The Molecular Basis of Phenotypic Convergence

Erica Bree Rosenblum,[1] Christine E. Parent,[1,2] and Erin E. Brandt[1]

[1]Department of Environmental Science, Policy, and Management, University of California, Berkeley, California 94720; email: rosenblum@berkeley.edu, eebrandt@berkeley.edu

[2]Department of Biological Sciences, University of Idaho, Moscow, Idaho 83844; email: ceparent@uidaho.edu

Annu. Rev. Ecol. Evol. Syst. 2014. 45:203–26

First published online as a Review in Advance on September 29, 2014

The *Annual Review of Ecology, Evolution, and Systematics* is online at ecolsys.annualreviews.org

This article's doi: 10.1146/annurev-ecolsys-120213-091851

Keywords

convergent evolution, parallelism, genetics, nonmodel species

Abstract

Understanding what aspects of evolution are predictable, and repeatable, is a central goal of biology. Studying phenotypic convergence (the independent evolution of similar traits in different organisms) provides an opportunity to address evolutionary predictability at different hierarchical levels. Here we focus on recent advances in understanding the molecular basis of convergence. Understanding when, and why, similar molecular solutions are used repeatedly provides insight into the constraints that shape biological diversity. We first distinguish between convergence as a phenotypic pattern and parallelism as a shared molecular basis for convergence. We then address the overarching question: What factors influence when parallel molecular mechanisms will underlie phenotypic convergence? We present four core determinants of convergence (natural selection, phylogenetic history, population demography, and genetic constraints) and explore specific factors that influence the probability of molecular parallelism. Finally, we address frontiers for future study, including integration across different systems, subfields, and hierarchical levels.

INTRODUCTION

> I am inclined to believe that in nearly the same way as two men have sometimes independently hit on the very same invention, so natural selection, working for the good of each being and taking advantage of analogous variations, has sometimes modified in very nearly the same manner two parts in two organic beings, which owe but little of their structure in common to inheritance from the same ancestor. —Darwin (1859, pp. 193–94)

Convergence: Evolution Repeats Itself

A central challenge in the biological sciences is determining the ways in which evolutionary change is repeatable. Thought experiments that "replay the tape of life" are a recurrent theme not only in biology but also in cosmology, philosophy, the arts, and even popular culture. However, we cannot, and need not, replay the Earth's history to study evolutionary repeatability. Instead we can compare evolutionary outcomes along different branches on the tree of life. Studying convergent evolution, the independent evolution of similar phenotypes, allows us to understand when, how, and why organisms arrive at similar solutions to similar evolutionary problems and, thus, address repeatability, and predictability, in evolution.

Interest in convergent evolution has deep roots in evolutionary biology. Darwin himself was puzzled by convergence and dedicated a passage in the "Difficulties of Theory" chapter of *The Origin of Species* to it (Darwin 1859). To Darwin, it seemed unlikely that unrelated organisms would evolve similar structures via natural selection, which he considered a diversifying process. Today, studies of convergent evolution span most biological disciplines, from paleontology to cognitive science to molecular biology (e.g., Emery & Clayton 2004, Whittall et al. 2006, Zhou et al. 2008). These contemporary studies of convergent evolution use many different methods, look over all evolutionary timescales, and focus on many different types of traits (e.g., molecular, cellular, anatomical, physiological, and behavioral).

Here we focus on recent advances in understanding the molecular basis of evolutionary convergence. In the past several decades, research on the molecular mechanisms of convergence has expanded rapidly (e.g., Gompel & Prud'homme 2009, Christin et al. 2010, Elmer & Meyer 2011). Technological breakthroughs in molecular biology, advances in phylogenetic inference, and the application of these tools in nonmodel systems have led to a flurry of research activity. Researchers can now ask not only if two similar traits have convergently evolved but also when and why similar molecular solutions evolve independently. Understanding the molecular mechanisms of repeated evolution will contribute to a more general and predictive formulation of evolutionary theory.

Levels: Convergence Is Hierarchical

Questions about the molecular mechanisms of phenotypic convergence are inherently hierarchical. Convergent phenotypes may or may not share a similar molecular basis. Even within the molecular realm, similarities can occur at many levels (e.g., nucleotide, allele, gene, network, pathway, function). Similarity at one hierarchical level does not necessarily imply similarity at another level. For example, different mutations, even in the same gene, can have different functional effects but still have a similar phenotypic outcome (e.g., Rosenblum et al. 2010). To understand the causes of convergent evolution, we must distinguish clearly among levels, and thus we advocate an explicit incorporation of hierarchical thinking into the study of convergent evolution. To reflect this hierarchical perspective, we advocate using the following definitions for two important terms: convergence and parallelism.

Convergence describes a phenotypic pattern. Convergence is the independent evolution of similar phenotypes and does not specify whether these phenotypes are found in close or distant

relatives or are caused by similar or different genetic underpinnings (**Figure 1**). Although outside our focus on molecular mechanism, defining convergence at the phenotypic level also requires clear and hierarchical thinking. For example, similarity can be found in morphological structure, in resulting function (e.g., different morphological structures recruited for similar functions), and in evolutionary trajectory (e.g., taxa in different regions of morphospace evolving in similar directions).

Parallelism describes a shared molecular explanation. Parallelism is the use of a shared mechanism to produce convergent phenotypes and can occur at different hierarchical levels (**Figure 1**). Parallelism should only be reported when strong evidence links a specific molecular mechanism to a convergent phenotype. When parallelism is reported, the level of mechanistic similarity should always be specified (e.g., nucleotide, allele, gene, network, pathway, function).

We explore the history and alternative definitions for the terms convergence and parallelism in **Figure 1** and the sidebar, Distinguishing Between Convergence and Parallelism. We also discuss methods for uncovering cases of convergence and parallelism. Using clarified terminology and a hierarchical perspective, we can now ask what factors influence the probability that similar molecular mechanisms (i.e., parallelism) are responsible for the independent evolution of similar phenotypic patterns (i.e., convergence).

DISTINGUISHING BETWEEN CONVERGENCE AND PARALLELISM

Arguments about the meaning and usage of the terms convergence and parallelism go back well over 100 years. A more detailed analysis of these debates can be found elsewhere (Gould 2002, Pearce 2012). Here, we briefly summarize the history of usage and several issues that are relevant to our exploration of the molecular basis of phenotypic convergence.

Over the past century, two primary usages of the terms convergence and parallelism have been common. The first usage distinguishes convergence from parallelism based on phylogenetic relatedness (**Figure 1*a***, *left*). In this case, parallelism is defined as independently evolved phenotypic similarity in closely related taxa, and convergence is defined as independently evolved phenotypic similarity in distantly related taxa. These definitions correspond to the earliest usage of these terms (e.g., Scott 1891, Osborn 1905) and are still common today (e.g., Conte et al. 2012). A phylogenetic distinction between convergence and parallelism dovetails with the expectation that more closely related taxa may be more likely to exhibit shared solutions to evolutionary challenges. The hypothesis that close relatives more often have a shared molecular basis for convergent phenotypes is supported by a number of studies in natural and experimental systems (e.g., Bollback & Huelsenbeck 2009; Conte et al. 2012). However, there are also problems with using phylogenetic similarity to distinguish between convergence and parallelism. First, there are many empirical examples that run counter to phylogenetic expectations [distant relatives that share a molecular basis for convergence (e.g., Manceau et al. 2010) and close relatives that do not (e.g., Steiner et al. 2009)]. Second, phylogenetic relatedness itself is not a mechanism and can be a poor proxy for a deeper question about evolutionary repeatability. Third, many authors, even at the turn of the last century, felt that using phylogenetic relatedness to distinguish convergence from parallelism was unsatisfying because taxonomic relatedness provides a distinction "of degree rather than of kind" (Scott 1891), as there are no absolute criteria to differentiate "close" from "distant" relatives.

The second usage distinguishes convergence from parallelism based on underlying developmental genetic mechanisms (**Figure 1*a***, *right*). This usage is common in the genetics of adaptation literature and also in studies that focus on questions of homology across broader segments of the tree of life. For example, Wake et al. (2011) defined convergence as "similarity resulting from different developmental genetic mechanisms" and parallelism as "similarity resulting from the same developmental genetic mechanisms." Similarly, Gould (2002) defined parallelism as independent phenotypic evolution "channeled from within by homologous generators." The developmental genetic

usage shares a similar limitation as the phylogenetic usage above, namely that it is difficult to unambiguously define whether genetic mechanisms are similar or dissimilar. Causal changes can be identical (e.g., the same nucleotide changes) or partial (e.g., different mutations in the same gene) and can occur at different hierarchical levels (e.g., nucleotide, gene, pathway, function).

In addition to the two most common usages, there have been other semantic proposals. First, researchers studying trait evolution often define convergence versus parallelism based on the geometry of trait change in morphospace (e.g., Stayton 2008). Second, Arendt & Reznick (2008) advocated use of the term convergence for all cases of independent evolution of similar phenotypes, entirely removing the term parallel from usage in this subfield. Third, other workers have avoided the debate by using the broader terminological umbrella of "homoplasy" or "repeated evolution" for all nondivergent evolutionary change.

Given the circuitous history and multiple conflicting usages of the terms convergent and parallel in evolutionary biology, we advocate the following usage convention. Our usage is most similar to that of Scotland (2011), who suggested that the term parallelism should be restricted to defining genetic mechanisms such that only some examples of phenotypic convergence are explained by molecular parallelism.

Phenotypic level. Convergence describes a phenotypic pattern. Convergence is the independent evolution of similar phenotypes and does not specify whether these phenotypes are found in close or distant relatives or have similar or different genetic underpinnings (**Figure 1*b***).

