory sequence variation on transcription factor binding and protein expression (Gertz et al. 2009). Cell sorters can quickly measure many thousands of cells for any phenotype that can be coupled to a fluorescent marker. Thus practical tools for creating and phenotyping molecular diversity on a large scale exist. To verify which phenotypes were targets of selection, we must use these tools to explore phenotype space in ways that inform us about the accessibility of a phenotypic change (Fig. 3).

### ***Tractable representations of phenotype space.***

To create the most general model of phenotypic evolution, we would need to study every phenotype created by every mutation of the genome, and then consider evolutionary trajectories in this vast phenotypic space (Maynard Smith 1978). This is clearly impossible and will remain so – the number of possible phenotypes and evolutionary trajectories is simply too large to be ever explored completely. If we hope to gain an empirical measure of accessibility, we must reduce the combinatorial complexity and the size of phenotype space. To reduce the number of evolutionary trajectories to be considered, we often assume that selection eliminates or fixes mutations more quickly than the waiting time between mutations (the strong selection, weak mutation, or SSWM, regime). Therefore mutations fix one at a time, and evolution works on the set

of phenotypes made available by single mutations of each successively fixed allele (Gillespie 1984; Weinreich et al. 2005). Thus we can explore the neighborhood of single mutations to estimate the null probability of a particular phenotypic change (note that this null is shaped by correlations between traits, but we do not need to explicitly specify the correlated traits). The size of this neighborhood depends on the number of mutations in the genome that influence our trait of interest. Are mutations of small phenotypic effect spread out evenly, as is typically assumed in quantitative trait genetics, or are there a few loci that influence a trait much more than others (Fig. 2)? The approach described in this paper is only needed when genetic constraints significantly influence the course of phenotypic evolution, which is more likely when traits have simple genetic causes (Grafen 1984; Springer et al. 2011). In cases with simple genetics, we can focus on coding and regulatory sequences of the genes known (or suspected) to form the genetic basis of the trait of interest. In the opposite limit of numerous mutations with small phenotypic effects, existing methods from quantitative genetics are capable of distinguishing direct and indirect selection (Lande and Arnold 1983; Shoval et al. 2012).

### **3. The mutational accessibility test.**

To illustrate the accessibility test, we will use one of the many examples of positive selection whose phenotypic effects can be determined but whose exact evolutionary cause is unknown (Stolz et al. 2003; Nielsen 2009). The ventral and dorsal light organs of Jamaican click beetles (*Pyrophorus plagiophthalmus*) fluoresce in a variety of colors because of variation in their luciferase proteins (Wood et al. 1989). The inferred ancestral luciferase emits green light. Alleles encoding yellow and orange fluorescence have evolved by positive selection, as confirmed by both McDonald-Kreitman and dN/dS tests (Stolz et al. 2003). But what was the phenotypic target that caused this adaptive molecular evolution? Is emitting orange an adaptation, in the sense that alleles producing more orange light conferred higher fitness? Or is the color shift inevitable, a indirect consequence of selection on an unmeasured trait that evolved because most mutations of the luciferase protein change the phenotype toward orange?

We can answer this question by determining the probability of evolving toward orange in the absence of direct selection. If most mutations in the neighborhood of the ancestral click beetle luciferase cause it to glow orange, we are not forced to invoke direct selection on color to explain the switch to orange, as it could have evolved with or without direct selection. However, if these mutations are non-neutral, high mutational accessibility of orange implies that this phenotypic change could have been driven by indirect selection. Selection may have targeted on another trait encoded by luciferase – for example, anti-oxidant activity or protease sensitivity (Thompson et al. 1997; Barros and Bechara 2000). Alternatively, the switch to orange could be both beneficial and highly accessible.

### ***Single-mutation scenario.***

If there were only a single mutation separating ancestral and derived alleles, the test would be conceptually simple. Specifically, we could sample a neighborhood of possible mutations of the ancestral sequence by mutagenesis. When expressed in *E. coli*, luciferase proteins fluoresce at the same wavelength as their natural equivalents (Stolz et al. 2003). We can thus measure the effects of individual mutations by cloning a library of artificial mutations *en masse* into individual *E. coli* cells, and phenotyping each variant with a fluorescence activated cell sorter (FACS). The aim is to compare the phenotype of the natural mutation with the phenotypes in the neighborhood of possible single mutations to determine if the natural mutation is atypical. Specifically, we could use the neighborhood to assign a p-value to the probability of a random mutation occurring with a phenotypic change at least as extreme as the one observed. A low p-value would thus imply that a mutation with the observed effect occurs rarely in the absence of direct selection.

Empirical maps of mutational neighborhoods allow us to estimate accessibility (i.e., the probability of moving from one phenotypic state to another) in our null model of evolution by indirect selection, and therefore to assess whether or not we need to invoke direct selection to explain an observed phenotypic difference (Fig. 4).

