

Genetic Draft, Selective Interference, and Population Genetics of Rapid Adaptation

Richard A. Neher

Max Planck Institute for Developmental Biology, Tübingen 72070, Germany; email: richard.neher@tuebingen.mpg.de

Annu. Rev. Ecol. Evol. Syst. 2013. 44:195–215

First published online as a Review In Advance on October 9, 2013

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

This article's doi:
10.1146/annurev-ecolsys-110512-135920

Copyright © 2013 by Annual Reviews.
All rights reserved

Keywords

Hill-Robertson effect, clonal interference, genetic diversity, genealogies

Abstract

To learn about the past from a sample of genomic sequences, one needs to understand how evolutionary processes shape genetic diversity. Most population genetics inferences are based on frameworks assuming that adaptive evolution is rare. But if positive selection operates on many loci simultaneously, as has recently been suggested for many species, including animals such as flies, then a different approach is necessary. In this review, I discuss recent progress in characterizing and understanding evolution in rapidly adapting populations, in which random associations of mutations with genetic backgrounds of different fitness, i.e., genetic draft, dominate over genetic drift. As a result, neutral genetic diversity depends weakly on population size but strongly on the rate of adaptation or more generally the variance in fitness. Coalescent processes with multiple mergers, rather than Kingman's coalescent, are appropriate genealogical models for rapidly adapting populations, with important implications for population genetics inference.

Genetic drift:

stochastic changes in allele frequencies due to nonheritable variation in offspring number

Purifying selection:

selection against deleterious mutations

Positive selection:

selection for novel beneficial mutations

Hitchhiking:

rapid rise in frequency through an association with a very fit background

Genetic draft:

changes in allele frequencies due to (partly) heritable random associations with genetic backgrounds

Multiple-merger coalescent:

coalescent process with simultaneous merging of more than two lineages

1. INTRODUCTION

Neutral diffusion or coalescent models (Kingman 1982, Kimura 1964) predict that genetic diversity at unconstrained sites is proportional to the population size, N , for a simple reason: Two randomly chosen individuals have a common parent with a probability of order $1/N$, and the first common ancestor of two individuals lived of order N generations ago. This neutral coalescence has its origin in genetic drift. Nevertheless, the observed correlation between genetic diversity and population size is rather weak (Lewontin 1974, Leffler et al. 2012), implying that processes other than genetic drift drive coalescence in large populations. This notion is reinforced by the observation that pesticide resistance in insects can evolve independently on multiple genetic backgrounds (Karasov et al. 2010, Labbé et al. 2007) and can involve several adaptive steps in rapid succession (Schmidt et al. 2010). This high mutational input suggests that the short-term effective population size of *Drosophila melanogaster* is greater than 10^9 and that conventional genetic drift should be negligible. Possible forces that accelerate coalescence and reduce diversity are purifying and positive selection. Historically, the effects of purifying selection have received the most attention (reviewed by Charlesworth 2012), and so my focus here is on the role of positive selection.

A selective sweep reduces nearby polymorphisms through hitchhiking. Polymorphisms linked to the sweeping allele increase to a higher frequency, whereas others are driven out (Maynard Smith & Haigh 1974). Linked selection not only reduces diversity but also slows down adaptation in other regions of the genome—an effect known as Hill-Robertson interference (Hill & Robertson 1966). Hill-Robertson interference has been intensively studied in two-locus models (Barton 1994), in which the effect is quite intuitive: Two linked beneficial mutations arising in different individuals compete, and the probability that both mutations fix increases with the recombination rate between the loci. Pervasive selection, however, requires many-locus models. Here, I review recent progress in understanding how selection at many loci limits adaptation and shapes genetic diversity. Linked selection is most dramatic in asexual organisms. The theory of asexual evolution is partly motivated by evolution experiments with microbes, which have provided us with detailed information about the spectrum of adaptive molecular changes and their dynamics. I then turn to facultatively sexual organisms, which include many important human pathogens, such as HIV and influenza, as well as some plants and nematodes. Finally, I discuss obligately sexual organisms, in which linked selection is mostly due to neighboring loci on the chromosome.

The common aspect of all these models is the source of stochastic fluctuations: random associations with backgrounds of different fitness. In contrast to genetic drift, such associations persist for many generations, amplifying their effect. To emphasize the analogy to genetic drift, Gillespie (2000) called fluctuations in allele frequencies due to linked selection genetic draft. The (census) population size N determines how readily adaptive mutations and combinations thereof are discovered, but it has little influence on coalescent properties and genetic diversity. Instead, selection determines genetic diversity and sets the timescale of coalescence. This reduced coalescence timescale is often used to define an effective population size, N_e . This is unfortunate, because such rebranding suggests that a rescaled neutral model is an accurate description of reality. In fact, many features are qualitatively different. Negligible drift does not imply that selection is efficient and that only beneficial mutations matter. On the contrary, deleterious mutations can reach high frequency through linkage to favorable backgrounds, and the dynamics of genotype frequencies in the population remains very stochastic. Genealogies of samples from populations governed by draft do not follow the standard binary coalescent process. Instead, coalescent processes allowing for multiple mergers seem to be appropriate approximations that capture the large and anomalous fluctuations associated with selection. Those coalescent models thus form the basis for a population genetics of rapid adaptation and serve as null models to analyze data when

Kingman's coalescent is inappropriate. To illustrate clonal interference, draft, and genealogies in the presence of selection, a collection of scripts based on FFPopSim (Zanini & Neher 2012) is available at <http://webdav.tuebingen.mpg.de/interference>.

2. ADAPTATION OF LARGE AND DIVERSE ASEXUAL POPULATIONS

Evolution experiments (reviewed in Burke 2012, Kawecki et al. 2012) have demonstrated that adaptive evolution is ubiquitous among microbes. Experiments with RNA viruses have shown that the rate of adaptation increases only slowly with the population size (de Visser et al. 1999, Miralles et al. 1999), suggesting that adaptation is limited by competition among different mutations and not by the availability of beneficial mutations. The competition between clones, also known as clonal interference, was directly observed using fluorescent markers in *Escherichia coli* populations (Hegreness et al. 2006). Similar observations have been made in Rich Lenski's experiments, in which *E. coli* populations were followed for more than 50,000 generations (Barrick et al. 2009). A different experiment selecting >100 *E. coli* populations for heat tolerance has shown that thousands of sites are available for adaptive substitutions, that extensive parallelism exists among lines in the genes and pathways bearing mutations, and that mutations frequently interact epistatically (Tenaillon et al. 2012). By following the frequencies of microsatellite markers in populations of *E. coli*, Perfeito et al. (2007) estimated the beneficial mutation rate to be $U_b \approx 10^{-5}$ per genome and generation, with average effects of approximately 1%. Similarly, beneficial mutations are readily available in yeast and compete with one another in the population for fixation (Desai et al. 2007, Kao & Sherlock 2008, Lang et al. 2011). At any given instant, the population is thus characterized by many segregating clones, giving rise to a broad fitness distribution (Desai et al. 2007). The fate of a novel mutation is mainly determined by the genetic background on which it arises (Lang et al. 2011). Similar rapid adaptation and competition are observed in global populations of influenza, which experience several adaptive substitutions per year (Bhatt et al. 2011, Smith et al. 2004, Strelkowa & Lässig 2012), mainly driven by immune responses of the host. In summary, evolution of asexual microbes does not seem to be limited by finding the necessary single-point mutations but rather by overcoming clonal interference and combining multiple mutations.

These observations have triggered intense theoretical research on clonal interference and adaptation in asexuals. In the models studied, rare events, e.g., the fittest individual acquiring additional mutations, dramatically affect the future dynamics. Intuition is a poor guide in such situations, and careful mathematical treatment is warranted. Nevertheless, rationalizing the results in a simple and intuitive way with hindsight is often possible, and I try to present the important aspects in accessible form.

This discussion assumes that fitness is a unique function of the genotype. I thereby ignore the possibility of frequency-dependent selection. A diverse population with many different genotypes can then be summarized by its distribution along this fitness axis (**Figure 1a,b**). Fitness distributions are shaped by a balance between injection of variation via mutation and the removal of poorly adapted variants. Most mutations have detrimental effects on fitness, whereas only a small minority of mutations are beneficial. The distribution of mutational effects in RNA viruses has been estimated by mutagenesis (Lalić et al. 2011, Sanjuán et al. 2004). Roughly half of random mutations are effectively lethal, whereas 4% were found to be beneficial in this experiment. More generally, fitness effects, s , of de novo mutations come from a distribution, $U(s)$, like the one shown in **Figure 1c**. I define $U(s)$ such that the integral of $U(s)$ over s equals the total mutation rate. General properties of $U(s)$ are largely unknown and depend on the environment.

Deleterious mutations rarely reach high frequencies but are numerous, whereas beneficial mutations are rare but amplified by selection. But to spread and fix, a beneficial mutation has to

Kingman's coalescent: basic coalescence process in which random pairs of individuals merge

Clonal interference: competition between well-adapted asexual subpopulations from which only one subpopulation emerges as winner

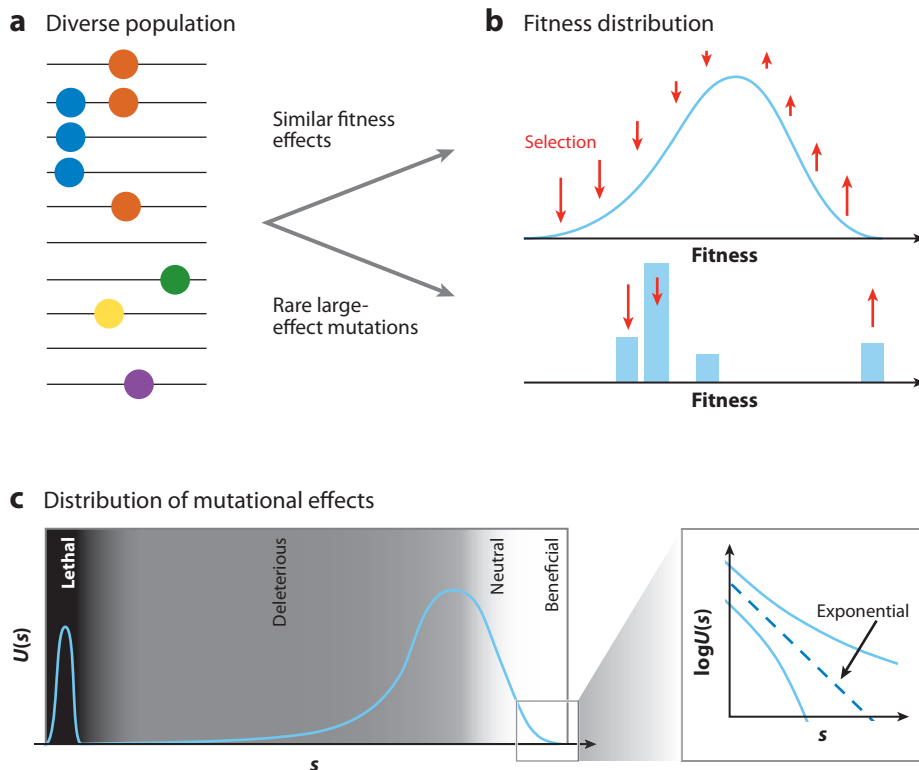


Figure 1

Fitness and fitness effect distributions. (a) A genetically diverse population (mutations indicated by *colored dots*) will typically harbor variation in fitness. (b, *top*) If many mutations have comparable fitness effects, then the resulting fitness distribution is smooth and roughly normal. (b, *bottom*) If few large-effect mutations exist, the distribution is multimodal. (c) Fitness effects, s , of de novo mutations across the genome may follow a distribution, $U(s)$, roughly like that shown here. A small fraction of mutations are beneficial, but most are neutral or deleterious, and some are lethal. In models of adaptive evolution, the high-fitness tail of $U(s)$, shown in the inset, is the most important part. If it falls off faster than exponentially, then the fitness distribution tends to be smooth. Otherwise, the distribution is often dominated by a few large-effect mutations.

