Polygenic adaptation to an environmental shift: temporal dynamics of variation under Gaussian stabilizing selection and additive effects on a single trait

Kevin R. Thornton[1]

[1]Department of Ecology and Evolutionary Biology, University of California Irvine

December 23, 2018

Abstract

Predictions about the effect of natural selection on patterns of linked neutral variation are largely based on models involving the rapid fixation of unconditionally-beneficial mutations. However, polygenic traits are expected to evolve by small allele frequency changes at many loci. Here, I use explicit forward simulations of a single trait with additive-effect mutations adapting to an optimum shift. In this model, sweeps of large-effect mutations occur from a nearly equal number of new mutations and standing variants, except when the trait is highly polygenic and fixations from standing variation dominate. Large-effect mutations contribute to adaptation even when the trait is not mutation-limited. However, such mutations do not

sweep to fixation on the time scale of adaptation for polygenic traits. The signature of hitchhiking, measured using standard summaries of linked variation, are approximately identical for the two classes of sweeps, which is due to the equilibrium frequencies of standing variants at mutation-selection equilibrium at the time of the environmental shift. I also show that the mean times to adaptation and patterns of linked selection are robust to certain details of the distribution of effect sizes, and that the proportion of new mutations of large effect is a critical parameter affecting patterns of linked selection.

Introduction

Empirical population genetics seeks to understand the evolutionary histories of natural populations by analyzing genome-wide patterns of polymorphism. The interpretation of observed patterns relies heavily on mathematical models, accompanied by various simulation methods, which make concrete predictions about the effect of evolutionary forces (natural selection, demographic events, etc.) on patterns of variation.

The models of natural selection used to interpret data come primarily from what we may call "standard population genetics" models. In these models, mutations have a direct effect on fitness (a "selection coefficient"). The fitness effects of mutations are most often assumed to be constant over time. For example, background selection is a model of unconditionally deleterious mutations resulting in strong purifying selection (Charlesworth *et al.*, 1993, 1995; Hudson and Kaplan, 1995; Cvijović *et al.*, 2018). The model of a selective sweep from a new mutation similarly posits that the variant is unconditionally beneficial with a constant effect on fitness over time (Smith and Haigh, 1974; Kaplan *et al.*, 1989; Braverman *et al.*, 1995; Durrett and Schweinsberg, 2004), and a similar assumption is made in models of selection from standing genetic variation (Hermisson and Pennings, 2005; Berg and Coop, 2015).

The standard population-genetic models of adaptation involve selective sweeps, or the fixation of unconditionally-beneficial variants. However, the models commonly used in evolutionary quantitative genetics predict that adaptation should generally proceed via subtle changes in allele frequencies at loci underlying variation in highly-polygenic traits. In recent years, several authors have begun to reconcile these disparate models of adaptation. Chevin and Hospital (2008) considered the case of a single mutation with a large effect on fitness in a highly polygenic background evolving according to an infinitesimal model. The authors found that sweeps "stall" at intermediate frequencies because frequency shifts in the polygenic background contributes to adaptation.

Recent theoretical work has attempted to clarify when sweeps should happen and when adaptation should proceed primarily via subtle allele frequency shifts. Using a model of Gaussian stabilizing selection and a single, sudden, shift in the optimal trait value, Jain and Stephan (2017b) explored the dynamics of adaptation in an additive-effects model with neither drift nor additional mutation after the optimum shift, and assuming linkage equilibrium amongst mutations affecting the trait. The authors identify the conditions required for a sweep to occur on a short time scale during which the population's mean trait value rapidly approaches the new optimum (reviewed in Jain and Stephan (2017a), and see Simons *et al.* (2018) for a similar derivation for the case of a highly-pleiotropic trait). When most mutations have small effects on fitness, using the definition from de Vladar and Barton (2014), the time to reach the new optimum depends on the number of loci contributing to variation in the trait (Jain and Stephan, 2017b). When most effects are large, the time to reach the new optimum does not depend on the number of loci, as the genetic variance for the trait is low and dominated by rare variants.

Under a model of a trait with a small number of phenotypic classes, Hoellinger *et al.* (2018) describe the dynamics of mutations following an optimum shift for mutation-limited and for highly-polygenic traits. The key parameter in their model is $\Theta = 4N\mu$, where μ

is the mutation rate relevant to the trait. When $\Theta \lesssim 1$, adaptation primarily occurs via complete sweeps. At intermediate values ($\Theta \approx 10$), partial and complete sweeps occur by the time the population has adapted. When $\Theta \approx 100$, adaptation is polygenic, due to allele frequency changes at many loci.

While the work of de Vladar and Barton (2014), Jain and Stephan (2017a), and Hoellinger *et al.* (2018) identify a definition of a large-effect mutation and the parameters under which selective sweeps are expected, we do not have a picture of the dynamics of linked selection during adaptation to an optimum shift. In large part, the difficulty of analyzing models of continuous phenotypes with partial linkage amongst sites has been an impediment to a theoretical description of the process. In general, the standard model of a single trait with additive effect mutations and Gaussian stabilizing selection assumes linkage equilibrium (or quasi-linkage equilibrium) (Turelli, 1984; Barton, 1986; de Vladar and Barton, 2014; Jain and Stephan, 2015, 2017b). Hoellinger *et al.* (2018) were able to accommodate partial linkage by simplifying how mutations affect phenotype. In their simplest model, an individual is either mutant or non-mutant, and therefore there are only two phenotypes possible.

Here, I use explicit forward-time simulations to describe the average dynamics of linked selection during the adaptation of a single trait under "real" stabilizing selection (Johnson and Barton, 2005) as it adapts to a single, sudden, shift in the optimum trait value. These simulations accommodate genetic drift and partial linkage, and are also able to track the dynamics of neutral variants over time. Thus, I am able to carry out the typical simulations done under standard population genetic models, where the properties of linked neutral variation may be obtained as a function of increasing genetic distance from a locus, or loci, where mutations contributing to adaptation arise (e.g. Kaplan *et al.*, 1988, 1989; Przeworski, 2002, and many others) under the classic model of evolutionary quantitative genetics.

By looking at several common summary statistics of genetic variation over time, I show

that the average deviation from the expectations of a neutral Wright-Fisher model is most pronounced for mutation limited traits (small Θ). This result is, of course, expected, given that these tests were designed to detect "hard" selective sweeps (Tajima, 1989; Fay and Wu, 2000; Zeng *et al.*, 2006). I find that summary statistics designed to detect "soft" sweeps from standing genetic variation (e.g. Garud *et al.*, 2015) show the patterns expected from hard sweeps. The reason is that, at mutation-selection balance, any large-effect mutations that may contribute to adaptation will be rare (Turelli, 1984). In other words, this standard model of a quantitative trait does not naturally give rise to intermediate-frequency variants of large effect with any appreciable probability.

Large-effect mutations contribute to adaptation for all of the parameter space explored here. However, the fixation time for large-effect variants increases with the mutation rate affecting the trait. Thus, consistent with the simpler model of Hoellinger *et al.* (2018), traits with high mutation rates adapt primarily via allele frequency changes, yet a subset of those variants eventually do go on to reach fixation when the trait is subject to Gaussian stabilizing selection. For large-effect mutations contributing to adaptation, they will become fixed or lost after the optimum shift because they contribute to the genetic load. Thus, their expected equilibrium frequencies are near zero and one de Vladar and Barton (2014), and they will typically become fixed or lost in order of decreasing effect size, as predicted by Robertson (1956). The time to fixation for such mutations is on the order of the population size in generations. Finally, I show that many of the average dynamics of linked selection are robust to the distribution of effect sizes on the trait, provided that the mutation rate is sufficiently large ($\Theta \geq 20$) and that the proportion of large-effect mutations is held constant.

An important distinction between this work and many previous studies (but comparable to Hoellinger *et al.* (2018)) is that quantities like the number of sweeps from new mutations and from standing variation are random variables. At the time of the optimum shift, the relationship between allele frequency and effect size is the skewed inverse relation-

ship expected for a trait under Gaussian stabilizing selection (Turelli, 1984). The model thus differs considerably from the standard approach to modeling selection from standing variation in population genetics, which is to assign the frequency at which positive selection begins (Przeworski *et al.*, 2005; Hermisson and Pennings, 2005; Innan and Kim, 2004; Berg and Coop, 2015).

In the literature, a distinction is often made between "hard" and "soft" sweeps. The former refers to the fixation of a new mutation, and its effect on linked variation. The latter refers to a *subset* of selection from standing genetic variation, in which a mutation at intermediate frequency becomes unconditionally and strongly advantageous. In general, soft sweeps do not results in strong effects on linked mutation frequencies. Instead, they result in unique patterns of haplotype structure that may be used to distinguish them from hard sweeps (Garud *et al.*, 2015; Schrider *et al.*, 2015; Schrider and Kern, 2016). A key observation here is that adaptation at loci experiencing the fixation of large-effect mutations is dominated by patterns expected under hard sweeps, even if a standing variant was fixed. This result is a consequence of the mutation frequencies at mutation-selection balance prior to the optimum shift. In the Discussion, I discuss limitations of the model with respect to the generation of soft sweep patterns, and highlight that pleiotropic effects and/or gene-by-environment interactions are fruitful avenues of future investigation in this regard.