### ***Multiple-mutation scenario.***

Now let us consider a more common case. Typically, we have a set of several substitutions that cause a phenotypic difference between the ancestral and derived phenotypes (the positively-selected mutations of



the luciferase protein in this example). We have inferred (or know) the ancestral state, but we do not know the temporal order of these substitutions, or the evolutionary paths taken by the population. In particular, we do not know whether the number of substitutions observed between the ancestral and derived sequences is typical, given the amount of evolutionary time that elapsed in the course of the observed shift from green to orange. Given that the actual path taken from the ancestral to the derived allele is difficult or impossible to reconstruct, we propose the following extension of the single-mutation accessibility test. Consider phenotypic measurements for a random subset of genotypes separated by the same number of mutations from the ancestral genotype as the derived genotype. For simplicity, only mutations at the sites found to be different between the ancestral and derived alleles will be taken into account; mutations at all other sites are assumed not to change the phenotype (this assumption can be tested experimentally, and if it proves incorrect, mutations at additional sites can be considered as well). These measurements of phenotypic states will be used to estimate a phenotype probability distribution, which in turn enables a p-value assignment to the derived state's phenotype in the absence of direct selection.

This type of reasoning, based on collecting information about likely and unlikely changes in the value of the observed trait, can also be extended to the situations when information about evolutionary paths is available. This may be the case in laboratory evolution experiments, where evolving populations can be monitored at regular intervals (Szendro et al. 2013; Jiang et al. 2013; González et al. 2015). In this scenario, the likelihood of each phenotypic change along the path (a member of the ensemble of paths connecting the ancestral and derived alleles) can be assessed using probability distributions constructed from observations of phenotypic states of mutational neighborhoods for each state along the observed path (Fig. 5). Even a single low-likelihood move anywhere along the evolutionary path will constitute strong evidence for the direct selection scenario on the trait under observation. Furthermore, weaker evidence from several steps along the path can be combined by considering their joint probability, which should increase the overall statistical power of the null (indirect selection only) hypothesis test.

#### **4. Challenges facing molecular tests of adaptation.**

##### ***Phenotypic measurements.***

Any test of phenotypic evolution rests on our ability to measure phenotypic effects accurately. Phenotypes measured in the lab must resemble their natural counterparts and must be independent of historical, biological, or experimental context in ways that obscure their natural effects.

##### ***Size of phenotype space.***

The approach we have described is only useful for genetically simple traits, whose variation depends on a small number of genomic loci (i.e., mutations elsewhere do not change the phenotype under study). Many natural adaptations have simple genetic underpinnings, and many examples of molecular, regulatory, and network evolution can be studied mutation-by-mutation (Stern and Orgogozo 2008; Bell 2009; Conte et al. 2012). Indeed it is simple traits that benefit most from explicit genetic analysis, as their evolution may be constrained by their genetics. In these traits, evolutionary processes can be subject to explicit molecular investigations of phenotype space provided one can devise appropriate large-scale phenotype assays.

##### ***Reconstruction of ancestral states and paths.***

Accurate inference of ancestral states and corresponding evolutionary trajectories is a serious challenge for molecular tests of adaptation. Outside of lineages whose evolution can be observed in real time, the ancestral sequence and the temporal order of mutations could be difficult or impossible to reconstruct (Maynard Smith and Haigh 1974; Gillespie 2000). Ancestral alleles can sometimes remain in a population, but we will only find these cases circumstantially; they cannot underpin a general approach (Springer and Crespi 2007; Rebeiz et al. 2009). Tests of direct and indirect selection in molecular adaptation will therefore have to allow for multiple ancestral sequence candidates and multiple evolutionary paths.

##### ***Population genetics assumptions.***

Our approach is based on the SSWM assumption – mutations arise one at a time and either fix or disappear, keeping the population monomorphic at all times. Depending on the strength of selection, either the first beneficial mutation discovered will always fix or, somewhat more realistically, the fixation probabilities of

beneficial mutations will be proportional to their selection coefficients (Gillespie 1984). In reality, populations may be polymorphic, which could significantly complicate the analysis (although a certain degree of polymorphism can be tolerated by considering the most common sequence variants found in the population). Furthermore, in small populations substitutions may occur due to genetic drift rather than selective advantage, although standard tests to distinguish selection from neutrality (such as McDonald-Kreitman and  $dN/dS$ ) should help address this problem. Finally, the effects of fluctuating selection and, more generally, past environmental changes may be difficult to capture in artificial conditions.