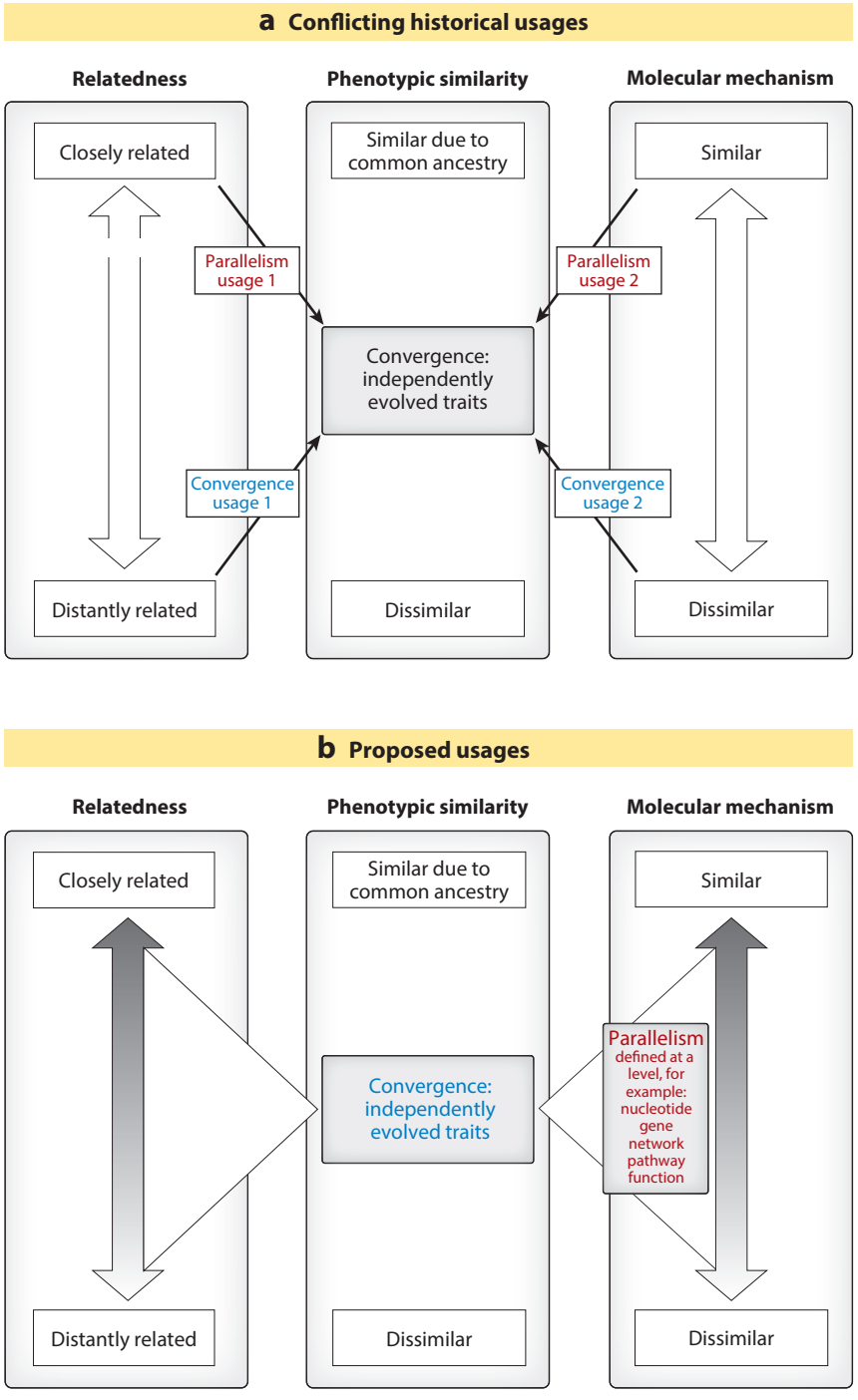
Molecular level. Parallelism describes a shared molecular explanation. Parallelism is the use of a shared mechanism to produce convergent phenotypes and can occur at different hierarchical levels (**Figure 1*b***). Parallelism should only be reported when a molecular basis for phenotypic convergence is found and when the hierarchical level of mechanistic similarity must be specified (e.g., nucleotide, allele, gene, network, pathway, function).

Phylogenetic level. Existing terminology can describe the degree of phylogenetic relatedness without constructing discrete categories of "closely" versus "distantly" related, which will differ subjectively across studies. Therefore we advocate specifying the taxonomic level of the comparison or using a less subjective proxy for phylogenetic distance like divergence time. We recognize that taxonomic delimitations can be in flux and that regions of genomes evolve at different rates (e.g., measures of genetic distance vary across the genome). However, divergence depths can be reported with confidence intervals and are robust enough for broad comparisons (e.g., distinguishing convergence within a genus versus convergence within an order).

Identifying cases of phenotypic convergence and molecular parallelism requires a rigorous and integrative approach. This is reviewed in the sidebar, Best Practices for Empirical Studies of the Molecular Basis of Convergence.

Figure 1

(*a*) Two common uses of the terms convergence and parallelism are depicted conceptually. The first usage distinguishes convergence from parallelism based on phylogenetic relatedness. In this case, parallelism and convergence are defined as independently evolved phenotypic similarity in closely versus distantly related taxa, respectively. The second usage distinguishes convergence from parallelism based on underlying developmental genetic mechanisms. In this case, parallelism and convergence are defined as independently evolved phenotypic similarity with similar and different developmental genetic mechanisms, respectively. The sidebar, Distinguishing Between Convergence and Parallelism, discusses the history of, and difficulties with, these two frameworks. (*b*) To facilitate conceptual clarity, we advocate the following terminological conventions. Convergence is the independent evolution of similar phenotypes and does not specify whether these phenotypes are found in closely or distantly related taxa or are caused by similar or different genetic mechanisms. Parallelism is the use of a shared mechanism to produce convergent phenotypes and can occur at different hierarchical levels. The term parallelism is therefore used only when a molecular explanation for phenotypic convergence is uncovered.



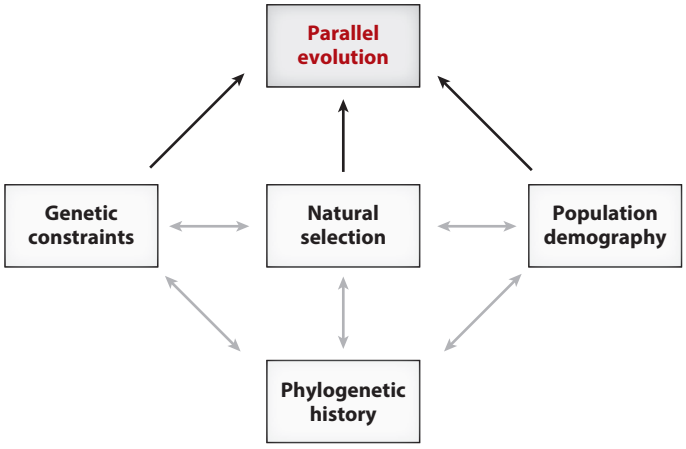


Figure 2

The four general determinants of convergent and parallel evolution. Natural selection, population demography, and genetic constraints directly influence (*black arrows*) the probability of parallel evolution. The dynamics of natural selection determine whether populations are evolving toward similar adaptive optima. Demographic characteristics (e.g., population size and gene flow) influence the mutational input to a population and the probability of allele sharing across populations. Myriad genetic constraints influence whether all evolutionary paths are equally accessible. Phylogenetic history indirectly influences the probability of parallel evolution through the other determinants (e.g., to the extent that closer relatives are more likely to share environmental, demographic, and genetic characteristics). All of the determinants interact (*gray arrows*), for example, when natural selection shapes genetic architecture or population demography modulates the efficacy of selection.

THE DETERMINANTS OF CONVERGENT AND PARALLEL EVOLUTION

Here we address the overarching question: What factors influence when parallel molecular mechanisms will underlie phenotypic convergence? We present four general determinants that influence the probability of parallel and convergent evolution (**Figure 2**). First, the dynamics of natural selection determine whether populations are evolving toward phenotypically similar adaptive optima. Second, population demography modulates the efficacy of selection and also influences the amount of genetic variation in a population. Third, myriad genetic constraints determine whether all evolutionary paths are equally accessible. Fourth, phylogenetic history shapes the other determinants because organisms that are close relatives are more likely to share environmental, demographic, and genetic characteristics.

To enumerate how natural selection, population demography, genetic constraints, and phylogenetic history influence the probability of parallel evolution, we draw on the recent literature from several complementary lines of evolutionary research. We focus largely on studies in natural systems. Empirical studies in natural systems generally compare taxa with convergent phenotypes and their close relatives with divergent phenotypes. Thus they shed light on the molecular mechanisms of dramatic and compelling examples of convergent evolution in the wild. We also draw on laboratory experimental evolution studies, which generally employ model systems (e.g., viruses, bacteria, yeast, worms, and fruit flies). Experimental evolution studies dissect the evolution of replicate populations in controlled selective regimes over multiple time points. Thus, rather than retrospectively studying the "finished products" of evolution, experimental studies provide insight into the repeatability of evolutionary trajectories. Finally, we draw on theoretical studies, which rely on mathematical models and simulations. Few theoretical studies directly investigate

the probability of evolutionary convergence. However, many theoretical studies focus on the conditions promoting evolutionary predictability at the molecular level and can thus shed light on factors that influence convergent and parallel evolution.

Our goal is not to be exhaustive; there are many additional factors and studies that we do not have space to present (e.g., **Table 1**). Rather, we highlight studies that have led to important insights in the four theme areas, and we underscore the importance of considering a diversity of factors in studies of the molecular basis of phenotypic convergence.

Selection: Defining the Adaptive Landscape

Natural selection has long been considered a primary determinant of phenotypic convergence (e.g., Simpson 1953). When selection is responsible for convergent evolution, similar traits evolve repeatedly because different lineages experience similar environments that favor similar adaptive solutions. Shared selection regimes not only increase the probability of convergence but also increase the likelihood of molecular parallelism. When independent taxa are adapting to the same environment and climbing the same fitness peak, the opportunity for molecular parallelism increases. Theoretical work supports this intuition clearly. For example, Orr (2005) explored the probability of fixation of the same mutation in two independent populations and found the probability of parallelism at the nucleotide level is greater under a model including natural selection compared with a purely neutral model.

