æ

5. EVOLUTION AS A POPULATION-GENETIC PROCESS

8 November 2022

With knowledge on rates of mutation, recombination, and random genetic drift in hand, we now consider how the magnitudes of these population-genetic features dictate the paths that are open vs. closed to evolutionary exploitation in various phylogenetic lineages. Because historical contingencies exist throughout the Tree of Life, we cannot expect to derive from first principles the evolutionary source of every molecular detail of cellular diversification. We can, however, use established theory to address more general issues, such as the degree of attainable molecular refinement, rates of transition from one state to another, and the degree to which nonadaptive processes (mutation and random genetic drift) contribute to phylogenetic diversification.

Much of the field of evolutionary theory is concerned with the mechanisms maintaining genetic variation within populations, as this ultimately dictates various aspects of the short-term response to selection (Charlesworth and Charlesworth 2010; Walsh and Lynch 2018). Here, however, we are primarily concerned with long-term patterns of phylogenetic diversification, with a focus on mean phenotypes. This still requires knowledge of the principles of population genetics, as evolutionary divergence is ultimately a manifestation of the accrual of genetic modifications at the population level. All evolutionary change initiates as a transient phase of genetic polymorphism, during which mutant alleles navigate the rough sea of random genetic drift, often being evaluated on diverse genetic backgrounds, with some paths being more accessible to natural selection than others.

The primary goals of this most technical of chapters is to summarize some of the more general challenges to understanding how evolutionary change is accomplished and to endow the reader with an appreciation for why the population-genetic details matter. With a specific focus on the ways in which selection acts to promote novel adaptive changes, emphasis will be placed on how the efficiency of selection is compromised or enhanced in different population-genetic environments, sometimes in counter-intuitive ways. Special attention will be given to the ways in which evolutionary rates and outcomes are expected to vary with the effective sizes of populations (N_e).

Most of the theory presented here will be discussed in a generic way, focusing for example on a mutation with selective advantage or disadvantage s, with no connection to the actual underlying trait(s). Such an approach is a necessary prelude to more explicit exploration of particular traits where genotypes can be directly connected to phenotypes and then to fitness. Specific examples to be presented in later chapters include the evolution of protein-protein interfaces, the coevolution of

transcription factors and their binding sites, and the evolution of maximum growth-rate potential. Although the rudimentary technical level of presentation here may be disappointing to high-level theoreticians, the goal is not to overwhelm the reader with a litany of equations and formal derivations (some of which can be found in the Foundations sections), but to facilitate understanding as to how population-genetic theory can transform comparative cell biology into evolutionary cell biology.

The Perils of the Adaptive Paradigm

Ever since Darwin, most discussions with any connection to evolutionary thought have started with the implicit assumption that all organismal traits are products of natural selection. Under this extreme view, the genetic details are irrelevant, as it is believed that natural selection is capable of finding the optimal solution to any environmental challenge, extinction being the alternative. Such logic underlies virtually every study in the field of evolutionary ecology. Closer to the subject material herein, a massive number of papers in cell biology end with a speculative paragraph on why the trait being studied (and its sometimes arcane structure) must have been refined to its current state by selective forces, almost always in the absence of any direct evidence or even an awareness that such evidence ought to be sought.

An appreciation for the power of natural selection is one of the great advances of the life sciences over the past century. However, problems arise when the wand of natural selection is deemed to be the only mechanism relevant to evolutionary change, as this eliminates any hope for broader understanding of evolutionary processes, and often leads to false narratives. Starting with the conclusion that the phenotype under investigation is a necessary product of natural selection, the only remaining challenge is to identify the actual agent of selection. If one hypothesis fails, one moves on to another possibility, but always with unwavering certainty that selection must somehow be involved. Many biologists have spent entire careers wandering down such paths in search of an adaptive explanation for a particular biological feature, and sometimes never finding it.

This is not to say that optimization thinking has completely mislead us with respect to the evolution of alternative behavioral and/or life-history strategies, and theory certainly explains how natural selection is often able to bring most phenotypes within the vicinity of adaptive peaks (Fisher 1930; Rice et al. 2015). However, as touched upon in Chapter 4, the evolutionary outcomes that are achievable by natural selection depend critically on levels of mutation, drift, and recombination. Moreover, owing to the stochastic nature of these processes, even under constant selective pressures, the phenotypic states of populations are expected to wander over time. Finally, depending on the bias and granularity of mutational effects, the most common phenotype need not even be the optimum for a given environment.

To begin to explore these ideas, this chapter will close with an overview of the concept of long-term steady-state distributions of mean phenotypes. This will also provide a more formal analysis of the drift-barrier hypothesis, introduced in the previous chapter in the context of mutation-rate evolution, demonstrating the conditions under which traits are expected to exhibit gradients in performance scaling

with N_e. The points being made here are not just arcane technical nuances. As will unfold in subsequent chapters focused on particular cellular traits, owing to the population-genetic and molecular features of biology, many aspects of evolution at the cellular level are best understood not by invoking the all-powerful guiding hand of natural selection, but by appreciating the factors that limit the reach of selection.

The Fitness Effects of New Mutations

Before proceeding with the theory, an overview of the fitness effects of mutations is necessary, as this defines the landscape and potential granularity of evolutionary change accessible to natural selection. As pointed out in Chapter 4, mutations with an absolute selective advantage/disadvantage (s) much smaller than the reciprocal of the effective population size ($1/N_e$) are essentially invisible to the eyes of natural selection. Thus, for random genetic drift to impose significantly different barriers to the evolution of a trait in different lineages, there must be a substantial pool of mutations with small enough deleterious effects that they can drift to fixation in species with small but not large N_e. This process will be further facilitated if mutations are biased in the negative direction. As outlined below, numerous lines of evidence are consistent with both conditions.

First, studies of serially bottlenecked mutation-accumulation (MA) lines across diverse species consistently reveal a slow per-generation decline in growth rate and other fitness traits (Keightley and Eyre-Walker 1999; Lynch et al. 1999; Baer et al. 2007; Katju and Bergthorsson 2019). Such experiments start with a set of isogenic lines, which are then maintained for large numbers of generations by propagation of just one (clone or selfer) or two (full-sib) individuals per generation. With an N_e this small, natural selection is incapable of promoting or removing individual mutations with fitness effects $< 25\%$ in these experiments, so the data are in full accord with a strong bias of mutations towards deleterious effects. Furthermore, statistical inferences based on the distribution of MA-line performances imply highly skewed distributions of fitness effects. The modes for such distributions are often indistinguishable from zero, with the bulk of mutations having absolute effects $< 1\%$ (almost all negative), although mean deleterious effects can sometimes be as high as 1 to 10% owing to the presence of rare mutations with large negative effects (Keightley 1994; Robert et al. 2018; Böndel et al. 2019).

Second, indirect inferences derived from allele-frequency distributions in natural populations of diverse multicellular species commonly suggest that 10 to 50% of mutations have deleterious effects smaller than 10^{-5}, with the inferred distribution sometimes being bimodal, but always with one mode being near 0.0 (Keightley and Eyre-Walker 2007; Bataillon and Bailey 2014; Huber et al. 2015; Kim et al. 2017; Lynch et al. 2017; Booker and Keightley 2018; Johri et al. 2020). Because many of these studies focus only on the nonsynonymous (amino-acid substitution) sites in protein-coding genes, the full distributions of effects (which would include the often much more numerous synonymous sites, introns, and intergenic DNA) can be expected to be even more skewed towards near-zero values. Substantial evidence also supports the idea that even silent (synonymous) sites in protein-coding genes

are subject to weak selection for usage of particular nucleotides, such that the scaled strength of selection (ratio of the power of selection relative to drift) favoring G/C content is typically in the range of $N_e s = 0.1$ to 1.0, i.e., on the edge of the domain of effective neutrality (Long et al. 2017).

Third, although the preceding inferences are based on indirect extrapolations from statistical distributions, the costs of some kinds of mutations can be derived from first principles. For example, from a knowledge of the total energy budget of a cell and the biosynthetic costs of its building blocks, it is possible to estimate the fractional reduction in cell growth rates resulting from various kinds of mutations (Chapter 17; Foundations 5.1). Bioenergetic considerations of the costs of small nucleotide insertions, which typically comprise $\sim 10\%$ of *de novo* mutations (Sung et al. 2016), imply fractional reductions in fitness far below 10^{-5} (Lynch and Marinov 2015). Likewise, from both the standpoint of biosynthetic expenditures and elemental (e.g., C, N, or S) composition (Chapter 18), the costs of using alternative amino acids or nucleotides imply that s associated with such substitutions is generally $\ll 10^{-5}$ unless there are additional functional consequences. Broad surveys of single amino-acid substitutions in a range of proteins generally imply that most such changes influence protein performance by $< 1\%$, with a large class indistinguishable from zero, and a secondary peak with very substantial effects (Chapter 12). Single-residue changes in protein-protein interfaces or DNA binding sites are also expected to have small effects (Chapters 13 and 21). This diverse set of observations makes clear that there are large fractions of deleterious mutations with fitness effects $< 10^{-5}$ in all species.

Drawing from these observations, as well as other considerations regarding genome content outlined in Foundations 5.1, the genome-wide distribution of fitness effects (DFE) of new mutations is expected to take on different forms in large- and small-N_e species (Figure 5.1). In the former, prokaryotes in particular, $> 95\%$ of the genome consists of coding DNA, of which $\sim 75\%$ are amino-acid substitution sites. A small fraction of silent-site mutations that retain identical nucleotides but alter their strand associations (e.g., A:T vs. T:A) may be absolutely neutral, but otherwise the lower bound of the fitness effect of a silent-site mutation is based on the differential cost of (C+G) vs. (A+T) nucleotide pairs relative to the total cost of building a cell (in ATP equivalents). An expected central peak in the idealized prokaryotic DFE in Figure 5.1 is associated with the substitution of amino acids with biosynthetic costs at sites that are otherwise insensitive to amino-acid identity, whereas the potential lower peak to the right is associated with amino-acid altering mutations with additional functional effects on protein performance.

The form of the DFE is expected to shift in eukaryotic species, owing to the dramatic expansion of genome content. Greater than 95% of nucleotide sites are noncoding and nonfunctional (with mutational effects limited to small differential biosynthetic costs of nucleotides) in many multicellular species, with the genome structures of unicellular eukaryotes typically being intermediate between this extreme and that of prokaryotes (Lynch 2007). Combined with a several order-of-magnitude increase in the cost of construction of large eukaryotic cells relative to prokaryotes, which dilutes the relative biosynthetic costs of individual mutations (Foundations 5.1), this shifts the DFE to the left and creates a more L-shaped form, owing to the elevated number of nonfunctional sites dominating the overall

distribution (Figure 5.1).

Taking all of these observations into consideration, the existence of large pools of mutations with deleterious effects small enough to allow fixation in some lineages but large enough to ensure removal by selection in others is not in doubt. This supports the contention that traits under identical selection pressures in species with different N_e will evolve different mean phenotypes, as discussed in the context of mutation-rate evolution in the previous chapter. Nonetheless, significant caveats remain.

First, statistical and empirical limitations prevent us from knowing with certainty the full form of the distribution of mutations of small effects, e.g., the fraction of mutations with effects $< 10^{-5}$, $< 10^{-6}$, $< 10^{-7}$, $< 10^{-8}$, etc. Thus, the exact form of these regions of the hypothetical DFEs presented in Figure 5.1 should not be taken too seriously. This is a significant concern. As we know that N_e ranges from $\sim 10^4$ to 10^9 (Chapter 4), these are the mutations that can be utilized / purged in some lineages but not in others. In addition, and most importantly, although most of the observations noted above address the general fitness properties of random mutations, they are disconnected from the actual cellular traits that we will wish to eventually explore. Further genetic dissection is essential to inform us as to the precise molecular targets and phenotypic effects of fitness-altering mutations.

The Classical Model of Sequential Fixation

A common conceptual starting point for thinking about the temporal dynamics of adaptive evolution invokes the case in which a trait is under persistent directional selection, with the pace of evolution being slow enough that each consecutive adaptive mutation is fixed before the next beneficial mutation destined to fix arises. In principle, such a scenario can exist if the supply of adaptive mutations is quite limited owing to either a relatively small population size, a low mutation rate to beneficial variants, or both. In this limiting situation, recombination is irrelevant because no two loci are ever simultaneously segregating polymorphisms at meaningful frequencies.

This sequential model of molecular evolution (sometimes also called the strong selection/weak mutation model or the origin-fixation model; Gillespie 1983; McCandlish and Stoltzfus 2014) may only rarely represent reality as it assumes a constant march towards higher fitness. Nonetheless, it serves as a useful heuristic for thinking about several issues concerning the limits to rates of adaptive evolution and how they might scale with population size.

Under this model, the long-term rate of adaptation is equal to the product of three terms: the rate of introduction of beneficial mutations, the fixation probability of such mutations, and their fitness effects. 1) The population-wide rate of origin of new beneficial mutations is NU_b for haploids and $2NU_b$ for diploids, where U_b is the mutation rate (per generation) to beneficial alleles. Depending on the focus, U_b can represent a single nucleotide site, a single gene, or an entire haploid genome. 2) The probability of fixation of a new beneficial mutation is $\simeq 2s(N_e/N)$, where N and N_e are the actual and effective population sizes, and s is the selective advantage

relative to the ancestral allele (Chapter 4). As N_e/N is almost always less than one, this ratio can be thought of as the efficiency with which selection promotes new beneficial mutations, with the maximum fixation rate being $2s$. 3) The increase in fitness per fixation is s for haploids, but $2s$ for diploids. (Note that for diploids, it is assumed here that mutational effects are additive, with heterozygotes having a fitness advantage $1+s$ intermediate to that for the two homozygotes, 1 and $1+2s$).

The expected rate of evolution in terms of fixation numbers is the product of the first two terms,

$$r_e = NU_b \cdot 2s(N_e/N) = 2N_eU_bs, \tag{5.1a}$$

for haploids, and twice this for diploids. The increase in fitness is then equal to

$$\Delta W = r_e \cdot s = 2N_eU_bs^2, \tag{5.1b}$$

for haploids, and $4\times$ this for diploids. (Note that this difference between haploids and diploids is merely a function of how the fitness effects are annotated in diploids. If s is redefined so that mutant heterozygotes and homozygotes have fitnesses $1+(s/2)$ and $1+s$, respectively, then Equations 5.1a,b hold for diploids).

This model for the speed limit to the rate of adaptation is idealized in many ways, as it assumes long-term persistent selection in one direction, constant beneficial mutation rates, and constant effect sizes of fixed mutations. Nonetheless, this simple approach highlights the key roles that the individual population-genetic parameters play in dictating the potential for evolutionary change. For example, Equations 5.1a,b suggest that, all other things being equal, the rate of adaptive evolution should scale linearly with the effective population size (which need not be equivalent to the absolute population size) and with the genome-wide beneficial mutation rate.