arise on an already fit genetic background or have a sufficiently large effect on fitness to outcompete all others. Two lines of theoretical works have put emphasis on either the large-effect mutations (clonal interference theory) or coalitions of multiple mutations of similar effect. Both approaches, sketched in **Figure 2**, are good approximations, depending on the distribution of fitness effects.

2.1. Clonal Interference

Consider a homogeneous population in which mutations with an effect on fitness between s and $s + ds$ arise with rate $U(s) ds$, as sketched in **Figure 1c**. In a large population, many beneficial mutations arise in every generation. To fix, a beneficial mutation has to outcompete all others (**Figure 2a**). In other words, a mutation fixes only if no mutation with a larger effect arises before the former has reached high frequencies in the population. This idea is the essence of clonal interference theory by Gerrish & Lenski (1998). The Gerrish-Lenski theory of clonal interference is an approximation

because it ignores the possibility that two or more mutations with moderate effects can combine to outcompete a large-effect mutation—a process I discuss below. The theory's accuracy depends on the functional form of $U(s)$ and the population size (Park & Krug 2007). One central prediction of clonal interference is that the rate of adaptation increases only slowly with the population size, N , and the beneficial mutation rate, U_b . Larger NU_b results in more competing beneficial mutations, which reduces the fixation probability of individual mutations and results in a sublinear dependence of the rate of adaptation on NU_b . This basic prediction has been confirmed in evolution experiments with viruses (de Visser et al. 1999; Miralles et al. 1999, 2000). How the rate of adaptation depends on N and U_b is sensitive to the distribution of fitness effects, $U(s)$. Generally, one finds that the rate of adaptation is $\propto (\log NU_b)^\alpha$, where α depends on the properties of $U(s)$ (Park et al. 2010).

Clonal interference theory places all the emphasis on the mutation with the largest effect and ignores variation in genetic background or, equivalently, the possibility that multiple mutations can accumulate in one lineage. It should therefore work if the distribution of effect sizes has a long tail allowing for mutations of widely different sizes. It fails if most mutations have similar effects on fitness. Park et al. (2010) present a careful discussion of the theory of clonal interference and its limitations.

2.2. Genetic Background and Multiple Mutations

If most beneficial mutations have similar effects, then a lineage cannot fix by acquiring a mutation with very large effect but instead has to accumulate more beneficial mutations than the competing lineages. If population sizes and mutation rates are large enough that many mutations segregate, then the distribution of fitness x in the population at time t , $n(x, t)$, is roughly Gaussian (**Figure 2b**), and the problem becomes tractable (Desai & Fisher 2007, Rouzine et al. 2003, Tsimring et al. 1996). More precisely, $n(x, t)$ is governed by the deterministic equation

$$\frac{d}{dt}n(x, t) = (x - \bar{x})n(x, t) + \int U(s)[n(x - s, t) - n(x, t)] ds, \quad 1.$$

where $(x - \bar{x})n(x, t)$ accounts for amplification by selection of individuals fitter than the fitness mean, \bar{x} , and elimination of the less fit ones. The second term accounts for mutations that move individuals from $x - s$ to x at rate $U(s)$. Integrating this equation over fitness x yields Fisher's fundamental theorem of natural selection, which states that the rate of increase in mean fitness is

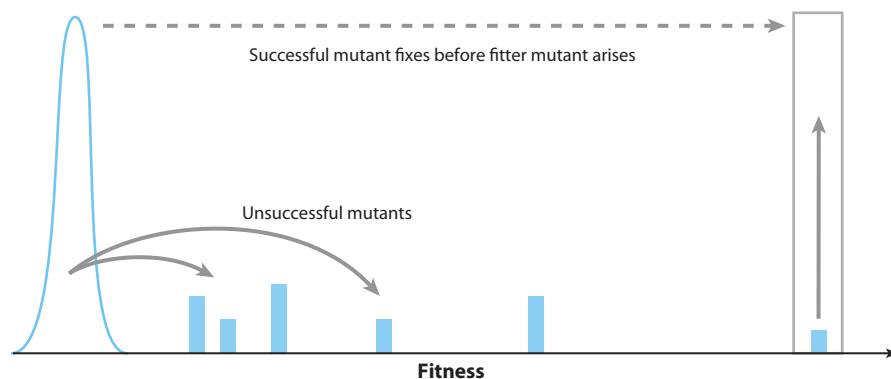
$$\frac{d}{dt}\bar{x} = v = \sigma^2 - \Delta_U, \quad 2.$$

where σ^2 is the variance in fitness, and Δ_U is the average mutation load a genome accumulates in one generation. A steadily moving mean fitness, $\bar{x} = vt$, suggests a traveling wave solution of the form $n(x, t) = n(\chi)$, where $\chi = x - \bar{x}$ is the fitness relative to the mean. Equation 2 is analogous to the breeder's equation, which links the response to selection to additive variances and covariances. In quantitative genetics, the trait variances are determined empirically and often assumed constant, whereas here we try to understand how σ^2 is determined by a balance between selection and mutation.

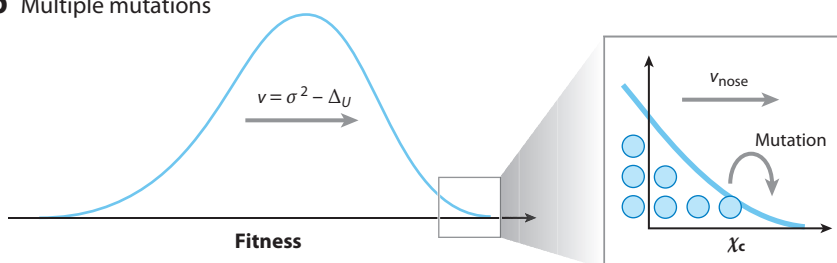
To determine the average v , we need an additional relation between v and the mutational input. To this end, we need to realize that the population is thinning out at higher and higher fitness and that only very few individuals are expected to be present above some χ_c , as sketched in **Figure 2b**. The dynamics of this high-fitness nose is very stochastic and not accurately described by Equation 1. Yet the nose is the most important part and is where most successful mutations arise. There have been two strategies to account for the stochastic effects and derive an additional relation for the velocity. First, the average velocity of the nose, v_{nose} , is determined by a detailed

study of the stochastic dynamics of the nose. At steady state, this velocity has to equal the average velocity of the mean fitness given by Equation 2, which produces the additional relation required to determine v (Brunet et al. 2008, Cohen et al. 2005a, Desai & Fisher 2007, Goyal et al. 2012, Rouzine et al. 2003, Tsimring et al. 1996). Second, if we assume additivity of mutations, then v has to equal the average rate at which fitness increases because of fixed mutations (Good et al. 2012,

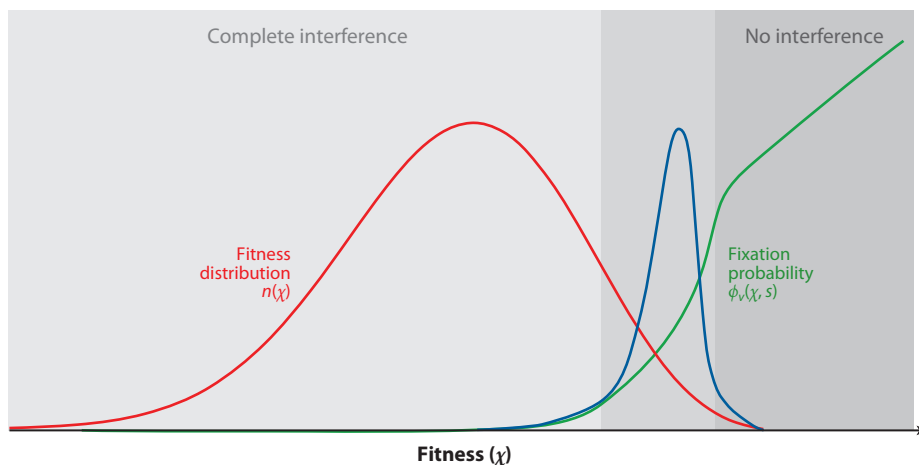
a Clonal interference



b Multiple mutations



c Fixation probability



Neher et al. 2010; see Hallatschek 2011 for a related idea). I largely focus on this latter approach, as it generalizes to sexual populations (discussed in Sections 3 and 4). In essence, we need to calculate the fixation probability of mutations with effect size s that arise in random individuals in a population that adapts with velocity v . Fixation probability $\Phi_v(s)$ depends on v and implicitly on the traveling fitness distribution, $n(x - vt)$. Using this notation, we can express v as the sum of effects of mutations that fix per unit of time:

$$v = \frac{d}{dt} \bar{x} = N \int U(s) \Phi(s, v) s \, ds. \quad 3.$$

The mutational input is proportional to the census population size, N . To solve Equation 3, we first have to calculate the fixation probability $\Phi_v(s)$, which in turn is a weighted average of the fixation probability, $\phi_v(\chi, s)$, given that the mutation appears on a genetic background with a relative fitness of χ . The latter probability can be approximated by branching processes (Good et al. 2012, Neher et al. 2010). The supplement of Good et al. (2012) gives a detailed derivation of $\phi_v(\chi, s)$, whereas Fisher (2013) discusses the subtleties associated with approximations. The qualitative features of $\phi_v(\chi, s)$ are shown in **Figure 2c**.