Methods

Modeling stabilizing selection

I model a single trait under "real" stabilizing selection (Johnson and Barton, 2005). Mutations affecting trait values arise at rate μ per gamete per generation according to an infinitely-many sites scheme (Kimura, 1969). For the majority of results, the effect sizes of

new mutations on trait values, γ, are drawn from a Gaussian distribution with mean zero and standard deviation σ_μ. Mutations have additive effects on trait value and therefore a diploid's genetic value, z, is the sum of all nonzero effect sizes in that individual.

Traits are under Gaussian stabilizing selection, such that fitness, w, is $w = e^{-\frac{(z-z_o)^2}{2V_S}}$, where z_o is the optimal trait value and V_S is the sum of the variance in fitness plus the environmental variance in phenotype (Bürger, 2000, p. 166). Figure 1 shows a schematic of the model.

I model an environmental challenge as sudden "optimum shift", where the optimum trait value changes from $z_o = 0$ to $z_o > 0$.

Forward simulation schemes

I ran all simulations using two different Python packages (see Software availability below) based on the C++ library `fwdpp` (Thornton, 2014). For a given diploid population size, N, I simulated for $10N$ generations with $z_o = 0$, at which point the optimum shifts and evolution continued for another $10N$ generations.

Simulating large genomic regions with only selected variants To study the dynamics of mutations affecting trait values over time, I evolved populations of size $N = 5,000$ diploids where mutations affecting trait values occur uniformly (at rate μ) in a continuous genomic interval in which recombination breakpoints arise according to a uniform Poisson process with a mean of 0.5 recombination breakpoints per diploid. The mutation rates used were 2.5×10^{-4}, 10^{-3} and 5×10^{-3}. The three post-shift optima used were $z_o = 0.1$, 0.5, and 1. For all combination of μ and z_o, $\sigma_\mu = 0.25$.

For all combinations of μ and z_o, various summaries of the genetic variation ($V_G, \bar{z}$, etc.) in the population were recorded every generation. In total, I ran 1,024 replicates of each parameter combination. For the first 256 replicates, the frequency trajectory of all mutations

were recorded.

Simulating a ten locus system with neutral and selected variants For multi-locus simulations, a locus has scaled neutral mutation rate $\theta = 4N\mu_n = 1,000$ and scaled recombination rate $\rho = 4Nr = 1,000$, where μ_n is the neutral mutation rate per gamete and r is the mean number of recombination events per diploid. Mutation and recombination events occur uniformly along a locus and each locus is separated by 50 centiMorgans. For these simulations, I performed 256 simulation experiments per parameter combination.

Figure 2 shows how a locus is broken up into windows for analysis. Mutations affecting the trait occur in the sixth out of 11 equal-sized windows in a locus. For "genome scan" analyses, I analyze each window separately. Thus, each window has $\theta = \rho \approx 90$ and mutations affecting trait values are clustered in the middle of each locus (and are intermixed with neutral mutations). In these simulations, the total mutation rate affecting the trait, μ, is the sum over loci, and the rate per locus is equal ($\mu/10$).

In a second set of multilocus simulations, I vary the distribution of effect sizes such that the $Pr(|\gamma| \geq \hat{\gamma})$, the probability that a new mutation is of large effect (de Vladar and Barton, 2014) takes on a specific value. For Gaussian distributions of effect sizes, the mean γ is zero, as above, and σ_μ is found by numerical optimization using `scipy` (Jones *et al.*, 2001–) to give the desired $Pr(|\gamma| \geq \hat{\gamma})$. I also use Gamma distributions with shape parameters equal to either one or one-half, and then find a value for the mean of the distribution using `scipy`. These shape parameters give probability density functions that are "exponential-like" in shape. For simulations with Gamma DES, I use an equal mixture of Gamma distributions with mean γ and $-\gamma$ such that the DES is symmetric around a value of zero. I performed 100 simulation replicates for each parameter combination.

Genome scan statistics from multi-locus simulations

The ten-locus simulations described above were used to look at the temporal dynamics of several population-genetic summaries of a sample. Each of the ten loci consists of 11 non-overlapping windows (Figure 2) and all summary statistics are calculated on a per-window basis. I used `pylibseq` version 0.2.1 (https://github.com/molpopgen/pylibseq), which is a Python interface to `libsequence` (Thornton, 2003), to calculate all genome scan statistics.

Coalescent simulation I used `msprime` (Kelleher *et al.*, 2016) version 0.5.0 for all coalescent simulations under neutral models and `discoal` (Kern and Schrider, 2016) version 0.1.1 for all simulations of selective sweeps. All simulation output were processed using `pylibseq` version 0.2.1.

Data storage and analyses

I stored all numerical data resulting from temporal analyses of simulated populations as `sqlite3` database files. The majority of downstream analyses were done in `R` (R Core Team, 2016) using the `dplyr` package (Wickham and Grolemund, 2017), which is able to process the database files out-of-memory without the need to write complex SQL queries. The remaining analyses were done with the Python package `pandas` (McKinney, 2017). To generate graphics, I used a combination of the `R` `lattice` package (Sarkar, 2008), and the Python `matplotlib` library.

Software availability I used `fwdpy` version 0.0.4 (http://molpopgen.github.io/fwdpy) compiled against `fwdpp` version 0.5.4 (http://molpopgen.github.io/fwdpp) for single-region simulations. I used `fwdpy11` versions 0.1.4 and 0.2.1 (http://molpopgen.github.io/fwdpy11) for all multi-region simulations. `fwdpy11` is also based on `fwdpp`, and includes that library's

source code for ease of installation. Both packages were developed for the current work, but only the latter will be maintained.

I used the Python package pylibseq version 0.2.1 (https://pypi.python.org/pypi/pylibseq/0.2.1), which is a Python interface to libsequence (Thornton, 2003), to calculate population-genetic summary statistics.

All of these packages are available under the terms of the GNU Public License from http://www.github.com/molpopgen. The specific software versions used here are available pre-compiled for Linux via Bioconda (Dale *et al.*, 2017), with the exception of `fwdpy11` 0.2.1, which must be installed from source. I have made all Python and R (R Core Team, 2016) scripts for this work available at http://github.com/molpopgen/qtrait_paper.

Results

Review of theory

The effect of natural selection on linked neutral variation has been extensively studied for the case of directional selection on mutations with direct effects on fitness (*e.g.* Stephan *et al.*, 1992; Wiehe and Stephan, 1993; Kaplan *et al.*, 1989). This framework leads to a natural simulation scheme using the structured coalescent (Kaplan *et al.*, 1988) which has been widely-used to study the power of various approaches to detect recent sweeps from new mutations (Fay and Wu, 2000; Kim and Nielsen, 2004), from standing variation (Przeworski *et al.*, 2005; Innan and Kim, 2004; Hermisson and Pennings, 2005), from new mutations occurring at a fixed rate in the genome (Braverman *et al.*, 1995; Przeworski, 2002), or for testing methods to distinguish between various models of adaptation (Garud *et al.*, 2015; Schrider and Kern, 2016). Hereafter, I refer to this as the "standard" model of natural selection considered in population genetics, as it has been the most influential in guiding our intuition about the effect of linked selection on patterns of genomic variation.

The model of Gaussian stabilizing selection around an optimal trait value differs from the standard model in that mutations affect fitness *indirectly* via their effects on trait values. For the additive model of gene action considered here, and considering a single segregating mutation affecting the trait, the mode of selection is under- or over- dominant in a frequency-dependent manner (Robertson, 1956). This model has been extended to multiple mutations in linkage equilibrium by Barton (1986); de Vladar and Barton (2014); Jain and Stephan (2015, 2017b), amongst others.

The equilibrium conditions of models of stabilizing selection on traits have been studied extensively (Bürger, 2000, chapters 4 and 5). In general, the dynamics are quite complicated, with many possible equilibria existing for the case of many biallelic loci with equal effect sizes and no linkage disequilibrium (Barton, 1986). Recently, de Vladar and Barton (2014) used numerical simulations to approximate the dynamics of the polygenic model of Barton (1986). Their approximation is to assume that the mean change in trait value per generation is constant, which allowed them to investigate allele frequency dynamics under models with both constant effect sizes across loci and effect sizes drawn from a distribution. A key result from de Vladar and Barton (2014) is the identification of a mutational effect size threshold, $\hat{\gamma}$, below which a mutation is of small effect and above which it is of large effect. Here, $\hat{\gamma}$ is in units of the mutational effect on trait value, but "large" and "small" effect are with respect to the genetic load of the population (de Vladar and Barton, 2014).

In large populations, and assuming distributions of mutational effects symmetric around $\gamma = 0$, alleles of small effect will have equilibrium frequencies close to 1/2 while alleles of large effect will tend to be near the boundaries (de Vladar and Barton, 2014; Jain and Stephan, 2017b). In the notation used here, $\hat{\gamma} = 2\sqrt{2}\sqrt{\mu V_S}$ (Jain and Stephan, 2017b). Under this model, phenotypic adaptation to a sudden optimum shift occurs rapidly and the probability that a population adapts via sweeps of large-effect mutations or by small allele

frequency changes at many loci depends on whether most new mutations are of small or large effect (Jain and Stephan, 2017a,b). These results have been obtained under assumptions of linkage equilibrium and the results of Jain and Stephan (2017b) assume an infinitely-large population. Here, simulations of a finite population with linkage are used to investigate the effects of polygenic adaptation on patterns of linked neutral variation. The mutation rates used here correspond to the total scaled mutation rates used in Hoellinger *et al.* (2018), $\Theta = 4N\mu$, of 5, 20, and 100, respectively. For the lower values of Θ, adaptation is expected to proceed via a mix of complete and partial selective sweeps, while $\Theta = 100$ corresponds to a domain of polygenic adaptation via shifts in allele frequencies.