However, there is room for caution in interpreting expressions like Equation 5.1b. First, the conditions under which mutations are likely to fix in a stepwise manner are limited. Sequential fixation requires that the average time between fixations (the inverse of $r_e = 2N_eU_bs$ for haploids) be greater than the mean time required for each mutation to fix, which is $\simeq (2/s)\ln(N)$ generations for a haploid population (Walsh and Lynch 2018, Equation 8.4c). It then follows that for sequential fixation to be the rule, $2N_eU_b$ must be smaller than $1/[2\ln(N)]$. Because $\ln(N)$ falls in the narrow range of 9 to 46 over a range of $N = 10^4$ to 10^{20}, as a first-order approximation, the sequential model will only hold if the effective number of beneficial mutations arising per generation is $N_eU_b < 0.01$, i.e., if no more than one beneficial mutation for the trait arises per 100 generations at the effective-population size level.

How likely is this condition to be met? Recall from Chapter 4 that the product of $2N_e$ and the mutation rate per nucleotide site per generation (u) generally falls in the range of $2N_eu = 10^{-3}$ to 10^{-2}. Multiplying $2N_eu$ by the number of selected sites in a chromosomal region of interest and the fraction of mutations that are beneficial converts this quantity to $2N_eU_b$. With a moderate-sized region of 10^5 fitness-relevant sites and just 0.01% of mutations being beneficial, then $2N_eU_b$ for such a region would be in the range of 10^{-2} to 10^{-1}. As this is on the edge of the strict cutoff for the sequential model, it follows that the sequential model cannot be assumed to be generally valid unless the fraction of beneficial mutations is very small. These issues are evaluated in great detail in Weissman and Barton (2012), Weissman and Hallatschek (2014), and Lynch (2020).

The primary reason for concern with violations of the sequential model is that simultaneously segregating mutations (both beneficial and deleterious) at different nucleotide sites interfere with each other in the selection process, diminishing the probabilities of fixation of the good and the purging of the bad (Campos and Wahl 2009, 2010; Frenkel et al. 2014; Pénisson et al. 2017). As new beneficial mutations arise, their fates are determined by the fitness of the linked backgrounds in which they appear, which will vary in large populations (Figure 4.2). Cosegregating deleterious mutations play a particularly prominent role in reducing the rate of adaptive evolution, as the majority of *de novo* mutations are deleterious. If tightly linked to a segregating deleterious mutation, a beneficial mutation will not experience its full intrinsic advantage, and in some cases will be completely overshadowed by the linked background load (Good and Desai 2014).

The central point here is that although the effects of background variation do not alter the expectation that the rate of adaptive evolution will scale positively with N_e, the gradient of scaling is expected to decline substantially with increasing N. Theoretical work suggests that selective interference reduces the scaling of the rate of adaptation from linear with N in the sequential domain to as weakly as logarithmic in N beyond that point (Weissman and Barton 2012; Neher 2013; Weissman and Hallatschek 2014).

Selective interference is commonly observed in long-term laboratory evolution experiments involving microbes. For example, Figure 5.2 illustrates the trajectories of genome-wide mutant-allele frequencies in three replicate populations of *E. coli* grown in just 10 ml of medium. Over a period of 60,000 generations, these populations experienced an average 30% increase in fitness in the culture conditions, albeit at a diminishing rate. The norm in these kinds of experiments is for many mutations to be simultaneously polymorphic, and more often than not, groups of mutations increase (and sometimes decline) in a coordinated manner. As noted earlier for asexual populations of yeast (Figure 4.9), this is a simple consequence of the clonal nature of the experimental populations, as a single, positively selected mutation driving to fixation sweeps along all linked "passenger" mutations (some of which themselves have beneficial or deleterious fitness effects).

Examples can also be seen in Figure 5.2 of mutations reaching very high frequencies in a short period of time, followed by a subsequent decline to 0.0 as other more fit mutant clones take over. In one population, two major clones, each containing multiple mutations, appear to reach equilibrium frequencies, with neither going to fixation (middle panel); this may be a result of some form of frequency-dependent selection, with each clone providing a metabolic product beneficial to the other (Behringer et al. 2018). In another case, there is a massive accumulation of mutations near the midpoint of the experiment (lower panel), owing to the appearance of a mutator strain, which may have hitchhiked to fixation in linkage with a beneficial mutation that it promoted.

Compensating for these constraining effects from selective interference is the fact that a larger fraction of beneficial mutations is exploitable in larger populations. Owing to the fact that efficient selection requires $|N_e s| > 1$ (Chapter 4), populations with larger N_e have access to mutations with smaller s. However, although this expansion in the pool of available beneficial mutations will further tip the balance in favor of higher rates of evolution in larger populations, from Equation 5.1 it can be

seen that the contribution of beneficial mutations to increases in fitness scales with the square of the selective advantage, s^2. Thus, because average s is typically $\ll 1$, broadening the window of mutational availability will have a less than linear effect on the rate of adaptation unless the pool of beneficial mutations is very strongly skewed towards those with small fitness effects.

Surveys of the clonal dynamics in long-term evolution experiments illustrate these principles in a more general way (Nguyen Ba et al. 2019). Clonal competition can lead to a "rich get richer"' scenario, whereby currently successful clones ride a wave to high abundance, and in doing so can more efficiently utilize additional beneficial mutations arising during their clonal expansion. On the other hand, currently low fit clones can occasionally vault their way to success via a fortuitous secondary mutation, but only if the latter has a very high fitness effect.

With respect to the scaling of the rate of adaptation with population size, there are still additional issues to consider. Recall from Chapter 4 that, at the interspecies level, there is a nearly inverse relationship between N_e and the mutation rate (u). From Equation 5.1, such a compensatory effect would yield near independence of the rate of adaptation and N_e, although the potential expansion of the window of exploitable mutations can modify this result.

Finally, there is the matter of time scale. The previous derivations consider the rate of adaptive evolution on a per-generation basis. However, smaller organisms typically have shorter generation times, which will elevate the rate of evolution on an absolute time scale. For example, if the generation length scaled inversely with N_e, this would render the expected rate of evolution (Equation 5.1) proportional to N_e^2 on an absolute time scale. The following simple argument supports the idea of such a generation-time effect. Recall from Figure 4.3 that across the Tree of Life, N_e varies by about five orders of magnitude, from $\sim 10^4$ for some vertebrates to nearly 10^9 for some bacteria. Generation lengths in bacteria tend to be on the order of 0.1 to 1.0 days, whereas those in vertebrates and land plants are generally on the order of 10^2 to 10^4 days, thus spanning nearly five orders of magnitude in the opposite direction of N_e. There can, of course, be considerable variation in generation lengths among organisms with the same N_e, but there is a clear general negative relationship between N_e and generation length.

Taken together, for simple adaptations involving mutations with additive effects, the above observations point toward the potential rate of evolutionary change (per absolute time unit) being greater for organisms with small size, short generation times, and large N_e. However, whether the scaling of adaptive evolutionary potential with N_e is sublinear, linear, or superlinear remains uncertain. Moreover, if both mutation rates and generation lengths scale inversely with N_e, the two effects will cancel out, and the scaling of rates of evolution with N_e on an absolute time scale will follow the expectations of Equations 5.1a,b). Finally, as will be discussed is the following section, for changes involving interactions among loci, the expected scaling of evolutionary rates with N_e can deviate from that described above.

Before proceeding, one final point merits discussion. Although the quantity $r_e = 2N_eU_bs$ is a measure of the expected number of long-term fixations per unit time under the ideal model, because mutation and fixation events are stochastic, considerable variation is expected around this expectation. For a Poisson process, where each rare event is independent of the others, the variance in the amount of

long-term change among replicate populations is equal to the expectation. If, for example, the time interval under consideration is long enough that one beneficial mutation is expected per lineage, the probability that one event actually accrues is just 0.368, but the probability of no fixations is also 0.368, of two is 0.182, of three is 0.061, and of four or more is 0.021. The central point is that considerable variation in observed rates of evolution is expected among lineages exposed to identical selection pressures, and that such dispersion should not be taken as evidence of adaptive differentiation, of the adaptive emergence of mutators/antimutators, or of intrinsic differences in evolutionary potential. Foundations 5.2 provides an even more dramatic example of how divergence among populations exposed to identical selection pressures can exceed that expected under neutral drift.

Vaulting Barriers to More Complex Adaptations

To this point, we have been assuming that the fitness effect of an allele is independent of the genetic background on which it resides. Under this view, Equations 5.1a,b provide the simplest possible model for the rate of adaptation by new mutations, as prior fixations have no bearing on subsequent events. However, this assumption can be violated for at least two reasons. First, for the case of stabilizing selection for an intermediate optimum phenotype or directional selection up the edge of a fitness plateau, each mutation fixation will alter the selection coefficients of future mutations by moving the mean phenotype closer to the optimal state, reducing the capacity for further improvement.

Second, when mutations have epistatic effects (i.e., interact in a nonadditive fashion), the possibility exists for neutral or even deleterious mutations to become beneficial in certain contexts (Figure 5.3). Multilocus traits exhibiting the latter types of genetic behavior will be referred to here as complex adaptations, as the paths for their evolution and the rapidity with which they are acquired are much less straight-forward than under conditions of additive fitness effects. Mutations that pave the way for an increase in fitness via secondary mutations that otherwise would be deleterious are sometimes referred to as enabling mutations (Tóth-Petróczy and Tawfik 2013).

One broad category of complex-trait evolution involves compensatory mutations, wherein single mutations at either of two loci cause a fitness reduction, while their joint appearance can restore or even elevate fitness beyond the ancestral state. Such epistatic interactions play a prominent role in Wright's (1931, 1932) shifting-balance theory of evolution, which postulates that adaptive valleys between fitness peaks are typically traversed in small subpopulations that facilitate drift through a deleterious intermediate state in an effectively neutral fashion. Compensatory mutations play important roles in protein-sequence evolution (Chapter 12), in the composition of nucleotides in the stems of RNA molecules (Stephan and Kirby 1993; Kondrashov et al. 2002; Kulathinal et al. 2004; Azevedo et al. 2006; DePristo et al. 2007; Breen et al. 2012; Wu et al. 2016), and in the coevolution of complexes consisting of components encoded in nuclear and organelle genomes (Chapter 23). Many other scenarios are possible for the transitions from one high fitness state to another. For example, situations likely exist in which mutations that are effectively

neutral in isolation yield an increase in fitness when combined, and some pathways with multiple steps may allow the bypass of shorter but more deleterious pathways (Figure 5.3).

As complex adaptations are expected to evolve over relatively long periods, with many background mutations accumulating in parallel, ascertaining the molecular paths of establishment from observations on evolutionary endpoints alone is challenging. However, clear examples observed in real time do exist. For example, in a long-term ($> 40,000$ generation) evolution experiment with *E. coli* selected for growth in flasks on a defined medium, the novel ability to utilize citrate as a carbon source emerged in one of twelve cultures (Blount et al. 2008; Quandt et al. 2014). Drawing from the historical record of evolution by resurrecting frozen samples, it was found that a weak variant for citrate utilization arising from a promoter-region mutation provided a potential mutational target for further refinement of the trait. While this initial mutation was still infrequent in the population (and possibly effectively neutral), a linked mutation appearing at a second locus conveyed a much greater ability to take up citrate, conferring a substantial increase in fitness that drove the double mutant to fixation.

Sequential fixation vs. stochastic tunneling. Even when only two loci are involved, ascertaining the population-genetic conditions under which complex adaptations are likely to occur brings in challenges not discussed above for the ideal single-site model. This is because unlike the situation in which a single mutation fixes at a rate depending only on its own initial frequency, the success of a mutation involved in an interlocus interaction depends on the frequency of alleles at the interacting locus, on the fitnesses associated with all possible multi-locus genotypes, and on the recombination rate between the two loci.

The focus here will be on the rate of establishment of a complex adaptation, defined to be the inverse of the expected arrival time of the ultimate multi-mutation configuration destined to be fixed in the population. Although this excludes the additional time required for fixation, the latter is generally considerably smaller than the time to establishment, and ignoring it does not influence the following conclusions.

Population size alone can dictate the kinds of evolutionary pathways that are open to the establishment of complex traits (Figure 5.4). For populations of sufficiently small size, the path toward adaptation almost always involves sequential fixations of the contributing mutations, owing to the extreme rarity of occasions in which multiple mutations are simultaneously segregating at key sites. This issue was explored in the previous section for the situation in which mutational effects are independent, but we now consider an adaptation requiring two genetic changes, the first of which is a neutral enabling mutation (Walsh 1995; Lynch and Abegg 2010).

We first consider the conditions under which a population is constrained to acquiring the two-site adaptation in two sequential steps. This is a function of three factors. Assuming a first-step (neutral) mutation is destined to fixation, its mean time to fixation is $4N_e$ generations in a diploid population (Kimura 1983). Because the frequency increases from near 0.0 to 1.0, the average frequency of this mutation during its sojourn to fixation is 0.5, and this implies the presence of an average $0.5 \cdot 2N = N$ copies during the polymorphic phase. Letting μ_2 be the rate

of second-step mutations linked to the first-step background, from the theory outlined above, the rate of appearance of second-step mutations destined to fix is then $\mu_2 \cdot 2s(N_e/N)$ per mutational target per generation. The product of these three quantities gives the approximate probability of a secondary mutation arising on a segregating first-step mutation background (and also being destined to fix) of $4N_e \cdot N \cdot \mu_2 \cdot 2s(N_e/N) = 8N_e^2\mu_2 s$. This shows that there is a negligible chance of arrival of a successful secondary mutation before fixation of the first if $N_e \ll (8\mu_2 s)^{-1/2}$. With $\mu_2 = 10^{-9}$ and $s = 10^{-4}$, for example, the critical effective population size is $\simeq 10^6$. For N_e below this threshold value, selection is restricted to exploring the fitness landscape by sequential mutational steps.

In contrast, in large populations, key secondary (and even tertiary) mutations can arise prior to the fixation of earlier-step contributors (Figure 5.3). This raises the possibility of the joint, simultaneous fixation of combinations of mutations as single haplotypes without any enabling mutation having been previously common in isolation as heterozygotes. For example, a conditionally beneficial secondary mutation may arise in linkage with a low-frequency deleterious first-step mutation, with the joint fixation of the double-mutation haplotype in effect rescuing the first-step mutation otherwise destined to be lost. Such a process, often referred to as stochastic tunneling (Komarova et al. 2003; Iwasa et al. 2004), provides a smooth route for the establishment of complex adaptations, allowing large populations to explore the fitness surface more broadly than possible by single-step mutations. Most notably, stochastic tunneling allows progression through intermediate deleterious alleles without the population ever experiencing the transient decline in fitness that would necessarily occur with sequential fixation (Gillespie 1984; Weinreich and Chao 2005; Gokhale et al. 2009; Weissman et al. 2009, 2010; Lynch and Abegg 2010; Lynch 2010). This shows how deleterious mutations with conditional effects can play a central role in evolutionary diversification.