### **5. Conclusions: Testing adaptation itself.**

When we call a trait an adaptation, we imagine selection on a particular aspect of phenotypic variation (“hummingbird-pollinated flowers are red because...” or “cave fish lose their eyes because...”). Adaptive explanations invoke specific historical scenarios – both a particular form of selection and, as importantly, assumptions about evolutionary constraints or lack thereof. Adaptation is the center of biology, and determining which phenotypes were direct targets of selection is the key to testing adaptation. We cannot verify an adaptive hypothesis without eliminating the alternative: evolution as a response to selection on a correlated trait (Williams 1966; Gould and Lewontin 1979; Nielsen 2009). A growing number of studies are using artificial mutagenesis to uncover the range of functional phenotypic variation available by mutation. These methods could in principle also be used to verify direct selection, if we could identify unambiguous hallmarks of adaptation. The evolution of a phenotype away from mutationally accessible states is a feature of adaptation by direct selection that could potentially be estimated. Measuring mutational accessibility is not trivial, but can be achieved using existing molecular biology tools. Conveniently, the only traits that truly need genotype-phenotype analysis are genetically the simplest and most tractable (Grafen 1984; Springer et al. 2011). Thus molecular genetics can overcome the problems posed by pleiotropy and phenotypic correlations, and satisfy the deepest goal of evolutionary research – verifying that a phenotypic difference was a direct target of selection and can therefore be considered an adaptation.

### **Competing interests**

The authors declare no competing interests.

### **Acknowledgements**

SAS is grateful to WJ Swanson, J Tyerman, F Breden and HD Bradshaw for helpful discussions. An NSERC PGSD International Graduate Fellowship to SAS supported this work. MM was supported by NIH grant F32 GM116217. AVM and SAS are grateful to Pierre Pontarotti for his extraordinary hospitality, which enabled the exchange of ideas that ultimately made this work possible.

## References

- Barros MP, Bechara E (2000) Luciferase and urate may act as antioxidant defenses in larval *Pyrearinus termitilluminans* (Elateridae : Coleoptera) during natural development and upon 20-hydroxyecdysone treatment. *Photochemistry and Photobiology* 71:648–654.
- Bell G (2009) The oligogenic view of adaptation. *Cold Spring Harb Symp Quant Biol* 74:139–144. doi: 10.1101/sqb.2009.74.003
- Bloom JD, Arnold FH (2009) In the light of directed evolution: pathways of adaptive protein evolution. *Proc Natl Acad Sci USA* 106 Suppl 1:9995–10000. doi: 10.1073/pnas.0901522106
- Bradshaw HD, Schemske DW (2003) Allele substitution at a flower colour locus produces a pollinator shift in monkeyflowers. *Nature* 426:176–178. doi: 10.1038/nature02106
- Bridgham JT, Ortlund EA, Thornton JW (2009) An epistatic ratchet constrains the direction of glucocorticoid receptor evolution. *Nature* 461:515–519. doi: 10.1038/nature08249
- Clark NL, Gasper J, Sekino M, et al (2009) Coevolution of interacting fertilization proteins. 5:e1000570. doi: 10.1371/journal.pgen.1000570
- Colosimo PF, Hosemann KE, Balabhadra S, et al (2005) Widespread parallel evolution in sticklebacks by repeated fixation of *Ectodysplasin* alleles. *Science* 307:1928–1933. doi: 10.1126/science.1107239
- Conte GL, Arnegard ME, Peichel CL, Schluter D (2012) The probability of genetic parallelism and convergence in natural populations. *Proc Roy Soc Lond B* 279:5039–5047. doi: 10.1098/rspb.2012.2146
- Cresko WA, Amores A, Wilson C, et al (2004) Parallel genetic basis for repeated evolution of armor loss in Alaskan threespine stickleback populations. *Proc Natl Acad Sci USA* 101:6050–6055. doi: 10.1073/pnas.0308479101
- Crespi BJ (1990) Measuring the Effect of Natural-Selection on Phenotypic Interaction Systems. *Am Nat* 135:32–47.
- Dean AM, Thornton JW (2007) Mechanistic approaches to the study of evolution: the functional synthesis. *Nat Rev Genet* 8:675–688. doi: 10.1038/nrg2160
- Dobler S, Dalla S, Wagschal V, Agrawal AA (2012) Community-wide convergent evolution in insect adaptation to toxic cardenolides by substitutions in the Na, K-ATPase. *Proc Natl Acad Sci USA* 109:13040–13045. doi: 10.2307/41685674?ref=no-x-route:9c57945a3d654bf0c27836b49e38d101
- Feldman CR, Brodie ED, Pfrender ME (2012) Constraint shapes convergence in tetrodotoxin-resistant sodium channels of snakes. *Proceedings of the National Academy of Sciences* 109:4556–4561. doi: 10.1073/pnas.1113468109
- Field SF, Matz MV (2010) Retracing Evolution of Red Fluorescence in GFP-Like Proteins from *Faviina* Corals. *Mol Biol Evol* 27:225–233. doi: 10.1093/molbev/msp230
- Frankel N, Wang S, Stern DL (2012) Conserved regulatory architecture underlies parallel genetic changes and convergent phenotypic evolution. *Proc Natl Acad Sci USA* 109:20975–20979. doi: 10.1073/pnas.1207715109
- Gallant JR, Imhoff VE, Martin A, et al (2014) Ancient homology underlies adaptive mimetic diversity across butterflies. *Nat Commun* 5:4817. doi: 10.1038/ncomms5817