The product $\phi_v(\chi, s)$ describes the distribution of backgrounds on which successful mutations arise. This distribution is often narrowly peaked right below the high-fitness nose (**Figure 2c**). Mutations that arise in individuals with background fitness below this peak are doomed, whereas very few individuals have fitness above the peak. The greater s is, the broader is this region in which successful mutations arise.

To determine the rate of adaptation, one has to substitute the results for $\Phi_v(s)$ into Equation 3 and solve for v (Desai & Fisher 2007, Good et al. 2012). A general consequence of the form of the self-consistency condition in Equation 3 is that if $\Phi_v(s)$ is weakly dependent on v , then v is proportional to N . In this case, the speed of evolution is proportional to the mutational input. With increasing fitness variance, σ^2 , the genetic background fitness starts to influence fixation probabilities, such that v eventually increases only slowly with N . For models in which beneficial mutations of fixed effect s arise at rate U_b , the rate of adaptation in large populations is given by

$$v \propto \begin{cases} s^2 \frac{\log N s}{(\log U_b/s)^2} & s \gg U_b \\ (U_b s^2)^{2/3} (\log N D^{1/3})^{1/3} & s \ll U_b \end{cases} \quad 4.$$

Figure 2

Adaptation in asexual populations. (a) If the distribution of beneficial mutations has a long tail, then the population consists of a few large clones, and only the mutations with the largest effects have a chance of fixing. (b) If many mutations of similar effect contribute to fitness diversity, then the bulk of the fitness distribution, $n(\chi)$, can be described by a smooth function that is roughly Gaussian in shape. There exists a fittest genotype in the population with no individuals (blue circles) to its right. Only mutations close to this high-fitness nose have an appreciable chance of fixing. The stochastic dynamics at the nose determines the evolution of the entire population, and the speed of the entire population, v , has to match the speed of the nose, v_{nose} , in a quasi-steady state. The fixation probability of a mutation with effect s , $\phi_v(\chi, s)$, increases with increasing background fitness, χ , as shown in panel c. A mutant in the bulk of the fitness distribution has essentially zero chance of taking over the population because many fitter individuals exist. In the opposite case, when the mutant is the fittest in the population, $\phi_v(\chi, s)$ is proportional to $\chi + s$, as we would expect in the absence of interference. Because very few individuals with very high fitness exist, most mutations that fix come from a narrow region (medium gray), where the product of $n(\chi)$ and $\phi_v(\chi, s)$ (blue line) peaks. Note that χ is Malthusian or log fitness. Scripts to illustrate interference and fixation can be found at <http://webdav.tuebingen.mpg.de/interference>.

Branching process: stochastic model of reproducing and dying individuals without a constraint on the overall population size

(Cohen et al. 2005a, Desai & Fisher 2007). Equation 4 assumes that s is constant, but these expressions hold for more general models with a short-tailed distribution $U(s)$ with suitably defined and effective U_b and s (Good et al. 2012).

2.3. Synthesis

Clonal interference and multiple-mutation models both predict diminishing returns as the population increases, but the underlying dynamics is rather different. In the clonal interference picture, population takeovers are driven by single mutations, and the genetic background on which they occur is largely irrelevant [$\phi_v(\chi, s)$ depends little on χ]. The mutations that are successful, however, have the largest effects. In the multiple-mutation regime, the effect of the mutations is not that crucial, but the mutations have to occur in very fit individuals to be successful [$\phi_v(\chi, s)$ increases rapidly with χ]. In both models, the speed of adaptation continues to increase slowly with the population size, and no hard speed limit exists. Distinguishing a speed limit from diminishing returns in experiments is hard (de Visser et al. 1999, Miralles et al. 2000).

Whether one picture or the other is more appropriate depends on the distribution of fitness effects of de novo mutations, $U(s)$. If $U(s)$ falls off faster than exponentially, then adaptation occurs via many small steps (Desai & Fisher 2007, Good et al. 2012); if the distribution is broader, then the clonal interference picture is a reasonable approximation (Park & Krug 2007, Park et al. 2010). Investigators have examined the borderline case of an exponential fitness distribution more closely, finding that large-effect mutations on a pretty good background make the dominant contributions (Good et al. 2012, Schiffels et al. 2011); i.e., hallmarks of both models are combined in this borderline case.

Empirical observations favor this intermediate situation. Strelkowa & Lässig (2012) have analyzed influenza evolution in great detail and found that a few rather than a single mutation drive the fixation of a particular strain. Similarly, evolution experiments suggest that the genetic background is important, but a moderate number of large-effect mutations account for most of the observed adaptation (Lang et al. 2011).

Note the somewhat unintuitive dependence of v on parameters in Equation 4. Instead of depending on the mutational input NU_b and the fitness effect s , v depends on Ns and U_b/s for $U_b \ll s$. In large populations, the dominant timescale of population turnover is governed by selection and is of order s^{-1} . Ns and U_b/s measure the strength of reproduction noise (drift) and mutations relative to s^{-1} , respectively [see Neher & Shraiman (2012) for a discussion of this issue in the context of deleterious mutations]. In large populations, the infinite-sites model starts to break down, and the same mutations can occur independently in several lineages, limiting interference (Bollback & Huelsenbeck 2007, Kim & Orr 2005).

3. EVOLUTION OF FACULTATIVELY SEXUAL POPULATIONS

Competition between beneficial mutations in asexuals results in a slow (logarithmic) growth of the adaptation speed with the population size, N (Equation 4). How does gradually increasing the outcrossing rate alleviate this competition? The associated advantages of sex and recombination have been studied extensively (Charlesworth 1993, Crow & Kimura 1965, Fisher 1930, Muller 1932, Rice & Chippindale 2001). It is instructive to consider facultatively sexual organisms that outcross at rate r and that have many independently segregating loci in the event of outcrossing. Facultatively sexual species are common among RNA viruses, yeasts, nematodes, and plants.

Most of our theoretical understanding of evolution in large facultatively mating populations comes from models similar to those introduced above for asexual populations. In addition to

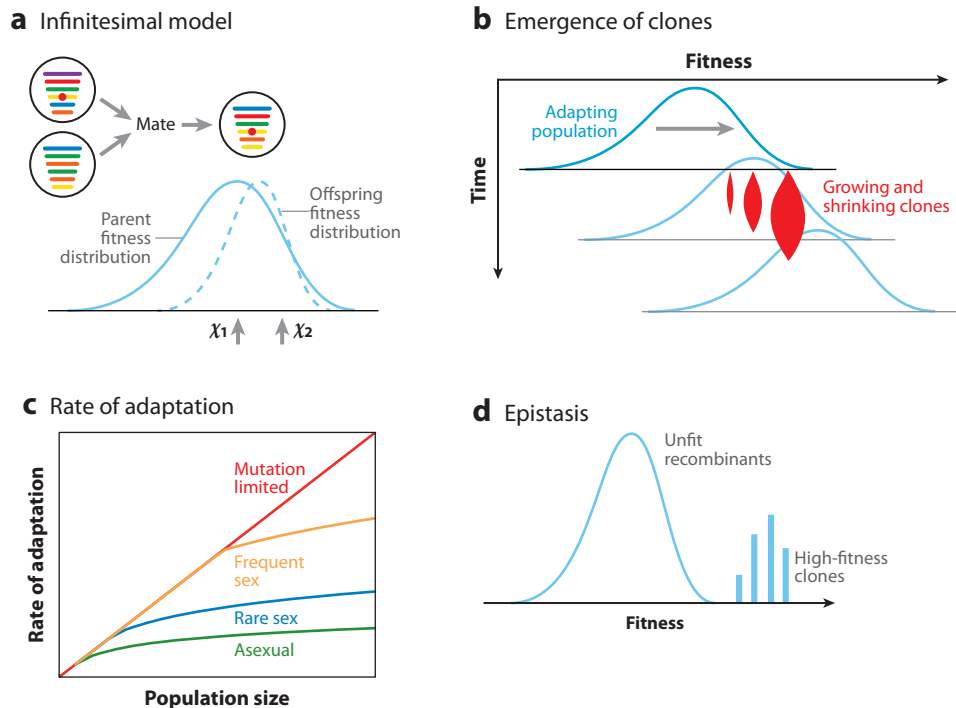


Figure 3

A facultatively sexual life cycle is common among many pathogens, plants, and some groups of animals. (a) If many loci segregate independently, then the infinitesimal model accurately describes the fitness distribution of offspring. Given two parents with fitness χ_1 and χ_2 sampled from the parental distribution with variance σ^2 , offspring fitness is symmetrically distributed around the parental mean with variance $\sigma^2/2$. A mutation (red dot) can thereby hop from an individual with one background fitness to an individual with a very different one. (b) If the outcrossing rate is lower than the fitness of some individuals, then clones (red) can grow at rate $\chi - r$, where r is the outcrossing rate. As the population adapts, the growth rate of the clones decreases and eventually becomes negative, and the clone disappears. The beneficial mutation, however, persists on other backgrounds. (c) In small populations, the rate of adaptation increases linearly with the population size. For each outcrossing rate, there is a point beyond which interference starts to be important. (d) Epistasis causes condensation of the population into a few very fit genotypes. Crosses between these genotypes result in unfit individuals. In the absence of forces that stabilize different clones, one clone will rapidly take over if $\chi > r$. Scripts illustrating the evolution of facultatively sexual populations can be found at <http://webdav.tuebingen.mpg.de/interference>.

mutation, we have to introduce a term that describes how an allele can move from one genetic background to another by recombination (Figure 3a). Given the fitness values of the two parents, χ_1 and χ_2 , and assuming many independently segregating loci, the offspring fitness, χ , is symmetrically distributed around the mid-parent value with half the population variance (Figure 3a) (Bulmer 1980, Turelli & Barton 1994). To understand the process of fixation in such a population, the following intuition is useful: An outcrossing event places a beneficial mutation onto a novel genotype, which is amplified by selection into a clone whose size grows rapidly with the fitness of the founder (Figure 3b). These clones are transient because even an initially fit clone will fall behind the increasing mean fitness. But large clones produce many recombinant offspring (daughter clones), thereby greatly enhancing the chance of fixing the mutations they

Epistasis: background dependence of the effect of mutations; can result in rugged fitness landscapes

Selective interference: reduction of fixation probability through competition with other beneficial alleles

carry. Because clone size increases rapidly with founder fitness, the fixation probability, $\phi_v(\chi, s)$, is still a very steep function of the background fitness and qualitatively similar to that of the asexual case (**Figure 2c**). With an increasing outcrossing rate, the fitness window from which successful clones originate becomes broader and broader.