The following analyses will focus on the domain in which stochastic tunneling dominates, i.e., populations of moderate to very large size, as will often be the case in single-celled organisms. However, before proceeding, it should be emphasized that the theory explored in the following paragraphs starts with the premise that each mutation contributing to a final adaptation arises independently of all others. Recall from Chapter 4 that mutations sometimes arise in clusters, which means that adaptations involving two or three, and perhaps even more, site-specific mutations will occasionally arise spontaneously in a single individual on realistic time scales. In such cases, assuming negligible recombination between these sites, the rate of fixation of the mutant haplotype follows directly from the single-mutation theory noted in the preceding chapter, Equations 4.1a-c, i.e., we simply inquire as to the probability of fixation of a newly arisen beneficial mutation with the specified haplotype, $\simeq 2s(N_e/N)$.

Two-locus transitions. We start with a simple selection scenario, first explored by Kimura (1985), in which haplotypes Ab and aB have reduced but equivalent fitness $(1 - s_d)$ relative to AB and ab, both of which have fitnesses of 1.0 (Figure 5.3, upper left). In this case, two-step transitions between pure population states of AB and ab render no gain in fitness, but do involve an intermediate deleterious genotype. Initially, the two sites will be assumed to be completely linked, and μ_d

and μ_b will denote the rate of mutation to first (potentially deleterious) and second (potentially conditionally beneficial) step variants. (If $s_d = s_b = 0$, the intermediate states are neutral).

As an explicit example, AB and ab might represent beneficial pairs of amino acids involved in protein folding or stability, with Ab and aB representing nonmatching combinations. The Watson-Crick pairs in the stems of ribosomal RNAs provide another compelling example. Although the overall stem/loop structure of rRNAs is highly conserved across species, orthologous complementary nucleotide pairs in stems often have different states in different species. Barring a rare double mutation, such shifts require passage through an intermediate deleterious state, e.g., A:T $\rightarrow$ A:C $\rightarrow$ G:C. Provided the overall secondary structure is maintained, which is presumably essential for proper ribosome assembly, rRNAs from different bacterial species with up to 20% divergence can substitute for each other with only small effects on fitness (Kitahara et al. 2012).

Throughout this section, we assume a large enough population size to be in the stochastic-tunneling domain. Starting with a population in state AB, we wish to determine the mean time for the population to reach an alternative state of fixation at both loci, i.e., with haplotype ab. Mutation will recurrently introduce new Ab and aB haplotypes, and provided the strength of selection exceeds that of mutation, both the aB and Ab haplotypes will then be expected to have steady-state frequencies of μ_d/s_d. This condition results from the balance between the rates of mutational input μ_d (from the ancestral AB haplotype) and selective removal s_d (Walsh and Lynch 2018). In a diploid population, there will be a total of $2N \cdot 2 \cdot \mu_d/s_d$ of these low-frequency aB and Ab haplotypes, as the $2N$ chromosomes each contain two loci with expected deleterious-allele frequency μ_d/s_d. These intermediate types then serve as reliable substrates for secondary mutations to the ab type, which arise at rate μ_b from each intermediate background. However, even though the ab type has an advantage over its deleterious parental Ab or aB haplotype, ab mutations fix in an essentially neutral fashion with probability $1/(2N)$. This follows under the assumption that $\mu_d/s_d \ll 1$, which ensures that almost all resident haplotypes are of type AB, which have equivalent fitness to ab. Thus, the rate of establishment of the ab type by stochastic tunneling from AB is

$$r_e \simeq \frac{4N\mu_d}{s_d} \cdot \mu_b \cdot \frac{1}{2N} = \frac{2\mu_d\mu_b}{s_d} \tag{5.2}$$

(Gillespie 1984; Stephan 1996). This rate is independent of population size (provided the conditions for selection-mutation balance, $4N_e s_d \gg 1$, are met, and ignoring for the moment any population-size dependence of the mutation rate).

Next, suppose that the secondary mutation has advantage $s_b > 0$, such that the fitness of the AB and ab haplotypes are 1 and $1 + s_b$, respectively (Figure 5.3, upper right). Modifying Equation 5.2 to account for the fact that the fixation probability of the double mutant is $\simeq 2s_b(N_e/N)$ leads to

$$r_e \simeq \frac{4N\mu_d}{s_d} \cdot \mu_b \cdot 2s_b(N_e/N) = \frac{8N_e\mu_d\mu_b s_b}{s_d}. \tag{5.3}$$

As in the case of selectively equivalent end states, the rate of establishment scales with the square of the mutation rate, but now also with the strength of positive selection scaled relative to that of drift, $4N_e s_b$.

Finally, consider the special situation in which first-step mutations are effectively neutral, so that again there should be a nonzero probability of their being present at some low frequency at the outset. Denoting this initial frequency as p_0, and simply substituting this for μ_d/s_d in Equation 5.3 yields

$$r_e = 4Np_0 \cdot \mu_b \cdot 2s_b(N_e/N) = 8N_e p_0 \mu_b s_b. \tag{5.4}$$

This is a potentially much higher rate of tunneling than implied by Equations 5.2 and 5.3, owing to the expectation that mutations at neutral sites will have much higher equilibrium frequencies than deleterious mutations. If, for example, the genetic substrate here is one particular nucleotide at a genomic site, assuming no mutational bias, the long-term average frequency of each of the four nucleotide types is 0.25 (even a small population harboring little polymorphism will be fixed for the enabling mutation with probability 0.25). This then yields $r_e = 2N_e\mu_b s_b$, which is just half the expectation for the single-site model, Equation 5.1a, because one of the two potential starting nucleotides (out of four) is present at the outset at each of the two sites. This kind of scenario would apply to the situation in which a codon requires two changes for a transition to a more beneficial amino-acid.

What can be inferred about the likely scaling of two-site adaptations from these results? As noted above, a key issue is that the algebraic scaling implied by the preceding expressions is confounded by the nonindependent behavior of the biological components. For example, as outlined in Chapter 4, there is a roughly inverse scaling between the mutation rate per nucleotide site per generation and N_e across the Tree of Life. Thus, treating the product of N_e and the mutation rate as an approximate constant may provide a more realistic view of how the per-generation rate of evolution scales with population size.

Consider, for example, Equation 5.2 for the rate of transition between two equivalent fitness states via deleterious intermediates. Although this expression suggests that the evolutionary rate scales with $\mu_d\mu_b$, independent of N_e, because both mutation rates scale inversely with N_e, the expectation is that r_e will actually scale inversely with N_e^2, i.e., $r_e \propto 2/(N_e^2 s_d)$. Extending this logic to Equation 5.3 for the rate of transition to a higher fitness state through deleterious intermediates implies an inverse scaling of r_e with N_e. On the other hand, Equation 5.4 for the rate of transition to a higher fitness state through neutral intermediates implies a scaling that is independent of N_e, as in the case of the single-site model, as $N_e\mu_b$ is approximately constant.

The key point here is that the rate of exploitation of various kinds of evolutionary pathways can depend critically on the nature of the adaptive change, on the effective population size, and on mutation rates. Equations 5.2, 5.3, and 5.4 ignore the possible relationship of generation lengths with N_e. However, if as suggested above, this is an inverse relationship, then combined with the near constancy of $N_e\mu$, the scaling of the rate of evolution with N_e on an absolute time scale is obtained by multiplying the scalings noted in the previous paragraph by N_e. The rate then becomes: 1) inversely related to N_e when deleterious intermediates lead to alternative end states with equivalent fitness (Equation 5.2); 2) independent of N_e when deleterious intermediates lead to one end state with elevated fitness (Equation 5.3); and 3) positively related with N_e when neutral intermediates enable the evolution of an end state with elevated fitness (Equation 5.4).

Finally, recall that simultaneous mutations arise at two sites much more rapidly than expected by chance. Denoting the mutation rate of AB to ab haplotypes in single events as μ_2, the rate of establishment of a beneficial two-locus variant would then be identical to that defined for the single-site expression, Equation 5.1. Provided $\mu_2 \gg \mu_b\mu_d$, the rate of this direct route would be much higher than that expected by tunneling.

More complex scenarios. While the above analyses assume an evolutionary path to a final adaptation through just a single intermediate step, the routes to many molecular/cellular modifications can be more complex, potentially involving pathways through any number of mutations, e.g., Figure 5.3 (bottom). The rates of establishment under these alternative scenarios have been examined in some detail, again with the primary focus on situations in which the intermediate states are neutral or deleterious (Gokhale et al. 2009; Weissman et al. 2009; Lynch and Abegg 2010; Santiago 2015). The mathematics necessarily becomes more complex, and simple analytical approximations have been found in only a few cases. Two of these are considered below to illustrate how the establishment rate r_e scales with the underlying features of population size, mutation rate, and selection intensity.

For the case of deleterious intermediates, suppose that all haplotypes involving 1 to $d-1$ mutations are equally deleterious (with fitness $1-s_d$), with the final mutation conferring an advantage s_b ($d = 2$ represents the two-step case noted above). First-step mutations then arise from mutation-free individuals at rate $2Nd\mu_d$, but owing to selection have an expected survivorship time of $1/s_d$ generations, during which period $d-2$ additional intermediate step mutations must be acquired, followed by the appearance of a final-step mutation destined to fixation. This leads to a rate of establishment via tunneling of

$$r_e \simeq 4N_e d!(\mu_d/s_d)^{d-1}\mu_b s_b \tag{5.5}$$

which reduces to Equation 5.3 when $d = 2$. Here, we see that the rate of establishment scales with the dth power of the mutation rate, owing to the limited opportunities for secondary mutations during the short sojourn times of deleterious mutations. Thus, the acquisition of a novel adaptation involving multiple, deleterious intermediate steps is a very low probability event, potentially diminishingly so for populations with large N_e, as assuming an inverse relationship between the mutation rate and N_e, the expected scaling is now with N_e^{1-d} (on a per-generation time scale). Again, however, multinucleotide mutations can, in principle, dramatically accelerate this process.

For the case of neutral intermediates with d mutations required for the final adaptation (and the order of events assumed to be irrelevant), there is again the conceptual issue of the starting conditions. The worst-case scenario is the one in which all contributing mutations are absent at time zero, with establishment then requiring a series of nested tunneling events. Consider first the special situation in which $d = 2$ (the two-step model introduced above) and enabling neutral mutations leading to the favorable genotype arise at rate μ_n per site (Figure 5.3, upper middle). With two sites in a diploid population, $4N\mu_n$ neutral first-step mutations arise per generation.To obtain the expected rate of tunneling in this case, we also require the

probability that tunneling occurs within a descendant lineage of a first-step mutation before the latter becomes lost from the population. Assuming complete linkage, this probability is approximately $\sqrt{2\mu_b s_b(N_e/N)}$ in large populations (Komarova et al. 2003; Iwasa et al. 2004; Weissman et al. 2009, 2010; Lynch and Abegg 2010). Thus, the rate of establishment via tunneling becomes

$$r_{e,2} \simeq 4N\mu_n\sqrt{2\mu_b s_b(N_e/N)} = 4\mu_n\sqrt{2\mu_b s_b N_e N} \tag{5.6a}$$

The key observation here is that when the intermediate steps are neutral, but the first-step mutations are initially absent from the population, the probability of tunneling scales positively with the square root of the product of absolute and effective population sizes.

Now consider the case of $d = 3$ (with two neutral enabling mutations required before the final adaptation is assembled with a third mutation). In this situation, a secondary (conditionally neutral) mutation must arise on a haplotype lineage containing the first such mutation, and before being lost by drift, the still smaller two-mutation lineage must acquire a third mutation destined to fixation. This nested set of events expands Equation 5.6a to

$$r_{e,3} = 3N\left(2\mu_n\sqrt{r_{e2}/(2N)}\right) = 6N\mu_n\sqrt{2\mu_n\sqrt{2\mu_b s_b(N_e/N)}} \tag{5.6b}$$

Note that the first term is now $6N\mu_n$ because first-step mutations can arise at three diploid sites. The next step then initiates at either of the two remaining sites, bringing in the additional $2\mu_n$ term, with the final stage being initiated at the one remaining site and fixing at the usual rate for a single beneficial mutation. With $d = 4$, this expression expands one step further to

$$r_{e,4} = 4N\left(2\mu_n\sqrt{r_{e3}/(2N)}\right) = 8N\mu_n\sqrt{3\mu_n\sqrt{2\mu_n\sqrt{2\mu_b s_b(N_e/N)}}}, \tag{5.6c}$$

and so on.

These results show that with neutral intermediates, the rate of establishment of complex adaptations can be much more rapid than expected under the naive assumption that independently arising mutations would lead to a scaling with the dth power of the mutation rate. Regardless of the number of sites involved with neutral intermediates, the rate of establishment by tunneling scales with no more than the square of the mutation rate. Again, adhering to the empirical observation of an approximately inverse relationship between mutation rates and population size, this implies that the expected rate of establishment of beneficial features dependent on a constellation of new enabling mutations is nearly inversely proportional to N_e on a per-generation time scale; and if the generation length also scales inversely with N_e, the rate of establishment is N_e-independent. If it is further assumed that $(d-1)$-stage mutants are present at some low frequency p_0 in the base population, Equation 5.4 applies, again implying N_e-independent scaling.

Finally, it should be noted that the above examples are just a small sample of the kinds of evolutionary pathways that can exist between two complex genotypes. In principle, multiple pathway types may connect two presumed endpoints, including those with mixtures of neutral and deleterious intermediates, different numbers

of links (Figure 5.3, lower right), and so on. The kind of theory just outlined can be used to evaluate the relative probabilities of such alternative routes, as well as the possibility of becoming transiently trapped at points with suboptimal fitness, and subsequently back-tracking and exploring alternative paths (McCandlish 2018). Experimental-evolution studies with microbes are being increasingly exploited to evaluate these issues (e.g., Lind et al. 2019; Rodrigues and Shakhnovich 2019; Zheng et al. 2019). Although much work remains to be done, the key point here is that complex adaptations can arise in large asexual populations (e.g., unicellular organisms) much more rapidly than one might imagine under the assumption of sequential fixation of interactive mutations.

Effects of recombination. All of the above analyses assume an absence of recombination. In the sequential-fixation regime, recombination can be ignored simply because multiple polymorphic sites are never present simultaneously. However, in the stochastic-tunneling domain, opportunities may exist for both the recombinational creation and breakdown of optimal haplotypes. Examination of this problem with a broad class of models leads to the conclusion that recombination is likely to have either a minor advantageous or a strong inhibitory effect on the *de novo* establishment of a complex adaptation (Higgs 1998; Lynch 2010; Weissman et al. 2010; Santiago 2015).