Beyond the intuitive result that selection increases the probability of parallel evolution, we know little about the influence of different selection regimes. Few studies have asked how the strength, mode, and stability of selection affect the dynamics of parallel evolution. Theoretical studies of evolutionary predictability generally assume that natural selection is identical and constant in all populations (i.e., mutations have the same fitness effects across replicates; e.g., Maynard Smith 1970, Gillespie 1991, Orr 2005; but see Whitlock & Gomulkiewicz 2005 and Uecker & Hermisson 2011 as examples of studies that allow selection to vary in space or time). Experimental evolution studies also tend to focus on replicate and constant selection environments (e.g., Lenski & Travisano 1994). Studies in natural systems rarely provide an opportunity to hold all other factors constant while evaluating the effects of nonidentical selection regimes on the probability of parallelism. Therefore it is not yet possible to review specific characteristics of selection regimes that affect the likelihood of parallelism at different hierarchical levels. Instead, we reiterate the importance of selection as a foundational process that shapes the adaptive landscape on which the other constraints operate. We also highlight the need for additional work on the dynamics of natural selection, and the potential for nonadaptive convergence (e.g., Losos 2011), in the section on Frontiers below.

History: Influencing the Probability of Shared Constraints

Much attention has been given to whether close relatives with convergent phenotypes are more likely to exhibit parallel genetic changes than distant relatives. More closely related taxa might be expected to share more similar environments and more similar genetic architectures. For example, at the closest phylogenetic level, similar solutions to similar adaptive problems may be common because closely related populations have a shared pool of standing genetic variation. Even if gene flow or shared ancestral variation does not unite populations, close relatives are likely to share genetic variance-covariance matrices and have similar genetic backgrounds for gene interactions. In contrast, distant relatives may have less opportunity for parallelism at the molecular level for many reasons, including the possibility that functional pathways may not contain orthologous

Table 1 Examples of genetic and demographic factors that influence the probability of parallel evolution[a]

Factor	Effect on the probability of parallelism
Population size	Larger population size increases total mutational input and the efficacy of selection, thus increasing the probability of parallelism
Gene flow	Higher levels of gene flow can constrain local adaptation but can also facilitate adaptive allele sharing, thus increasing the probability of parallelism
Source of adaptive alleles	Whether adaptive alleles come from new mutations, standing genetic variation, or horizontal gene transfer influences the probability that the same alleles will be reused across populations
Clonal interference	When multiple beneficial mutations arise independently in large asexual populations, mutations with the largest beneficial effect fix first, increasing the probability of parallelism at the nucleotide and adaptive walk levels
Recombination	Organisms and genomic regions vary in recombination rate and, thus, their ability to decouple beneficial mutations from the background on which they arose, affecting the probability of parallelism
Mutation rate	Organisms and genomic regions vary in their mutation rate; a higher mutation rate increases the potential for new beneficial mutations, the likelihood that all mutations will be tested and, thus, the probability of parallelism
Mutational target size	Smaller genes and genomes have fewer nucleotides that can be affected by mutations and, thus, a higher probability of parallel beneficial mutations occurring over a short time frame
Epistasis	The accessibility of adaptive paths depends on the genetic background in which mutations appear—regions of the genome that are less constrained by genetic interactions are more likely to exhibit parallelism
Pleiotropy	Regions of the genome that are less constrained by negative pleiotropy are more likely to be reused in adaptive evolution and, thus, seen as shared targets for parallelism
Position in network	Genes that are less deeply embedded in genetic networks are less likely to be constrained and more likely to exhibit parallelism
Gene expression pattern	Genes that are expressed in few tissues and/or fewer developmental stages are less likely to be constrained and more likely to exhibit parallelism
Loss versus gain of function	Functions are easier to lose than to gain; thus losses of function are more likely to be involved in parallelism
Coding versus regulatory changes	Regulatory and coding regions may differ in their average evolutionary rates, mutational target size, and degree of negative pleiotropy in ways that influence the probability of parallelism, but variation in these characteristics within each category of genomic element must also be considered
Functional redundancy	Genomic regions with functional redundancy are less constrained and more likely to exhibit parallelism
Number of beneficial mutations	When fewer beneficial mutations are available, the probability of parallelism increases
Mutational effect size	Beneficial mutations with larger effect sizes have a higher probability of fixation and, thus, of contributing to parallelism
Length of adaptive walk	Shorter adaptive walks, and those that share a first-step mutation, have a higher probability of being repeated in independent populations and contributing to parallelism

[a]Many of these factors have complex, or sometimes unknown, relationships with the likelihood of parallelism. Thus, we present only simple illustrations of how each factor might influence the probability of a shared molecular basis for phenotypic convergence.

genes. Determining whether close relatives exhibit a shared molecular basis for convergent phenotypes is a first step toward understanding why closer relatives might be more likely to exhibit molecular parallelism. Relatedness is not itself a mechanism to explain convergent and parallel evolution but rather a proxy for the environmental, demographic, and genetic similarities that might be experienced by close relatives. Thus, we briefly review the evidence that closer relatives

are more likely to share molecular mechanisms of phenotypic convergence and then turn to the demographic and genetic factors that influence the probability of molecular parallelism and can be studied directly.

Relatedness. The relationship between relatedness and molecular parallelism has been discussed extensively in studies of natural systems. Empirical studies of wild populations are rarely designed to explicitly test the relationship between relatedness and parallelism, but posthoc synthesis shows that all patterns are possible. Sometimes closely related taxa have a shared basis for convergent phenotypes (e.g., Reed et al. 2011), and sometimes they do not (e.g., Steiner et al. 2009). Similarly, sometimes distantly related taxa have a shared basis for convergent phenotypes (e.g., Yokoyama & Yokoyama 1990), and sometimes they do not (e.g., Shen et al. 2012). In general, the hypothesis that more closely related taxa tend to exhibit a shared molecular basis for convergent phenotypes appears to be supported in studies of natural systems (Conte et al. 2012); however, additional work is needed to determine the generality of this pattern.

The correlation between relatedness and molecular parallelism has also been addressed in experimental evolution studies. Experimental evolution studies generally hold phylogenetic history constant because they establish experimental replicates from a single line. However, some experimental evolution studies address hypotheses about relatedness and parallel evolution by prediverging different lab lines before imposing a common selection regime. For example, Bollback & Huelsenbeck (2009) found an inverse relationship between parallel evolution at the nucleotide level and divergence distance, consistent with the hypothesis that more closely related taxa are more likely to exhibit parallel genetic mechanisms underlying convergent phenotypes. Although it is more difficult to experimentally evolve diverged lines for eukaryotic organisms with longer generation times, it is possible to include relatedness as a covariate in animal experimental evolution studies by bringing wild strains of varying levels of relatedness into the lab (e.g., Matos et al. 2004).

Demography: Determining the Population Context

Demography has not classically been recognized as a determinant of convergent evolution. However, demographic factors (e.g., population size, migration rate) strongly influence the efficiency of selection and the potential for adaptive phenotypic convergence. Demography also affects the potential for stochastic processes to lead to nonadaptive convergence. Moreover demographic factors directly influence the probability of parallelism at the molecular level by modulating the balance among selection, migration, and drift. Here we provide two examples of the importance of population demography for the probability of parallel evolution.

Source of adaptive alleles. The source of adaptive alleles influences the probability of shared genetic underpinnings to convergent phenotypes. Basic population genetic models often assume that adaptation is caused by a single new mutation sweeping rapidly to fixation in a panmictic population. However, real populations exhibit spatial structure and can exchange genes. Therefore any mechanism that allows the recruitment of the same adaptive allele in multiple populations increases the probability of parallelism. Here we briefly address three of these potential mechanisms.

First, shared ancestral variation increases the probability that populations reuse the same genetic variation when confronted with novel environments (Barrett & Schluter 2008). For example, stickleback fish have repeatedly colonized and adapted to freshwater environments, and these replicate populations have independently been exposed to similar selection environments (Colosimo et al. 2005). However, the genetic underpinnings of adaptation are not always independent. For example, the ancestral marine population contains low-frequency alleles (in the

Eda gene controlling body armor) that facilitate adaptation to freshwater environments (Colosimo et al. 2005), and the same allele has independently increased in frequency in replicate populations. Second, adaptive alleles can cross population boundaries via gene flow, and the probability of parallelism increases with the likelihood of adaptive allele sharing across populations. For example, Grant et al. (2004) provide evidence that introgression (i.e., gene flow) contributed to convergence in beak shape and body size in two species of Darwin's finches over a 20-year period. Third, horizontal or lateral gene transfer (movement of genetic material between species) can facilitate adaptive alleles moving across diverse species. Evidence that lateral gene transfer can explain similar phenotypes in distinct taxa is well established in bacteria (e.g., transfer of antibiotic resistance genes; Ochman et al. 2000) and also observed in eukaryotes (e.g., animal taxa have acquired genes to produce carotenoid pigments from fungi; Cobbs et al. 2013).

In all of these cases (i.e., shared standing genetic variation, gene flow, lateral gene transfer), the ultimate source of adaptive alleles is shared (i.e., adaptive alleles are identical by descent). The contribution of alleles that are identical by descent to convergent evolution is controversial because the causal variants are not independently generated. However, when selection acts in parallel on the same source material in different populations, both allele frequencies and population trait values change independently over time. Moreover the independent increase in frequency of an adaptive allele may occur on different genetic backgrounds and in nonidentical selection contexts, providing opportunities to develop a more sophisticated understanding of the dynamics of parallel evolution. Fundamentally, because changes in allele frequency over time occur independently, we consider the source of adaptive alleles as an important factor influencing the probability of parallel evolution. Of course if taxa do not have the potential to share adaptive alleles and adaptation occurs only from new mutations, the probability of parallelism, and the speed at which it occurs, decreases.