If outcrossing rates are fast enough that recombination disassembles genotypes faster than selection can amplify them, then $\phi_v(\chi, s)$ is essentially flat, and the genetic background does not matter much. Neher et al. (2010) examined this transition:

$$v \approx \begin{cases} \frac{2r^2 \log(NU_b)}{(\log r/s)^2} & r \ll \sqrt{NU_b s^2} \\ NU_b s^2 & r \geq \sqrt{NU_b s^2} \end{cases} \quad 5.$$

The essence of this result is that recombination limits adaptation whenever r is smaller than the standard deviation in fitness in the absence of interference. In this regime, v depends weakly on N but increases rapidly with r (**Figure 3c**). Similar results were reported by Weissman & Barton (2012). The above analysis assumes that recombination is rare but still frequent enough to ensure that mutations that rise to high frequencies are essentially in linkage equilibrium. This requires that $r \gg s$. Rouzine & Coffin (2005, 2010) studied selection on standing variation at intermediate and slow recombination rates. Cohen et al. (2005b), Wylie et al. (2010), and Neher et al. (2010) investigated adaptation in the presence of horizontal gene transfer.

In contrast to asexual evolution, epistasis can dramatically affect the evolutionary dynamics in sexual populations. Epistasis implies that the effect of mutations depends on the state at other loci in the genome. In the absence of sex, the only quantity that matters is the distribution of fitness effects of available mutations, $U(s)$. The precise nature of epistasis is not crucial. In sexual populations, however, epistasis can affect the evolutionary dynamics dramatically: When different individuals mix their genomes, it matters whether mutations acquired in different lineages are compatible. Because selection favors well-adapted combinations of alleles, recombination is expected to be, on average, disruptive, and recombinant offspring have, on average, lower fitness than their parents (the so-called recombination load). This competition between selection for good genotypes and recombination can result in a condensation of the population into fit clones (**Figure 3d**) (Neher & Shraiman 2009, Neher et al. 2013).

4. SELECTIVE INTERFERENCE IN OBLIGATELY SEXUAL ORGANISMS

Selective interference has historically received the most attention in the obligately sexual organisms most relevant to crop and animal breeding. Farmers and breeders have performed artificial selection for thousands of years with remarkable success (Hill & Kirkpatrick 2010). Evolution experiments with diverse species, including chicken, mice, and *Drosophila*, have shown that standing variation at many loci responds to diverse selection pressures (Burke et al. 2010, Chan et al. 2012, Johansson et al. 2010, Turner et al. 2011, Zhou et al. 2011; see Burke 2012 for a recent review). In obligately sexual populations, distant loci can respond independently to selection and remain in approximate linkage equilibrium. The frequencies of different alleles change according to their effect on fitness averaged over all possible background fitnesses in the population. Small deviations from linkage equilibrium can be accounted for perturbatively using the so-called quasi-linkage equilibrium approximation (Barton & Turelli 1991, Kimura 1965, Neher & Shraiman 2011b).

This approximate independence, however, does not hold for loci that are tightly linked. Hill & Robertson (1966) observed that interference between linked competing loci can slow down the response to selection—an effect now termed Hill-Robertson interference (Felsenstein 1974).

Felsenstein realized that interference is not restricted to competing beneficial mutations; linked deleterious mutations also impede fixation of beneficial mutations (see Section 5.4). The term Hill-Robertson interference is now used for any reduction in the efficacy of selection caused by linked fitness variation. In the 1990s, Barton (1994, 1995b) provided a deeper understanding of selective interference. His key insight was to calculate the fate of a novel mutation, considering all possible genetic backgrounds on which it can arise and summing over all possible trajectories it can take through the population. For a few loci, the equations describing the probability of fixation can be integrated explicitly.

Weakly linked sweeps cause a cumulative reduction of the fixation probability at a focal site that depends on the ratio of additive variance in fitness and the squared degree of linkage (Barton 1995b, Santiago & Caballero 1998). Barton (1994) further identified a critical rate of strong selective sweeps that effectively prevents the fixation of mutations with an advantage smaller than s_c . If sweeps are too frequent, the weakly selected mutation has little chance of spreading before the next strong sweep reduces its frequency again.

At short distances, selective sweeps impede each other's fixation more strongly. This interference is limited to a time interval of the order of s^{-1} generations in which one of the sweeping mutations is at intermediate frequencies. During this time, a new beneficial mutation will often fall onto the wild-type background and become lost again if it is not rapidly recombined onto the competing sweep. This rescue by recombination is likely only if it is farther than s/ρ nucleotides away from the competing sweep, where ρ is the crossover rate per base pair (Barton 1994). In other words, a sweeping mutation with effect s prevents other sweeps in a region of width s/ρ and occupies this chromosomal real estate for a time s^{-1} (**Figure 4a**) (Weissman & Barton 2012). Hence, strong sweeps briefly interfere with other sweeps in a large region, whereas weak sweeps affect a narrow region for a longer period of time. The amount of interference is therefore roughly independent of the strength of the sweeps, and the total number of sweeps per unit time is limited by the map length, $R = \int \rho(y) dy$, where the integral is over the entire genome and $\rho(y)$ is the crossover rate at position y . Larger populations can squeeze slightly more sweeps into R (Weissman & Barton 2012). In most obligately sexual organisms, sweeps rarely cover more than a few percent of the total map length, such that recombination does not limit adaptation unless sweeps cluster in certain regions (Sella et al. 2009). However, as I discuss below, even rare selective sweeps have dramatic effects on neutral diversity.

5. GENETIC DIVERSITY, DRAFT, AND COALESCENCE

Interference between selected mutations reduces the fixation probability of beneficial mutations, slows adaptation, and weakens purifying selection. These effects are very important but hard to observe because significant adaptation often takes longer than our window of observation. Typically, data consist of a sample of sequences from a population. These sequences differ by single-nucleotide polymorphisms, insertions, or deletions, and we rarely know the effect of these differences on the organism's fitness.

From a sequence sample of this sort, the genealogy of the population is reconstructed and compared with models of evolution—in most cases, a neutral model governed by Kingman's coalescent (Kingman 1982). From this comparison, we hope to learn about evolutionary processes. But linked selection, be it in asexual, facultatively sexual, or obligately sexual organisms, has dramatic effects on the genealogies. Substantial effects on neutral diversity are observed at rates of sweeps that have not yet caused strong interference between selected loci for the simple reason that neutral alleles segregate for longer periods of time (Weissman & Barton 2012).

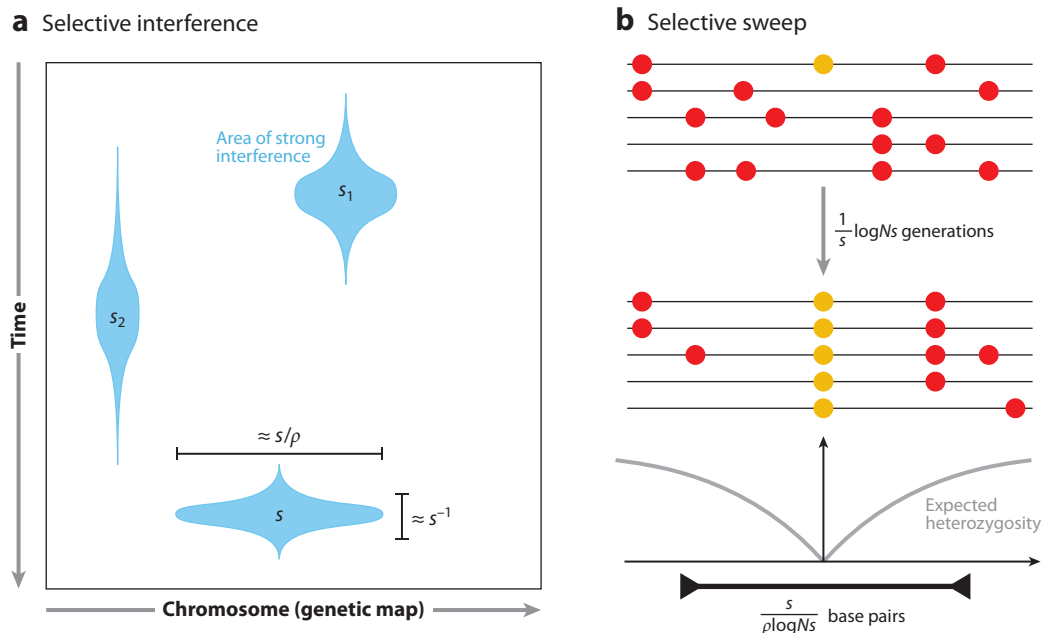


Figure 4

Interference in obligately sexual populations. (a) The interference effects of selective sweeps through time (*vertical axis*) and along the genome (*horizontal axis*). A sweeping mutation with fitness effect s interferes with other mutations in a region of width s/ρ over a time s^{-1} , where ρ is the crossover rate per base. The extent of interference is indicated by blue bulges, each of which corresponds to a mutation that fixed. Interference starts to be important when the bulges overlap. Because the area of the bulges, roughly height \times width, is approximately independent of s , interference depends on ρ and the rate of sweeps rather than the effect size. The rate of adaptation is therefore primarily a function of the map length, R . (b) A selective sweep of a beneficial mutation (yellow circle) reduces neutral genetic variation (red circles) in a region of width $s / (\rho \log Ns)$. The effect of sweeps on neutral diversity is explored at <http://webdav.tuebingen.mpg.de/interference>.

5.1. Genetic Draft in Obligately Sexual Populations

Selective sweeps have strong effects on linked neutral diversity and genealogies (Barton 1998, Barton & Etheridge 2004, Kaplan et al. 1989, Maynard Smith & Haigh 1974, Stephan et al. 1992, Wiehe & Stephan 1993). A sweeping mutation takes approximately $t_{\text{sw}} \approx s^{-1} \log Ns$ generations to rise to high frequency. Linked neutral variation is preserved only if substantial recombination happens during this time. Given a crossover rate of ρ per base, recombination will separate the sweep from a locus at distance l with probability $r = \rho l$ per generation (assuming $r \ll 1$). Hence, a sweep leaves a dip of width $l = (\rho t_{\text{sw}})^{-1} \approx s / (\rho \log Ns)$ in the neutral diversity (Figure 4b). Within this region, selection causes massive and rapid coalescence, and only a fraction of the lineages continue into the ancestral population (Figure 5a). Durrett & Schweinsberg (2005) have further investigated this phenomenon and showed that the effect of recurrent selective sweeps is well approximated by a coalescent process that allows for multiple mergers (Berestycki 2009, Pitman 1999): Each sweep forces the almost simultaneous coalescence of many lineages (a fraction $e^{-rt_{\text{sw}}}$). Gillespie (2000) had made similar arguments previously, calling the stochastic force responsible for coalescence genetic draft. Coop & Ralph (2012) extended the analysis of Durrett & Schweinsberg to partial sweeps that could be common in structured populations, at loci with overdominance, or in scenarios with frequency-dependent selection.