Simulation parameters

In standard population genetic models of adaptation, the scaled strength of selection on a beneficial mutation ($2Ns$) and its initial frequency at the onset of selection determine its effect on variation at linked sites (*e.g.* Przeworski *et al.*, 2005; Kim and Nielsen, 2004; Barton, 1986; Hermisson and Pennings, 2005).In this section, I describe some back of the envelope calculations relating the parameters of the stabilizing selection model simulated here to properties of new mutations. These calculations are intended to guide the reader's intuition about the expected patterns of selection on linked variation. As μ, the mutation rate to variants affecting the trait, is changed, the mutational variance of the trait is affected, which in turn affects $\hat{\gamma}$ (de Vladar and Barton, 2014). Thus, as we change mutation rates but hold V_S and the distribution of effect sizes (DES) constant, the proportion of large- vs small- effect mutations changes.

Consider a population with a single mutation with effect size $\gamma > 0$ existing in a single copy. Under an additive model with trait values $0, \gamma$ and 2γ for the AA, Aa, and aa genotypes, respectively, the mean trait value is $\bar{z} = 2q\gamma$. Suppose that the optimum trait value is $z_o > 0$. The probability of the mutant heterozygote being chosen as a parent is $e^{-(\gamma-\Delta z)^2/2V_S}$, where

$\Delta z = z_o - 2q\gamma$ is the difference between the population mean trait value and the optimal trait value. Thus, the strength of selection on the new mutation in the first generation is the difference in the probability that a heterozygote versus a homozygote is chosen as a parent, or $s = 1 - e^{-\frac{\gamma}{2V_S}(2\Delta z - \gamma)}$ (Robertson, 1956). Note that s is frequency-dependent due to dependency on Δz (Robertson, 1956).

Given that much of the focus of modern population genetics has been on mutations that fix due to selection, we note that a new mutation with effect size $0 < \gamma \leq (z_o - \bar{z})/2$ is able to fix without overshooting the new optimum. Whether such a mutation is of large or small effect ($\gamma > \hat{\gamma}$ or $\gamma < \hat{\gamma}$, respectively) depends on the mutation rate to variants affecting the trait and on the strength of stabilizing selection (de Vladar and Barton, 2014). For the mutation rates used here, and assuming $V_S = 1$ and a Gaussian distribution of effect sizes with mean zero and standard deviation σ_μ, Table 1 shows how the mutation rate relates to $Pr(|\gamma| \geq \hat{\gamma})$, which is the probability that a new mutation has a large effect on fitness (de Vladar and Barton, 2014). For these mutation rates and distribution of effect sizes, a large fraction of new mutations have large effects. When a large proportion of new variants have large effects, the genetic variance will initially increase during adaptation during a short-term phase dominated by directional selection (Jain and Stephan, 2017b; de Vladar and Barton, 2014). Following this initial phase, the genetic variance will decrease as large-effect mutations approach their equilibria near the boundaries and, in finite populations, eventually become fixed or lost in the population.

Table S1 details how these mutation rates and distribution of effect sizes affect the proportion of large-effect mutations with $\gamma > 0$ (*i.e.*, adaptive mutations in the direction of the new optimum). The table also shows the mean initial scaled strength of selection, $2Ns_0$, assuming that the population consists of $N - 1$ mutation free diploids and a single diploid heterozygous for an average small- or large- effect mutation, with the latter having $\gamma < z_o/2$. Table S1 also shows the proportion of newly-arising small-effect mutations in the range

$0 \leq \gamma \leq \hat{\gamma}$ and large-effect mutations in the range of $\hat{\gamma} \leq \gamma \leq z_o/2$. Table S1 focuses on these ranges of effect sizes because they reflect adaptive mutations ($\gamma > 0$) that may fix without overshooting ($\gamma \leq z_o/2$). However, it is important to note that overshooting the optimum does not necessarily imply a large fitness cost. The reason is that the curvature of the Gaussian fitness function is somewhat forgiving. Imagine the fixation of a mutation with $\gamma = 0.6$ in an otherwise monomorphic population, and the optimum has shifted to $z_o = 1.0$ and $V_S = 1$. Upon fixation, the population mean fitness will be $\bar{w} = e^{-\frac{(1.2-1)^2}{2}} \approx 0.98$, which is a two percent reduction in mean fitness when the optimum is overshot by twenty percent.

When the trait is relatively mutation-limited (small μ), the proportion of large-effect mutations increases (Table 1). Intuitively, the mutation rate affects $\hat{\gamma}$ through its effect on the standing genetic variance, V_G. A new mutation for a trait with a low mutation rate is more likely to have a larger effect on the segregation load than is a new mutation in a high-mutation rate trait. As the magnitude of the optimum shift increases (larger z_o), the initial strength of selection on new mutation ($2Ns_0$) increases (Table S1). Because $\hat{\gamma}$ increases with the mutation rate, there are no mutations of large effect possible whose effect sizes are $\leq z_o/2$ if z_o is sufficiently small, implying that we only expect complete sweeps to be of small-effect mutations for these scenarios. The parameters in Table 1 vary $\hat{\gamma}$ in order to vary the potential for a population to adapt from new mutations of large effect or from standing mutations of smaller effect. Although simplistic, these calculations predict stronger sweep signals when μ is smaller and z_o is larger.

Single region results

In this section I use simulations of a large contiguous region with recombination events occurring at a Poisson rate of 0.5 per diploid per generation. Mutations affecting trait values occur uniformly throughout the region according to the parameters in Table 1. No neutral mutations occur. Rather, these simulations are used to explore the dynamics of selected

mutations and properties of the population. I evolved populations for $10N$ generations to mutation-selection equilibrium around $z_o = 0$, at which point z_o is changed to 0.1, 0.5, or 1.0 (see Table 1) and evolution continued for another $10N$ generations. These simulations may be viewed as similar to the numerical calculations in de Vladar and Barton (2014) and Jain and Stephan (2017b), but with loose linkage between selected variants whereas the previous studies assumed linkage equilibrium and we allow for new mutation after the optimum shift.

Figure 3 shows how $\bar{z}$ and the genetic variance evolve following an optimums shift. Figure 3a shows that the mean trait value rapidly approaches the new optimum (see also de Vladar and Barton, 2014; Jain and Stephan, 2017b). For the population size of $N = 5,000$ diploids in Figure 3a, the new optimum is reached in less than 100 generations (0.02 time units). Figure 3a reveals that populations will typically slightly overshoot z_o when the parameters are in the range where adaptation proceeds via complete and partial selective sweeps (Hoellinger *et al.*, 2018).

Prior to the optimum shift, the average genetic variance is given by $4\mu V_S$ (Turelli (1984), Figure 3b). Following the optimum shift, the genetic variance spikes as the population adapts (see also de Vladar and Barton, 2014; Jain and Stephan, 2017b) and then recovers to a value near $4\mu V_S$ within ≈ 0.02 time units when the mutation rate is small and taking longer to return to equilibrium when the mutation rate is higher.

Figure 4 shows examples of the dynamics of the mean trait value ($\bar{z}$), of the genetic variance (V_G), and of mutation frequencies following the optimum shift. Each example is a single simulation replicate. The top row of plots shows that $\bar{z}$ quickly reaches z_o for the individual replicates. The approach of $\bar{z}$ to z_o corresponds to a substantial increase in the genetic variance similar to what is shown for the average genetic variance over time in Figure 3b. The middle row of panels in Figure 4 shows the frequency dynamics of mutations that eventually fix. Importantly, $\bar{z}$ typically reaches z_o before the first fixation has occurred

(see Figure S1 for details over a shorter time scale).

The legend of of each panel in Figure 4 contains the effect size for large effect variants ($|\gamma| > \hat{\gamma}$) and their origin times, o, in units of N generations, scaled such that the optimum shift corresponds to $o = 0$. For these examples, mutations with large effects on trait value fix first as predicted by Robertson (1956). In Figure 4, fixations of large effect typically have origin times very close to zero, meaning that the mutation arose close to the time of the optimum shift. This observation is expected as such mutations contribute significantly to genetic load and thus their equilibrium frequencies prior to the optimum shift should be near the boundaries (de Vladar and Barton, 2014; Jain and Stephan, 2015, 2017b).

The final row of plots in Figure 4 shows the dynamics of mutations that reached a frequency of at least one percent but were eventually lost from the population. Again, the legend includes γ and o for mutations where $|\gamma| > \hat{\gamma}$. Such large effect mutations only exist for a relatively brief period of time (relative to N) after the optimum shift, after which most segregating variation reaching appreciable derived allele frequencies are of relatively small effect. An important observation in the final row of Figure 4 is that, for a short time following the optimum shift, several intermediate-frequency variants with large effects on trait values may be segregating. Many of these variants are adaptive ($\gamma > 0$) but will only make short-term contributions to adaptation prior to their loss. The dynamics of these mutations recapitulate results from de Vladar and Barton (2014)–due to epsistatic effects on fitness, some mutations that are initially beneficial later become deleterious and are removed. Figure S1 shows the data from Figure 4 over a shorter time scale, allowing a more detailed look at the dynamics of mutations during adaptation.