Population size. Populations sometimes share adaptive alleles as described above, but often populations do not recruit shared ancestral variation or exchange genes. When populations evolve independently, the probability of parallelism at the nucleotide level depends on the likelihood that the same mutations arise independently and subsequently increase in frequency. In this case, population size strongly affects the probability of convergence by influencing the dynamics of genetic drift, natural selection, and mutation. Most intuitively, population size affects the impact of genetic drift, and thus evolutionary predictability. Because the role of chance in allele frequency change is so pronounced in small populations, natural selection is less efficient in fixing beneficial mutations. The heightened effect of drift in small populations decreases the probability that the same beneficial mutation will be selected in independent populations (Jain & Krug 2007). In addition, population size affects the rate at which mutations appear. Smaller populations have lower potential for de novo beneficial mutations than larger populations, simply because fewer individuals provide fewer total genomic targets for mutations to occur. Similarly, smaller populations maintain less standing genetic variation than larger populations, again because there are fewer individuals to contribute to allelic diversity. Therefore the probability of parallelism is expected to be higher in large populations: More mutational input increases the chances the same beneficial mutation will eventually arise in multiple independent populations, and decreased stochastic effects allow for the efficient selection of beneficial mutations.

Few studies have directly manipulated population size and evaluated the resulting probability of parallel evolution. However, experimental evolution studies support the general prediction that organisms with larger population sizes exhibit increased parallelism. At the extreme end of the population-size spectrum, Wichman et al. (2000) found 62% of nucleotide substitutions were identical between two lines of bacteriophage experimentally adapted to high heat. Other viral evolution studies have found a large range of parallelism at the nucleotide level (25–50%; e.g.,

Bull et al. 1997, Ferris et al. 2007), but rates of parallelism at the nucleotide level for viruses are still substantially higher than those observed for multicellular organisms with smaller population sizes. For example, Denver et al. (2010) conducted a mutation rescue study in replicated lines of the eukaryotic worm *Caenorhabditis elegans* and reported much lower levels of parallelism at the nucleotide level (7%). Broad conclusions are premature given that these studies were not designed to explicitly test for probability of parallelism as a function of population size and that some studies find no evidence for genetic parallelism even for organisms with relatively large population sizes [e.g., *Escherichia coli* (Fong et al. 2005)].

It is also important to note that large population size is often correlated with other demographic and genetic factors that may increase the probability of molecular parallelism. Specifically, organisms with large population sizes tend to have short generation times, another demographic factor that can affect the probability of parallelism. In addition, organisms with large population sizes tend to have small genomes, a genetic factor discussed below that also increases the likelihood of identical mutations occurring independently in replicate populations.

Genetics: Altering the Accessibility of Evolutionary Trajectories

Genetic constraints have long been recognized as determinants of convergent and parallel evolution. Similar traits can evolve repeatedly because different lineages have similar "building blocks" to employ. In other words, genetic architecture determines whether all evolutionary paths are equally accessible and, thus, the probability of convergent evolution. Genetic factors also strongly influence the probability that parallel molecular mechanisms underlie convergent phenotypes. In fact, genetic factors provide some of the clearest expectations for when similar molecular mechanisms will underlie phenotypic convergence. For example, differences in mutation and recombination rates (across organisms and across genomic regions) affect the probability of parallel evolution. Similarly, variation in levels of functional constraint across genomic regions explains why some genomic elements are observed as shared targets for parallel evolution across lineages. Thus, of the four determinants of convergent and parallel evolution, we focus most extensively on predictions generated from studies of the genetic architecture of convergence. We provide examples of how characteristics of genes, genomic regions, and whole genomes can influence the probability of parallelism.

Genotype-phenotype degeneracy. Some traits exhibit more degeneracy in their genotype-phenotype map than others. In other words, there may be one or many molecular paths to a particular phenotype. More "degenerate" pathways (i.e., many-to-one genotype-phenotype mapping) should decrease the probability of shared genetic solutions to common selection regimes. On one end of the spectrum are examples in which constraints in a molecular pathway led to parallelism at the nucleotide level across divergent taxa. One compelling example is the evolution of tetrodotoxin (TTX) resistance. Multiple lineages of snake have independently evolved TTX-resistant sodium channels as an adaptation to ingest poisonous prey. Some TTX-resistant lineages are very closely related (e.g., multiple species of North American garter snakes of the genus *Thamnophis* that eat toxic newts; Feldman et al. 2009), but others are more distantly related (e.g., *Liophis epinephelus* (Latin America), *Rhabdophis tigrinus* (Asia), and *Amphiesma pryeri* (Asia); Feldman et al. 2012). Not only do all of these snake species employ the same gene, *Na(v)1.4*, for TTX resistance but they have mutations in the same regions of the protein, often at the same specific nucleotide positions. This degree of molecular parallelism is remarkable because experimental work shows that additional mutations should theoretically be able to facilitate TTX resistance (Feldman et al. 2012). Thus, there are likely constraints, for example, strong purifying selection on ion channel function, that restrict the set of actualized changes in natural systems.

Even more amazing is the reuse of the same domains of the same gene in other very distantly related species that are also highly resistant to TTX (or similar toxins), including several species of puffer fishes (Jost et al. 2008) and a species of clam (Bricelj et al. 2005). Thus, extreme predictability at the molecular level is observed in nature. Such examples of genetic parallelism in distantly related species reflect reduced degeneracy where only a restricted number of changes can produce a specific phenotype.

At the other end of the spectrum are examples where multiple different genetic changes can lead to a similar phenotype. This is elegantly illustrated by the evolution of antifreeze glycoproteins in ice fishes. Antarctic notothenioid fishes and Arctic gadid fishes have independently evolved astoundingly similar mechanisms for surviving in freezing waters: Both groups have freeze-avoidance systems powered by antifreeze glycoproteins. The glycoproteins recruited in the two cases are from the same family of glycoproteins and have nearly identical primary protein structures. However, the specific genes involved differ between the Antarctic and Arctic fishes. The antifreeze glycoprotein genes exhibit dramatic coding sequence differences and evolved independently from different evolutionary precursors (Chen et al. 1997). Despite the independent evolutionary origins of the antifreeze glycoprotein genes, there is an amazing amount of functional convergence between the Antarctic and Arctic fishes' antifreeze systems, including the anatomical sites of synthesis of glycoproteins, the fine-scale localization of glycoproteins in gastrointestinal and circulatory systems, and the pathway facilitating antifreeze glycoprotein recycling (Evans et al. 2012). Thus the degree of degeneracy influences the hierarchical level on which parallelism is likely to occur. In the fish antifreeze case, many-to-one genotype-phenotype mapping means that parallel evolution is found not at the nucleotide or gene level but rather at the molecular pathway leading to a shared functional basis for phenotypic convergence.

Propensity for duplication. Gene duplication can be a powerful source of novelty and functional fine-tuning. We looked at one example above in which different members of a glycoprotein gene family were independently recruited as "antifreeze" genes in cold-water fishes. However, there are some gene families that exhibit high rates of evolution and duplication and are reused in adaptive evolution across lineages. For example, at nested levels of the animal tree, changes in opsin genes are associated with shifts in photopigment sensitivity. Often these shifts are considered adaptive when spectral tuning could increase fitness in particular environments (e.g., to see in dim light conditions or to detect UV reflectance of intraspecific communication ornaments). Even within this system, parallel evolution occurs at different hierarchical levels. For example, Nagai et al. (2011) found parallel evolution at the nucleotide level in the RH1 opsin gene in cichlid fish adapting to different water depth conditions. One specific amino acid replacement was repeatedly associated with the transition from shallow to deep water, and the reverse replacement was repeatedly associated with the transition from deep to shallow water. In other cases, different solutions are found within a particular gene family. For example, Carvalho et al. (2007) found evidence that UV sensitivity has evolved numerous times in avian taxa. Although the evolution of UV sensitivity often involves a particular visual pigment gene (SWS1), at least two different mechanisms are implicated in increased short wavelength visual capability. Even if details vary across systems, the molecular mechanisms driving opsin evolution are often similar: Tandem duplications often lead to a release from functional constraint, and subsequent coding and regulatory changes allow fine-tuning of spectral sensitivity. Therefore the propensity for gene duplications with adaptive value in particular gene families can influence the probability of reuse of particular genes and, thus, the probability of molecular parallelism.

In addition to the importance of tandem duplication in convergent evolution, chromosome- or genome-scale duplications can also be important mechanisms of parallel evolution. Experimental

evolution studies provide some of the most concrete examples of adaptive parallelisms at the level of genome structure. For example, convergent aneuploidies (copy number changes in particular chromosomes) were found in different lines of glucose-limited *Saccharomyces cerevisiae* (Dunham et al. 2002). Specifically, increased copy number of carbohydrate-processing genes occurred via chromosomal duplication (Dunham et al. 2002). Although not necessarily linked to higher-level phenotypic convergence, parallel gene rearrangements and parallel shifts in ploidy have also been observed in experimental evolution studies (e.g., Bollback & Huelsenbeck 2009, Gerstein et al. 2006). Thus, when rapid changes in genome structure are possible and potentially adaptive, increased parallelism can be observed at different levels of genomic architecture.