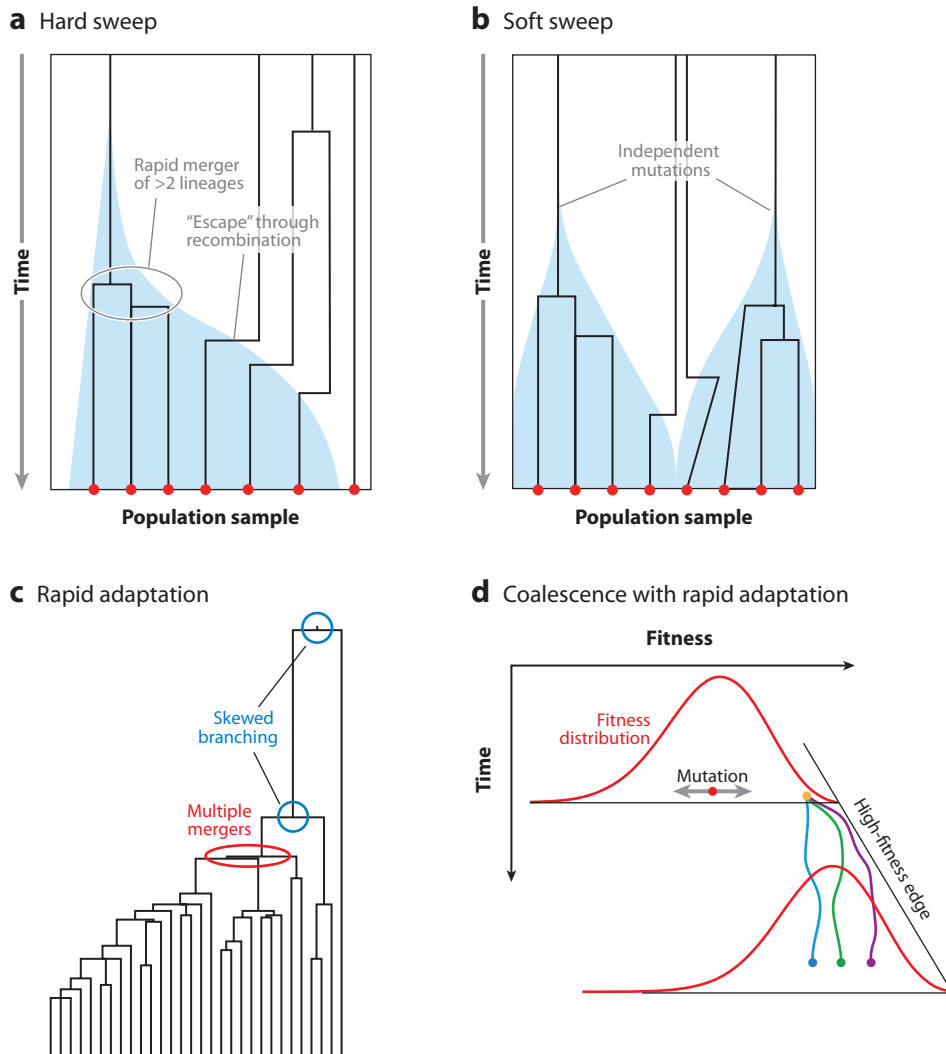


Figure 5

Coalescence driven by selection. (a) A selective sweep (*blue region*) causes rapid coalescence of lineages at a nearby locus. Each sweep causes a fraction of lineages to merge, whereas the remaining lineages recombine onto an ancestral background. (b) Soft sweep refers to a scenario in which single mutations arise multiple times independently in response to environmental change. This scenario is expected as soon as the product of N (the population size) and the per-site mutation rate exceeds one and can result in multiple bursts of coalescence at almost the same time. (c) A genealogical tree drawn from a simulation of a model of rapidly adapting asexual organisms. Coalescence often occurs in bursts. Furthermore, branching is often uneven. Skewed branching and bursts of coalescence are well-known features of multiple-merger coalescent processes such as the Bolthausen-Sznitman coalescent. (d) Coalescence and fitness classes. Most population samples consist of individuals from the center of the fitness distribution, even though their distant ancestors were among the fittest. In large populations, most coalescence happens in the high-fitness nose, and the length of time until ancestral lineages arrive in the nose corresponds to long terminal branches (compare with panel c). We can study how genealogies depend on selection by using simulations; see <http://webdav.tuebingen.mpg.de/interference>. Panels c and d adapted from Neher & Hallatschek 2013.

Bolthausen-Sznitman coalescent (BSC): a special multiple-merger coalescent that approximates genealogies in many models of adaptation

The rapid coalescence of multiple lineages is unexpected in the standard neutral coalescent (a merger of p lineages occurs with probability $\propto N^{-p}$). In coalescence induced by a selective sweep, however, multiple mergers are common and dramatically change the statistical properties of genealogies. A burst of coalescence corresponds to a portion of the tree with an almost star-like shape (Slatkin & Hudson 1991). Alleles that arose before the burst are common, whereas those that arose after the burst are rare. Hence, such bursts cause a relative increase of rare alleles as well as of alleles very close to fixation (Braverman et al. 1995, Fay & Wu 2000, Gillespie 2000).

The degree to which linked selective sweeps reduce genetic diversity depends primarily on the rate of sweeps per map length (Weissman & Barton 2012). In accord with this expectation, diversity has been found to increase with the recombination rate and decrease with the density of functional sites (Begun et al. 2007, Shapiro et al. 2007). In addition to occasional selective sweeps, genetic diversity and the degree of adaptation can be strongly affected by many weakly selected sites (e.g., weakly deleterious mutations) that generate a broad fitness distribution (McVean & Charlesworth 2000).

5.2. Soft Sweeps

Soft sweeps refer to selective sweeps that originate from multiple genomic backgrounds (Hermisson & Pennings 2005, Pennings and Hermisson 2006), either because the favored allele arose independently multiple times or because it had been segregating for a long time prior to environmental change. Soft sweeps have recently been observed in the pesticide resistance of *Drosophila* (Karasov et al. 2010) and are a common phenomenon in viruses with high mutation rates.

A genealogy of individuals sampled after a soft sweep is illustrated in **Figure 5b**. Most of the individuals trace back to one of two or more ancestral haplotypes on which the selected mutation arose. Hence, coalescence is again dominated by multiple-merger events, except several of those events happen almost simultaneously. Schweinsberg (2000) has described this type of coalescent process.

Despite dramatic effects on genealogies, soft sweeps can be difficult to detect by standard methods that scan for selective sweeps. Those methods use local reductions in genetic diversity, which can be modest if the population traces back to several ancestral haplotypes. The number of ancestral haplotypes in a sample after a soft sweep depends on the product of N and the per-site mutation rate, μ , and selection against the allele before the sweep (Pennings & Hermisson 2006). Detecting soft sweeps requires methods that explicitly search for signatures of rapid coalescence into several lineages in linkage disequilibrium data or haplotype patterns (Messer & Neher 2012, Pennings & Hermisson 2006).

5.3. The Bolthausen-Sznitman Coalescent and Rapidly Adapting Populations

Individual selective sweeps have an intuitive effect on genetic diversity, but what do genealogies look like when many mutations are competing in asexual or facultatively sexual populations? Brunet et al. (2007) studied a model in which a population expands its range and found that the genealogies of individuals at the front are described by a special type of multiple-merger coalescent known as the Bolthausen-Sznitman coalescent (BSC) (for a review, see Berestycki 2009). More recently, Neher & Hallatschek (2013) showed that a modified BSC describes the genealogies in a model of adaptation in a panmictic population. Whereas Neher & Hallatschek considered a limit of high mutation rates, Desai et al. (2013) showed that similar conclusions hold with a limit of low mutation rates.

Figure 5c shows a tree sampled from a model of a rapidly adapting population. A typical sample from a rapidly adapting population consists of individuals from the center of the fitness distribution (**Figure 5d**). Their ancestors tend to have been among the fittest in the population

(Hermisson et al. 2002, Rouzine & Coffin 2007). Substantial coalescence happens only once the ancestral lineages have reached the high-fitness tip, resulting in long terminal branches of the trees. Once at the tip, coalescence is driven by the competition of lineages against one another and happens in bursts whenever one lineage jumps ahead of all others. These bursts correspond to a large fraction of the population descending from one particular individual. These coalescent events have approximately the same statistics as neutral coalescent processes with very broad but nonheritable offspring distributions (Der et al. 2011, Eldon & Wakeley 2006, Schweinsberg 2003).

In the case of rapidly adapting asexual populations, the effective distribution of the number of offspring, n , is given by $P(n) \sim n^{-2}$, which gives rise to the BSC. This type of distribution seems to be universal to populations in which individual lineages are amplified while they diversify and is found in facultatively sexual populations (Neher & Shraiman 2011a), asexual populations adapting by small steps, and populations in a dynamic balance between deleterious and beneficial mutations. Asymptotic features of the site frequency spectrum can be derived analytically (Berestycki 2009, Desai et al. 2013, Neher & Hallatschek 2013). One finds that the frequency spectrum diverges when $f(v) \sim v^{-2}$ at low frequencies, resulting in many singletons. Furthermore, neutral alleles close to fixation are common, with $f(v)$ diverging again as $v \rightarrow 1$. This excess of rare and very common alleles relative to intermediate frequencies is a consequence of multiple mergers, which produce star-like subtrees and the very asymmetric branching at nodes deep in the tree (compare with **Figure 5c**).

The timescale of coalescence and, with it, the level of genetic diversity are mostly determined by the strength of selection and only weakly increase with population size. Essentially, the average time to a common ancestor of two randomly chosen individuals is given by the time it takes until the fittest individuals dominate the population. In most models, this time depends only logarithmically on the population size, N .

5.4. Background Selection and Genetic Diversity

Background selection refers to the effect of purifying selection on linked loci, which is particularly important if linked regions are long. If deleterious mutations incur a fitness decrement of s and arise with a genome-wide rate of U_d , then a sufficiently large population settles in a state in which the number of mutations in individuals follows a Poisson distribution with a mean of $\lambda = U_d/s$ (Haigh 1978). Individuals loaded with many mutations are selected against but are continually produced by de novo mutations. All individuals in the population ultimately descend from individuals carrying the least deleterious mutations. Within this model, the least loaded class has size $N \exp(-U_d/s)$, and coalescence in this class is accelerated by $\exp(U_d/s)$, compared with a neutrally evolving population of size N (Charlesworth et al. 1993). For large ratios of U_d/s , the Poisson distribution of background fitness spans many fitness classes, and this heterogeneity substantially reduces the efficacy of selection (McVean & Charlesworth 2000).