Consider, for example, the case of a two-site adaptation involving a deleterious intermediate, starting with a population fixed for the suboptimal ab haplotype. The overall influence of recombination on the rate of establishment of the AB haplotype will then be a function of two opposing effects. On the one hand, the rate of origin of AB haplotypes by recombination between the two single-mutation haplotypes (aB and Ab) will be proportional to the rate of recombination between the sites (c). On the other hand, recombinational breakdown discounts the net selective advantage of resultant AB haplotypes from s_b to $s_b - c$. This occurs because in the early stages of AB establishment, ab haplotypes still predominate, and hence are the primary partners in recombination events with AB, generating the maladaptive Ab and aB products. Thus, because the product $c(s_b - c)$ is maximized at $c = s_b/2$, two-site adaptations are expected to emerge most rapidly in chromosomal settings where the recombination rate is half the selective advantage of the final adaptation. This is a rather specific requirement, as the optimal recombination rate depends on the advantage of the final adaptation. Moreover, even in the case of neutral intermediates and at the optimal recombination rate, the rate of establishment is generally enhanced by much less than an order of magnitude relative to the situation with complete linkage (Lynch 2010).

The situation is particularly bleak when first-step mutations are deleterious. In this case, if the rate of recombination exceeds the selective advantage of the AB haplotype (i.e., $s_b - c < 0$), recombination presents an extremely strong barrier to establishment of the AB haplotype, as most recombinant (Ab and aB) intermediates are removed by selection more rapidly than the AB type can be promoted. Thus, because the role that recombination plays in the origin of specific adaptations depends on both the selective advantage of the final product, the selective disadvantages of the intermediate states, and the physical distance between the genomic sites of the underlying mutations, recombination is far from universally advantageous.

The Phylogenetic Dispersion of Mean Phenotypes

The theory discussed above provides insight into the rapidity with which populations can respond to novel and/or persistent directional selective challenges. Such scenarios might be encountered in a continuous coevolutionary arms race between a host cell and a pathogen, or in situations involving a sudden environmental shift. However, numerous cellular traits may have been under very similar selective pressures across phylogenetic lineages since their origin. This is likely to be true, for example, for enzymes whose sole function has always been to convert a specific substrate into a specific product, membrane channels specialized to admitting and/or excluding specific ions, or polymerases responsible for generating complementary base-pair matches.

In such cases, the dynamical response to changing selection pressures is no longer the key issue. The more relevant evolutionary focus is the long-term steady-state probability distribution of alternative genotypes. Although natural selection relentlessly strives to improve trait performance, there are numerous reasons why perfection will seldom, if ever, be achievable. First, absolute limits to the refinements of chemical and physical processes are dictated by factors such as diffusion rates and effectively discrete processes such as the energy associated with individual hydrogen bonds. Second, the stochastic processes of mutation and drift can result in the dispersion of mean phenotypes around an expected value, to a degree that depends on the range of effectively neutral parameter space. Third, mutation pressure will almost never be perfectly aligned with the goals of selection, and this will cause the average phenotype to deviate from an optimum, even in the absence of mutation bias.

The latter two points have significant implications for interpreting patterns of phenotypic divergence. Most notably, populations under identical selection pressures will not necessarily have identical mean phenotypes, but instead will exhibit a dispersion of such measures. Moreover, the most common phenotype need not be the optimum phenotype. Finally, gradients of mean phenotypes with respect to N_e are likely to be molded by differences in the power of random genetic drift across the Tree of Life. To reduce the likelihood of evolutionary cell biology succumbing to the common practice of interpreting all phylogenetic variation in phenotypes as necessary reflections of differences in selective environments, these basic principles will first be sketched out for the case of a very simple two-state trait, and then further explored for traits encoded by multiple genetic loci.

Two-state traits. Consider a single biallelic locus with one allele (denoted as state 1) having a fractional selective advantage s over the other allele (denoted as state 0). Although allele 1 has the highest achievable fitness, this does not ensure that once fixed, allele 1 will be immune to replacement by the suboptimal type. Denoting the mutation rate from allele 0 to 1 as μ_{01}, with μ_{10} being the reciprocal mutation rate, and assuming a haploid population, we wish to determine the long-term mean frequency of allele 1.

The simplest situation involves a population with a small enough effective size that the waiting times between mutations destined to fixation are large enough that

the population is nearly always fixed for one allele or the other, with probabilities $\widetilde{p}_0$ and $\widetilde{p}_1$. The domain of N_e and mutation rates necessary for such a situation is equivalent to that noted above for the sequential-fixation model, with the more general model (allowing for any N_e) being outlined in Foundations 5.3.

Under these weak-mutation, strong-selection conditions, a lineage spends a long period of time in one particular monomorphic state before making a stochastic shift to another. The intervening intervals (waiting times for transitions) are functions of the relative strengths of selection, mutation, and random genetic drift, but over a very long time period, the rate of movement from state 0 to 1 must equal that in the opposite direction (a principle known as detailed balance in the statistical physics literature). This is because, whether favorable or encouraged by mutation pressure, the most abundant state provides more opportunities for transitions, which are individually less likely to proceed to fixation; in contrast, the rarer state provides fewer transition opportunities, but when these arise they are more likely fix. It then follows that

$$\widetilde{p}_0 \cdot (N\mu_{01}) \cdot \phi_{01} = \widetilde{p}_1 \cdot (N\mu_{10}) \cdot \phi_{10}, \tag{5.7}$$

where ϕ_{01} and ϕ_{10} denote the probabilities of fixation of newly arising beneficial and deleterious alleles, defined by Equation 4.1b. Each side of this equation is the product of the expected frequency of a state, the rate of origin of mutations to the opposite state, and the probability of fixation of the mutant allele. Using the useful identity $\phi_{01}/\phi_{10} = e^S$, where $S = 2N_e s$, and the fact that $\widetilde{p}_1 = 1 - \widetilde{p}_0$, Equation 5.7 rearranges to

$$\widetilde{p}_0 = \frac{1}{1 + \beta e^S}, \tag{5.8}$$

where $\beta = \mu_{01}/\mu_{10}$ is the ratio of the mutation rates in both directions (mutation bias being implied when $\beta \neq 1$).

This simple model illustrates three key points. First, unless completely lethal, the low-fitness state has a non-zero probability of occurrence. Thus, despite constant selection pressure, a lineage is not expected to remain in a stable fixed state forever. Second, the two alleles approach equal probabilities as $\beta e^S \to 1$. This composite parameter is simply equal to the product of the mutation and selection biases in favor of state 1, so that even if state 1 is selectively favored ($S > 0$), state 0 will be more common than state 1 if there is mutation bias in the opposite direction of sufficient strength that $\beta e^S < 1$. The relevance of this point is that maximum divergence occurs when $\beta e^S = 1$, again demonstrating the potential for substantial variation in the face of uniform selection. Third, if the effective population size and/or strength of selection is sufficiently small that $S \ll 1$ (the domain of effective neutrality), the equilibrium frequency of the disfavored allele will be entirely driven by mutation pressure. In this case, because $e^S \simeq 1$, $\widetilde{p}_0 \simeq 1/(1+\beta)$, which is a function of the relative (but not absolute) mutation rates.

These expectations are altered when the population-level mutation rate exceeds the limits of the domain of the sequential model (Foundations 5.3). Most notably, Equation 5.8 defines the lower bound to the expected frequency of the low-fitness allele. The expected frequency of the beneficial allele declines once the population-level mutation rate ($N_e\mu_{01}$) exceeds ~ 0.01 (i.e., a new mutation enters the population at least every 100 generations), asymptotically approaching the neutral expectation $\widetilde{p}_1 \simeq \beta/(1+\beta)$ as $N_e\mu_{01}$ exceeds 1.0 (Figure 5.5). The latter condition arises when

mutation brings in allelic variants faster than natural selection can promote beneficial over detrimental alleles. In this case, the population is also almost always represented by a polymorphic collection of both alleles, rather than by a state of fixation.

Multistate-traits and the drift-barrier hypothesis. Extension of the preceding single-locus model to an arbitrary number of L sites (factors) yields additional insights into the limits to what natural selection can accomplish. To appreciate the fundamental points in a relatively simple manner, it will be assumed that all genetic factors are equivalent, with two alternative (+ and −) allelic states contributing positively and negatively to the trait. Depending on the context, these factors may be viewed as single nucleotides, amino-acid codons, or entire genes.

For all but the two most extreme genotypes (all + or all − alleles), a multiplicity of functionally equivalent classes exists with respect to the number of + alleles, i. These are defined by the binomial coefficients. As an example, for the case of $L = 4$, there are $2^4 = 16$ possible genotypes, but just five genotypic classes (having 0, 1, 2, 3, 4, and 5 + alleles, with respective multiplicities 1, 4, 6, 4, and 1) (Figure 5.6). With equivalent fitness for all members (haplotypes) within a particular class, this variation in multiplicity of states plays a significant role in determining the long-term evolutionary distribution of alternative classes, as classes with higher multiplicities are more mutationally accessible. This type of biallelic model has been widely exploited in theoretical studies of the genetic structure of quantitative (multilocus) traits (Walsh and Lynch 2018), and will be encountered in a number of different contexts in subsequent chapters, including the evolution of protein-protein interfaces, transcription-factor binding sites, and growth rate. Here, we assume a haploid, nonrecombining population of N individuals. The site-specific per-generation mutation rates from the − to the + state, and vice versa, are again defined as μ_{01} and μ_{10}, respectively.

As with the single-factor model, the multiple-factor model has a long-term steady-state probability distribution of population residence in the $L + 1$ alternative states. Again, starting with the assumption that population-level mutation rates are low enough that transitions only occur between adjacent classes (satisfying the conditions for a sequential-fixation scenario), the relative flux rates between classes are equal to the expressions on the arrows in Figure 5.6. These rates are proportional to the products of rates of mutational production and probabilities of subsequent fixation, with the numerical coefficients being defined by the numbers of − and + sites within each class. The absolute population size N defines the number of mutational targets per generation, but because N influences all mutational flux rates in the same way, it is omitted as a prefactor, although both N and N_e still influence the equilibrium solution via the fixation probabilities (Equation 4.1b).

This linear sequential model has a relatively simple solution (Berg et al. 2004; Sella and Hirsh 2005; Lynch and Hagner 2014; Lynch 2020). As a reference point, consider first the extreme case of effective neutrality. In this situation, $\eta = \mu_{01}/(\mu_{01} + \mu_{10}) = \beta/(1+\beta)$ is the expected equilibrium frequency of + alleles at each site, which arises when the net flux of $+ \rightarrow -$ and $- \rightarrow +$ mutations is balanced. With no selection for particular combinations of alleles, each site evolves in an independent fashion, so the steady-state distribution of phenotypes under neutrality is simply

equal to the binomial probability distribution,

$$\begin{aligned}\widetilde{p}_{n,i} &= \binom{L}{i}\eta^i(1-\eta)^{L-i} \\ &= C\cdot\binom{L}{i}\beta^i \end{aligned} \tag{5.9}$$

where n denotes the neutral condition, and i denotes the number of + alleles in a genotypic class. The term in large parentheses is the binomial coefficient $L!/[i!(L-i)!]$ (which is the multiplicity of alternative orderings of + and − positions within a particular class i), and $C = (1+\beta)^{-L}$. Equation 5.9 defines the long-term probability of a population residing in each of the $L+1$ possible genotypic classes, i.e., the fractional time wandering over the evolutionary landscape that is spent in each class. Note that the neutral steady-state distribution depends only on the ratio of mutation rates (β), not on their absolute values. Denoting the overall genotypic state as the sum of + alleles, the long-term mean and variance of the trait are $L\eta$ and $L\eta(1-\eta)$, respectively.

Under the sequential model, selection transforms the neutral distribution in a remarkably simple way, with each class being weighted by the exponential function of the scaled strength of selection $S_i = 2N_e s_i$,

$$\widetilde{p}_i = C\cdot\widetilde{p}_{n,i}\cdot e^{S_i}, \tag{5.10}$$

where C is a new normalization constant that ensures that the $\widetilde{p}_i$ sum to 1.0. The selection coefficients associated with each class are generally defined as deviations from some benchmark in the population (say the optimum type), but this does not matter, as the reference is a constant that simply enters the normalization constant. The utility of this approach is that, provided there are mutational connections between all adjacent states, Equation 5.10 can be applied to any fitness function describing the relationship between s and i.

Taken together, Equations 5.9 and 5.10 show that the equilibrium frequencies of the genotypic classes are functions of three factors: 1) the multiplicity of configurations, as defined by the binomial coefficients; 2) the ratio of mutation rates; and 3) the strength of selection scaled by the power of random genetic drift. All other things being equal, the within-class multiplicity magnifies the likelihood of residing in such a state. This demonstrates that mutation need not be directionally biased to have an impact on the overall distribution. All that is required is that neutral distribution deviates from the expectations under selection alone, which will almost always be the case.

Two examples will now be explored to illustrate how these three factors jointly define the distribution of phenotypes within and among alternative population-genetic environments. First, consider the simple case in which a trait determined by just $L = 2$ factors is under stabilizing selection, such that there is an optimum phenotype θ, with fitness (W) dropping off at a rate determined by the width of the fitness function ω (analogous to the standard deviation of a normal distribution). This Gaussian (bell-shaped) function is defined by

$$W_i = e^{-(i-\theta)^2/(2\omega^2)}. \tag{5.11}$$

With two sites, there are three possible genotypic classes, $i = 0, 1$, and 2, with the phenotypically equivalent $+-$ and $-+$ states being lumped into the $i = 1$ class. Selection is purely directional if the optimum is at or beyond an end state, i.e., $\theta \leq 0$ or $\geq L$, and neutrality is approached as the fitness function becomes flatter, i.e., $\omega \to \infty$. Although i is confined to integer values, θ need not be, and if θ has a value other than 0, 1, or 2, the optimum phenotype is unattainable. The selection coefficients can be simply defined as deviations of fitness from the maximum value of 1, $s_i = 1 - W_i$. The mean phenotype (in this case, the average genotypic value of i) is $\widetilde{p}_1 + 2\widetilde{p}_2$, which reduces to 2η in the case of neutrality.

This expansion to a second site introduces complexities not encountered with the one-site model (Figure 5.7a). For example, for the case of $\theta = 1.5$, where the optimum is straddled by the class 1 and 2 genotypes, assuming mutation bias towards $-$ alleles ($\beta < 1$), the long-term genotypic mean never reaches the optimum, even at very large N_e, and instead remains much closer to 1. This bias results because although the class 1 and 2 genotypes have equivalent fitness, mutation pressure towards $-$ alleles weights the frequency of class 1 by a factor of 2β (the two being the multiplicity of this class), but class 2 by the smaller factor of β^2 (from Equation 5.10)

For the case in which $\theta = 2$ (pure directional selection), there is a progressive succession of the prevailing genotype classes with increasing N_e (Figure 5.7b). When N_e is sufficiently low to impose effective neutrality, class 0 predominates owing to the mutation bias towards $-$ alleles. With increasing N_e, selection becomes more effective at promoting class 1, but there remains effective mutation pressure against class 2. Finally, with very large N_e, selection becomes efficient enough to drive class 2 to near fixation, thereby decreasing the incidence of class 1. These results show that in the face of a constant pattern and strength of selection, the genotypic mean can exhibit a considerable gradient with N_e owing to changes in the power of drift. They also show that appreciable incidences of all three genotypic classes can be expected over time in lineages with intermediate N_e, i.e., the population mean phenotype is expected to wander temporally between alternative states despite the imposition of constant selection.