A striking feature of Figure 4 is that the fixation times appear to be rather long, on the order of N generations, even for mutations of large effect. The long fixation times seen in Figure 4 are in fact typical. Table 2 shows the mean fixation time, $\bar{t}$ and the mean numbers of sweeps per simulation replicate. Typical fixation times are $O(3N)$ generations

for mutations of small effect, not much less than the expectation for a neutral mutation. For mutations of large-effect, the fixation time decreases as the magnitude of the optimum shift increases. This effect is expected, because large-effect mutations are expected to be more likely to fix when the optimum shift is large. Note that the times to fixation of such large-effect variants are $O(N)$ generations, which is long compared to the expectation for strongly-selected sweeps of new mutations Stephan *et al.* (1992). Table 2 also shows that the mean number of fixations from standing variation and the mean number of fixations from new mutations after the optimum shift are not appreciably different.

The effect sizes of fixations is time-dependent, such that substitutions that arose near the time of the optimum shift have larger effect sizes (Figure 5). Further, given that a fixation was of large effect ($|\gamma| > \hat{\gamma}$, corresponding to the solid lines in Figure 5), its origin time was very close to the time of the optimum shift (time zero in Figure 5). These large-effect fixations arising near the time of the optimum shift typically show faster times to fixation (Figure S2). These two patterns are most pronounced when the trait is mutation-limited and the optimum shift is large. For such parameters, may fixations occur that would over-shoot the optimum if they were the only mutation present in the population (points with γ values above the dot-dashed lines in Figure 5). For these mutation-limited models, fixations with $\gamma < -\hat{\gamma}$ occur, and such fixations are from mutations arising shortly after the optimum shift (Figure 5).

Figure 5 shows fixations from standing variation in red and those from new mutations in blue. At the time of the optimum shift, these large effect mutations that eventually fix are typically rare (Figure 6). For models where fixations with $|\gamma| > \hat{\gamma}$ occur, there are a considerable number of fixations of large effect mutations from standing variation, and the mean number of fixations from new mutations is typically less than the number from standing variation, even when the trait is mutation limited and the optimum shift is large (Table 2). The time-dependence of the effects sizes of substitutions is a key difference between this

model and the standard models of recurrent selective sweeps (Kaplan *et al.*, 1989, *e.g.*), in which the rate of substitution is constant over time. Here, the genetic background imposes epistasis on the fitness effects of future mutations as the population's mean trait value approaches z_o.

Under the standard models of selective sweeps, mutations are unconditionally beneficial and fix on time scales $\ll N$ generations Smith and Haigh (1974); Stephan *et al.* (1992). Here, fixations times can be much longer than under the standard model (Figure S2, Table 2). Importantly, a rapid time to fixation is not required to generate a "hitch-hiking" effect on linked neutral variation. Rather, it is the time taken to go from low frequencies on the order of the reciprocal of the population size to intermediate frequency (Barton, 1998). A rapid initial increase in frequency implies fewer opportunities for recombination, meaning that the genealogies of regions linked to the selected mutation are more likely to coalesce back to relatively few initial lineages, which is ultimately the cause of reduced variation and skewed allele frequencies. Thus, many of the large-effect mutations that quickly reach intermediate frequency in Figure 4 may generate hitch-hiking-like patterns even though they will never fix.

Taken together, Figures 5, S2, and 6 suggest that large-effect fixations occur from both new mutations and from standing variation, with more large-effect fixations occurring when the trait is more mutation-limited and/or the optimum shift is larger. Thus, we may predict that large-effect fixations from new mutations may show the signals of "hard sweeps", such as an excess of high-frequency derived neutral variants (Fay and Wu, 2000; Zeng *et al.*, 2006). Further, given that large-effect fixations from standing variation are typically rare (Figure 6), we may also expect them to have an affect on linked neutral variation (Przeworski *et al.*, 2005; Berg and Coop, 2015). Finally, given that fixations from variants that are common at the time of the optimum shift have small effects on trait values (Figure 6), such mutations are unlikely to generate the patterns of haplotype diversity associated with "soft sweeps"

because such patterns require strong selection (Garud *et al.*, 2015).

Dynamics of linked selection in a multi-locus system

I now describe the temporal dynamics of genetic variation over time in a ten locus system. The layout of each individual locus is shown in Figure 2. At each locus, mutations affecting the trait occur in the central window, localizing mutations into something "gene-like". Neutral mutations occur over a larger region, and the total neutral mutation rate and recombination rate per locus are given by $\theta = 4N\mu$ and $\rho = 4Nr$, respectively. Here, $\theta = \rho = 1,000$. Each of the ten loci assort independently. The trait parameters remain the same as in the previous section. The mutation rate affecting the trait (μ) is defined here as the sum across all loci, and the rate per locus is equal across loci.

Table 3 shows the mean number of sweeps from new mutations and from standing variation per locus in these simulations as well as the mean time to fixation for each class. The table further subdivides fixations by whether or not they involved large effect mutations ($|\gamma| \geq \hat{\gamma}$) or not ($|\gamma| < \hat{\gamma}$). In the table, sweeps from new mutations only include those arising within one hundred generations of the optimum shift. The reason for this cutoff is that populations typically reach the new optimum within that time (Figure 3a). Table 3 shows that the mean number of large-effect sweeps per locus from standing variation and from new mutations are very similar for most of the parameter space, as are mean times to fixation. Overall, fixations of small-effect mutations become more common as μ increases and, in general, the mean fixation time for these mutations is $O(N)$ generations. The behavior of sweeps from large-effect mutations can be explained using results from the previous section. Most large-effect mutations that eventually fix arose close to the time of the optimum shift (Figure 5) and large-effect standing variants are rare when the optimum changes (Figure 6). Thus, sweeps from large-effect standing variants will be qualitatively similar to sweeps from new mutations. Figure S3 shows the mean number of mutations affecting the trait over time

in these simulations. The mutation rates used here result in a range from ≈ 10 segregating variants to several hundred and that there is a transient increase in the number of segregating large-effect variants for a short period of time after the optimum changes.

Although each locus is identical with respect to mutation and recombination rates, the contribution of individual loci to adaptation is unequal. In general, there is a single locus at which segregating variation and fixations make the largest contribution towards the optimum trait value (Figure 7). However, as the trait becomes more polygenic, this "leading" locus makes less of a contribution and the remaining loci make non-trivial contributions on average.

For a given recombination rate, the strength of the "hitch-hiking" effect on linked neutral variation depends upon both the strength of selection on a sweeping mutation and the number of distinct haplotypes carrying the beneficial mutation at the onset of selection (Kaplan *et al.*, 1989; Smith and Haigh, 1974; Durrett and Schweinsberg, 2004; Berg and Coop, 2015). Thus, while the terms "hard" and "soft" sweep are common in the literature, it is more helpful to distinguish sweeps involving a mutation on very few genetic backgrounds from a sweep of a mutation found on many backgrounds at the onset of selection (Berg and Coop, 2015). When beneficial mutations are strongly-selected, the former case will give "harder" sweep signals, such as a very negative D, H, or H' statistics (Tajima, 1989; Fay and Wu, 2000; Zeng *et al.*, 2006). While the power to detect a sweep from standing variation is often presented as a function of its initial frequency at the start of selection (Przeworski *et al.*, 2005; Innan and Kim, 2004), that frequency is a proxy for the number of haplotypes carrying the mutation, which is a function of the mutation's age, the neutral mutation rate in the region, and the local recombination rate (see Berg and Coop (2015) for an approximation of the initial number of haplotypes for the case of a neutral mutation). Here, the standing variants at the time of the optimum shift are not neutral. Rather, they are deleterious and large-effect mutations will be at low frequencies (Figure 6). Theory predicts that sweeps from previously-deleterious variants to be associated with fewer haplotypes than

sweeps from neutral standing variation (Orr and Betancourt, 2001), which is exactly what we observe in our multi-locus simulations that include neutral mutations (Figure 8). Where sweeps of large effect from standing variation do occur, they are typically associated with a single haplotype (Figure 8).

The small number of haplotypes containing standing variants of large effect at the time of the optimum shift that eventually sweep to fixations (Figure 8) predicts that these fixations will generate "hard sweep"-like patterns at linked regions. In these simulations, sweeps from standing variants of large effect initially associated with multiple haplotypes did not occur (Figure 8). These results predict that patterns of haplotype variation unique to sweeps from standing variants of large effect (Garud *et al.*, 2015; Schrider *et al.*, 2015; Schrider and Kern, 2016) will not occur with appreciable probability under the model considered here.

Reduction in diversity

A "hitch-hiking" effect of a strongly-selected variant will reduce levels of diversity at flanking regions (Smith and Haigh, 1974) by an amount that depends on the ratio of the strength of selection to the local recombination rate (Stephan *et al.*, 1992; Kaplan *et al.*, 1989; Durrett and Schweinsberg, 2004). Further, this reduction in variation will be most pronounced, in both magnitude and physical scale, for mutations that start sweeping from a low frequency (Hermisson and Pennings, 2005; Przeworski *et al.*, 2005; Innan and Kim, 2004; Orr and Betancourt, 2001; Berg and Coop, 2015). For the ten-locus systems simulated here, the mean reduction in diversity at a randomly-chosen window of a randomly-chosen locus is modest (Figure 9). At most, diversity is reduced by $\approx 12\%$ on average in the central window of a locus where mutations affecting the trait arise and when the trait is mutation-limited and the optimum shift is large (upper left panel of Figure 9). The reduction in diversity is less severe in flanking windows and slowly recovers to pre-shift values over the next $\approx 4N$ generations.