The effect of background selection is best appreciated in a genealogical picture. Genetic backgrounds sampled from the population tend to come from the center of the distribution. Because the deleterious mutations they carry accumulated in the recent past, lineages shed mutations as we trace them back in time until they arrive in the mutation-free class akin to **Figure 5d**. This resulting genealogical process, a fitness-class coalescent, has been described by Walczak et al. (2012). A recent study on the genetic diversity of whale lice (Seger et al. 2010) suggests that purifying selection and frequent deleterious mutations can severely distort genealogies. O'Fallon et al. (2010) present methods for the analysis of sequence samples under purifying selection.

The fitness-class coalescent is appropriate as long as U_d is low enough that Muller's ratchet does not click frequently. More generally, fixation of deleterious mutations, adaptation, and

environmental change will approximately balance out. A small fraction of beneficial mutations can be sufficient to halt Muller's ratchet (Goyal et al. 2012). In this dynamic balance between frequent deleterious and rare beneficial mutations, the genealogies tend to be similar to genealogies under rapid adaptation, as discussed in Section 5.3.

6. CONCLUSIONS AND FUTURE DIRECTIONS

Contradicting neutral theory, genetic diversity correlates only weakly with population size (Leffler et al. 2012), suggesting that linked selection or genetic draft are more important than conventional genetic drift. Draft is most severe in asexual populations, for which models predict that the fitness differences rather than the population size determine the level of neutral diversity. As outcrossing becomes more frequent, the strength of draft decreases, and diversity increases. With increasing coalescence times, selection becomes more efficient, as more time is available to differentiate deleterious from beneficial alleles. In obligately sexual populations, most interference is restricted to tightly linked loci, and the number of sweeps per map length and generation determines genetic diversity.

Because interference slows adaptation, adaptation is expected to select for higher recombination rates (Charlesworth 1993). Indeed, positive selection results in indirect selection on recombination modifiers (Barton 1995a, Barton & Otto 2005, Hartfield et al. 2010, Otto & Barton 1997). Changing frequencies of outcrossing have been observed in evolution experiments (Becks & Agrawal 2010), but the evolution of recombination and outcrossing rates in rapidly adapting populations remains poorly understood, both theoretically and empirically.

The traveling wave models discussed above assume many polymorphisms with similar effects on fitness and a smooth fitness distribution, which are drastic idealizations. More typically, one finds a handful of polymorphisms with a distribution of effects (Barrick et al. 2009, Lang et al. 2011, Strelkowa & Lässig 2012). Simulations indicate, however, that statistical properties of genealogies are rather robust regarding model assumptions, as long as draft dominates over drift (Neher & Hallatschek 2013). Appropriate genealogical models are prerequisite for demographic inference. If, for example, a neutral coalescent model is used to infer the population size history of a rapidly adapting population, then one would conclude that the population has been expanding. Incidentally, this conclusion is inferred in most cases. Investigators have recently made progress toward incorporating the effect of purifying selection into estimates from reconstructed genealogies (Nicolaisen & Desai 2012, O'Fallon 2011). Alternative genealogical models accounting for selection should be included in popular analysis programs such as BEAST (Drummond & Rambaut 2007).

Assigning an effective size, N_e , to various populations is still common. In most cases, N_e is a proxy for genetic diversity, which depends on the time to the most recent common ancestor. Given that coalescence times depend on linked selection and genetic draft, rather than the population size and genetic drift, the term should be avoided and replaced by T_c , the timescale of coalescence. Defining N_e suggests that the neutral model is valid as long as N_e is used instead of N . We have seen multiple times that drift and draft are of rather different natures and that this difference cannot be captured by simple rescaling: Each quantity then requires its own private N_e , rendering the concept essentially useless. Some quantities, such as site frequency spectra, are qualitatively different, and no N_e maps them to a neutral model. The (census) population size is nevertheless important in discovering beneficial mutations. For this reason, large populations are expected to respond more quickly to environmental change, as we are painfully aware in the case of antibiotic resistance of pathogens. Large populations might therefore track phenotypic optima more closely, resulting in beneficial mutations with smaller effect, which in turn might explain their greater diversity.

Most models discussed assume a time-invariant fitness landscape. This assumption reflects our ignorance regarding the degree and timescale of environmental fluctuations [for work on selection in time-dependent fitness landscapes, see Mustonen & Lässig (2009)]. Time-variable selection pressures, combined with spatial variation, could potentially have strong effects. Similarly, frequency-dependent selection and, more generally, the interaction of evolution with ecology are important avenues for future work. The challenge consists of choosing useful models that are tractable, appropriate, and predictive.

DISCLOSURE STATEMENT

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

ACKNOWLEDGMENTS

I would like to thank Boris Shraiman, Sally Otto, Frank Chan, Philipp Messer, Aleksandra Walczak, and Michael Desai for stimulating discussions and providing helpful comments on the manuscript. This work is supported by the European Research Council starting grant HIVEVO 260686.

LITERATURE CITED

- Barrick JE, Yu DS, Yoon SH, Jeong H, Oh TK, et al. 2009. Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. *Nature* 461(7268):1243–47
- Barton NH. 1994. The reduction in fixation probability caused by substitutions at linked loci. *Genet. Res.* 64(3):199–208
- Barton NH. 1995a. A general model for the evolution of recombination. *Genet. Res.* 65(2):123–44
- Barton NH. 1995b. Linkage and the limits to natural selection. *Genetics* 140(2):821–41
- Barton NH. 1998. The effect of hitch-hiking on neutral genealogies. *Genet. Res.* 72(2):123–33
- Barton NH, Etheridge AM. 2004. The effect of selection on genealogies. *Genetics* 166(2):1115–31
- Barton NH, Otto SP. 2005. Evolution of recombination due to random drift. *Genetics* 169(4):2353–70
- Barton NH, Turelli M. 1991. Natural and sexual selection on many loci. *Genetics* 127(1):229–55
- Becks L, Agrawal AF. 2010. Higher rates of sex evolve in spatially heterogeneous environments. *Nature* 468(7320):89–92
- Begun DJ, Holloway AK, Stevens K, Hillier LW, Poh Y-P, et al. 2007. Population genomics: whole-genome analysis of polymorphism and divergence in *Drosophila simulans*. *PLoS Biol.* 5(11):e310
- Berestycki N. 2009. Recent progress in coalescent theory. *Ens. Mat.* 16:1–193
- Bhatt S, Holmes EC, Pybus OG. 2011. The genomic rate of molecular adaptation of the human influenza A virus. *Mol. Biol. Evol.* 28(9):2443–51
- Bollback JP, Huelsenbeck JP. 2007. Clonal interference is alleviated by high mutation rates in large populations. *Mol. Biol. Evol.* 24(6):1397–406
- Braverman JM, Hudson RR, Kaplan NL, Langley CH, Stephan W. 1995. The hitchhiking effect on the site frequency spectrum of DNA polymorphisms. *Genetics* 140(2):783–96
- Brunet E, Derrida B, Mueller AH, Munier S. 2007. Effect of selection on ancestry: an exactly soluble case and its phenomenological generalization. *Phys. Rev. E* 76(4):041104
- Brunet E, Rouzine I, Wilke C. 2008. The stochastic edge in adaptive evolution. *Genetics* 179(1):603–20
- Bulmer MG. 1980. *The Mathematical Theory of Quantitative Genetics*. Oxford: Oxford Univ. Press
- Burke MK. 2012. How does adaptation sweep through the genome? Insights from long-term selection experiments. *Proc. R. Soc. B* 279(1749):5029–38
- Burke MK, Dunham JP, Shahrestani P, Thornton KR, Rose MR, Long AD. 2010. Genome-wide analysis of a long-term evolution experiment with *Drosophila*. *Nature* 467(7315):587–90

- Chan YF, Jones FC, McConnell E, Bryk J, Bünger L, Tautz D. 2012. Parallel selection mapping using artificially selected mice reveals body weight control loci. *Curr. Biol.* 22(9):794–800
- Charlesworth B. 1993. Directional selection and the evolution of sex and recombination. *Genet. Res.* 61(3):205–24
- Charlesworth B. 2012. The effects of deleterious mutations on evolution at linked sites. *Genetics* 190:5–22
- Charlesworth B, Morgan MT, Charlesworth D. 1993. The effect of deleterious mutations on neutral molecular variation. *Genetics* 134(4):1289–303
- Cohen E, Kessler DA, Levine H. 2005a. Front propagation up a reaction rate gradient. *Phys. Rev. E* 72(6):066126
- Cohen E, Kessler DA, Levine H. 2005b. Recombination dramatically speeds up evolution of finite populations. *Phys. Rev. Lett.* 94(9):098102
- Coop G, Ralph PL. 2012. Patterns of neutral diversity under general models of selective sweeps. *Genetics* 192(1):205–24
- Crow J, Kimura M. 1965. Evolution in sexual and asexual populations. *Am. Nat.* 99(909):439–50
- de Visser JAGM, Zeyl CW, Gerrish PJ, Blanchard JL, Lenski RE. 1999. Diminishing returns from mutation supply rate in asexual populations. *Science* 283(5400):404–6
- Der R, Epstein CL, Plotkin JB. 2011. Generalized population models and the nature of genetic drift. *Theor. Popul. Biol.* 80(2):80–99
- Desai MM, Fisher DS. 2007. Beneficial mutation—selection balance and the effect of linkage on positive selection. *Genetics* 176(3):1759–98
- Desai MM, Fisher DS, Murray AW. 2007. The speed of evolution and maintenance of variation in asexual populations. *Curr. Biol.* 17(5):385–94
- Desai MM, Walczak AM, Fisher DS. 2013. Genetic diversity and the structure of genealogies in rapidly adapting populations. *Genetics* 193(2):565–85
- Drummond A, Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* 7(1):214
- Durrett R, Schweinsberg J. 2005. A coalescent model for the effect of advantageous mutations on the genealogy of a population. *Stoch. Process. Appl.* 115(10):1628–57
- Eldon B, Wakeley J. 2006. Coalescent processes when the distribution of offspring number among individuals is highly skewed. *Genetics* 172(4):2621–33
- Fay JC, Wu CI. 2000. Hitchhiking under positive Darwinian selection. *Genetics* 155(3):1405–13
- Felsenstein J. 1974. The evolutionary advantage of recombination. *Genetics* 78(2):737–56
- Fisher DS. 2013. Asexual evolution waves: fluctuations and universality. *J. Stat. Mech.* 2013(1):P01011
- Fisher RA. 1930. *The Genetical Theory of Natural Selection*. Oxford: Clarendon
- Gerrish PJ, Lenski RE. 1998. The fate of competing beneficial mutations in an asexual population. *Genetica* 102–103(1–6):127–44
- Gillespie JH. 2000. Genetic drift in an infinite population: the pseudohitchhiking model. *Genetics* 155(2):909–19
- Good BH, Rouzine IM, Balick DJ, Hallatschek O, Desai MM. 2012. Distribution of fixed beneficial mutations and the rate of adaptation in asexual populations. *Proc. Natl. Acad. Sci. USA* 109(13):4950–55
- Goyal S, Balick DJ, Jerison ER, Neher RA, Shraiman BI, Desai MM. 2012. Dynamic mutation—selection balance as an evolutionary attractor. *Genetics* 191:1309–19
- Haigh J. 1978. The accumulation of deleterious genes in a population—Muller’s ratchet. *Theor. Popul. Biol.* 14(2):251–67
- Hallatschek O. 2011. The noisy edge of traveling waves. *Proc. Natl. Acad. Sci. USA* 108(5):1783–87
- Hartfield M, Otto S, Keightley P. 2010. The role of advantageous mutations in enhancing the evolution of a recombination modifier. *Genetics* 184(4):1153–64
- Hegreness M, Shores N, Hartl D, Kishony R. 2006. An equivalence principle for the incorporation of favorable mutations in asexual populations. *Science* 311(5767):1615–17
- Hermisson J, Pennings P. 2005. Molecular population genetics of adaptation from standing genetic variation. *Genetics* 169(4):2335–52
- Hermisson J, Redner O, Wagner H, Baake E. 2002. Mutation-selection balance: ancestry, load, and maximum principle. *Theor. Popul. Biol.* 62(1):9–46