As a second example, consider the case in which fitness declines with the number of $-$ alleles ($L - i$) in a multiplicative fashion,

$$W_i = (1 - s)^{(L-i)}, \tag{5.12}$$

where i is the number of beneficial ($+$) alleles in the genome. Three examples are shown in Figure 5.8 for different numbers of completely linked loci, $L = 10^4$, 10^5, and 10^6. In each case, the overall performance is equally subdivided across all L factors, such that $s = 1/L$. With such large numbers of loci, the analytical solution noted above is not reliable at large N_e, as the number of mutations arising per generation may exceed the limit of the domain of the sequential model, although results have been obtained by computer simulations (Lynch 2020).

Three general features are clear (Figure 5.8). First, as already noted, when N_e is sufficiently small that the mutational effects are rendered effectively neutral, the mean fraction of $+$ alleles is defined by the neutral expectation, with the probability of a $+$ allele being η at each locus. Second, with increasing N_e, the mean fraction of $+$ alleles progressively increases, converging on 1.0 as N_e becomes much greater than $1/(2s)$. This gradient in trait means with N_e is a result of the drift barrier, which

increasingly compromises the ability of natural selection to alter the frequencies of mutations as the population size declines. The exact location of the drift barrier is defined by the relative power of drift and selection, becoming lower with smaller s, but also by the mutation bias and by the multiplicity effect. Third, the observed gradients are much shallower than the expectations under free recombination. This illustrates the point made in Chapter 4 that owing to selective interference among linked loci, populations with large absolute sizes behave genetically as though they are much smaller. For example, in the absence of selective interference, sites with selection coefficients equal to 10^{-5} are expected to be nearly fixed for + alleles once the absolute population size exceeds 500,000, but with linked loci at the same absolute population size, the vast majority of alleles are of the − type owing to the combination of mutation pressure and random genetic drift (Figure 5.8).

These results highlight the riskiness of an evolutionary biology that assumes that all phenotypes simply reflect optimal outcomes dictated by natural selection. In addition to the pervasive influence of drift, mutation can cause mean phenotypes to deviate from the optimum in substantial and often unexpected ways that are not simply functions of the magnitude of mutation bias. Rather, when alternative, functionally equivalent underlying genotypes exist for a trait, the multiplicity of certain intermediate combinations can result in a mutational pull of the mean phenotype away from the optimum. This effect becomes particularly significant when the phenotypic optimum is far from the expected mean under neutrality alone, especially if the level of multiplicity for the optimum is relatively small relative to other phenotypic states. Moreover, cases may even exist in which the opposing pressures of selection and mutation are sufficiently strong that the equilibrium mean-phenotype distribution can have two peaks, one driven by selection and the other by mutation (Lynch and Hagner 2014; Tuğrul et al. 2015; Lynch 2018).

Summary

- Critical to understanding the evolutionary potential and limitations of phylogenetic lineages is information on the distribution of fitness and phenotypic effects of new mutations. Based on bioenergetic considerations alone, few mutations can have absolutely zero fitness effects. Multiple lines of evidence indicate that the vast majority of mutations are deleterious, with most being very mildly so and the mode being near zero.

- The frequency of mutations with very small effects almost certainly increases in organisms of larger size, and there is strong evidence for a substantial pool of mutations that are only sensitive to selection in species with large effective population sizes (N_e). These mutations play a key role in defining the limits to adaptation in different phylogenetic lineages.

- Despite the common view that populations under identical selection pressures

will tend to be highly similarly phenotypically, many plausible situations exist in which uniform selection combined with random genetic drift and/or mutation bias can lead to substantial interspecies divergence, sometimes more than expected under drift alone.

- Many molecular adaptations require the co-occurrence of two or more mutations to elicit a phenotype with elevated fitness. Theory suggests that the rapidity (on an absolute time scale) of acquiring such adaptations is roughly independent of N_e if the intermediate states are neutral, but scales negatively with N_e if the intermediate states are deleterious, and more rapidly so with increasing numbers of intermediate steps. Thus, the likelihood of alternative paths of adaptive evolution can be strongly modulated by changes in N_e. However, sufficiently high rates of multinucleotide mutation can instantaneously embark a complex genotype on a more rapid path to establishment, the more so in larger populations.

- Many cellular traits have retained the same function for hundreds of millions of years, and may have been under nearly invariant selection pressures over this same time scale. This shifts the evolutionary focus away from dynamical changes in allele frequency under directional selection to the long-term steady-state probability distribution of alternative phenotypic states. The drift-barrier hypothesis predicts that, despite the operation of persistent directional or stabilizing selection, the mean phenotypes of such traits will commonly exhibit gradients with respect to N_e, with the level of functional refinement increasing with the latter.

- Aside from being the source of variation upon which natural selection operates, mutation impacts the expected distribution of mean phenotypes because genotypic states differ in the multiplicity of ways in which they can be constructed from the underlying set of genetic loci. Mutation bias further influences the evolutionary attraction towards a particular region of phenotypic space, in ways that may conflict with or reinforce prevailing selection pressures.

- Taken together, these results from evolutionary theory call into question the common practice of assuming that observed mean phenotypes provide a perfect reflection of prevailing selection pressures. Even under constant environmental conditions, the most common phenotype can be an unreliable indicator of the optimum defined by natural selection.

Foundations 5.1. The distribution of fitness effects of newly arising mutations. The success of natural selection depends critically on the presence of mutations with large enough effects to overcome the stochastic effects of random genetic drift. Yet, the distribution of fitness effects (DFE) of *de novo* mutations is one of the most poorly known aspects of genetics relevant to evolution. The central problem is that biology is structured in such a way that a large fraction of mutations have effects that are too small to be perceived as allele-frequency changes in laboratory experiments on reasonable time scales. Relying on results from subsequent chapters, here the case is made that the genome-wide DFE is highly skewed towards deleterious mutations with tiny effects, although many details remain to be worked out.

First, because the costs of biosynthesis of the four nucleotides differ, even among unexpressed mutations with no functional effects, only a very small fraction is likely to be absolutely neutral. From Foundations 17.2, the total cost of synthesizing the A and T in an A:T (or T:A) bond exceeds that for the C and G in a C:G (or G:C) bond by ~ 2 ATP equivalents. From Equation 8.2b, it is known that the cost of building a cell (in 10^9 ATPs) is a function of its volume (V, in μm^3), $\simeq 27V^{0.96}$. Thus, as a first-order approximation a transversion from A/T to C/G (or vice versa) alters the cell's energy budget (and the cell-division time) by a fraction 10^{-10} in a bacterial-sized cell of 1 μm^3 and by 10^{-12} in a medium-sized eukaryotic cell of 100 μm^3. The only truly neutral mutations may be those that alter A:T to T:A or C:G to G:C in nucleotide sites that are fully nonfunctional (i.e., are not transcribed and do not serve regulatory purposes).

Second, there are no absolutely neutral insertions and deletions, Because a complementary nucleotide pair in DNA has a total cost of ~ 100 ATP equivalents, a single base-pair insertion will impose a fractional reduction in fitness $\simeq 100/[(27 \times 10^9)V] \simeq 10^{-9}/V$ (again, rounding off to an order of magnitude), with an insertion x base pairs in length having a x-fold higher cost. These are the minimal costs of insertions, as would be expected in nonfunctional DNA. The total magnitude of effect would be substantially larger in functional regions owing to frameshifts in coding DNA and/or alterations in regulatory regions.

Third, owing to the additional energetic costs of gene expression (Lynch and Marinov 2015), nucleotide sites in protein-coding genes have average costs that are $\sim 100\times$ those for unexpressed nucleotides, and owing to the different costs of synthesizing alternative amino acids, an amino-acid altering mutation can further alter the overall cost per site by a factor of 0.1 to 6.0 (Foundations 17.2). This means that based on energetic costs alone, depending on the level of gene expression, amino-acid altering mutations will have average fitness effects on the order of $10\times$ to $600\times$ those for the most innocuous silent-site substitutions noted above, with the dispersion around these expectations following roughly a log-normal distribution with a range of at least an order of magnitude in both directions. For example, in a bacterial-sized cell, the fitness effects of amino-acid substitutions based on biosynthetic costs alone will be distributed over a range of 10^{-10} to 10^{-6}. In addition to these energetic consequences, amino-acid substitutions can have effects associated with protein function ranging from near zero to essentially lethal ($s = 1$), with most having $s \ll 0.1$ and the average being of order $s = 0.01$ (Chapter 12). Summing both types of effects, we expect the distribution of fitness effects of amino-acid altering mutations to have a very wide range, likely with a mode of at least 10^{-8} in bacteria, and perhaps several orders of magnitude higher.

These biological considerations allow a crude genome-wide view of the DFE for *de novo* mutations. In most bacteria, $\sim 95\%$ of the genome is coding DNA, and perhaps half of the intergenic sites and silent sites will have functional consequences associated with gene regulation. However, with increasing organism size, genomes become progressively more bloated with nonfunctional intergenic and intronic DNA as well as with more complex regulatory regions (Lynch 2007). As a consequence, in

many vertebrates and land plants, $< 2\%$ of the genome is coding DNA and $> 50\%$ of the remaining nucleotide sites are likely nonfunctional and so experience fitness-altering mutations only via small energetic effects. Owing to the nature of the genetic code, $\sim 25\%$ of coding-region sites are silent and the remaining 75% are amino-acid replacement sites (Lynch 2007), and assuming no mutation bias, one in three nucleotide substitutions at silent sites will have no energetic consequences. In addition, across the Tree of Life, $\sim 10\%$ of mutations involve insertions and deletions of one to several nucleotides, with the remainder being base substitutions (Chapter 4). Taking all of these observations into consideration leads to DFEs with the provisional forms outlined in Figure 5.1, as further discussed in the main text.

Foundations 5.2. Divergence under uniform selection. Although it is generally thought that selection will increase the evolutionary determinism of a system, causing pairs of populations under identical selection pressures to be more similar than expected on the basis of random assortment of variation, this is not necessarily the case (Cohan 1984; Lynch 1986). Consider a pair of populations exposed to identical conditions and starting with two alleles, A and a, with identical frequencies of p and $(1-p)$, respectively. Letting $\phi(p)$ be the probability of fixation of allele A, the probability that a pair of populations will ultimately experience fixation for different alleles is $\Delta = 2\phi(p)[1-\phi(p)]$, which reaches a maximum value when $\phi(p) = 0.5$. Under the naive view that the beneficial allele always fixes, one expects $\phi(p) = 1$ and $\Delta = 0$, but this is incorrect because the probability of fixation of beneficial alleles is < 1.

That populations can sometimes diverge to a greater extent under uniform selection than under pure neutral drift can be seen as follows. In the absence of selection, the probability of fixation of allele A is simply p, and the probability of alternative outcomes is $\Delta = 2p(1-p)$, which is maximized when $p = 0.5$. For the same reason, the probability of divergence is increased by selection if $\phi(p)$ under selection is closer to 0.5 than the initial frequency p. Because $\phi(p) > p$ for a selectively favored allele, it follows that a minimum requirement for increased divergence under pan-selection is that the starting frequency of the advantageous allele be < 0.5.

The conditions for excess divergence under drift plus selection to exceed that under drift alone are not very restrictive. Consider two replicate populations with identical initial frequencies of the A allele, $p = 0.25$. Under pure drift, the probability that one replicate becomes fixed for A and the other for a is $2 \cdot 0.25 \cdot (1-0.25) = 0.375$. Now suppose that A is weakly favored by selection, with $N_e s = 0.5$. Again assuming $p_0 = 0.25$, Equation 4.1a gives the fixation probability of A as 0.46, implying a probability of fixing alternative alleles of $2 \cdot 0.46 \cdot 0.54 = 0.496$, close to the maximum level of divergence expected under neutrality. Thus, even when experiencing identical directional selection pressures, populations that initiate with low-frequency, advantageous alleles can exhibit levels of divergence conventionally interpreted as being associated with diversifying selection. Of course, this analysis ignores the recurrent downstream introduction of mutations, which could ultimately lead to convergence.

Foundations 5.3. Mean probabilities of alternative alleles at steady state. A concern with the sequential model outlined in the text is that large populations are expected to reside in polymorphic states for significant amounts of time. This would not be a problem if the average frequencies of alleles when monomorphic were the same as those while polymorphic, but such a condition is only met in the special case of neutrality. This disparity arises because deleterious mutations that are strongly in-

hibited from going to fixation can nonetheless maintain measurable within-population frequencies owing to recurrent mutational input. As population sizes increase, the likelihood of residing in a polymorphic state necessarily increases, owing to the greater total influx of variation per generation. Under such conditions, one can still inquire as to the average frequency of a sampled allele over a long-term steady-state equilibrium, but this average must also factor in all possible polymorphic states, ranging from allele frequency $(1/N)$ to $[1-(1/N)]$ for haploids.

Let P_1, P_0, and P_p denote the steady-state probabilities of a population being monomorphic for the optimal allele (1), monomorphic for the suboptimal allele (0), or polymorphic (p). Under the sequential model, $P_1 + P_0 \simeq 1$. Here, we make use of a result from diffusion theory that describes the steady-state probability distribution of allele frequency x for the deleterious state 0 (which is equivalent to the beneficial allele 1 being present at frequency $1-x$), described more fully by Kimura et al. (1963), Wright (1969), and Charlesworth and Jain (2014). Although actual allele-frequency distributions are discrete, with large N, the probability that a population has allele frequency x can be accurately approximated by the continuous distribution

$$\phi(x) = Cx^{U_{10}-1}(1-x)^{U_{01}-1}e^{-Sx}, \tag{5.3.1a}$$

where $U_{01} = 2N_e\mu_{01}$, $U_{10} = 2N_e\mu_{10}$, and the normalization constant

$$C = \frac{\Gamma(U_{01}+U_{10})}{\Gamma(U_{01})\cdot\Gamma(U_{10})\cdot {}_1F_1(U_{10};(U_{01}+U_{10});-S)}, \tag{5.3.1b}$$

ensures that integration of the distribution over the full range of allele frequencies sums to 1.0. Γ denotes the gamma function, and ${}_1F_1$ is the confluent hypergeometric function. These two functions can be calculated numerically using series expansions defined respectively as Equations 6.1.2 and 13.1.2 in Abramowitz and Stegun (1964).

The probability of being monomorphic for state 1 can be approximated by integration of the end class, implying an absence of the deleterious allele,

$$P_1 = \int_0^{1/N} \phi(x)\cdot dx.$$

Because x is very small in this region, both $(1-x)$ and e^{-Sx} can be approximated as 1, leading to

$$P_1 \simeq \left(\frac{C}{U_{10}}\right)\left(\frac{1}{N}\right)^{U_{10}}. \tag{5.3.2a}$$

At the opposite end of the spectrum, using $x \simeq 1$ and $e^{-Sx} \simeq e^{-S}$ yields the probability of being monomorphic for state 0,

$$P_0 = \int_{1-(1/N)}^{1} \phi(x)\cdot dx \simeq \left(\frac{C}{U_{01}}\right)\left(\frac{1}{N}\right)^{U_{01}} e^{-S}. \tag{5.3.2b}$$

Here, it can be seen that the ratio P_1/P_0 obtained with this approach deviates from the prediction of the sequential model, $\widetilde{p}_1/\widetilde{p}_0 = (\mu_{01}/\mu_{10})e^S$ (inferred from Equation 5.8), unless $\mu_{01} = \mu_{10}$. Although the details will not be covered here, it can be shown that the probability of polymorphism, $P_p = 1 - P_0 - P_1$, is only weakly dependent on the magnitude of selection, and generally does not exceed 0.1 until $N_e\mu_{01} > 0.01$.