If we focus on loci where mutations of large effect sweep to fixation, the reduction in diversity is much larger (Figure 9b). On average, diversity is more reduced at loci experiencing sweeps from new mutations compared to sweeps from standing variation. At flanking windows, diversity is strongly reduced on average the more mutation-limited the trait is and the larger the optimum shift. Consistent with the standard models, we therefore expect the parts of the parameter space corresponding to the upper left corner of Figure 9b to show the most extreme departures from neutrality as measured by summaries of the site-frequency spectrum and haplotype diversity.

The site frequency spectrum over time

Figure 10 shows the average behavior of Tajima's (1989) D over time. Figure 10a shows the mean D per window, averaging across loci and across replicates. Prior to the optimum shift, the mean D is negative in the central window containing selected variants. For highly polygenic traits, the equilibrium D is ≈ -0.1 in this window due to a large number of rare deleterious alleles segregating. After the optimum shift, D becomes more negative when the optimum shift is large and the trait is mutation-limited. In linked windows, the magnitude of the change in averaged D decays rapidly with increasing genetic distance. The minimal mean D observed is ≈ -0.25, and D returns to equilibrium values near zero within 2N to 6N generations.

Averaging over loci experiencing large-effect fixations, Figure 10b shows a stronger hitchhiking pattern centered on the window containing selected variants. Although the deviation in D from equilibrium decays relatively quickly both along a chromosome and over time, large-effect substitutions generate sufficiently negative D values that such loci will be enriched in the tails of empirical distributions of the statistic.

The H' statistic (Zeng *et al.*, 2006), which is a normalized version of Fay and Wu's (2000) H, shows the same qualitative behavior as Tajima's D, but over a much shorter time

scale (Figure 11), consistent with previous results (Przeworski, 2002). To detect recent hitch-hiking, this statistic relies on detecting an excess of high-frequency derived variants (Fay and Wu, 2000). These variants have hitch-hiked with sweeping selected mutations and will themselves drift to fixation shortly after the end of the sweep, resulting in a rapid return to equilibrium compared to Tajima's D. As with D, the mean value of H is more negative when conditioning on loci where large-effect fixations occurred (Figure 11b).

A curious feature of Figures 10 and 11 is that large effect fixations of new mutations and those from standing variation have similar mean effects. Sweeps from new mutations (those arising after the optimum shift) result in mean D and H' values only slightly more negative than sweeps from standing variants predating the optimum shift. The reason for this effect is seen in Figure 8–large-effect sweeps from standing variation involve rare mutations.

Instead of splitting loci based on having experience a fixation of a large-effect variant or not, we may instead look at summaries of variation partitioned by the contribution of a locus to adaptation. Figure S4 shows the average values of the D and H' statistics corresponding to the curves in Figure 7. As one would likely expect, the loci contributing the most towards adapting to the new optimum show the strongest signals of hitch-hiking.

The composite likelihood ratio test of Nielsen *et al.* (2005) tests for the same excess of high-frequency variants as the H/H' tests (Fay and Wu, 2000; Zeng *et al.*, 2006), but provides a composite likelihood ratio statistic (CLRS) of the probability of the data under hitch-hiking to the probability under the null model, which is obtained from the site-frequency spectrum of neutral variants. This method is designed to detect "hard" selective sweeps, meaning fixations from mutations that were unconditionally beneficial since their origin. As expected, the mean CLRS (calculated using `SweeD` (Pavlidis *et al.*, 2013)) increases rapidly for large optimum shifts and mutation-limited traits (Figure S5a) and the effect is more pronounced for loci where large-effect mutations fixed (Figure S5b). However, at loci experiencing large-effect substitutions, the mean CLRS is nearly identical, both in

mean value and in behavior over time, for sweeps from new mutations versus sweeps from standing variation (Figure S5b).

Power to reject the null model

Summary statistics such as D and H' are not powerful statistics for rejecting neutral null models. For the specific case of a strong sweep from a new mutation in the very recent past, power may be high at significance levels such as 0.05 (Braverman *et al.*, 1995; Przeworski, 2002), but the p-values for individual hypothesis tests are unlikely to survive a genome-wide false-discovery rate correction. Figure S6a shows the power to detect a value of D more negative than expected under the standard neutral model, after applying a multiple testing correction such that the per-window rejection rate under the null model is 0.05. The overall power of the test is low due to the number of tests performed (one per window). However, the set of loci representing "significant" deviations from the null model are enriched for large-effect substitutions (Figure S6b). For a large optimum shift, the most mutation-limited trait simulated here shows over 50% of "significant" D values are due to large-effect fixations, which themselves are almost equally-divided between sweeps from new mutations and sweeps from standing variation, with the former slightly out-numbering the latter. As the trait becomes more polygenic and/or the optimum shift less drastic, the enrichment shifts towards sweeps from standing variation. As expected (see Przeworski (2002)), the power of H' to reject the null model decays rapidly with increasing time since the optimum shift (Figure S7a).

The power to reject the null model using the CLRS is low overall and decays rapidly with time (Figure S8a), much like the H' statistic (Figure S7). When the trait is mutation-limited and the optimum shift is large, "significant" CLR statistics occurred in loci with large-effect substitutions $\approx 65\%$ of the time shortly after the optimum shift, with a slight majority of such loci having experienced sweeps from new mutations (Figure S8b). For smaller optimum

shifts and more polygenic traits, at most one-half of "significant" loci experienced large-effect substitutions, and a significant CLRT was equally likely to be due to standing variation or a new mutation (Figure S8b).

Haplotype homozygosity

Rapid increases in allele frequency due to selection will result in long stretches of homozygosity flanking the selected mutation (Kim and Nielsen, 2004). The nS_L statistic (Ferrer-Admetlla *et al.*, 2014) measures the ratio of homozygosity on the ancestral allele to that on the derived allele for each variant in the data. Thus, a negative value of the statistic implies longer runs of homozygosity around the derived allele. Figure 12 shows the average behavior of z-scores obtained from binning nS_L scores by derived allele frequency (see Methods). The signal of strong positive selection, indicated by a negative z-score, is short-lived and only observed when three conditions hold. First, the trait is mutation-limited. Second, the statistic is calculated from genomic windows containing mutations affecting the trait. Third, the optimum shift is large. The mean z-score then becomes positive for approximately N generations on average before returning to its expected equilibrium value of zero. If the optimum shift is intermediate ($z_o = 0.5$), the average z-score is positive, including in the window with variants affecting fitness. Focusing only on loci experiencing fixations of large-effect, these patterns are more pronounced and are more extreme on average for sweeps from new mutations than for sweeps from standing variation (Figure 12b).

Shortly after the optimum shift, the mean z-score becomes positive (Figure 12). This temporal dynamic is qualitatively similar to what is seen under standard models of selective sweeps, as the time since the sweep moves further into the past. Figure S9 shows the dynamics of the mean z-score for models of a single, strongly-selected mutation sweeping to fixation from different starting frequencies. As the age of the sweep increases into the past, the mean z-score is positive. Further, sweeps from common variants do not lead to

appreciably negative z-scores. Thus, the positive z-scores in Figure 12 may be interpreted as either older sweeps from new mutations or strong sweeps from common variants. However, the latter class of sweeps does not occur in these simulations (Figure 8). This difficulty in interpretation is similar to the "soft shoulder" effect described by Schrider *et al.* (2015) where the patterns of variation linked to an older "hard" sweep overlap considerably with those of incomplete sweeps and strong sweeps from standing variation.

A related class of statistics designed to detect strong sweeps from standing variation are based on the overall haplotype diversity in a window (Garud *et al.*, 2015). The temporal patterns associated with these statistics are again short-lived and are all in the direction of reduced overall haplotype heterozygosity, which is a signal of strong sweeps from new mutations (Figures S10, S11, and S12).

Holding the fraction of large-effect mutations constant

The results described in the previous section are based on a Gaussian distribution of effect sizes whose standard deviation is held constant. Parameterizing a simulation in this fashion implies that the mutational variance is only affected by the mutation rate and that the proportion of new mutations of large effect varies accordingly (Table 1). In this section, I vary the distribution of effect sizes such that the fraction of large-effect mutations is held constant and compare the average dynamics of adaptation and patterns of hitch-hiking. I also compare a Gaussian distribution to a Gamma distribution with different shape parameters. Figure S13 shows an example of the DES used. In order to simplify the presentation, I only show results for the case of a large optimum shift ($z_o = 1$), which is the case resulting in the most extreme hitch-hiking signals.

Figure 13 shows the behavior of the mean genetic value in the population ($\bar{z}$) and the mean genetic value due to the largest-effect locus as both μ and $Pr(|\gamma| \geq \hat{\gamma})$ vary. For lower mutation rates, populations reach the new optimum more quickly when the fraction of

new, large-effect mutations is higher (Figure 13a). For the highest mutation rate considered there, the proportion of large-effect variants entering the population has only a subtle effect on the time required to adapt. Figure 13a further shows that $\bar{z}$ depends only weakly on the distribution of effect sizes except when the mutation rate is low ($\mu = 2.5^{-4}$) and $Pr(|\gamma| \geq \hat{\gamma}) \geq 0.25$. For these parameters, the Gamma DES approach the new optimum more quickly than the Gaussian DES (Figure 13a).