Gene interactions. Constraints imposed by gene interactions strongly influence the probability of shared molecular mechanisms for convergent phenotypes. Here we touch on two examples that play a critical role in determining what genomic segments are independently reused in convergent evolution. First, genes differ in their degree of pleiotropy (i.e., whether they have multiple phenotypic effects). For example, a gene that influences a single phenotypic trait is less constrained by pleiotropy than a gene that influences multiple traits. Negative pleiotropy can limit the number of available adaptive solutions because mutations with potentially beneficial effects on one trait can have deleterious effects on another trait and, thus, be selected against. The importance of negative pleiotropy has been demonstrated repeatedly in experimental evolution studies (e.g., Cooper & Lenski 2000, Ostrowski et al. 2008). Negative pleiotropy has also been discussed in studies of natural systems. For example, the melanocortin-1 receptor gene (*Mc1r*) is only one of many genes in the vertebrate melanin synthesis pathway, but it is repeatedly used as a mechanism for convergence in pigmentation in diverse groups (e.g., mammals, birds, lizards; e.g., Kronforst et al. 2012, Manceau et al. 2010, Rosenblum et al. 2010). Although there are alternative explanations for this observation, minimal pleiotropy is one possible cause of gene reuse (Manceau et al. 2010). It is often difficult to directly address the constraints imposed by pleiotropy in natural systems, but another pigmentation study provides an example of how pleiotropy can be minimized. Although not specifically a study of convergent evolution, Linnen et al. (2013) demonstrated that light coat color in deer mice (*Peromyscus maniculatus*) was composed of multiple traits, each associated with distinct regions within the *Agouti* locus. Therefore different mutations, even in the same gene, can have separate effects and, thus, minimize pleiotropy at the mutational level (Linnen et al. 2013).

Second, genes differ in their epistatic interactions (e.g., when gene expression at one locus is conditional on another locus). Because genes interact, the fitness effects of particular mutations often depend on the genetic background on which they arise (e.g., Weinreich et al. 2005). Again, experimental evolution studies provide the most direct evidence for the importance of epistasis. For example, in a bacterial antibiotic resistance study, negative epistasis was shown to constrain the order in which beneficial mutations fixed (Salverda et al. 2011), confirming that epistasis can considerably reduce the number of available evolutionary trajectories. Thus there are different ways that genetic constraints imposed by gene interactions can decrease the number of accessible evolutionary paths, and these constraints generally increase the probability of molecular parallelism.

Structural versus regulatory changes. There is a long-standing debate over whether structural or regulatory changes fuel adaptive evolutionary change (e.g., King & Wilson 1975, Hoekstra & Coyne 2007, Carroll 2008, Wittkopp & Kalay 2012). Much of the debate centers around whether *cis*-regulatory elements, by virtue of evolving more quickly, exhibiting less negative pleiotropy, and encompassing a larger mutational target size, are more likely to be targets of adaptive evolution than coding regions of genes (Stern & Orgogozo 2008). From the perspective of the genetics of convergent evolution, how differences in constraint in regulatory and coding regions translate

into the probability of parallelism remains unresolved. There are numerous examples of both coding mutations and regulatory changes contributing to convergent phenotypes. We have already discussed a number of systems in which parallelism in protein-coding genes is responsible for convergent phenotypes [e.g., parallel use of *Na(v)1.4* for toxin resistance in snakes (Feldman et al. 2012); parallel use of *Mc1r* for blanched coloration in White Sands lizards (Rosenblum et al. 2010)]. In addition, there are many examples of similar regulatory changes underlying convergent phenotypes. For example, Frankel et al. (2012) show that convergent evolution of trichome patterns in two species of *Drosophila* (*D. sechellia* and *D. ezoana*) result from independent but similar changes in enhancers in the *cis*-regulatory region of the *Shavenbaby* gene. Similarly, Reed et al. (2011) conclude that *cis*-regulatory changes associated with the gene *optix* underlie convergence in red wing pattern in mimetic species of *Heliconius* butterfly.

It is also possible for multiple types of molecular changes to be recruited in a single system. The anthocyanin synthesis pathway in flowering plants is responsible for conferring red, blue, and purple floral coloration, and modifications to this pathway have led to multiple instances of phenotypic convergence. For example, loss of floral anthocyanins has led to pale flowers multiple times in the columbine *Aquilegia* (Whittall et al. 2006), and shifts in anthocyanin synthesis are associated with the transition from blue to red flowers in multiple plant genera (Smith & Rausher 2011). Many of these convergent floral color transitions are associated with regulatory changes (e.g., Whittall et al. 2006, Smith & Rausher 2011). However, even within the realm of anthocyanin biosynthesis, a number of genomic processes, including *cis*-regulation (Yuan et al. 2013), *trans*-regulation (Streisfeld & Rausher 2009), gene duplication (Des Marais & Rausher 2008), and gene loss (Smith & Rausher 2011), contribute to convergent evolution. Thus, it seems likely that specific characteristics of different genomic elements (e.g., degree of pleiotropy, mutational target size), whether regulatory or coding, will ultimately determine the probability of molecular parallelism and the repeated use of specific segments of the genome during convergent evolution.

Mutational target size. All genes, and genomes, have a finite number of sites, and each site can assume a finite number of states. Therefore the raw genetic material for generating phenotypic diversity is inherently limited. Because genes differ in their length and genomes differ in their overall size, the probability of molecular convergence across populations is influenced by the mutational target size. Genome size is also often correlated with demographic factors discussed above that influence mutational input to a population, including population size and generation time. Therefore, we might predict that organisms with small genomes, which also tend to have large population sizes and fast generation times, are more likely to exhibit identical genetic solutions to adaptive problems. The combination of small mutational target size (e.g., due to small genome size) and large mutational input (e.g., due to large population size) means that all possible beneficial mutations are more likely to be "tested." As discussed above, experimental evolution studies generally support the prediction that organisms with small genome sizes exhibit increased parallelism at the nucleotide and gene level.

Within a genome, target size may be less important than other genetic constraints. For example, in a study of 35 replicate yeast lines adapting to a common environmental challenge (the fungicide nystatin), adaptive mutations were found in only four genes (Gerstein et al. 2012). All of the genetic targets were in the ergosterol biosynthesis pathway, demonstrating parallelism at the pathway level. There was less parallelism at the nucleotide level, with 20 different mutations found. However, there was parallelism at the gene level: 95% of the lines accumulated adaptive mutations in one of two genes. In this study, degree of parallelism at the gene level was not predicted by gene length, likely because loss-of-function mutations in some genes have catastrophic organismal effects regardless of their length (Gerstein et al. 2012). Thus,

factors like propensity for gain-of-function versus loss-of-function mutations and degree of negative pleiotropy may often be more influential than mutation target size.

Recombination rate. Recombination rates vary widely across the tree of life and across regions of the genome (e.g., Nachman 2002, Vos & Didelot 2009). A number of processes contribute to variation in recombination rates within and across genomes, but differences in reproductive mode (i.e., sexual versus asexual) largely determine baseline recombination rates. Recombination rate in turn influences the population genetics of convergent evolution and the probability of molecular parallelism. When recombination is absent or rare, beneficial mutations that arise simultaneously within a population, but in different individuals, must compete with one another. Because recombination cannot link mutations that arise in different asexual individuals, the mutation with the largest beneficial effect increases in frequency before subsequent beneficial mutations arise on this background. This process is termed clonal interference and is commonly observed in experimental evolution studies of asexual organisms (e.g., Lenski et al. 1991, Crozat et al. 2005). Assuming the same beneficial mutations arise in replicate asexual populations experiencing similar environmental conditions, clonal interference will increase the determinism of evolutionary trajectories (e.g., Cuevas et al. 2002) and, therefore, increase the probability of parallelism.

Although reproductive mode explains broadscale variation in recombination rates across taxa, recombination rates also vary within genomes. For example, position on chromosome and proximity to centromere influence recombination rates in sexually reproducing organisms (Nachman 2002). For sexually reproducing organisms (i.e., when clonal interference does not dominate), higher levels of recombination can potentially increase parallelism at the genetic level. Because recombination can allow beneficial mutations to escape their genetic backgrounds, different beneficial mutations that arise independently in the same population can ultimately be combined in the same individual. Therefore different processes in different groups of organisms (e.g., recombination in sexual organisms and clonal interference in asexual organisms) can affect the probability of molecular parallelism.

Number, effect size, and order of beneficial mutations. Theoretical studies suggest that there are several additional factors that are difficult to study in natural systems but likely influence the probability of molecular parallelism. For example, theoretical studies suggest that the number of possible beneficial mutations should affect the probability of parallelism at the nucleotide level. Building on work by Maynard Smith (1962, 1970) and Gillespie (1991), Orr (2005) explored the probability of fixation of the same beneficial mutation in two independent populations. Orr showed that the probability that two populations evolving in similar but independent selective environments will fix the same mutation is $2/(n + 1)$, where n is the number of different beneficial mutations available. Under this model, the probability of parallel evolution is independent of many additional factors, such as the distribution of fitness effects of alleles, size of the gene, patterns of epistasis, and recombination. Orr's 2005 model used a scenario of strong selection and weak mutation so that only one-step mutations were considered and only one mutant was present at any one time in the population. In natural populations, these assumptions may or may not hold. Therefore, other authors have tested the robustness of this model under different conditions, for example multiple-step adaptive walks (Rokyta et al. 2006, Beisel et al. 2007, Joyce et al. 2008). Orr's predictions for the probability of parallel evolution appear fairly robust to different distributions for the fitness effects of the beneficial mutations but are less robust in multiple-step walks. Thus in simple one-step scenarios it might be possible to predict parallel evolution based on the number of beneficial mutations available, but predictability at the nucleotide level is more challenging for multiple-step adaptive walks.

Although we expect predictability at the nucleotide level to decrease with increasing length of adaptive walks, theory suggests that similarity across populations will be observed in higher-level characteristics of the adaptive walk: the number, order, and effect size of substitutions. For example, mutations with larger beneficial fitness effect are predicted to fix early in the adaptive walk. Studies on this topic build from early work by Fisher (1930), who used a geometric model to predict the probability that mutations with different phenotypic effect sizes would be favorable. Orr (1998) evaluated the adaptive walk for populations off the phenotypic optimum (i.e., analogous to natural populations that recently experienced an environmental shift). He showed that the beneficial mutations fixed during an adaptive walk follow an approximately exponential distribution and that the first mutation fixed can have a relatively large effect size. In addition, theoretical studies suggest that there can be convergence across populations in the number of steps in adaptive walks. For example, Orr (2003) demonstrated that (starting from a random point on an adaptive landscape) there is a lower limit on the number of substitutions required to reach a local optimum. Therefore, adaptive walks may be of similar length across populations even if they do not include the same substitutions. Although similarities across populations in the adaptive walk may seem different than other types of parallelism (e.g., parallelism at the nucleotide or gene level), they reinforce the importance of hierarchical thinking and indicate the potential for deep process-oriented similarity during adaptive convergent evolution.