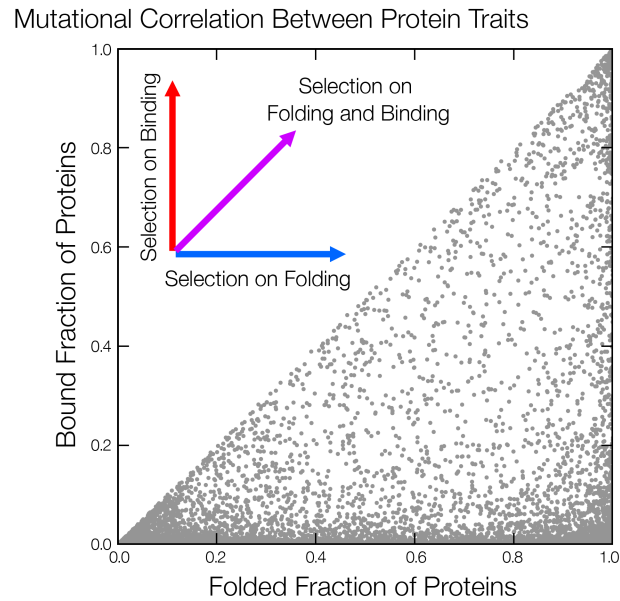
- Gertz J, Siggia ED, Cohen BA (2009) Analysis of combinatorial cis-regulation in synthetic and genomic promoters. *Nature* 457:215–218. doi: 10.1038/nature07521
- Gillespie J (1984) Molecular Evolution Over the Mutational Landscape. *Evolution* 38:1116–1129.
- Gillespie J (2000) Genetic drift in an infinite population. The pseudohitchhiking model. *Genetics* 155:909–919.
- González C, Ray JCJ, Manhart M, et al (2015) Stress-response balance drives the evolution of a network module and its host genome. *Mol Syst Biol* 11:827–827. doi: 10.15252/msb.20156185
- Gould SJ, Lewontin R (1979) The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. *Proc Roy Soc Lond B* 205:581–598.
- Grafen A (1984) Natural selection, kin selection and group selection. *Behavioural ecology: an evolutionary approach* 62–84.
- Harms MJ, Thornton JW (2013) Evolutionary biochemistry: revealing the historical and physical causes of protein properties. *Nat Rev Genet* 14:559–571. doi: 10.1038/nrg3540
- Hellberg ME, Dennis AB, Arbour-Reily P, et al (2012) The tegula tango: a coevolutionary dance of interacting, positively selected sperm and egg proteins. *Evolution* 66:1681–1694. doi: 10.1111/j.1558-5646.2011.01530.x
- Hoekstra HE, Hirschmann RJ, Bunday RA, et al (2006) A single amino acid mutation contributes to adaptive beach mouse color pattern. *Science* 313:101–104. doi: 10.1126/science.1126121
- Holder KK, Bull JJ (2001) Profiles of adaptation in two similar viruses. *Genetics* 159:1393–1404.
- Jiang P, Josue J, Li X, et al (2012) Major taste loss in carnivorous mammals. *Proc Natl Acad Sci USA* 109:4956–4961. doi: 10.1073/pnas.1118360109
- Jiang P-P, Corbett-Detig RB, Hartl DL, Lozovsky ER (2013) Accessible Mutational Trajectories for the Evolution of Pyrimethamine Resistance in the Malaria Parasite *Plasmodium vivax*. *J Mol Evol* 77:81–91. doi: 10.1007/s00239-013-9582-z
- Koufopanou V, Bell G (1991) Developmental Mutants of *Volvox*: Does Mutation Recreate the Patterns of Phylogenetic Diversity? *Evolution* 45:1806–1822.
- Lande R, Arnold SJ (1983) The measurement of selection on correlated characters. *Evolution* 1210–1226.
- Lewontin RC (1978) Adaptation. *Sci Am* 239:212–8– 220– 222 passim.
- Linnen CR, Poh Y-P, Peterson BK, et al (2013) Adaptive evolution of multiple traits through multiple mutations at a single gene. *Science* 339:1312–1316. doi: 10.1126/science.1233213
- Lunzer M, Miller SP, Felsheim R, Dean AM (2005) The biochemical architecture of an ancient adaptive landscape. *Science* 310:499–501. doi: 10.1126/science.1115649
- Manhart M, Morozov AV (2015) Protein folding and binding can emerge as evolutionary spandrels through structural coupling. *Proceedings of the National Academy of Sciences* 112:1797–1802. doi: 10.1073/pnas.1415895112
- Martin A, Orgogozo V (2013) The Loci of repeated evolution: a catalog of genetic hotspots of phenotypic variation. *Evolution* 67:1235–1250. doi: 10.1111/evo.12081