Both the DES and $Pr(|\gamma| \geq \hat{\gamma})$ affect the mean genetic value of the largest-effect locus in the ten-locus system (Figure 13b). The mean trait value at the largest-effect locus is higher when the probability of a large-effect mutation is higher and when the DES is a Gamma distribution. Further, when the DES is a Gamma distribution, a single locus is more likely to make a larger overall contribution to adaptation (Figure S14).

Figure 14 shows the mean value of Tajima's D and H' over time in the central genomic window where mutations affecting trait values arise (Figure 2). The overall pattern is that the average differences between DES are subtle, with Gamma distributions showing less extreme hitch-hiking patterns (negative values) on average than the Gaussian DES. However, this difference between DES is only observed when the mutation rate is small and $Pr(|\gamma| \geq \hat{\gamma})$ is large.

Overall, Figures 13 and 14 show that the mean time to reach the new optimum is longer (Figure 13a), and the magnitude of hitch-hiking patterns more extreme (Figure 14) when the fraction of new large-effect mutations is lower. In other words, holding μ constant, a trait that is mutation-limited with respect to the origin of mutations with effect sizes larger than $\hat{\gamma}$ will reach the new optimum more slowly on average and show larger departures from equilibrium patterns of variation. In the extreme case of a low mutation rate and a low probability of a large-effect mutation, rapid adaptation to the optimum shift does not occur.

Discussion

I have used simulations to describe the average behavior of selected and neutral mutations during the adaptation of a quantitative trait to a single, sudden, shift in the optimal trait value. The genotype-to-phenotype model considered here is the classic model of evolutionary quantitative genetics, assuming strictly additive mutational effects on trait value with fitness determined by Gaussian stabilizing selection (Turelli, 1984; Barton, 1986; Bürger, 2000). The primary goal here was to merge this model of a phenotype with the simulation methods commonly-used in population genetics to study the effect of natural selection on the dynamics of linked neutral variation (Kaplan *et al.*, 1988, 1989; Braverman *et al.*, 1995; Przeworski, 2002; Innan and Kim, 2004).

Under a wide range of details concerning the DES and of assumptions about the genotype-to-phenotype model, "real" stabilizing selection on a single trait (Johnson and Barton, 2005) results in a "genetic architecture" at equilibrium where much of the genetic variance is contributed by rare mutations of large effect (Turelli, 1984; Thornton *et al.*, 2013; Sanjak *et al.*, 2017)(Bürger, 2000, pp 229-239). When a population is at mutation-selection balance and the optimum is displaced, large-effect standing variants affecting the trait will be rare and common variants will have small effects on trait values. Fixations from standing variants of large effect will therefore have similar dynamics to fixations of new mutations, due to the rarity of the variant at the onset of adaptation (Figure 6, Berg and Coop (2015); Przeworski *et al.* (2005)). Thus, a key result of the simulations presented here is that sweeps from standing variation under the classic model of a quantitative trait result in "harder" signatures of selection (Figures 10-12, S10-S12), meaning a pronounced excess of high-frequency, derived variants (Fay and Wu, 2000; Zeng *et al.*, 2006), reduced heterozygosity of neutral mutations (Stephan *et al.*, 1992), and very few haplotypes reaching high frequency in the genomic interval containing the adaptive variant(Depaulis and Veuille,

1998; Garud *et al.*, 2015).

It is important to note a key methodological difference between this work and that of Hoellinger *et al.* (2018). Hoellinger *et al.* (2018) stopped their simulations when the population was close to the new optimum while the simulations conducted here allowed evolution to continue much longer. Thus, on the time scale during which the population adapts, fixations are not observed when Θ is high (see Figure 4 of Hoellinger *et al.* (2018)). Here, we observe fixations of large effect for the mutation rates corresponding to $\Theta = 4N\mu = 100$ (Figure 4), which Hoellinger *et al.* (2018) show is the parameter range where adaptation occurs primarily by small changes in allele frequency ("polygenic adaptation"). These results are consistent with the theoretical predictions from Hoellinger *et al.* (2018). In the rightmost column of Figure 4, the population has adapted very quickly, with the fixations occurring over a much large time scale (Figure S1) Likewise, the leftmost column of Figure 4 corresponds to $\Theta = 5$, where we observe a mixture of partial and complete selective sweeps by the time the new optimum is reached, which is expected from the theory presented in Hoellinger *et al.* (2018).

The distribution of effect sizes has different effects on properties of the trait than on patterns of hitch-hiking. The mutation rates used here span the parameter space from partial and complete sweeps being most common to polygenic adaptation via allele frequency shifts at many mutations (Figure 4, Hoellinger *et al.* (2018)). As expected (Hoellinger *et al.*, 2018), mutation-limited traits show stronger patterns of hitch-hiking at linked sites during adaptation to the new optimum. These patterns will, on average, be difficult to detect unless the optimum shift is large, even when the trait is mutation-limited. The different DES used here show similar average patterns, except for very mutation-limited traits (Figure 14a,Figure 14b). Likewise, the mean trait value over time is similar for the different DES (Figure 13a), but the effect size due to the largest-effect locus appears strongly-dependent on the DES (Figure 13b).

When considering the pattern of hitch-hiking at a locus, the presence or absence of a large-effect fixation at a locus is a reliable predictor or the magnitude of hitch-hiking patterns. On the one hand, this result is not surprising, as large-effect variants experience stronger selection, which is required to have a large effect on linked variation. On the other hand, fixation times under the model used here are long with respect to the mean time to reach the new optimum. However, when traits are mutation-limited, fixations of large-effect variants are accompanied by rapid initial increases in frequency (Figure 4, Figure S1) and relatively rapid times to fixation (Table 2). The rapid increase in frequency (from a low initial frequency) is primarily responsible for the hitch-hiking patterns (Barton, 1998; Berg and Coop, 2015). When the population adapts primarily via polygenic adaptation (large Θ, Hoellinger *et al.* (2018)), fixation times are long for large-effect mutations (Figure 4, Table 2) and there is little appreciable effect on variation at linked sites.

Holding the fraction of newly-arising, large-effect mutations ($Pr(|\gamma| \geq \hat{\gamma})$) constant results in more extreme departures from the standard neutral model when $Pr(|\gamma| \geq \hat{\gamma})$ is small. Effectively, the trait becomes more mutation-limited with respect to the types of mutations that generate strong patterns of hitch-hiking at linked sites. When the trait is more polygenic, the average patterns of variation are not strongly-dependent on the DES nor on the proportion of new variants with large effect (Figure 14). Here, I simulate $Pr(|\gamma| \geq \hat{\gamma})$ down to one percent, meaning that 99 percent of new mutations are of small effect. However, the dynamics of the genetic variance suggest that all of the parameter space explored here fall into the domain described by Jain and Stephan (2017b) as being dominated by mutations of large-effect. In that domain, the genetic variance "spikes" during adaptation, as seen in Figure 3, as rare mutations of large effect rapidly increase in frequency. Once the mean trait value reaches the new optimum, the genetic variance decreases because segregating variants of large effect move towards their new equilibrium values, which are near zero and one (de Vladar and Barton, 2014; Jain and Stephan, 2017b), and eventually become

fixed or lost in a finite population. In the domain where most mutations have small effects, the genetic variance actually decreases during adaptation (de Vladar and Barton, 2014; Jain and Stephan, 2017b). The decrease is because, at mutation-selection balance prior to the optimum shift, the equilibrium frequency of a small-effect variant is one-half, which maximizes the genetic variance. Adaptation results in allele frequency shifts, which moves mutations away from one-half, thus reducing the variance. Figure S17 shows the mean genetic variance over time, and a spike is seen for all values of $Pr(|\gamma| \geq \hat{\gamma})$.

There are several key differences from this work and that of Jain and Stephan (2017b) and de Vladar and Barton (2014). The simulations used here incorporate genetic drift and also allow for further mutations to occur after the optimum shift. Additionally, the frequencies of mutations affecting the trait are at mutation-selection equilibrium here at the onset of adaptation, in contrast to random frequencies as in de Vladar and Barton (2014). It will be useful for future work to investigate the role of population size on the genetic variance as a function of $Pr(|\gamma| \geq \hat{\gamma})$. Doing so efficiently will require the use of tree sequence recording Kelleher *et al.* (2018), which results in order-of-magnitude improvements run times when simulating large populations. Finally, the simulations here involve a $V_S = 1$, which lets us view γ in units of fitness effects due to deviation from the optimum. In the numerical simulations of the "most effects are small" domain, de Vladar and Barton (2014) and Jain and Stephan (2017b) show results for V_S of $O(10)$ and $O(1,000)$, respectively, with very small mean γ of segregating variants. It is likely that the difference in model parameters drives the temporal dynamics of the genetic variance for the trait seen here.

Here, I have focussed on standard summaries of variation data that have widely been applied to detect selection from sequence data. The behavior of the majority of such summary statistics have only been tested using coalescent simulations of strong selection on a single sweeping variant, which is the dominant generative model used for making predic-

tions about linked selection. Thus, it is unsurprising that these statistics show the strongest departures from equilibrium neutrality for mutation-limited traits. However, an important observation here is that the mean behavior of these statistics are very similar for sweeps from new mutations and sweeps from standing genetic variation, which is a consequence of the standing variants being rare at the onset of selection (Figure 6, also see Berg and Coop (2015); Przeworski *et al.* (2005); Orr and Betancourt (2001); Hermisson and Pennings (2005)). The only test statistic that I am aware of for detecting polygenic adaptation is the singleton density score (Field *et al.*, 2016). I have not explored this statistic here, as it would be more fruitful to do so using simulations of much larger genomic regions applying tree sequence recording (Kelleher *et al.*, 2018) so that the temporal dynamics of the population's entire genealogy can be studied along with the statistic itself. It also appears that the magnitude of selective effects on phenotypes attributable to changes in the singleton density by Field *et al.* (2016) were substantially overestimated due to uncontrolled-for population structure in the GWAS data, and there is little evidence for selection on height when the analysis is done using effect sizes from the UK Biobank data (Berg *et al.*, 2018).