FRONTIERS

The grand challenge for studies of the molecular basis of phenotypic convergence is now to build a unified understanding of convergent evolution across systems and across hierarchical levels. We highlight several particularly important areas for future research.

Expectations: Generating Null Models

First, it is important to formulate rigorous expectations for the conditions that promote convergent and parallel evolution. For example, null models are sometimes used when studying phenotypic patterns of convergence (e.g., Stayton 2008), but they are rarely used when studying the molecular basis of phenotypic convergence. Null models are needed that describe the probability of parallel evolution under different conditions (e.g., in the absence of selection, with identical selection regimes in replicate populations, with differences in strength of selection in replicate populations). In addition, researchers should formulate clear expectations about the relative contributions of the four determinants of convergence under different conditions. For example, some factors may dominate over certain timescales. Specifically, when closely related populations have the potential to share adaptive alleles, demographic factors (e.g., rates of gene flow) may have a stronger influence over the probability of parallelism than genetic factors (e.g., mutational target size). Mathematical modeling and experimental approaches will thus be essential for refining expectations about the factors that govern the probability of parallel evolution.

Interactions: Understanding Multiple Factors Simultaneously

Second, an important avenue of research is to study the nonlinear and interactive effects of the factors that facilitate convergent evolution. We have enumerated a number of factors that influence the probability of convergent and parallel evolution. However, little is known about how these factors interact. For example, we know that strong and shared selection regimes increase the probability of phenotypic convergence, but we do not know how strong selection must

be to facilitate convergent evolution in populations of different sizes or in organisms with genomes of different sizes. In addition, there is no reason that the factors discussed above should be expected to have linear or additive effects on the probability of convergence and parallelism. For example, the predictability of a population's evolutionary path initially scales with increasing population size (less stochastic effects of drift) and mutation rate (greater likelihood of the same beneficial mutation occurring). However, these effects become nonlinear as population size and mutation rate continue to increase (increased potential of multiple mutations in the same individual; Szendro et al. 2013). Therefore it is important to evaluate factors over a wide range of parameter space and to study multiple factors simultaneously.

Integration: Synthesizing Across Levels, Systems, and Subfields

Third, an important frontier in studies of the molecular basis of convergent evolution is to integrate data across different systems and across hierarchical levels. It remains difficult to quantitatively compare patterns of convergence and parallelism. For example, researchers might be interested in whether the degree of similarity is concordant between phenotypic and molecular levels or

BEST PRACTICES FOR EMPIRICAL STUDIES OF THE MOLECULAR BASIS OF CONVERGENCE

To develop a general understanding of the molecular basis of convergent evolution, integrative studies of pattern and process in natural systems are needed. Many studies of the genetics of convergence endeavor primarily to link genotype to phenotype. However, understanding the historical, environmental, and demographic context is also essential. Thus, we suggest the following best practices for studies of the molecular mechanisms of phenotypic convergence in natural systems.

Linking phenotype and phylogeny. Before investigating the molecular mechanisms of convergence, researchers must first ensure that the phenotype of interest is convergent. First, researchers should sample not only the putatively convergent taxa but also closely related taxa with divergent phenotypes. Second, researchers should quantify phenotypes rigorously even if they appear qualitatively similar. Third, researchers should define convergent evolution in a phylogenetic context. This requires an explicit integration of phenotypic data with a molecular phylogeny and should incorporate uncertainty in the phylogeny and the model. For example, the independent evolution of similar phenotypes can be identified from ancestral state reconstruction, comparisons of phylogenetic and genetic distance, or inferred shared selective regime (e.g., Muschick et al. 2012, Ingram & Mahler 2013). Researchers should also explicitly consider the null expectation that patterns reflect chance alone, for example, by simulating random evolution of quantitative characters along phylogenies (e.g., Stayton 2006, 2008).

Finally, authors should provide a quantitative index of relatedness for the focal taxa. The appropriate relatedness index will vary depending on the type of organisms under study and the type of study conducted. For example, time since divergence, percent sequence difference, and number of generations can all be used to express the phylogenetic scale over which convergence is observed. Defining relatedness in a quantitative framework will help resolve the binary, and somewhat arbitrary, distinction between close and distant relatives. It can also allow refined predictions about the effect of shared history on convergence and parallelism.

Linking phenotype and selection. Before undertaking a study of the molecular mechanisms of convergence, authors should also determine whether there is adaptive significance to convergent phenotypes. As discussed in other reviews (e.g., Losos 2011), there are multiple alternative explanations for convergent phenotypes, not all of

which are adaptive. Convergent phenotypes can also arise due to chance, due to correlated selection on other traits, or due to exaptation (Losos 2011). Studying the molecular mechanisms of convergence can be fruitful whether the phenotypes are adaptive or not (e.g., comparing patterns of molecular evolution for adaptive versus nonadaptive convergent phenotypes). Regardless, researchers should not take phenotypic patterns of convergence as a priori evidence of natural selection. Rather, authors should test whether convergently evolved traits are adaptive.

Testing the adaptive value of convergent phenotypes requires quantitative assessment of environmental variables and of the fitness consequences of focal traits. The specific methods used to assess the phenotype-environment-fitness link will vary across organisms and studies. There are numerous well-established methods to measure environmental characteristics (e.g., remote sensing, microhabitat surveys, predation surveys). Similarly, there are many approaches to measuring fitness (e.g., population growth estimates, performance assays, mark-recapture studies).

Increased attention to the environmental and demographic context for adaptation is also critically important. Studies of convergent evolution in natural systems often treat selection as a presence-absence variable. However, the mode, strength, direction, and stability of selection all influence the probability of convergent evolution. Studies that explicitly consider the temporal and spatial dynamics of selection and compare selection regimes for replicated taxa will provide important insights. Similarly, most studies of natural systems do not explicitly consider demographic factors, such as population size and migration rate. However, demographic factors directly modulate response to selection and, thus, strongly influence the probability of convergence. There is tremendous opportunity for studies that explicitly address how the interaction between population demography and natural selection influences the dynamics of convergent and parallel evolution.

Linking phenotype and genotype. Before attempting to study the underlying genetic basis of a convergent trait, researchers should also test, rather than assume, that target phenotypes are heritable. For example, the contribution of phenotypic plasticity to convergent evolution is understudied but should not be dismissed. Phenotypic plasticity can, at least initially, allow independent lineages to solve similar environmental challenges without requiring genetic adaptation. Therefore, researchers should assess the heritability of traits of interest (e.g., using breeding, common garden, or reciprocal transplant experiments; mid-parent offspring regression; or genomic data among siblings).

Having determined that the phenotypes of interest are heritable, there are a number of approaches that can be used to link genotype to phenotype. The utility of different methods (e.g., candidate gene studies, quantitative trait locus analysis, population genomic scans for selection, whole genome sequencing, gene expression analysis, functional assays) have been reviewed elsewhere (e.g., Nielsen 2005, Arendt & Reznick 2008, Stinchcombe & Hoekstra 2008, Springer et al. 2011). Of course, determining the genetic and developmental basis of convergent phenotypes is no small task. Understanding the molecular basis of convergent traits is more complicated than evaluating the genetic basis of a focal phenotype in a single lineage. Studies must sample not only the taxa with convergent phenotypes but also relatives with divergent traits. Moreover, it remains difficult to take a truly unbiased approach that provides conclusive results and does not predispose a researcher to find certain kinds of genotype-phenotype links. For example, if candidate gene studies uncover a genotype-phenotype association at a particular locus, it may not be the only contributing gene and, thus, may cause bias toward genes of large effect. Even if whole genome sequencing approaches are taken, there may still be a bias toward finding genes of large effect and coding rather than regulatory changes (e.g., Rockman 2011). Similarly, even if gene expression studies are used, results are highly dependent on the tissue and developmental time point sampled. Thus all methods have important strengths and weaknesses. Matching the methods to the focal question and using multiple approaches in each system is wise. Similarly, researchers should approach as many hierarchical levels as possible, from causal mutations to functional effects to physiology. By studying multiple levels, researchers can more confidently define parallelism at a particular hierarchical level and ultimately facilitate a more general synthesis of the conditions promoting convergence and parallelism.

whether particular systems show more convergence than others. Developing new indices of convergence would potentially provide a way to make these comparisons more explicit. Some metrics to quantify convergent evolution (e.g., the CONVEVOL software package; Stayton 2013) and the strength of convergent evolution (e.g., Arbuckle et al. 2014) are already available. However, the challenge remains to define convergence at multiple hierarchical levels in a way that is amenable to downstream statistical analysis. Thus the statistical analysis of convergent evolution, to facilitate integration of data across systems and hierarchical levels, is an area ripe for additional attention.

To address the leading-edge questions in this field, there is also tremendous opportunity for integration across subfields. Observational and experimental approaches in natural systems are essential for growing the "catalog" of convergence (Martin & Orgogozo 2013), supplying data for meta-analyses, and providing inspiration (explored in the sidebar, Best Practices for Empirical Studies of the Molecular Basis of Convergence). Theoretical and experimental evolution approaches will also play a key role in testing hypotheses about the interactions among factors at different levels. In addition to traditional mathematical, statistical, and experimental approaches, digital evolution platforms (using artificial computer code "organisms") offer a novel way to address questions about convergent evolution (O'Neill 2003). Critically important is a commitment to interaction and integration across subfields with theory informed by empirical observations, experiments designed to test theory, and analytical methods developed to address data integration challenges.