The average frequencies of the two alleles over the stationary distribution can be obtained by weighting the frequency classes by their densities, Equation 5.3.1a, and integrating over (0,1), which yields

$$\overline{p}_0 = \frac{\mu_{10}\cdot {}_1F_1[(U_{10}+1);(U_{01}+U_{10}+1);-S]}{(\mu_{01}+\mu_{10})\cdot {}_1F_1[U_{10};(U_{01}+U_{10});-S]}. \tag{5.3.3}$$

Foundations 5.4. The detailed-balance solution for the evolutionary distribution of alternative molecular states. Here we assume a linear array of alternative molecular states, with population-level transitions only occurring between adjacent states (Figure 5.6). For the latter condition to be met, each transition rate must be sufficiently small that a population generally resides in one state for an extended period of time before fixation of a subsequent mutation leads to a switch between states. Under these conditions, a relatively simple model defines the probability of residing in each class after a sufficient amount of time has elapsed to ensure potential occupancy over the entire distribution of states. At this equilibrium, for any particular state, the rates of entry and exit must be equal, a condition known as detailed balance. The overall form of the steady-state distribution, which depends on the full set of transition rates, is reached regardless of the starting conditions.

Letting $m_{i,j}$ denote the rate of evolutionary transition from state i to state j, we have a system of $L+1$ simultaneous equations (where L denotes the final state in the series, which starts with index 0),

$$\begin{aligned} p_0(t+1) &= (1-m_{0,1})p_0(t) + m_{1,0}p_1(t), \\ p_i(t+1) &= m_{i-1,i}p_{i-1}(t) + (1-m_{i,i-1}-m_{i,i+1})p_i(t) + m_{i+1,i}p_{i+1}(t), \\ p_L(t+1) &= m_{L-1,L}p_{L-1}(t) + (1-m_{L,L-1})p_L(t). \end{aligned}$$

Assuming nonzero transition rates between all adjacent classes in this linear array, the equilibrium solution (the steady-state probability of being in state i) takes on a simple form (Lynch 2013),

$$\widetilde{p}_i = \frac{\left(\prod_{j=0}^{i-1} m_{j,j+1}\right)\left(\prod_{k=i+1}^{L} m_{k,k-1}\right)}{C} \tag{5.4.1}$$

where the first term in the numerator is equal to the product of all transition rates pointing up toward the class, the second term is the product of all transition rates pointing down toward the class, and C is simply a normalization constant that ensures that all of the $\widetilde{p}_i$ sum to one (it is equal to the sum of numerators for all i, and is generally called the partition function).

As an example, with four alternative states (indexed 0, 1, 2, 3), the equilibrium probabilities become

$$\begin{aligned} \widetilde{p}_0 &= m_{1,0}m_{2,1}m_{3,2}/C, \\ \widetilde{p}_1 &= m_{0,1}m_{2,1}m_{3,2}/C, \\ \widetilde{p}_2 &= m_{0,1}m_{1,2}m_{3,2}/C, \\ \widetilde{p}_3 &= m_{0,1}m_{1,2}m_{2,3}/C, \end{aligned}$$

where C is the sum of the numerators in all four expressions. The steady-state probabilities, $\widetilde{p}_i$, can be equivalently viewed as the proportion of time a specific lineage spends in state i over a long evolutionary period, or as the fraction of independent populations experiencing identical population-genetic environments that are expected to reside in class i at any specific time.

In the context of the model introduced in the text, each of the m coefficients can be viewed as the product of the number of mutations arising in the population per generation and the probability of fixation. Letting N be the population size, $\mu_{i,j}$ be the mutation rate from allelic state i to j (where j can be only $i-1$ or $i+1$), and $\phi_{i,j}$ be the probability of fixation of a newly arisen j allele in a population currently occupied

by allele i, the transition rates are equal to the products of the relevant numbers of new mutant alleles arising per generation and their probabilities of fixation, i.e., $m_{i,j} = 2N\mu_{i,j}\phi_{i,j}$ assuming diploidy (or one half that assuming haploidy). Because every coefficient has the same prefactor, $2N$ or N, this can be ignored, reducing the coefficients to $m_{i,j} = \mu_{i,j}\phi_{i,j}$. (In the text, the $\mu_{i,j}$ are functions of per-site mutation rates and the number of sites relevant to the particular transition).

A second key simplification arises from the behavior of the probability of fixation in opposite directions between adjacent states. Letting s_i denote the selective disadvantage of allele i, measured relative to a perfect fitness of 1.0, then $s_{i+1} < s_i$ implies that allele $i+1$ is beneficial compared to allele i. Assuming mutations with additive effects on fitness, application of the formula for the fixation probability of new mutations, Equation 4.1b, yields the convenient result that $\phi_{i,i+1}/\phi_{i+1,i} = e^{4N_e(s_i - s_{i+1})}$ for diploids (with a 2 substituted for the 4 in haploids) (Berg et al. 2004; Sella and Hirsh 2005; Lynch 2013; Lynch and Hagner 2014).

As a simple application of the preceding methods, consider the situation in which there are just two alternative states, A and a, with the mutation rate from A to a being u, from a to A being v, and s being the selective advantage of a (negative if a is disadvantageous). The combined mutation/selection pressure towards a is then $u\phi_{A,a}$, while that towards A is $v\phi_{a,A}$, implying that

$$\widetilde{p}_a = \frac{u\phi_{A,a}}{u\phi_{A,a} + v\phi_{a,A}}, \tag{5.4.2a}$$

where the denominator is the partition function. Dividing all terms by $v\phi_{a,A}$, and using the relationship just noted for ratios of opposite fixation probabilities for diploids, leads to the simplification

$$\widetilde{p}_a = \frac{(u/v)e^{4N_e s}}{1 + (u/v)e^{4N_e s}}. \tag{5.4.2b}$$

Literature Cited

Abramowitz, M., and I. A. Stegun (eds.) 1964. Handbook of Mathematical Functions. Dover Publ., Inc., New York, NY.

Azevedo, L., G. Suriano, B. van Asch, R. M. Harding, and A. Amorim. 2006. Epistatic interactions: how strong in disease and evolution? Trends Genet. 22: 581-585.

Baer, C. F., M. M. Miyamoto, and D. R. Denver. 2007. Mutation rate variation in multicellular eukaryotes: causes and consequences. Nat. Rev. Genet. 8: 619-631.

Bataillon, T., and S. F. Bailey. 2014. Effects of new mutations on fitness: insights from models and data. Ann. N. Y. Acad. Sci. 1320: 76-92.

Behringer, M. G., B. I. Choi, S. F. Miller, T. G. Doak, J. A. Karty, W. Guo, and M. Lynch. 2018. *Escherichia coli* cultures maintain stable subpopulation structure during long-term evolution. Proc. Natl. Acad. Sci. USA 115: E4642-E4650.

Berg, J., S. Willmann, and M. Lässig. 2004. Adaptive evolution of transcription factor binding sites. BMC Evol. Biol. 4: 42.

Blount, Z. D., C. Z. Borland, and R. E. Lenski. 2008. Historical contingency and the evolution of a key innovation in an experimental population of *Escherichia coli.* Proc. Natl. Acad. Sci. USA 105: 7899-7906.

Böndel, K. B., S. A. Kraemer, T. Samuels, D. McClean, J. Lachapelle, R. W. Ness, N. Colegrave, and P. D. Keightley. 2019. Inferring the distribution of fitness effects of spontaneous mutations in *Chlamydomonas reinhardtii.* PLoS Biol. 17: e3000192.

Booker, T. R., and P. D. Keightley. 2018. Understanding the factors that shape patterns of nucleotide diversity in the house mouse genome. Mol. Biol. Evol. 35: 2971-2988.

Breen, M. S., C. Kemena, P. K. Vlasov, C. Notredame, and F. A. Kondrashov. 2012. Epistasis as the primary factor in molecular evolution. Nature 490: 535-538.

Campos, P. R. A., and Wahl, L. M. 2009. The effects of population bottlenecks on clonal interference, and the adaptation effective population size. Evolution 63: 950-958.

Campos, P. R. A., and Wahl, L. M. 2010. The adaptation rate of asexuals: deleterious mutations, clonal interference and population bottlenecks. Evolution 64: 1973-1983.

Charlesworth, B., and D. Charlesworth. 2010. Elements of Evolutionary Genetics. Roberts and Co. Publ., Greenwood Village, CO.

Charlesworth, B., and K. Jain. 2014. Purifying selection, drift, and reversible mutation with arbitrarily high mutation rates. Genetics 198: 1587-1602.

Cohan, F. M. 1984. Can uniform selection retard random genetic divergence between isolated populations? Evolution 38: 495-504.

DePristo, M. A., D. L. Hartl, and D. M. Weinreich. 2007. Mutational reversions during adaptive protein evolution. Mol. Biol. Evol. 24: 1608-1610.

Fisher, R. A. 1930. The Genetical Theory of Natural Selection. Oxford Univ. Press, Oxford, UK.

Frenkel, E. M., B. H. Good, and M. M. Desai. 2014. The fates of mutant lineages and the distribution of fitness effects of beneficial mutations in laboratory budding yeast populations. Genetics 196: 1217-1226.

Gillespie, J. H. 1983. Some properties of finite populations experiencing strong selection and weak mutation. Amer. Natur. 121: 691-708.

Gillespie, J. H. 1984. Molecular evolution over the mutational landscape. Evolution 38: 1116-1129.

Gokhale, C. S., Y. Iwasa, M. A. Nowak, and A. Traulsen. 2009. The pace of evolution across fitness valleys. J. Theor. Biol. 259: 613-620.

Good, B. H., and M. M. Desai. 2014. Deleterious passengers in adapting populations. Genetics 198: 1183-1208.

Good, B. H., M. J. McDonald, J. E. Barrick, R. E. Lenski, and M. M. Desai. 2017. The dynamics of molecular evolution over 60,000 generations. Nature 551: 45-50.

Higgs, P. G. 1998. Compensatory neutral mutations and the evolution of RNA. Genetica 102/103: 91-101.

Huber, C. D., B. Y. Kim, C. D. Marsden, and K. E. Lohmueller. 2015. Determining the factors driving selective effects of new nonsynonymous mutations. Proc. Natl. Acad. Sci. USA 114: 4465-4470.

Iwasa, Y., F. Michor, and M. A. Nowak. 2004. Stochastic tunnels in evolutionary dynamics. Genetics 166: 1571-1579.

Johri, P., B. Charlesworth, and J. D. Jensen. 2020. Towards an evolutionarily appropriate null model: jointly inferring demography and purifying selection. Genetics 215: 173-192. .

Katju, V., and U. Bergthorsson. 2019. Old trade, new tricks: insights into the spontaneous mutation process from the partnering of classical mutation accumulation experiments with high-throughput genomic approaches. Genome Biol. Evol. 11: 136-165.

Keightley, P. D. 1994. The distribution of mutation effects on viability in *Drosophila melanogaster.* Genetics 138: 1315-1322.

Keightley, P. D., and A. Eyre-Walker. 1999. Terumi Mukai and the riddle of deleterious mutation rates. Genetics 153: 515-523.

Keightley, P. D., and A. Eyre-Walker. 2007. Joint inference of the distribution of fitness effects of deleterious mutations and population demography based on nucleotide polymorphism frequencies. Genetics 177: 2251-2261.

Kim, B. Y., C. D. Huber, and K. E. Lohmueller. 2017. Inference of the distribution of selection coefficients for new nonsynonymous mutations using large samples. Genetics 206: 345-361.

Kimura, M. 1983. The Neutral Theory of Molecular Evolution. Cambridge Univ. Press, Cambridge, UK.

Kimura, M. 1985. The role of compensatory neutral mutations in molecular evolution. J. Genetics 64: 7-19.

Kimura, M., T. Maruyama, and J. F. Crow. 1963. The mutation load in small populations. Genetics 48: 1303-1312.

Kitahara, K., Y. Yasutake, and K. Miyazaki. 2012. Mutational robustness of 16S ribosomal RNA, shown by experimental horizontal gene transfer in *Escherichia coli.* Proc. Natl. Acad. Sci. USA 109: 19220-19225.

Komarova, N. L., A. Sengupta, and M. A. Nowak. 2003. Mutation-selection networks of cancer

initiation: tumor suppresser genes and chromosomal instability. J. Theor. Biol. 223: 433-450.

Kondrashov, A. S., S. Sunyaev, and F. A. Kondrashov. 2002. Dobzhansky-Muller incompatibilities in protein evolution. Proc. Natl. Acad. Sci. USA 99: 14878-14883.

Kulathinal, R. J., B. R. Bettencourt, and D. L. Hartl. 2004. Compensated deleterious mutations in insect genomes. Science 306: 1553-1554.

Lind, P. A., E. Libby, J. Herzog, and P. B. Rainey. 2019. Predicting mutational routes to new adaptive phenotypes. eLife 8: e38822.

Long, H., W. Sung, S. Kucukyildirim, E. Williams, S. W. Guo, C. Patterson, C. Gregory, C. Strauss, C. Stone, C. Berne, et al. 2017. Evolutionary determinants of genome-wide nucleotide composition. Nature Ecol. Evol. 2: 237-240.

Lynch, M. 1986. Random drift, uniform selection, and the degree of population differentiation. Evolution 40: 640-643.

Lynch, M. 2007. The Origins of Genome Architecture. Sinauer Assocs., Inc., Sunderland, MA.

Lynch, M. 2010. Scaling expectations for the time to establishment of complex adaptations. Proc. Natl. Acad. Sci. USA 107: 16577-16582.

Lynch, M. 2013. Evolutionary diversification of the multimeric states of proteins. Proc. Natl. Acad. Sci. USA 110: E2821-E2828.

Lynch, M. 2018. Phylogenetic diversification of cell biological features. eLife 7: e34820.

Lynch, M. 2020. The evolutionary scaling of cellular traits imposed by the drift barrier. Proc. Natl. Acad. Sci. USA 117: 10435-10444.

Lynch, M., and A. Abegg. 2010. The rate of establishment of complex adaptations. Mol. Biol. Evol. 27: 1404-1414.

Lynch, M., M. Ackerman, K. Spitze, Z. Ye, and T. Maruki. 2017. Population genomics of *Daphnia pulex.* Genetics 206: 315-332.

Lynch, M., J. Blanchard, D. Houle, T. Kibota, S. Schultz, L. Vassilieva, and J. Willis. 1999. Spontaneous deleterious mutation. Evolution 53: 645-663.