I have only considered the equilibrium Wright-Fisher model here. However, it is well-understood that departures from this demographic model affect patterns of neutral variation and thus the detection of regions affected by linked selection (Thornton and Andolfatto, 2006; Thornton and Jensen, 2007; Thornton *et al.*, 2007; Jensen *et al.*, 2007, 2008). Demographic departures from constant population size indeed affect the prevalence of sweeps and the rate of phenotypic adaptation in optimum shift models Stetter *et al.* (2018). Here, we are primarily interested in how the parameters affecting the trait's "architecture", mainly the parameters affecting the mutational variance of the trait, impact patterns of linked selection.

It is crucial to restate the assumptions of the genetic model assumed here, which involves strictly additive effects on a single trait under "real" stabilizing selection (Johnson and

Barton, 2005). This model is the standard model of evolutionary quantitative genetics (Turelli, 1984; Barton, 1986; Bürger, 2000), which is why it is the focus of this work. However, a more thorough understanding of the dynamics of linked selection during polygenic adaptation will require investigation of models with pleiotropic effects (*e.g.*, Zhang and Hill, 2002; Simons *et al.*, 2018). Because the adaptation to the new optimum is very rapid when traits are not mutation-limited, the allele frequency changes involved are also small when mutational effects are pleiotropic (Simons *et al.*, 2018). The question in a pleiotropic model is the role that large-effect mutations may play, which is an unresolved question.

The genetic model assumed here does not lead to sweeps of large-effect mutations from standing variation. Importantly, it is not possible to tune the model parameters to obtain such events with high probability. Changing the strength of stabilizing selection (V_S) preserves the rank orders of fitnesses for all genotypes, merely changing how fit they are in an absolute sense. One could randomly reassign effect sizes at the time of the optimum shift in an attempt to approximate a gene by environment interaction. However, such a procedure would be arbitrary, and thus not represent a principled model for generating "soft sweep" patterns. Rather, it is tempting to invoke a need for pleiotropic effects in order to have large-effect mutations segregating at intermediate frequencies at the time of the optimum shift, with the shift itself accompanied by a change in the covariance between trait values and fitness.

It is clearly of interest to explore non-additive genetic models in future work. In particular, models of non-complementing recessive effects within genes are a specific class of models with epistasis that deserve consideration due to their connection with poly-allelic bases of phenotypic variation (Thornton *et al.*, 2013; Sanjak *et al.*, 2017; Gruber and Long, 2009; King *et al.*, 2014; Long *et al.*, 2014; McClellan and King, 2010; Chakraborty *et al.*, 2018). Acknowledging the focus on the standard additive model, the current work is best viewed as an investigation of a central concern in molecular population genetics (the

effect of natural selection on linked neutral variation) having replaced the standard model of that sub-discipline with the standard model of evolutionary quantitative genetics. As laid out by several authors (Jain and Stephan, 2017b,a; Messer and Petrov, 2013), there are considerable theoretical and empirical challenges remaining in the understanding of the genetics of rapid adaptation. For models of phenotypic adaptation, our standard "tests of selection" are likely to fail, and are highly under-powered even when the assumptions of the phenotype model are close to that of the standard model (*e.g*, the trait is mutation-limited). The results here also have some implication for the ability to map genes underlying adaptive genetic variation. Even when the trait is mutation-limited, two loci may make substantial contributions to adaptation and the number of loci contributing depends on the magnitude of the optimum shift (Figure 7). Thus, the number of mappable loci is likely to depend on the strength of recent selection experienced by natural populations.

Acknowledgements

The author thanks Jaleal Sanjak, David Lawrie, Jeffrey Ross-Ibarra, Markus Stetter, and Tony Long for valuable discussions. This work is supported by NIH Grant R01-GM115564 to the author.

References

Barton, N. H., 1986 The maintenance of polygenic variation through a balance between mutation and stabilizing selection. Genet. Res. **47**: 209–216.

Barton, N. H., 1998 The effect of hitch-hiking on neutral genealogies. Genet. Res. **72**: 123–133.

Berg, J. J. and G. Coop, 2015 A coalescent model for a sweep of a unique standing variant. Genetics **201**: 707–725.

Berg, J. J., A. Harpak, N. Sinnott-Armstrong, A. M. Joergensen, H. Mostafavi *et al.*, 2018 Reduced signal for polygenic adaptation of height in UK biobank.

Braverman, J. M., R. R. Hudson, N. L. Kaplan, C. H. Langley and W. Stephan, 1995 The hitchhiking effect on the site frequency spectrum of DNA polymorphisms. Genetics **140**: 783–796.

Bürger, R., 2000 *The Mathematical Theory of Selection, Recombination, and Mutation*. Wiley.

Chakraborty, M., J. J. Emerson, S. J. Macdonald and A. D. Long, 2018 Structural variants exhibit allelic heterogeneity and shape variation in complex traits.

Charlesworth, B., M. T. Morgan and D. Charlesworth, 1993 The effect of deleterious mutations on neutral molecular variation. Genetics **134**: 1289–1303.

Charlesworth, D., B. Charlesworth and M. T. Morgan, 1995 The pattern of neutral molecular variation under the background selection model. Genetics **141**: 1619–1632.

Chevin, L.-M. and F. Hospital, 2008 Selective sweep at a quantitative trait locus in the presence of background genetic variation. Genetics **180**: 1645–1660.

Cvijović, I., B. H. Good and M. M. Desai, 2018 The effect of strong purifying selection on genetic diversity. Genetics **209**: 1235–1278.

Dale, R., B. Grüning, A. Sjödin, J. Rowe, B. A. Chapman *et al.*, 2017 Bioconda: A sustainable and comprehensive software distribution for the life sciences.

de Vladar, H. P. and N. Barton, 2014 Stability and response of polygenic traits to stabilizing selection and mutation. Genetics **197**: 749–767.

Depaulis, F. and M. Veuille, 1998 Neutrality tests based on the distribution of haplotypes under an infinite-site model. Mol. Biol. Evol. **15**: 1788–1790.

Durrett, R. and J. Schweinsberg, 2004 Approximating selective sweeps. Theor. Popul. Biol. **66**: 129–138.

Fay, J. C. and C. I. Wu, 2000 Hitchhiking under positive darwinian selection. Genetics **155**: 1405–1413.

Ferrer-Admetlla, A., M. Liang, T. Korneliussen and R. Nielsen, 2014 On detecting incomplete soft or hard selective sweeps using haplotype structure. Mol. Biol. Evol. **31**: 1275–1291.

Field, Y., E. A. Boyle, N. Telis, Z. Gao, K. J. Gaulton *et al.*, 2016 Detection of human adaptation during the past 2000 years. Science **354**: 760–764.

Garud, N. R., P. W. Messer, E. O. Buzbas and D. A. Petrov, 2015 Recent selective sweeps in north american drosophila melanogaster show signatures of soft sweeps. PLoS Genet. **11**: e1005004.

Gruber, J. D. and A. D. Long, 2009 Cis-regulatory variation is typically polyallelic in drosophila. Genetics **181**: 661–670.

Hermisson, J. and P. S. Pennings, 2005 Soft sweeps: molecular population genetics of adaptation from standing genetic variation. Genetics **169**: 2335–2352.

Hoellinger, I., P. Pennings and J. Hermisson, 2018 Polygenic adaptation: From sweeps to subtle frequency shifts.

Hudson, R. R. and N. L. Kaplan, 1995 Deleterious background selection with recombination. Genetics **141**: 1605–1617.

Innan, H. and Y. Kim, 2004 Pattern of polymorphism after strong artificial selection in a domestication event. Proc. Natl. Acad. Sci. U. S. A. **101**: 10667–10672.

Jain, K. and W. Stephan, 2015 Response of polygenic traits under stabilizing selection and mutation when loci have unequal effects. G3 **5**: 1065–1074.

Jain, K. and W. Stephan, 2017a Modes of rapid polygenic adaptation. Mol. Biol. Evol. .

Jain, K. and W. Stephan, 2017b Rapid adaptation of a polygenic trait after a sudden environmental shift. Genetics **206**: 389–406.

Jensen, J. D., K. R. Thornton and P. Andolfatto, 2008 An approximate bayesian estimator suggests strong, recurrent selective sweeps in drosophila. PLoS Genet. **4**: e1000198.

Jensen, J. D., K. R. Thornton, C. D. Bustamante and C. F. Aquadro, 2007 On the utility of linkage disequilibrium as a statistic for identifying targets of positive selection in nonequilibrium populations. Genetics **176**: 2371–2379.

Johnson, T. and N. Barton, 2005 Theoretical models of selection and mutation on quantitative traits. Philos. Trans. R. Soc. Lond. B Biol. Sci. **360**: 1411–1425.

Jones, E., T. Oliphant, P. Peterson *et al.*, 2001– SciPy: Open source scientific tools for Python. [Online; accessed ¡today¿].