CONCLUSION

With a renewed focus on the factors that influence convergent and parallel evolution, we can address fundamental questions about what processes shape evolutionary trajectories and when evolutionary paths are repeatable and predictable. Moreover, advances and synthesis in the study of the molecular basis of phenotypic convergence are relevant to additional subfields. For example, studies of the molecular mechanisms of convergence can have practical utility in applied fields such as conservation biology and biomedicine. Understanding the factors that promote convergence and parallelism can facilitate a predictive approach (e.g., to predict which bacterial strains are more likely to develop resistance to particular antibiotics and what interventions are likely to succeed). The framework developed to study molecular parallelism and phenotypic convergence can also be applied to other disciplines. For example, the core processes that affect evolutionary trajectories have analogs in community ecology (Vellend 2010), and ecological communities can exhibit convergence in trophic structure, species composition, species ecology, or species traits (e.g., Melville et al. 2006, Helsen et al. 2012, Segar et al. 2013). Therefore, a deeper understanding of phenotypic convergence and molecular parallelism could be useful for understanding the factors that influence convergence in composition, assembly, and function of ecological communities. Thus studies of the molecular basis of phenotypic convergence provide a promising foundation to build a unified understanding of the processes that produce similarity across the tree of life, both in evolutionary biology and beyond.

SUMMARY POINTS

1. We define convergence as the independent evolution of similar phenotypes and parallelism as a shared molecular mechanism to produce convergent phenotypes. Defining these terms more clearly will help promote conceptual unity in the field.

2. Parallelism can occur at different hierarchical levels, from nucleotide to gene to pathway to function. Evidence for a specific level of mechanistic similarity should always be presented when reporting parallelism.
3. Four primary determinants influence the probability that convergent phenotypes have a shared molecular basis: natural selection, phylogenetic history, population demography, and genetic constraints.
4. Studies in natural systems should endeavor to not only link genotype to phenotype but also explicitly define the phylogenetic, demographic, and environmental context for convergent evolution.
5. Mathematical and statistical approaches are needed to formalize null models for convergent and parallel evolution, make predictions about interactions among factors that influence the probability of parallel evolution, and develop new metrics to integrate data across hierarchical levels and study systems.
6. Experimental evolution studies are essential for testing theoretical models, rigorously manipulating the core factors and their interactions, and setting expectations that can be evaluated in natural systems.
7. Future research should focus on integration across systems and hierarchical levels. We identify opportunities for a more integrative framework and a deeper understanding of the processes that influence repeatability, and predictability, in evolution.

DISCLOSURE STATEMENT

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

ACKNOWLEDGMENTS

We thank H. Bradley Shaffer, Stevan Springer, Luke Harmon, Craig Miller, Christopher Martin, Eveline Diepeveen, and the Rosenblum Lab for constructive comments on an earlier version of the manuscript. We also thank the graduate students in the 2012 Patterns of Diversity seminar for fruitful discussions on this theme. We gratefully acknowledge support from the National Science Foundation to E.B.R. (DEB-1054062).

LITERATURE CITED

Arbuckle K, Bennett CM, Speed MP. 2014. A simple measure of the strength of convergence? *Methods Ecol. Evol.* 5(7):685–93

Arendt J, Reznick D. 2008. Convergence and parallelism reconsidered: What have we learned about the genetics of adaptation? *Trends Ecol. Evol.* 23(1):26–32

Barrett RDH, Schluter D. 2008. Adaptation from standing genetic variation. *Trends Ecol. Evol.* 23(1):38–44

Beisel CJ, Rokyta DR, Wichman HA, Joyce P. 2007. Testing the extreme value domain of attraction for distributions of beneficial fitness effects. *Genetics* 176(4):2441–49

Bollback JP, Huelsenbeck JP. 2009. Parallel genetic evolution within and between bacteriophage species of varying degrees of divergence. *Genetics* 181(1):225–34

Bricelj VM, Connell L, Konoki K, MacQuarrie SP, Scheuer T, et al. 2005. Sodium channel mutation leading to saxitoxin resistance in clams increases risk of PSP. *Nature* 434(7034):763–67

Bull JJ, Badgett MR, Wichman HA, Huelsenbeck JP, Hillis DM, et al. 1997. Exceptional convergent evolution in a virus. *Genetics* 147(4):1497–507
Carroll SB. 2008. Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* 134(1):25–36
Carvalho LS, Cowing JA, Wilkie SE, Bowmaker JK, Hunt DM. 2007. The molecular evolution of avian ultraviolet- and violet-sensitive visual pigments. *Mol. Biol. Evol.* 24(8):1843–52
Chen LB, DeVries AL, Cheng CHC. 1997. Convergent evolution of antifreeze glycoproteins in Antarctic notothenioid fish and arctic cod. *Proc. Natl. Acad. Sci. USA* 94(8):3817–22
Christin P-A, Weinreich DM, Besnard G. 2010. Causes and evolutionary significance of genetic convergence. *Trends Genet.* 26(9):400–5
Cobbs C, Heath J, Stireman JO, Abbot P. 2013. Carotenoids in unexpected places: gall midges, lateral gene transfer, and carotenoid biosynthesis in animals. *Mol. Phylogenet. Evol.* 68(2):221–28
Colosimo PF, Hosemann KE, Balabhadra S, Villarreal G, Dickson M, et al. 2005. Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. *Science* 307(5717):1928–33
Conte GL, Arnegard ME, Peichel CL, Schluter D. 2012. The probability of genetic parallelism and convergence in natural populations. *Proc. R. Soc. B-Biol. Sci.* 279(1749):5039–47
Cooper VS, Lenski RE. 2000. The population genetics of ecological specialization in evolving *Escherichia coli* populations. *Nature* 407(6805):736–39
Crozat E, Philippe N, Lenski RE, Geiselmann J, Schneider D. 2005. Long-term experimental evolution in *Escherichia coli*. xii. DNA topology as a key target of selection. *Genetics* 169(2):523–32
Cuevas JM, Elena SF, Moya A. 2002. Molecular basis of adaptive convergence in experimental populations of RNA viruses. *Genetics* 162(2):533–42
Darwin C. 1859. *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life*. London: John Murray
Denver DR, Howe DK, Wilhelm LJ, Palmer CA, Anderson JL, et al. 2010. Selective sweeps and parallel mutation in the adaptive recovery from deleterious mutation in *Caenorhabditis elegans*. *Genome Res.* 20(12):1663–71
Des Marais DL, Rausher MD. 2008. Escape from adaptive conflict after duplication in an anthocyanin pathway gene. *Nature* 454(7205):762–65
Dunham MJ, Badrane H, Ferea T, Adams J, Brown PO, et al. 2002. Characteristic genome rearrangements in experimental evolution of *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* 99(25):16144–49
Elmer KR, Meyer A. 2011. Adaptation in the age of ecological genomics: insights from parallelism and convergence. *Trends Ecol. Evol.* 26(6):298–306
Emery NJ, Clayton NS. 2004. The mentality of crows: convergent evolution of intelligence in corvids and apes. *Science* 306(5703):1903–7
Evans CW, Hellman L, Middleditch M, Wojnar JM, Brimble MA, Devries AL. 2012. Synthesis and recycling of antifreeze glycoproteins in polar fishes. *Antarct. Sci.* 24(3):259–68
Feldman CR, Brodie ED, Brodie ED, Pfrender ME. 2009. The evolutionary origins of beneficial alleles during the repeated adaptation of garter snakes to deadly prey. *Proc. Natl. Acad. Sci. USA* 106(32):13415–20
Feldman CR, Brodie ED, Brodie ED, Pfrender ME. 2012. Constraint shapes convergence in tetrodotoxin-resistant sodium channels of snakes. *Proc. Natl. Acad. Sci. USA* 109(12):4556–61
Ferris MT, Joyce P, Burch CL. 2007. High frequency of mutations that expand the host range of an RNA virus. *Genetics* 176(2):1013–22
Fisher RA. 1930. *The Genetical Theory of Natural Selection*. Oxford, UK: Oxford Univ. Press
Fong SS, Joyce AR, Palsson BO. 2005. Parallel adaptive evolution cultures of *Escherichia coli* lead to convergent growth phenotypes with different gene expression states. *Genome Res.* 15(10):1365–72
Frankel N, Wang S, Stern DL. 2012. Conserved regulatory architecture underlies parallel genetic changes and convergent phenotypic evolution. *Proc. Natl. Acad. Sci. USA* 109(51):20975–79
Gerstein AC, Chun H-JE, Grant A, Otto SP. 2006. Genomic convergence toward diploidy in *Saccharomyces cerevisiae*. *PLOS Genet.* 2(9):e145
Gerstein AC, Lo DS, Otto SP. 2012. Parallel genetic changes and nonparallel gene-environment interactions characterize the evolution of drug resistance in yeast. *Genetics* 192(1):241–52