- Maynard Smith J (1978) Optimization Theory in Evolution. *Ann Rev Ecol Syst* 9:31–56.
- Maynard Smith J, Haigh J (1974) The hitch-hiking effect of a favourable gene. *Genet Res* 23:23–35.
- Mayr E (1983) How to Carry Out the Adaptationist Program? *Am Nat* 121:324–334.
- McDonald MJ, Gehrig SM, Meintjes PL, et al (2009) Adaptive divergence in experimental populations of *Pseudomonas fluorescens*. IV. Genetic constraints guide evolutionary trajectories in a parallel adaptive radiation. *Genetics* 183:1041–1053. doi: 10.1534/genetics.109.107110
- Musset L, Le Bras J, Clain J (2007) Parallel evolution of adaptive mutations in *Plasmodium falciparum* mitochondrial DNA during atovaquone-proguanil treatment. *Mol Biol Evol* 24:1582–1585. doi: 10.1093/molbev/msm087
- Nachman MW, Hoekstra HE, D'Agostino SL (2003) The genetic basis of adaptive melanism in pocket mice. *Proc Natl Acad Sci USA* 100:5268–5273. doi: 10.1073/pnas.0431157100
- Nielsen R (2009) Adaptionism-30 years after Gould and Lewontin. *Evolution* 63:2487–2490. doi: 10.1111/j.1558-5646.2009.00799.x
- O'Maille PE, Malone A, Dellas N, et al (2008) Quantitative exploration of the catalytic landscape separating divergent plant sesquiterpene synthases. *Nat Chem Biol* 4:617–623. doi: 10.1038/nchembio.113
- Orr HA, Coyne JA (1992) The genetics of adaptation: a reassessment. *Am Nat* 140:725–742. doi: 10.1086/285437
- Overstreet CM, Yuan TZ, Levin AM, et al (2012) Self-made phage libraries with heterologous inserts in the Mtd of *Bordetella bronchiseptica*. *Prot Eng Des Sel* 25:145–151. doi: 10.1093/protein/gzr068
- Owen CR, Bradshaw H (2011) Induced mutations affecting pollinator choice in *Mimulus lewisii* (Phrymaceae). *Arthropod-Plant Interactions* 1–10.
- Parichy DM, Johnson SL (2001) Zebrafish hybrids suggest genetic mechanisms for pigment pattern diversification in *Danio*. *Dev Genes Evol* 211:319–328.
- Pál G, Kouadio J-LK, Artis DR, et al (2006) Comprehensive and quantitative mapping of energy landscapes for protein-protein interactions by rapid combinatorial scanning. *J Biol Chem* 281:22378–22385. doi: 10.1074/jbc.M603826200
- Pearson K (1903) Mathematical contributions to the theory of evolution. XI. On the influence of natural selection on the variability and correlation of organs. *Philosophical Transactions of the Royal Society of London Series A, Containing Papers of a Mathematical or Physical Character* 200:1–66.
- Rebeiz M, Pool JE, Kassner VA, et al (2009) Stepwise modification of a modular enhancer underlies adaptation in a *Drosophila* population. *Science* 326:1663–1667. doi: 10.1126/science.1178357
- Reed RD, Papa R, Martin A, et al (2011) *optix* drives the repeated convergent evolution of butterfly wing pattern mimicry. *Science* 333:1137–1141. doi: 10.1126/science.1208227
- Rosenblum EB, Parent CE (2014) The molecular basis of phenotypic convergence. *Annual Review of Ecology ....* doi: 10.1146/annurev-ecolsys-120213-091851
- Scott JK, Smith GP (1990) Searching for peptide ligands with an epitope library. *Science* 249:386–390.