Lynch, M., and K. Hagner. 2014. Evolutionary meandering of intermolecular interactions along the drift barrier. Proc. Natl. Acad. Sci. USA 112: E30-E38.

Lynch, M., and G. K. Marinov. 2015. The bioenergetic costs of a gene. Proc. Natl. Acad. Sci. USA 112: 15690-15695.

McCandlish, D. M. 2018. Long-term evolution on complex fitness landscapes when mutation is weak. Heredity 121: 449-465.

McCandlish, D. M., and A. Stoltzfus. 2014. Modeling evolution using the probability of fixation: history and implications. Quart. Rev. Biol. 89: 225-252.

Neher, R. A. 2013. Genetic draft, selective interference, and population genetics of rapid adaptation. Ann. Rev. Ecol. Evol. Syst. 44: 195-215.

Nguyen Ba, A. N., I. Cvijović, J. I. Rojas Echenique, K. R. Lawrence, A. Rego-Costa, X. Liu, S. F. Levy, and M. M. Desai. 2019. High-resolution lineage tracking reveals travelling wave of adaptation in laboratory yeast. Nature 575: 494-499.

Pénisson S., T. Singh, P. Sniegowski, and P. Gerrish. 2017. Dynamics and fate of beneficial mutations under lineage contamination by linked deleterious mutations. Genetics 205: 1305-1318.

Quandt, E. M., D. E. Deatherage, A. D. Ellington, G. Georgiou, and J. E. Barrick. 2014. Recursive genomewide recombination and sequencing reveals a key refinement step in the evolution of a metabolic innovation in *Escherichia coli.* Proc. Natl. Acad. Sci. USA 111: 2217-2222.

Rice, D. P., B. H. Good, and M. M. Desai. 2015. The evolutionarily stable distribution of fitness effects. Genetics 200: 321-329.

Robert, L., J. Ollion, J. Robert, X. Song, I. Matic, and M. Elez. 2018. Mutation dynamics and fitness effects followed in single cells. Science 359: 1283-1286.

Rodrigues, J. V., and E. I. Shakhnovich. 2019. Adaptation to mutational inactivation of an essential gene converges to an accessible suboptimal fitness peak. eLife 8: e50509.

Santiago, E. 2015. Probability and time to fixation of an evolving sequence. Theor. Popul. Biol. 104: 78-85.

Sella, G., and A. E. Hirsh. 2005. The application of statistical physics to evolutionary biology. Proc. Natl. Acad. Sci. USA 102: 9541-9546.

Stephan, W. 1996. The rate of compensatory evolution. Genetics 144: 419-426.

Stephan, W., and D. A. Kirby. 1993. RNA folding in *Drosophila* shows a distance effect for compensatory fitness interactions. Genetics 135: 97-103.

Sung, W., M. S. Ackerman, M. Dillon, T. Platt, C. Fuqua, V. Cooper, and M. Lynch. 2016. Evolution of the insertion-deletion mutation rate across the tree of life. G3 (Bethesda) 6: 2583-2591.

Tóth-Petróczy, A., and D. S. Tawfik. 2013. Protein insertions and deletions enabled by neutral roaming in sequence space. Mol. Biol. Evol. 30: 761-771.

Tuğrul, M., T. Paixão, N. H. Barton, and G. Tkačik. 2015. Dynamics of transcription factor binding site evolution. PLoS Genet. 11: e1005639.

Walsh, J. B. 1995. How often do duplicated genes evolve new functions? Genetics 139: 421-428.

Walsh, J. B., and M. Lynch. 2018. Evolution of Quantitative Traits. Sinauer Assocs., Inc., Sunderland, MA.

Weinreich, D. W., and L. Chao. 2005. Rapid evolutionary escape by large populations from local fitness peaks is likely in nature. Evolution 59: 1175-1182.

Weissman, D. B., and N. H. Barton. 2012. Limits to the rate of adaptive substitution in sexual populations. PLoS Genet. 8: e1002740.

Weissman, D. B., M. M. Desai, D. S. Fisher, and M. W. Feldman. 2009. The rate at which asexual populations cross fitness valleys. Theor. Pop. Biol. 75: 286-300.

Weissman, D. B., D. S. Fisher, and M. W. Feldman. 2010. The rate of fitness-valley crossing in sexual populations. Genetics 186: 1389-1410.

Weissman, D. B., and O. Hallatschek. 2014. The rate of adaptation in large sexual populations with linear chromosomes. Genetics 196: 1167-1183.

Wiser, M. J., N. Ribeck, and R. E. Lenski. 2013. Long-term dynamics of adaptation in asexual populations. Science 342: 1364-1367.

Wright, S. 1931. Evolution in Mendelian populations. Genetics 16: 97-159.

Wright, S. 1932. The roles of mutation, inbreeding, crossbreeding and selection in evolution. Proc. 6th Internat. Cong. Genetics 1: 356-366.

Wright, S. 1969. Evolution and Genetics of Populations. The Theory of Gene Frequencies. Univ. Chicago Press, Chicago, IL.

Wu, N. C., L. Dai, C. A. Olson, J. O. Lloyd-Smith, and R. Sun. 2016. Adaptation in protein fitness landscapes is facilitated by indirect paths. eLife 5: e16965.

Zheng, J., J. L. Payne, and A. Wagner. 2019. Cryptic genetic variation accelerates evolution by opening access to diverse adaptive peaks. Science 365: 347-353.

Figure 5.1. An idealized representation of the distribution of fitness effects (DFE) of new mutations in large-N_e species (with genomes consisting primarily of coding DNA) vs. small-N_e species (with genomes consisting primarily of noncoding DNA), based on arguments presented in Foundations 5.1. Although it is clear that most mutations have relatively small effects, there remain substantial uncertainties about the precise form of the DFE, and the absolute placements of the peaks are somewhat arbitrary. The small peak to the right is meant to represent a class of mutations of relatively large functional effects, as expected with amino-acid substitutions in catalytic sites of proteins, premature stop codons, and gene deletions. Other amino-acid altering mutations may have minor effects on protein function, while nonetheless imposing small biosynthetic costs. The distribution of beneficial effects is not shown, but in total this is expected to be $< 1\%$ of that for deleterious mutations, again with the mode being near zero. The dots on the left denote the fraction of absolutely neutral mutations, whereas a deleterious effect of $10^0 = 1$ denotes a lethal mutation. Note that the curves in both examples are normalized to have constant areas, as they are relative frequency distributions; if these were scaled to the absolute numbers of mutations, the total heights of the abundance distributions would depend on the mutation rate per nucleotide site, which tends to be lowest in unicellular eukaryotes, somewhat higher in prokaryotes, and highest in multicellular eukaryotes (Chapter 4). Hypothetical positions of the drift barrier, $\simeq 1/N_e$, denoted by the vertical red lines, are given for $N_e = 10^9$ and 10^6 (upper and lower panels, respectively); mutations to the left of this line cannot be perceived by natural selection.

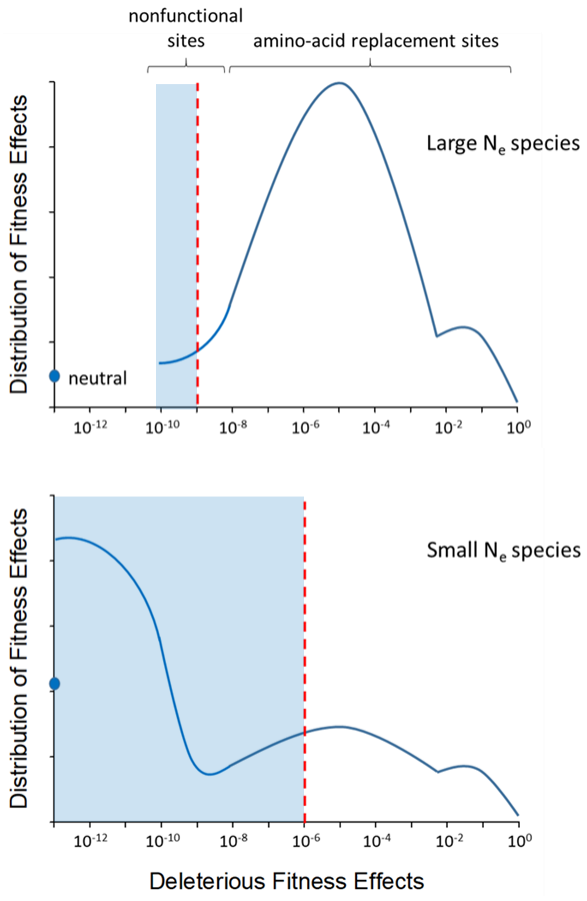

Figure 5.2 Upper panels). Allele-frequency trajectories in three replicate cultures of *E. coli* grown in 10-mL cultures serially diluted 100-fold on a daily basis. To estimate allele frequencies, the complete genomes of each mixed culture were subject to pooled-population sequencing, to an average of 50× depth of coverage, every 500 generations over a 60,000-generation period. Each individual line in the plots denotes a mutation that arose to frequency 0.1 on at least one occasion. Results are shown for three replicate populations. All cultures were genetically identical and monomorphic at time zero. From Good et al. (2017). **Lower panel)** Evolutionary trajectory of mean fitness (cell-division rate) relative to time-zero control averaged over multiple replicate populations. From Wiser et al. (2013).

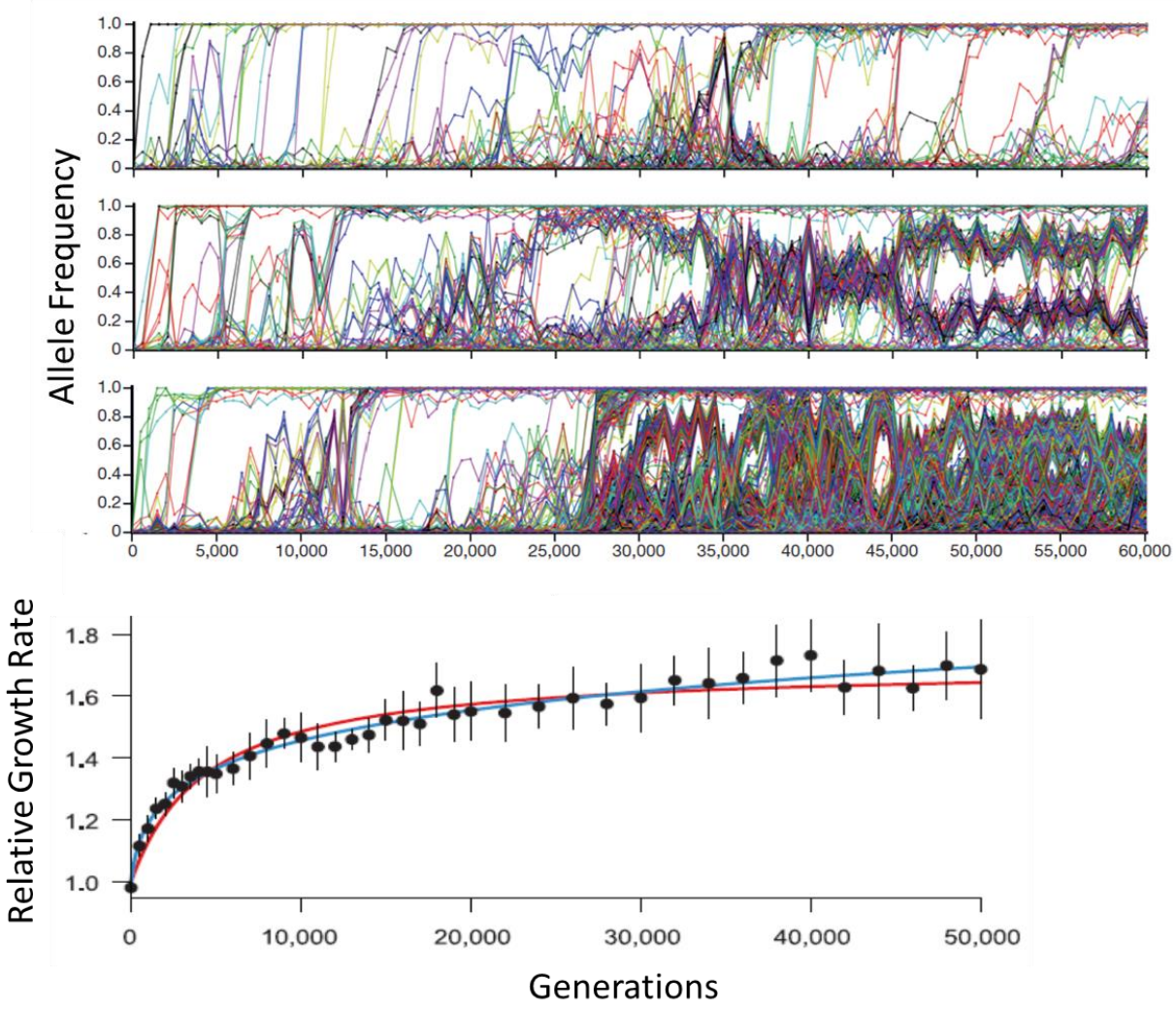

Figure 5.3. Some possible routes to the establishment of adaptations involving two or more mutations at different genetic loci (which might be two nucleotide sites within single loci). In each case, the starting genotype is on the left. Mutations that are deleterious with respect to their ancestral allele are denoted in blue, whereas neutral changes are in white, and beneficial changes are denoted in red. Intermediate states can also be beneficial, although such instances are not shown. The lower left denotes a case involving three loci, with all intermediate states being deleterious, and the final three-mutation genotype being beneficial. The lower right illustrates a situation in which the first site has three alternative allelic states, with one derived allele (a) being deleterious and the other (a′) providing a three-step neutral path to the final adaptation.