- Hill WG, Kirkpatrick M. 2010. What animal breeding has taught us about evolution. *Annu. Rev. Ecol. Evol. Syst.* 41:1–19
- Hill WG, Robertson A. 1966. The effect of linkage on limits to artificial selection. *Genet. Res.* 8(3):269–94
- Johansson AM, Pettersson ME, Siegel PB, Carlberg O. 2010. Genome-wide effects of long-term divergent selection. *PLoS Genet.* 6(11):e1001188
- Kao KC, Sherlock G. 2008. Molecular characterization of clonal interference during adaptive evolution in asexual populations of *Saccharomyces cerevisiae*. *Nat. Genet.* 40(12):1499–504
- Kaplan NL, Hudson RR, Langley CH. 1989. The “hitchhiking effect” revisited. *Genetics* 123(4):887–99
- Karasov T, Messer PW, Petrov DA. 2010. Evidence that adaptation in *Drosophila* is not limited by mutation at single sites. *PLoS Genet.* 6(6):1000924
- Kawecki TJ, Lenski RE, Ebert D, Hollis B, Olivieri I, Whitlock MC. 2012. Experimental evolution. *Trends Ecol. Evol.* 27(10):547–60
- Kim Y, Orr HA. 2005. Adaptation in sexuals vs. asexuals: clonal interference and the Fisher-Muller model. *Genetics* 171(3):1377–86
- Kimura M. 1964. Diffusion models in population genetics. *J. Appl. Probab.* 1(2):177–232
- Kimura M. 1965. Attainment of quasi linkage equilibrium when gene frequencies are changing by natural selection. *Genetics* 52(5):875–90
- Kingman J. 1982. On the genealogy of large populations. *J. Appl. Probab.* 19A:27–43
- Labbé P, Berthomieu A, Berticat C, Alout H, Raymond M, et al. 2007. Independent duplications of the acetylcholinesterase gene conferring insecticide resistance in the mosquito *Culex pipiens*. *Mol. Biol. Evol.* 24(4):1056–67
- Lalić J, Cuevas JM, Elena SF. 2011. Effect of host species on the distribution of mutational fitness effects for an RNA virus. *PLoS Genet.* 7(11):e1002378
- Lang GI, Botstein D, Desai MM. 2011. Genetic variation and the fate of beneficial mutations in asexual populations. *Genetics* 188(3):647–61
- Leffler EM, Bullaughey K, Matute DR, Meyer WK, Ségurel L, et al. 2012. Revisiting an old riddle: What determines genetic diversity levels within species? *PLoS Biol.* 10(9):e1001388
- Lewontin RC. 1974. *The Genetic Basis of Evolutionary Change*. New York: Columbia Univ. Press
- Maynard Smith J, Haigh J. 1974. The hitch-hiking effect of a favourable gene. *Genet. Res.* 23(1):23–35
- McVean GA, Charlesworth B. 2000. The effects of Hill-Robertson interference between weakly selected mutations on patterns of molecular evolution and variation. *Genetics* 155(2):929–44
- Messer PW, Neher RA. 2012. Estimating the strength of selective sweeps from deep population diversity data. *Genetics* 191:593–605
- Miralles R, Gerrish PJ, Moya A, Elena SF. 1999. Clonal interference and the evolution of RNA viruses. *Science* 285(5434):1745–47
- Miralles R, Moya A, Elena SF. 2000. Diminishing returns of population size in the rate of RNA virus adaptation. *J. Virol.* 74(8):3566–71
- Muller HJ. 1932. Some genetic aspects of sex. *Am. Nat.* 66(703):118–38
- Mustonen V, Lässig M. 2009. From fitness landscapes to seascapes: non-equilibrium dynamics of selection and adaptation. *Trends Genet.* 25(3):111–19
- Neher RA, Hallatschek O. 2013. Genealogies of rapidly adapting populations. *Proc. Natl. Acad. Sci. USA* 110(2):437–42
- Neher RA, Shraiman BI. 2009. Competition between recombination and epistasis can cause a transition from allele to genotype selection. *Proc. Natl. Acad. Sci. USA* 106:6866–71
- Neher RA, Shraiman BI. 2011a. Genetic draft and quasi-neutrality in large facultatively sexual populations. *Genetics* 188:975–96
- Neher RA, Shraiman BI. 2011b. Statistical genetics and evolution of quantitative traits. *Rev. Mod. Phys.* 83(4):1283–300
- Neher RA, Shraiman BI. 2012. Fluctuations of fitness distributions and the rate of Muller’s ratchet. *Genetics* 191:1283–93
- Neher RA, Shraiman BI, Fisher DS. 2010. Rate of adaptation in large sexual populations. *Genetics* 184:467–81
- Neher RA, Vucelja M, Mezard M, Shraiman BI. 2013. Emergence of clones in sexual populations. *J. Stat. Mech.* 2013(01):P01008

- Nicolaisen LE, Desai MM. 2012. Distortions in genealogies due to purifying selection. *Mol. Biol. Evol.* 29(11):3589–600
- O’Fallon BD. 2011. A method for accurate inference of population size from serially sampled genealogies distorted by selection. *Mol. Biol. Evol.* 28(11):3171–81
- O’Fallon BD, Seger J, Adler FR. 2010. A continuous-state coalescent and the impact of weak selection on the structure of gene genealogies. *Mol. Biol. Evol.* 27(5):1162–72
- Otto SP, Barton NH. 1997. The evolution of recombination: removing the limits to natural selection. *Genetics* 147(2):879–906
- Park S-C, Damien S, Joachim K. 2010. The speed of evolution in large asexual populations. *J. Stat. Phys.* 138:381–410
- Park S-C, Krug J. 2007. Clonal interference in large populations. *Proc. Natl. Acad. Sci. USA* 104(46):18135–40
- Pennings PS, Hermisson J. 2006. Soft sweeps III: the signature of positive selection from recurrent mutation. *PLoS Genet.* 2(12):e186
- Perfeito L, Fernandes L, Mota C, Gordo I. 2007. Adaptive mutations in bacteria: high rate and small effects. *Science* 317(5839):813–15
- Pitman J. 1999. Coalescents with multiple collisions. *Ann. Probab.* 27(4):1870–902
- Rice WR, Chippindale AK. 2001. Sexual recombination and the power of natural selection. *Science* 294(5542):555–59
- Rouzine IM, Coffin JM. 2005. Evolution of human immunodeficiency virus under selection and weak recombination. *Genetics* 170(1):7–18
- Rouzine IM, Coffin JM. 2007. Highly fit ancestors of a partly sexual haploid population. *Theor. Popul. Biol.* 71(2):239–50
- Rouzine IM, Coffin JM. 2010. Multi-site adaptation in the presence of infrequent recombination. *Theor. Popul. Biol.* 77:189–204
- Rouzine IM, Wakeley J, Coffin JM. 2003. The solitary wave of asexual evolution. *Proc. Natl. Acad. Sci. USA* 100(2):587–92
- Sanjuán R, Moya A, Elena SF. 2004. The distribution of fitness effects caused by single-nucleotide substitutions in an RNA virus. *Proc. Natl. Acad. Sci. USA* 101(22):8396–401
- Santiago E, Caballero A. 1998. Effective size and polymorphism of linked neutral loci in populations under directional selection. *Genetics* 149(4):2105–17
- Schiffels S, Szöllösi GJ, Mustonen V, Lässig M. 2011. Emergent neutrality in adaptive asexual evolution. *Genetics* 189(4):1361–75
- Schmidt JM, Good RT, Appleton B, Sherrard J, Raymant GC, et al. 2010. Copy number variation and transposable elements feature in recent, ongoing adaptation at the *Cyp6g1* locus. *PLoS Genet.* 6(6):e1000998
- Schweinsberg J. 2000. Coalescents with simultaneous multiple collisions. *Electron. J. Probab.* 5:12
- Schweinsberg J. 2003. Coalescent processes obtained from supercritical Galton-Watson processes. *Stoch. Process. Appl.* 106(1):107–39
- Seger J, Smith W, Perry J, Hunn J, Kaliszewska Z, et al. 2010. Gene genealogies strongly distorted by weakly interfering mutations in constant environments. *Genetics* 184(2):529–45
- Sella G, Petrov DA, Przeworski M, Andolfatto P. 2009. Pervasive natural selection in the *Drosophila* genome? *PLoS Genet.* 5(6):e1000495
- Shapiro JA, Huang W, Zhang C, Hubisz MJ, Lu J, et al. 2007. Adaptive genic evolution in the *Drosophila* genomes. *Proc. Natl. Acad. Sci. USA* 104(7):2271–76
- Slatkin M, Hudson RR. 1991. Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics* 129(2):555–62
- Smith DJ, Lapedes AS, de Jong JC, Bestebroer TM, Rimmelzwaan GF, et al. 2004. Mapping the antigenic and genetic evolution of influenza virus. *Science* 305(5682):371–76
- Stephan W, Wiehe THE, Lenz MW. 1992. The effect of strongly selected substitutions on neutral polymorphism: analytical results based on diffusion theory. *Theor. Popul. Biol.* 41(2):237–54
- Strelkowa N, Lässig M. 2012. Clonal interference in the evolution of influenza. *Genetics* 192(2):671–82
- Tenaillon O, Rodríguez-Verdugo A, Gaut RL, McDonald P, Bennett AF, et al. 2012. The molecular diversity of adaptive convergence. *Science* 335(6067):457–61