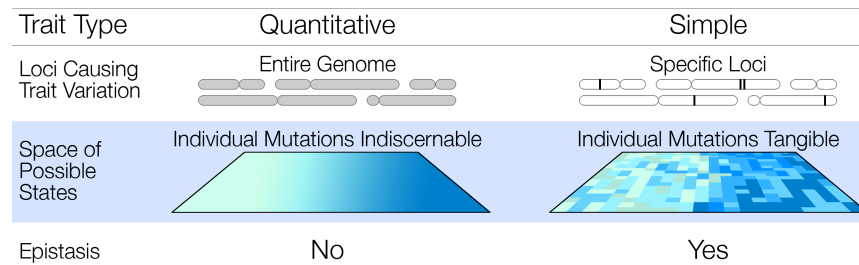
- Shapiro MD, Marks ME, Peichel CL, et al (2004) Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. *Nature* 428:717–723. doi: 10.1038/nature02415
- Shen Y-Y, Liang L, Li G-S, et al (2012) Parallel Evolution of Auditory Genes for Echolocation in Bats and Toothed Whales. *PLoS Genet* 8:e1002788. doi: 10.1371/journal.pgen.1002788.t001
- Shoval O, Sheftel H, Shinar G, et al (2012) Evolutionary trade-offs, Pareto optimality, and the geometry of phenotype space. *Science* 336:1157–1160. doi: 10.1126/science.1217405
- Springer S, Crespi BJ (2007) Adaptive gamete-recognition divergence in a hybridizing *Mytilus* population. *Evolution* 61:772–783. doi: 10.1111/j.1558-5646.2007.00073.x
- Springer S, Crespi BJ, Swanson WJ (2011) Beyond the phenotypic gambit: molecular behavioural ecology and the evolution of genetic architecture. *Mol Ecol* 20:2240–2257. doi: 10.1111/j.1365-294X.2011.05116.x
- Springer S, Diaz SL, Gagneux P (2014) Parallel evolution of a self-signal: humans and new world monkeys independently lost the cell surface sugar Neu5Gc. *Immunogenetics* 66:671–674. doi: 10.1007/s00251-014-0795-0
- Stadler BM, Stadler PF, Wagner GP, Fontana W (2001) The topology of the possible: formal spaces underlying patterns of evolutionary change. *J Theor Biol* 213:241–274. doi: 10.1006/jtbi.2001.2423
- Stern DL, Orgogozo V (2008) The loci of evolution: how predictable is genetic evolution? *Evolution* 62:2155–2177. doi: 10.1111/j.1558-5646.2008.00450.x
- Stolz U, Velez S, Wood KV, et al (2003) Darwinian natural selection for orange bioluminescent color in a Jamaican click beetle. *Proc Natl Acad Sci USA* 100:14955–14959. doi: 10.1073/pnas.2432563100
- Storz JF, Sabatino SJ, Hoffmann FG, et al (2007) The molecular basis of high-altitude adaptation in deer mice. *PLoS Genet* 3:e45. doi: 10.1371/journal.pgen.0030045
- Szendro IG, Franke J, de Visser JAGM, Krug J (2013) Predictability of evolution depends nonmonotonically on population size. *Proc Natl Acad Sci USA* 110:571–576. doi: 10.1073/pnas.1213613110
- Thompson JF, Geoghegan KF, Lloyd DB, et al (1997) Mutation of a protease-sensitive region in firefly luciferase alters light emission properties. *Journal of Biological Chemistry* 272:18766–18771. doi: 10.1074/jbc.272.30.18766
- Thorn KS, Bogan AA (2001) ASEdb: a database of alanine mutations and their effects on the free energy of binding in protein interactions. *Bioinformatics* 17:284–285.
- Tokuriki N, Stricher F, Schymkowitz J, et al (2007) The stability effects of protein mutations appear to be universally distributed. *J Mol Biol* 369:1318–1332. doi: 10.1016/j.jmb.2007.03.069
- Wagner GP, Zhang J (2011) The pleiotropic structure of the genotype-phenotype map: the evolvability of complex organisms. *Nat Rev Genet* 12:204–213. doi: 10.1038/nrg2949
- Weinreich DM, Delaney NF, Depristo MA, Hartl DL (2006) Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science* 312:111–114. doi: 10.1126/science.1123539
- Weinreich DM, Watson RA, Chao L (2005) Perspective: Sign epistasis and genetic constraint on evolutionary trajectories. *Evolution* 59:1165–1174.



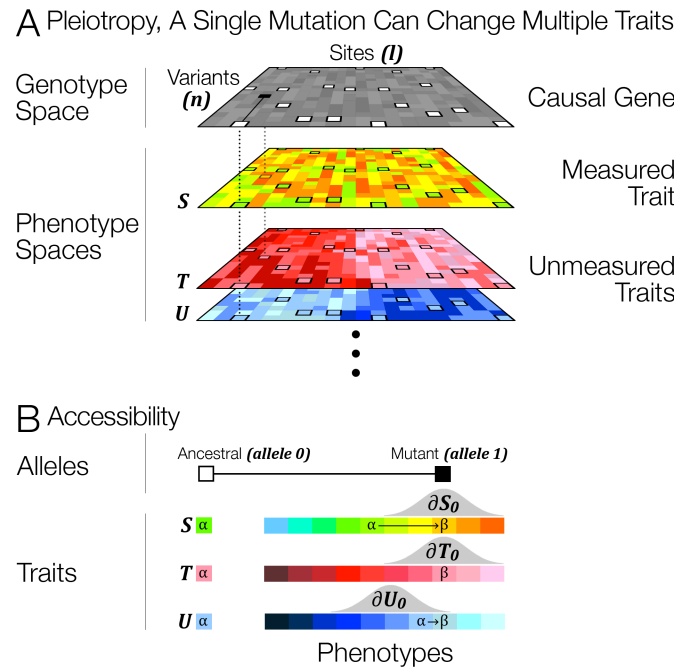
- Weiss GA, Watanabe CK, Zhong A, et al (2000) Rapid mapping of protein functional epitopes by combinatorial alanine scanning. *Proc Natl Acad Sci USA* 97:8950–8954. doi: 10.1073/pnas.160252097
- Wells JA (1990) Additivity of mutational effects in proteins. *Biochemistry* 29:8509–8517.
- Wessinger CA, Rausher MD (2015) Ecological transition predictably associated with gene degeneration. *Mol Biol Evol* 32:347–354. doi: 10.1093/molbev/msu298
- Wichman HA, Badgett MR, Scott LA, et al (1999) Different trajectories of parallel evolution during viral adaptation. *Science* 285:422–424.
- Williams GC (1966) *Adaptation and Natural Selection*. Princeton University Press
- Wood KV, Lam YA, McElroy WD, Seliger HH (1989) Bioluminescent click beetles revisited. *J Biolumin Chemilumin* 4:31–39. doi: 10.1002/bio.1170040110
- Zhang J (2006) Parallel adaptive origins of digestive RNases in Asian and African leaf monkeys. *Nature Genetics* 38:819–823. doi: 10.1038/ng1812
- Zhen Y, Aardema ML, Medina EM, et al (2012) Parallel molecular evolution in an herbivore community. *Science* 337:1634–1637. doi: 10.1126/science.1226630
- Zhu G, Golding GB, Dean AM (2005) The selective cause of an ancient adaptation. *Science* 307:1279–1282. doi: 10.1126/science.1106974