Gillespie JH. 1991. *The Causes of Molecular Evolution*. New York: Oxford Univ. Press
Gompel N, Prud'homme B. 2009. The causes of repeated genetic evolution. *Dev. Biol.* 332(1):36–47
Gould SJ. 2002. *The Structure of Evolutionary Theory*. Cambridge, MA: Belknap
Grant PR, Grant BR, Markert JA, Keller LF, Petren K. 2004. Convergent evolution of Darwin's finches caused by introgressive hybridization and selection. *Evolution* 58(7):1588–99
Helsen K, Hermy M, Honnay O. 2012. Trait but not species convergence during plant community assembly in restored semi-natural grasslands. *Oikos* 121(12):2121–30
Hoekstra HE, Coyne JA. 2007. The locus of evolution: evo devo and the genetics of adaptation. *Evolution* 61(5):995–1016
Ingram T, Mahler DL. 2013. SURFACE: detecting convergent evolution from comparative data by fitting Ornstein-Uhlenbeck models with stepwise AIC. *Methods Ecol. Evol.* 4(5):416–25
Jain K, Krug J. 2007. Deterministic and stochastic regimes of asexual evolution on rugged fitness landscapes. *Genetics* 175(3):1275–88
Jost MC, Hillis DM, Lu Y, Kyle JW, Fozzard HA, Zakon HH. 2008. Toxin-resistant sodium channels: parallel adaptive evolution across a complete gene family. *Mol. Biol. Evol.* 25(6):1016–24
Joyce P, Rokyta DR, Beisel CJ, Orr HA. 2008. A general extreme value theory model for the adaptation of DNA sequences under strong selection and weak mutation. *Genetics* 180(3):1627–43
King M, Wilson AC. 1975. Evolution at two levels in humans and chimpanzees. *Science* 188:107–16
Kronforst MR, Barsh GS, Kopp A, Mallet J, Monteiro A, et al. 2012. Unraveling the thread of nature's tapestry: the genetics of diversity and convergence in animal pigmentation. *Pigment Cell Melanoma Res.* 25(4):411–33
Lenski R, Rose M, Simpson S, Tadler S. 1991. Long-term experimental evolution in *Escherichia coli*. 1. Adaptation and divergence during 2,000 generations. *Am. Nat.* 138(6):1315–41
Lenski R, Travisano M. 1994. Dynamics of adaptation and diversification—a 10,000-generation experiment with bacterial populations. *Proc. Natl. Acad. Sci. USA* 91(15):6808–14
Linnen CR, Poh Y-P, Peterson BK, Barrett RDH, Larson JG, et al. 2013. Adaptive evolution of multiple traits through multiple mutations at a single gene. *Science* 339(6125):1312–16
Losos JB. 2011. Convergence, adaptation, and constraint. *Evolution* 65(7):1827–40
Manceau M, Domingues VS, Linnen CR, Rosenblum EB, Hoekstra HE. 2010. Convergence in pigmentation at multiple levels: mutations, genes and function. *Philos. Trans. R. Soc. B-Biol. Sci.* 365(1552):2439–50
Martin A, Orgogozo V. 2013. The loci of repeated evolution: a catalog of genetic hotspots of phenotypic variation. *Evolution* 67(5):1235–50
Matos M, Simoes P, Duarte A, Rego C, Avelar T, Rose MR. 2004. Convergence to a novel environment: comparative method versus experimental evolution. *Evolution* 58(7):1503–10
Maynard Smith J. 1962. The limitations of molecular evolution. In *The Scientist Speculates: An Anthology of Partly-Baked Ideas*, ed. IJ Gould, pp. 252–56. New York: Basic Books
Maynard Smith J. 1970. Natural selection and the concept of a protein space. *Nature* 225:563–64
Melville J, Harmon LJ, Losos JB. 2006. Intercontinental community convergence of ecology and morphology in desert lizards. *Proc. R. Soc. Lond. Ser. B-Biol. Sci.* 273(1586):557–63
Muschick M, Indermaur A, Salzburger W. 2012. Convergent evolution within an adaptive radiation of cichlid fishes. *Curr. Biol.* 22(24)2362–68
Nachman MW. 2002. Variation in recombination rate across the genome: evidence and implications. *Curr. Opin. Genet. Dev.* 12(6):657–63
Nagai H, Terai Y, Sugawara T, Imai H, Nishihara H, et al. 2011. Reverse evolution in *RH1* for adaptation of cichlids to water depth in Lake Tanganyika. *Mol. Biol. Evol.* 28(6):1769–76
Nielsen R. 2005. Molecular signatures of natural selection. *Annu. Rev. Genet.* 39:197–218
O'Neill B. 2003. Digital evolution. *PLOS Biol.* 1(1):11–14
Ochman H, Lawrence JG, Groisman EA. 2000. Lateral gene transfer and the nature of bacterial innovation. *Nature* 405(6784):299–304
Orr HA. 1998. The population genetics of adaptation: the distribution of factors fixed during adaptive evolution. *Evolution* 52(4):935–49
Orr HA. 2003. The distribution of fitness effects among beneficial mutations. *Genetics* 163:1519–26

Orr HA. 2005. The probability of parallel evolution. *Evolution* 59(1):216–20
Osborn HF. 1905. The ideas and terms of modern philosophical anatomy. *Science* 21:959–61
Ostrowski EA, Woods RJ, Lenski RE. 2008. The genetic basis of parallel and divergent phenotypic responses in evolving populations of *Escherichia coli*. *Proc. R. Soc. B-Biol. Sci.* 275(1632):277–84
Pearce T. 2012. Convergence and parallelism in evolution: a neo-Gouldian account. *Br. J. Philos. Sci.* 63:429–48
Reed RD, Papa R, Martin A, Hines HM, Counterman BA, et al. 2011. Optix drives the repeated convergent evolution of butterfly wing pattern mimicry. *Science* 333(6046):1137–41
Rockman M. 2011. Evolutionary genomics and systems biology. *Q. Rev. Biol.* 86(4):353–54
Rokyta DR, Beisel CJ, Joyce P. 2006. Properties of adaptive walks on uncorrelated landscapes under strong selection and weak mutation. *J. Theor. Biol.* 243(1):114–20
Rosenblum EB, Roempler H, Schoeneberg T, Hoekstra HE. 2010. Molecular and functional basis of phenotypic convergence in white lizards at White Sands. *Proc. Natl. Acad. Sci. USA* 107(5):2113–17
Salverda MLM, Dellus E, Gorter FA, Debets AJM, van der Oost J, et al. 2011. Initial mutations direct alternative pathways of protein evolution. *PLOS Genet.* 7(3):e1001321
Scotland RW. 2011. What is parallelism? *Evol. Dev.* 13(2):214–27
Scott WB. 1891. On the osteology of *Mesohippus* and *Leptomeryx*, with observations on the modes and factors of evolution in the Mammalia. *J. Morphol.* 5:301–406
Segar ST, Pereira RAS, Compton SG, Cook JM. 2013. Convergent structure of multitrophic communities over three continents. *Ecol. Lett.* 16(12):1436–45
Shen Y-Y, Liang L, Li G-S, Murphy RW, Zhang Y-P. 2012. Parallel evolution of auditory genes for echolocation in bats and toothed whales. *PLOS Genet.* 8(6):e1002788
Simpson GG. 1953. *The Major Features of Evolution*. New York: Columbia Univ. Press
Smith SD, Rausher MD. 2011. Gene loss and parallel evolution contribute to species difference in flower color. *Mol. Biol. Evol.* 28(10):2799–810
Springer SA, Crespi BJ, Swanson WJ. 2011. Beyond the phenotypic gambit: molecular behavioural ecology and the evolution of genetic architecture. *Mol. Ecol.* 20(11):2240–57
Stayton CT. 2006. Testing hypotheses of convergence with multivariate data: morphological and functional convergence among herbivorous lizards. *Evolution* 60(4):824–41
Stayton CT. 2008. Is convergence surprising? An examination of the frequency of convergence in simulated datasets. *J. Theor. Biol.* 252(1):1–14
Stayton CT. 2013. R Package CONVEVOL: quantify and assess the significance of convergent evolution. R Package version 0.1–1. **http://cran.r-project.org/src/contrib/Archive/convevol/**
Steiner CC, Roempler H, Boettger LM, Schoeneberg T, Hoekstra HE. 2009. The genetic basis of phenotypic convergence in beach mice: similar pigment patterns but different genes. *Mol. Biol. Evol.* 26(1):35–45
Stern DL, Orgogozo V. 2008. The loci of evolution: How predictable is genetic evolution? *Evolution* 62(9):2155–77
Stinchcombe JR, Hoekstra HE. 2008. Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits. *Heredity* 100(2):158–70
Streisfeld MA, Rausher MD. 2009. Altered *trans*-regulatory control of gene expression in multiple anthocyanin genes contributes to adaptive flower color evolution in *Mimulus aurantiacus*. *Mol. Biol. Evol.* 26(2):433–44
Szendro IG, Franke J, de Visser JAGM, Krug J. 2013. Predictability of evolution depends nonmonotonically on population size. *Proc. Natl. Acad. Sci. USA* 110(2):571–76
Uecker H, Hermisson J. 2011. On the fixation process of a beneficial mutation in a variable environment. *Genetics* 188(4):915–30
Vellend M. 2010. Conceptual synthesis in community ecology. *Q. Rev. Biol.* 85(2):183–206
Vos M, Didelot X. 2009. A comparison of homologous recombination rates in bacteria and archaea. *Int. Soc. Microb. Ecol. J.* 3(2):199–208
Wake DB, Wake MH, Specht CD. 2011. Homoplasy: from detecting pattern to determining process and mechanism of evolution. *Science* 331(6020):1032–35
Weinreich DM, Watson RA, Chao L. 2005. Perspective: sign epistasis and genetic constraint on evolutionary trajectories. *Evolution* 59(6):1165–74
Whitlock MC, Gomulkiewicz R. 2005. Probability of fixation in a heterogeneous environment. *Genetics* 171(3):1407–17

Whittall JB, Voelckel C, Kliebenstein DJ, Hodges SA. 2006. Convergence, constraint and the role of gene expression during adaptive radiation: floral anthocyanins in *Aquilegia*. *Mol. Ecol.* 15(14):4645–57

Wichman HA, Scott LA, Yarber CD, Bull JJ. 2000. Experimental evolution recapitulates natural evolution. *Philos. Trans. R. Soc. Lond. Ser. B-Biol. Sci.* 355(1403):1677–84

Wittkopp PJ, Kalay G. 2012. *Cis*-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. *Nat. Rev. Genet.* 13(1):59–69

Yokoyama R, Yokoyama S. 1990. Convergent evolution of the red- and green-like visual pigment genes in fish, *Astyanax fasciatus*, and human. *Proc. Natl. Acad. Sci. USA* 87(23):9315–18

Yuan Y-W, Sagawa JM, Young RC, Christensen BJ, Bradshaw HD. 2013. Genetic dissection of a major anthocyanin QTL contributing to pollinator-mediated reproductive isolation between sister species of *Mimulus*. *Genetics* 194(1):255–63

Zhou Z, Clarke J, Zhang F. 2008. Insight into diversity, body size and morphological evolution from the largest Early Cretaceous enantiornithine bird. *J. Anat.* 212(5):565–77

Annual Review of Ecology, Evolution, and Systematics

Volume 45, 2014

Contents