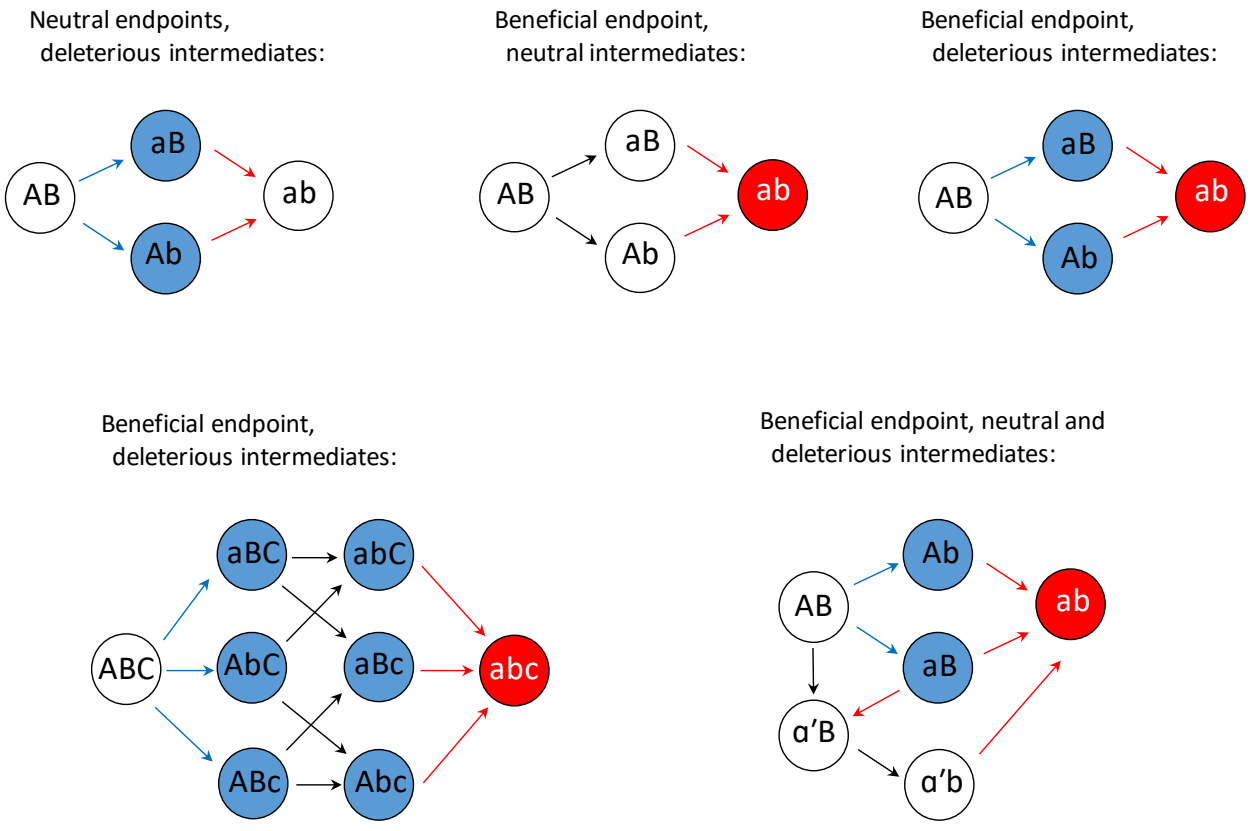

Figure 5.4. Origin of a complex adaptation involving three mutations (blue (1), red (2), and green (3)) in small and large populations. The vertical axis denotes the frequency in the population, with fixation implied when a vertical line through the temporal profile is all of one color. **Top)** Here, the population size is small enough and mutation sufficiently weak that each contributing mutation goes to fixation prior to the appearance of the next mutation destined to fix. This leads to sequential fixation of single-step changes, with potentially long waiting times between fixation events. As illustrated, a few secondary mutations may arise and go extinct (closed red polygons) before one reaches a high enough frequency to be susceptible to fixation. In the illustrated case, the final (green) mutation has not arisen yet. **Bottom)** In large populations, allelic variants will typically be segregating at low frequencies at all times, even if deleterious, owing to the recurrent input by mutation. Secondary and tertiary mutations can then arise on such backgrounds, occasionally generating a beneficial combination that leads to fixation of the entire linked haplotype by positive selection.

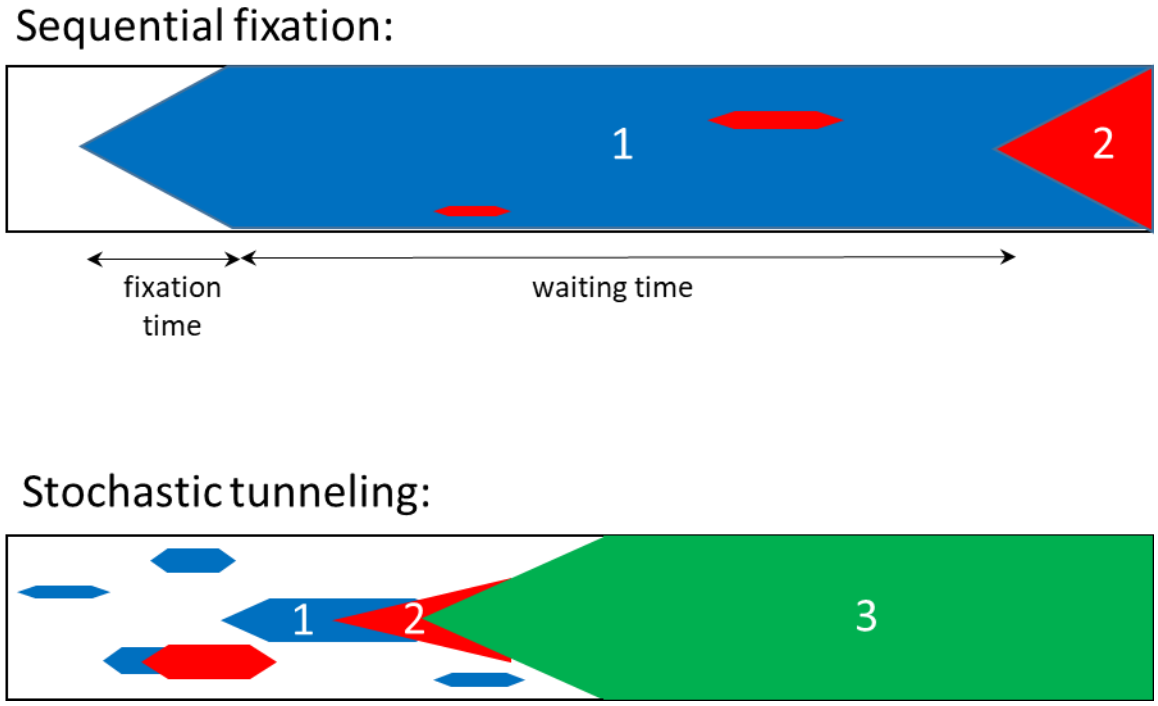

Figure 5.5. Expected frequency of the beneficial allele at a biallelic locus under the joint forces of drift, mutation, and selection. The mutation rate to deleterious alleles (μ_{10}) is assumed to be $3\times$ that to beneficial alleles (μ_{01}), as would be approximately the case for complementary nucleotides in an RNA stem. Results are given for four intensities of selection (with s denoting the selective advantage of the beneficial allele) relative to drift. Solid lines give the exact results from Foundations 5.3, whereas the dashed lines are the results under the sequential model, Equation 5.8, which assumes the population to be nearly always in a monomorphic state and becomes increasingly unreliable as $N_e\mu_{01}$ exceeds 0.01. Note that when $s = 0$ (neutrality), the expected frequency (0.25 in this case) is independent of the population mutation rate.

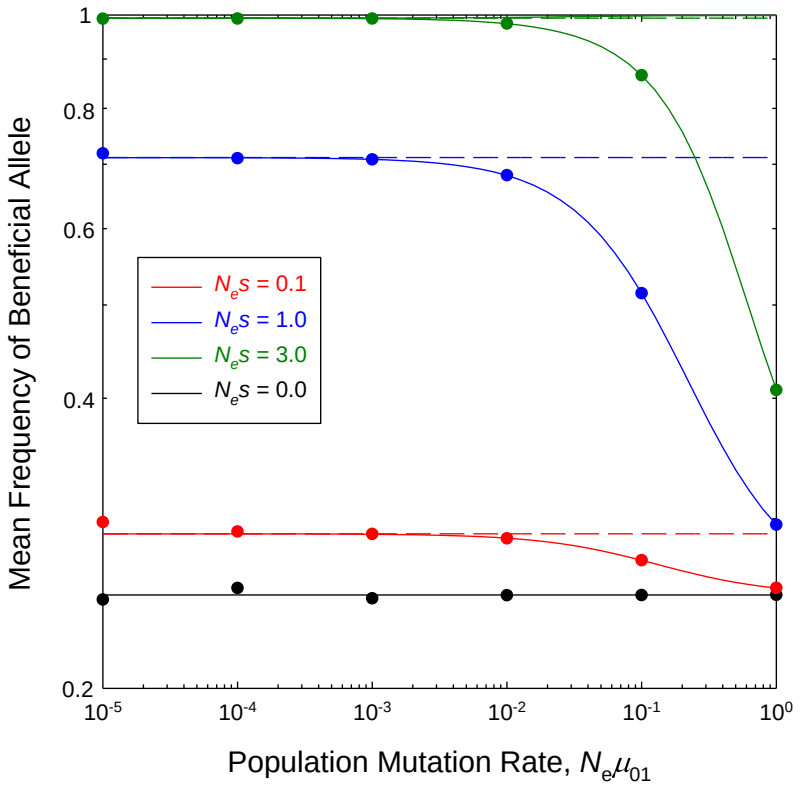

Figure 5.6. Schematic for the transition rates between adjacent genotypic classes under the sequential-fixation model for the case of $L = 4$ sites. This layout readily generalizes to any value of L. μ_{01} and μ_{10} are the mutation rates from $-$ to $+$ allelic states, and vice versa, and ϕ_{ij} is the probability of fixation of a newly arisen mutation to allele j from a background of i, defined by Equation 4.1b. The number of $-$ alleles in a class is denoted by i, and except for the two extreme classes ($i = 0$ and $i = 4$), there are multiple equivalent genotypic states within each class of genotypic values.

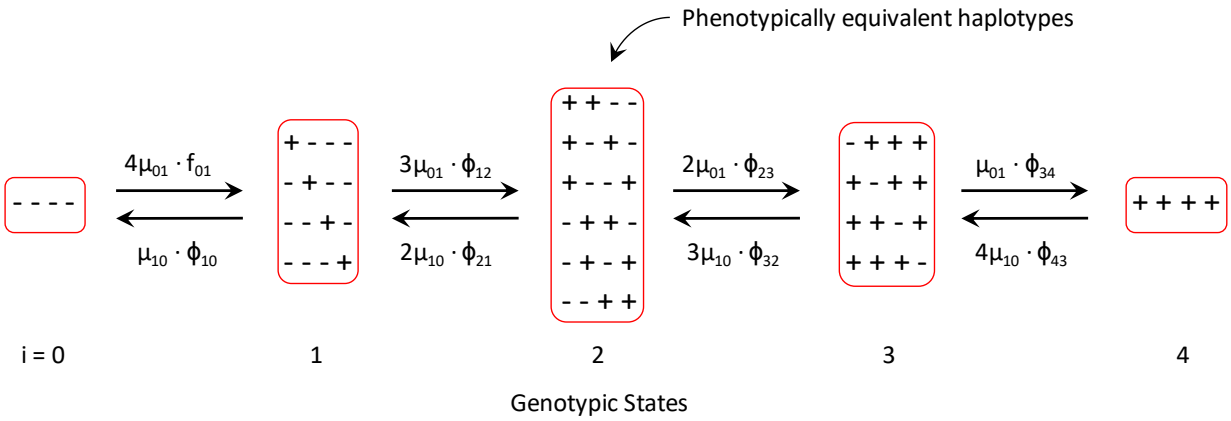

Figure 5.7. The response of the long-term mean genotypic state (number of + alleles; panel **a**) and the underlying mean class frequencies (for $i = 0$, 1, and 2 + alleles; panel **b**) over a gradient of effective population sizes for a two-locus, two-allele, sequential model. The mutation rate to beneficial alleles is 10% of that to deleterious alleles ($\beta = 0.1$). Results are given for three phenotypic optima, with the width of the fitness function $\omega = 5000$. For each color-coded optimum, the mean frequencies of the three genotypic classes ($i = 0$, 1, 2) are given as solid, short-dashed, and long-dashed lines, which for each N_e sum to 1.0. Results are derived from Equations 5.10 and 5.11. From Lynch (2020).

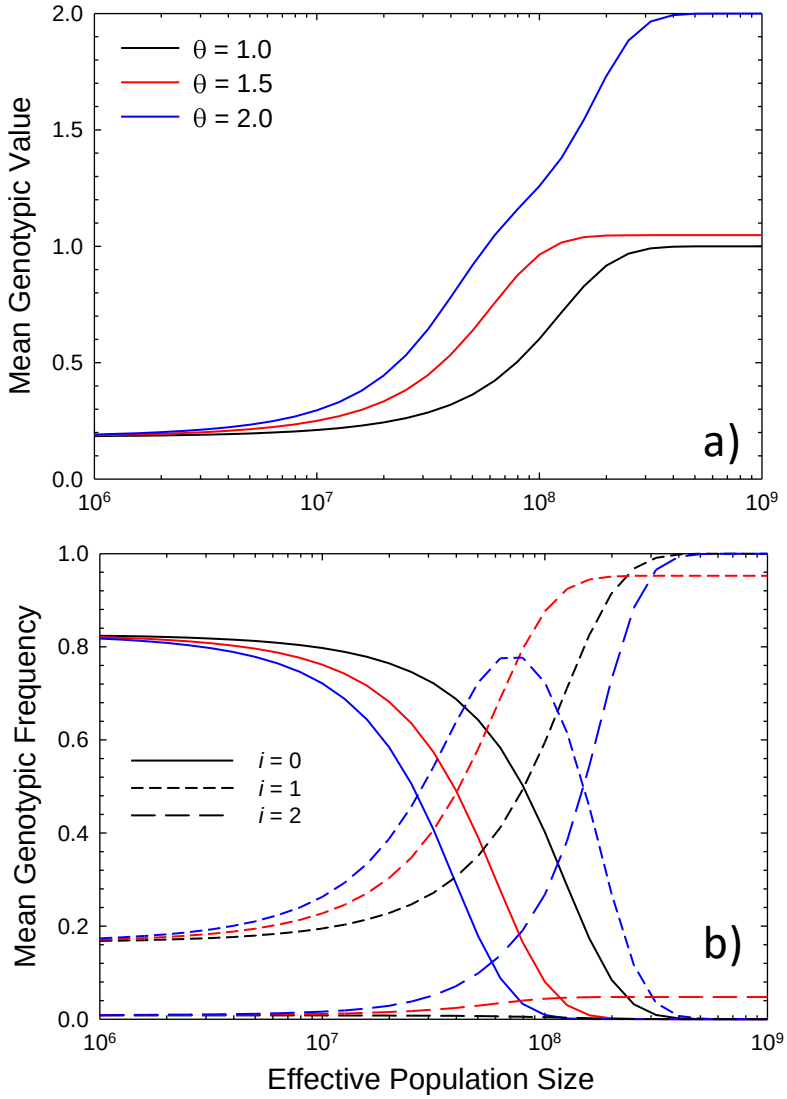

Figure 5.8. The equilibrium mean frequency of + alleles as a function of the absolute population size for the case in which the mutation bias to + alleles is 0.01. Results are given for three numbers of loci (L) with equivalent multiplicative effects on fitness. Analytical results (solid lines) are given for the case of free recombination, Equation 5.3.3, where simulation results (dashed lines and data points) are given for the case of complete linkage. Throughout, the selective disadvantage of a $-$ allele is $s = 1/L$, the inverse of the number of factors contributing to the trait: red, $L = 10^4, s = 10^{-4}$; green, $L = 10^5, s = 10^{-5}$; blue, $L = 10^6, s = 10^{-6}$, so that the total fitness effects of mutations is the same in all cases. From Lynch (2020).