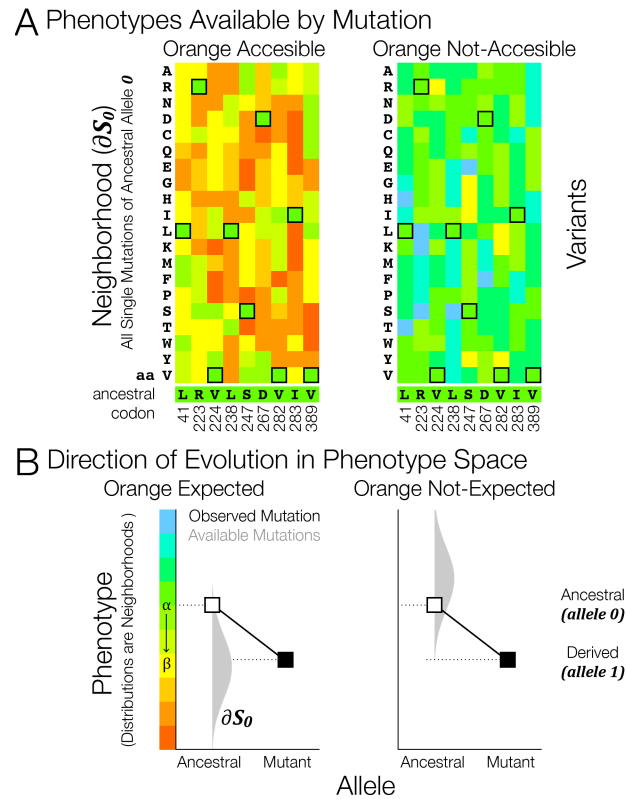
**Figure 1. Mutational correlations and indirect selection on protein traits.** Distribution of two possible protein traits – folding (quantified by the fraction of proteins that are folded) and binding (quantified by the fraction of proteins that are bound to a ligand) – in a simple thermodynamic model of protein kinetics (Manhart and Morozov 2015). Each residue makes an independent, additive contribution to the free energies of folding and binding (Wells 1990), which in turn determine the folding and binding probabilities via Boltzmann factors. Here we consider protein sequences varying only at 6 sites (e.g., at the binding interface), with a reduced alphabet of 5 amino acid types. The free energy contribution of each amino acid type at each site is randomly sampled from a Gaussian distribution with mean and standard deviation of 1 kcal/mol, consistent with observed distributions of mutational effects (Thorn and Bogan 2001; Tokuriki et al. 2007); total free energies are offset such that mean (over all possible sequences at the binding interface) free energy of folding is 2 kcal/mol and the mean free energy of binding is 5 kcal/mol. Each grey point represents the folding and binding traits of a different protein sequence. These two traits must be correlated because the fraction of folded proteins is always greater than, or equal to, the fraction of bound proteins (since the protein can only bind when properly folded). Arrows indicate different direct selection scenarios: direct selection on binding only (red arrow), direct selection on folding only (blue arrow), and direct selection on both binding and folding (magenta arrow).



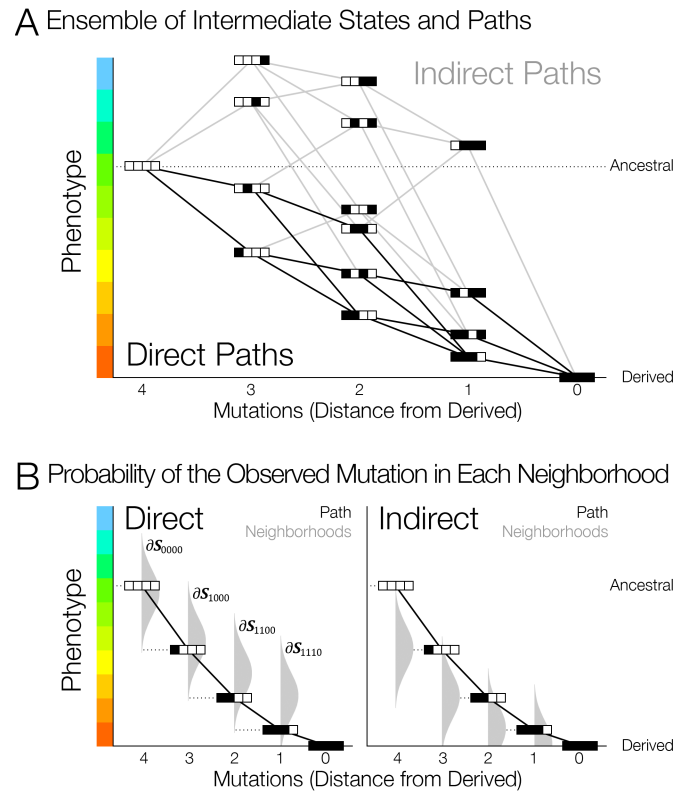
**Figure 2. The phenotype spaces of quantitative and simple traits.** Finding the phenotypic targets of selection in quantitative traits and genetically simple traits requires different assumptions about phenotype space. Classical optimality models reduce the complexity of phenotype space by assuming that any small phenotypic change can occur, that there is no epistasis, and that the structure of the phenotype space does not change from one mutation to the next. In contrast, the mutational accessibility test focuses on simple traits (traits dependent on just a few genomic loci). In simple traits the size of the phenotype space is reduced, allowing us to study variants of just those loci that cause an observed phenotypic difference. In simple traits, the effects of individual mutations may be epistatic and can significantly alter the space of alternatives available to evolution. Therefore, these traits will benefit most from an explicit genetic investigation.