- Tsimring L, Levine H, Kessler D. 1996. RNA virus evolution via a fitness-space model. *Phys. Rev. Lett.* 76(23):4440–43
- Turelli M, Barton NH. 1994. Genetic and statistical analyses of strong selection on polygenic traits: what, me normal? *Genetics* 138(3):913–41
- Turner TL, Stewart AD, Fields AT, Rice WR, Tarone AM. 2011. Population-based resequencing of experimentally evolved populations reveals the genetic basis of body size variation in *Drosophila melanogaster*. *PLoS Genet.* 7(3):e1001336
- Walczak AM, Nicolaisen LE, Plotkin JB, Desai MM. 2012. The structure of genealogies in the presence of purifying selection: a fitness-class coalescent. *Genetics* 190:753–79
- Weissman DB, Barton NH. 2012. Limits to the rate of adaptive substitution in sexual populations. *PLoS Genet.* 8(6):e1002740
- Wiehe TH, Stephan W. 1993. Analysis of a genetic hitchhiking model, and its application to DNA polymorphism data from *Drosophila melanogaster*. *Mol. Biol. Evol.* 10(4):842–54
- Wylie CS, Trout AD, Kessler DA, Levine H. 2010. Optimal strategy for competence differentiation in bacteria. *PLoS Genet.* 6(9):e1001108
- Zanini F, Neher RA. 2012. FFPopSim: an efficient forward simulation package for the evolution of large populations. *Bioinformatics* 28(24):3332–33
- Zhou D, Udpa N, Gersten M, Visk DW, Bashir A, et al. 2011. Experimental selection of hypoxia-tolerant *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA* 108(6):2349–54



Contents

Genomics in Ecology, Evolution, and Systematics Theme

Introduction to Theme “Genomics in Ecology, Evolution, and Systematics” <i>H. Bradley Shaffer and Michael D. Purugganan</i>	1
Genotype-by-Environment Interaction and Plasticity: Exploring Genomic Responses of Plants to the Abiotic Environment <i>David L. Des Marais, Kyle M. Hernandez, and Thomas E. Juenger</i>	5
Patterns of Selection in Plant Genomes <i>Josh Hough, Robert J. Williamson, and Stephen I. Wright</i>	31
Genomics and the Evolution of Phenotypic Traits <i>Gregory A. Wray</i>	51
Geographic Mode of Speciation and Genomic Divergence <i>Jeffrey L. Feder, Samuel M. Flaxman, Scott P. Egan, Aaron A. Comeault, and Patrik Nosil</i>	73
High-Throughput Genomic Data in Systematics and Phylogenetics <i>Emily Moriarty Lemmon and Alan R. Lemmon</i>	99
Population Genomics of Human Adaptation <i>Joseph Lachance and Sarah A. Tishkoff</i>	123
Topical Reviews	
Symbiogenesis: Mechanisms, Evolutionary Consequences, and Systematic Implications <i>Thomas Cavalier-Smith</i>	145
Cognitive Ecology of Food Hoarding: The Evolution of Spatial Memory and the Hippocampus <i>Vladimir V. Pravosudov and Timothy C. Roth II</i>	173
Genetic Draft, Selective Interference, and Population Genetics of Rapid Adaptation <i>Richard A. Neher</i>	195
Nothing in Genetics Makes Sense Except in Light of Genomic Conflict <i>William R. Rice</i>	217

The Evolutionary Genomics of Birds <i>Hans Ellegren</i>	239
Community and Ecosystem Responses to Elevational Gradients: Processes, Mechanisms, and Insights for Global Change <i>Maja K. Sundqvist, Nathan J. Sanders, and David A. Wardle</i>	261
Cytonuclear Genomic Interactions and Hybrid Breakdown <i>Ronald S. Burton, Ricardo J. Pereira, and Felipe S. Barreto</i>	281
How Was the Australian Flora Assembled Over the Last 65 Million Years? A Molecular Phylogenetic Perspective <i>Michael D. Crisp and Lyn G. Cook</i>	303
Introgression of Crop Alleles into Wild or Weedy Populations <i>Norman C. Ellstrand, Patrick Meirmans, Jun Rong, Detlef Bartsch, Atiyo Ghosh, Tom J. de Jong, Patsy Haccou, Bao-Rong Lu, Allison A. Snow, C. Neal Stewart Jr., Jared L. Strasburg, Peter H. van Tienderen, Klaas Vrieling, and Danny Hooftman</i>	325
Plant Facilitation and Phylogenetics <i>Alfonso Valiente-Banuet and Miguel Verdú</i>	347
Assisted Gene Flow to Facilitate Local Adaptation to Climate Change <i>Sally N. Aitken and Michael C. Whitlock</i>	367
Ecological and Evolutionary Misadventures of <i>Spartina</i> <i>Donald R. Strong and Debra R. Ayres</i>	389
Evolutionary Processes of Diversification in a Model Island Archipelago <i>Rafe M. Brown, Cameron D. Siler, Carl H. Oliveros, Jacob A. Esselstyn, Arvin C. Diesmos, Peter A. Hosner, Charles W. Linkem, Anthony J. Barley, Jamie R. Oaks, Marites B. Sanguila, Luke J. Welton, David C. Blackburn, Robert G. Moyle, A. Townsend Peterson, and Angel C. Alcala</i>	411
Perceptual Biases and Mate Choice <i>Michael J. Ryan and Molly E. Cummings</i>	437
Thermal Ecology, Environments, Communities, and Global Change: Energy Intake and Expenditure in Endotherms <i>Noga Kronfeld-Schor and Tamar Dayan</i>	461
Diversity-Dependence, Ecological Speciation, and the Role of Competition in Macroevolution <i>Daniel L. Rabosky</i>	481
Consumer Fronts, Global Change, and Runaway Collapse in Ecosystems <i>Brian R. Silliman, Michael W. McCoy, Christine Angelini, Robert D. Holt, John N. Griffin, and Johan van de Koppel</i>	503

Implications of Time-Averaged Death Assemblages for Ecology and Conservation Biology <i>Susan M. Kidwell and Adam Tomasovych</i>	539
Population Cycles in Forest Lepidoptera Revisited <i>Judith H. Myers and Jenny S. Cory</i>	565
The Structure, Distribution, and Biomass of the World's Forests <i>Yude Pan, Richard A. Birdsey, Oliver L. Phillips, and Robert B. Jackson</i>	593
The Epidemiology and Evolution of Symbionts with Mixed-Mode Transmission <i>Dieter Ebert</i>	623

Indexes

Cumulative Index of Contributing Authors, Volumes 40–44	645
Cumulative Index of Article Titles, Volumes 40–44	649

Errata

An online log of corrections to *Annual Review of Ecology, Evolution, and Systematics*
articles may be found at <http://ecolsys.annualreviews.org/errata.shtml>



ANNUAL REVIEWS

It's about time. Your time. It's time well spent.

New From Annual Reviews:

Annual Review of Statistics and Its Application

Volume 1 • Online January 2014 • <http://statistics.annualreviews.org>

Editor: **Stephen E. Fienberg**, *Carnegie Mellon University*

Associate Editors: **Nancy Reid**, *University of Toronto*

Stephen M. Stigler, *University of Chicago*

The *Annual Review of Statistics and Its Application* aims to inform statisticians and quantitative methodologists, as well as all scientists and users of statistics about major methodological advances and the computational tools that allow for their implementation. It will include developments in the field of statistics, including theoretical statistical underpinnings of new methodology, as well as developments in specific application domains such as biostatistics and bioinformatics, economics, machine learning, psychology, sociology, and aspects of the physical sciences.

Complimentary online access to the first volume will be available until January 2015.

TABLE OF CONTENTS:

- *What Is Statistics?* Stephen E. Fienberg
- *A Systematic Statistical Approach to Evaluating Evidence from Observational Studies*, David Madigan, Paul E. Stang, Jesse A. Berlin, Martijn Schuemie, J. Marc Overhage, Marc A. Suchard, Bill Dumouchel, Abraham G. Hartzema, Patrick B. Ryan
- *The Role of Statistics in the Discovery of a Higgs Boson*, David A. van Dyk
- *Brain Imaging Analysis*, F. DuBois Bowman
- *Statistics and Climate*, Peter Guttorp
- *Climate Simulators and Climate Projections*, Jonathan Rougier, Michael Goldstein
- *Probabilistic Forecasting*, Tilmann Gneiting, Matthias Katzfuss
- *Bayesian Computational Tools*, Christian P. Robert
- *Bayesian Computation Via Markov Chain Monte Carlo*, Radu V. Craiu, Jeffrey S. Rosenthal
- *Build, Compute, Critique, Repeat: Data Analysis with Latent Variable Models*, David M. Blei
- *Structured Regularizers for High-Dimensional Problems: Statistical and Computational Issues*, Martin J. Wainwright
- *High-Dimensional Statistics with a View Toward Applications in Biology*, Peter Bühlmann, Markus Kalisch, Lukas Meier
- *Next-Generation Statistical Genetics: Modeling, Penalization, and Optimization in High-Dimensional Data*, Kenneth Lange, Jeanette C. Papp, Janet S. Sinsheimer, Eric M. Sobel
- *Breaking Bad: Two Decades of Life-Course Data Analysis in Criminology, Developmental Psychology, and Beyond*, Elena A. Erosheva, Ross L. Matsueda, Donatello Telesca
- *Event History Analysis*, Niels Keiding
- *Statistical Evaluation of Forensic DNA Profile Evidence*, Christopher D. Steele, David J. Balding
- *Using League Table Rankings in Public Policy Formation: Statistical Issues*, Harvey Goldstein
- *Statistical Ecology*, Ruth King
- *Estimating the Number of Species in Microbial Diversity Studies*, John Bunge, Amy Willis, Fiona Walsh
- *Dynamic Treatment Regimes*, Bibhas Chakraborty, Susan A. Murphy
- *Statistics and Related Topics in Single-Molecule Biophysics*, Hong Qian, S.C. Kou
- *Statistics and Quantitative Risk Management for Banking and Insurance*, Paul Embrechts, Marius Hofert

Access this and all other Annual Reviews journals via your institution at www.annualreviews.org.

ANNUAL REVIEWS | Connect With Our Experts

Tel: 800.523.8635 (US/CAN) | Tel: 650.493.4400 | Fax: 650.424.0910 | Email: service@annualreviews.org