**Figure 3. Pleiotropy, accessibility, and evolution by indirect selection.** (A) A genotype space and its corresponding phenotype spaces. Each of the  $l$  sites has  $n$  mutational variants. Thus, the mutational neighborhood size (i.e., the number of single-point mutations available to a sequence) is  $l \times n$ , and the total number of sequences is  $n^l$ . For simplicity, we do not consider strongly deleterious mutations, which would result in null phenotypes. White boxes show the wild-type sequence; the black box shows a mutation, which will change the measured phenotype  $S$ , as well as unmeasured phenotypes  $T$  and  $U$ . It is difficult to verify the target of selection because a single mutation can have pleiotropic effects on multiple traits. (B) Mutational neighborhoods of individual alleles are denoted as  $\partial \text{Trait}_{\text{allele}}$ . Grey frequency distributions represent accessibility: they show how often a particular phenotypic change occurs by a single mutation of an ancestral allele (*allele 0*) for different traits ( $S$ ,  $T$ ,  $U$ ). Trait  $S$ : most mutations shift the ancestral phenotype green ( $\alpha$ ) toward orange ( $\beta$ ); the observed phenotypic change (from green to orange,  $\alpha \rightarrow \beta$ ) would occur frequently even in the absence of direct selection. Trait  $T$  is neutral with respect to this mutation; its phenotype did not change, nor was it expected to, given that mutations with phenotype  $\alpha = \beta$  are common in  $\partial T_0$ . Trait  $U$  evolved in an unexpected direction, away from mutationally accessible states. Selection on this trait may have caused the change observed in trait  $S$  because of their shared genetic basis.



**Figure 4. Distribution of phenotypic effects due to a single mutation.** (A) Two potential neighborhoods of a green luciferase allele. Amino acids are arranged alphabetically by their full name. The amino-acid states of the ancestral luciferase allele are outlined in boxes, and their identity and position in the sequence are shown below. All amino acid substitutions at the positions that are known to be variable in nature (Stolz et al. 2003) are shown. For simplicity we assume that none of the mutations result in non-viable, “no color” phenotypes. The distribution of phenotypic changes caused by all single amino-acid substitutions of the ancestral allele  $\theta$  is denoted as  $\Delta S_{\theta}$ . In the left panel, the orange color is accessible. A change to this luciferase for any reason is expected to shift its emission toward orange. In the right panel, most mutations move the phenotype toward blue. A shift toward orange would therefore not be expected by chance or by indirect selection on a trait that is *not* strongly correlated with the color trait under consideration. (B) Comparing the effect of an observed mutation with a neighborhood of possible mutations. White and black boxes indicate the ancestral and derived (mutant) alleles, respectively; their corresponding phenotypes are on the Y-axis. Grey distributions show frequencies of phenotypic changes estimated from each mutational neighborhood. These distributions can be used to estimate the probabilities that a random mutation of the ancestral allele will move the phenotype toward or away from the derived state for this particular trait.



**Figure 5. Using mutational accessibility along evolutionary paths to infer the phenotypic target of selection.** (A) The ensemble of paths showing all possible intermediate states connecting the ancestral allele and the derived allele. The value of the trait changes monotonically among some of the paths, while on others there is at least one reversal of the trait – a potential indication of indirect rather than direct selection if this type of path is commonly observed. Note that here we assume for simplicity that there are no strongly deleterious mutations, which would not produce a phenotype. (B) Estimating the probability of making the observed mutational substitution at each step along the observed path. For each intermediate allele, we calculate the probability of making the observed substitution using the current allele’s mutational neighborhood. In the left panel, all successive substitutions are unlikely, making direct selection the most probable scenario. In contrast, in the right panel each mutation is likely (according to the distribution of phenotypic changes in the mutational neighborhoods), meaning that the accessibility test gives no reason to reject the indirect selection scenario.