Kaplan, N. L., T. Darden and R. R. Hudson, 1988 The coalescent process in models with selection. Genetics **120**: 819–829.

Kaplan, N. L., R. R. Hudson and C. H. Langley, 1989 The "hitchhiking effect" revisited. Genetics **123**: 887–899.

Kelleher, J., A. M. Etheridge and G. McVean, 2016 Efficient coalescent simulation and genealogical analysis for large sample sizes. PLoS Comput. Biol. **12**: e1004842.

Kelleher, J., K. R. Thornton, J. Ashander and P. L. Ralph, 2018 Efficient pedigree recording for fast population genetics simulation. PLoS Comput. Biol. **14**: e1006581.

Kern, A. D. and D. R. Schrider, 2016 Discoal: flexible coalescent simulations with selection. Bioinformatics **32**: 3839–3841.

Kim, Y. and R. Nielsen, 2004 Linkage disequilibrium as a signature of selective sweeps. Genetics **167**: 1513–1524.

Kimura, M., 1969 The number of heterozygous nucleotide sites maintained in a finite population due to steady flux of mutations. Genetics **61**: 893–903.

King, E. G., B. J. Sanderson, C. L. McNeil, A. D. Long and S. J. Macdonald, 2014 Genetic dissection of the drosophila melanogaster female head transcriptome reveals widespread allelic heterogeneity. PLoS Genet. **10**: e1004322.

Long, A. D., S. J. Macdonald and E. G. King, 2014 Dissecting complex traits using the drosophila synthetic population resource. Trends Genet. **30**: 488–495.

McClellan, J. and M.-C. King, 2010 Genetic heterogeneity in human disease. Cell **141**: 210–217.

McKinney, W., 2017 *Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython*. O'Reilly Media, 2 edition edition.

Messer, P. W. and D. A. Petrov, 2013 Population genomics of rapid adaptation by soft selective sweeps. Trends Ecol. Evol. **28**: 659–669.

Nielsen, R., S. Williamson, Y. Kim, M. J. Hubisz, A. G. Clark *et al.*, 2005 Genomic scans for selective sweeps using SNP data. Genome Res. **15**: 1566–1575.

Orr, H. A. and A. J. Betancourt, 2001 Haldane's sieve and adaptation from the standing genetic variation. Genetics **157**: 875–884.

Pavlidis, P., D. Živković, A. Stamatakis and N. Alachiotis, 2013 SweeD: Likelihood-Based detection of selective sweeps in thousands of genomes. Mol. Biol. Evol. **30**: 2224–2234.

Przeworski, M., 2002 The signature of positive selection at randomly chosen loci. Genetics **160**: 1179–1189.

Przeworski, M., G. Coop and J. D. Wall, 2005 The signature of positive selection on standing genetic variation. Evolution **59**: 2312–2323.

R Core Team, 2016 *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Robertson, A., 1956 The effect of selection against extreme deviants based on deviation or on homozygosis. J. Genet. **54**: 236.

Sanjak, J. S., A. D. Long and K. R. Thornton, 2017 A model of compound heterozygous, Loss-of-Function alleles is broadly consistent with observations from Complex-Disease GWAS datasets. PLoS Genet. **13**: e1006573.

Sarkar, D., 2008 *Lattice: Multivariate Data Visualization with R (Use R!)*. Springer, 1st edition edition.

Schrider, D. R. and A. D. Kern, 2016 S/HIC: Robust identification of soft and hard sweeps using machine learning. PLoS Genet. **12**: e1005928.

Schrider, D. R., F. K. Mendes, M. W. Hahn and A. D. Kern, 2015 Soft shoulders ahead: spurious signatures of soft and partial selective sweeps result from linked hard sweeps. Genetics **200**: 267–284.

Simons, Y. B., K. Bullaughey, R. R. Hudson and G. Sella, 2018 A population genetic interpretation of GWAS findings for human quantitative traits. PLoS Biol. **16**: e2002985.

Smith, J. M. and J. Haigh, 1974 The hitch-hiking effect of a favourable gene. Genet. Res. **23**: 23–35.

Stephan, W., T. H. E. Wiehe and M. W. Lenz, 1992 The effect of strongly selected substitutions on neutral polymorphism: Analytical results based on diffusion theory. Theor. Popul. Biol. **41**: 237–254.

Stetter, M. G., K. Thornton and J. Ross-Ibarra, 2018 Genetic architecture and selective sweeps after polygenic adaptation to distant trait optima. PLoS Genet. **14**: e1007794.

Tajima, F., 1983 Evolutionary relationship of DNA sequences in finite populations. Genetics **105**: 437–460.

Tajima, F., 1989 Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics **123**: 585–595.

Thornton, K., 2003 Libsequence: a c++ class library for evolutionary genetic analysis. Bioinformatics **19**: 2325–2327.

Thornton, K. and P. Andolfatto, 2006 Approximate bayesian inference reveals evidence for a recent, severe bottleneck in a netherlands population of drosophila melanogaster. Genetics **172**: 1607–1619.

Thornton, K. R., 2014 A c++ template library for efficient forward-time population genetic simulation of large populations. Genetics **198**: 157–166.

Thornton, K. R., A. J. Foran and A. D. Long, 2013 Properties and modeling of GWAS when complex disease risk is due to non-complementing, deleterious mutations in genes of large effect. PLoS Genet. **9**: e1003258.

Thornton, K. R. and J. D. Jensen, 2007 Controlling the false-positive rate in multilocus genome scans for selection. Genetics **175**: 737–750.

Thornton, K. R., J. D. Jensen, C. Becquet and P. Andolfatto, 2007 Progress and prospects in mapping recent selection in the genome. Heredity **98**: 340–348.

Turelli, M., 1984 Heritable genetic variation via mutation-selection balance: Lerch's zeta meets the abdominal bristle. Theor. Popul. Biol. **25**: 138–193.

Wickham, H. and G. Grolemund, 2017 *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*. O'Reilly Media, 1 edition edition.

Wiehe, T. H. E. and W. Stephan, 1993 Analysis of a genetic hitchhiking model, and its application to DNA polymorphism data from drosophila melanogaster 1. Mol. Biol. Evol. **10**: 842–854.

Zeng, K., Y.-X. Fu, S. Shi and C.-I. Wu, 2006 Statistical tests for detecting positive selection by utilizing high-frequency variants. Genetics **174**: 1431–1439.

Zhang, X.-S. and W. G. Hill, 2002 Joint effects of pleiotropic selection and stabilizing selection on the maintenance of quantitative genetic variation at mutation-selection balance. Genetics **162**: 459–471.

μ	$\hat{\gamma}$	$Pr(\lvert\gamma\rvert \geq \hat{\gamma})$
2.5×10^{-4}	0.045	0.858
1×10^{-3}	0.089	0.721
5×10^{-3}	0.20	0.423

Table 1: Relationship between mutation rate and fraction of new mutations with large effects on fitness.

Table 2: Summaries of fixations.

			Standing variation			New mutations		
μ	z_o	Large effect?	Mean γ	Mean fixation time	Mean number	Mean γ	Mean fixation time	Mean number
0.00025	0.1	No	0.02055	2.8597	2.335	0.03296	1.4011	1.000
		Yes	0.06711	2.1455	1.132	0.07998	2.7357	1.000
	0.5	No	0.01775	3.2923	2.092	NA	NA	NA
		Yes	0.19011	0.6753	1.961	0.27441	0.2379	1.039
	1.0	No	0.01871	3.1813	2.228	0.02759	3.4474	1.000
		Yes	0.26503	0.5635	2.473	0.37561	0.2058	1.131
0.001	0.1	No	0.01892	3.2166	4.949	0.04601	2.2334	1.000
	0.5	No	0.01992	3.2526	5.189	0.04269	1.7240	1.222
		Yes	0.19998	0.3267	1.249	0.21418	0.3346	1.000
	1.0	No	0.01893	3.2677	4.987	0.05650	1.2714	1.000
		Yes	0.34665	0.2165	1.659	0.36372	0.2730	1.179
0.005	0.1	No	0.01461	3.5181	19.544	0.01669	3.7438	1.000
		Yes	0.22024	2.2806	2.000	NA	NA	NA
	0.5	No	0.02048	3.2653	23.223	0.04669	2.2231	1.000
		Yes	0.24799	1.6938	2.000	NA	NA	NA
	1.0	No	0.02392	3.1887	23.460	0.06282	2.0532	1.205
		Yes	0.28118	0.4081	1.081	0.27006	0.8358	1.000

Table 3: Mean number of sweeps per locus and mean fixation time (in units of N generations).

μ	z_o	$\lvert\gamma\rvert > \hat{\gamma}$	Standing variation	Fixation time	New mutation	Fixation time
2.5×10^{-4}	0.1	N	0.138	2.378	0.1	1.721
		Y	0.1	1.321	0.1	2.043
	0.5	N	0.133	2.95	0.1	2.019
		Y	0.14	0.183	0.112	0.198
	1	N	0.152	2.563	0.1	2.036
		Y	0.157	0.156	0.13	0.09
1×10^{-3}	0.1	N	0.251	2.97	0.113	1.91
		Y	0.1	1.799	0	NA
	0.5	N	0.28	2.979	0.112	1.495
		Y	0.126	0.326	0.1	0.428
	1	N	0.281	2.892	0.125	1.155
		Y	0.164	0.224	0.117	0.161
5×10^{-3}	0.1	N	0.818	3.709	0.118	2.595
		Y	0	NA	0	NA
	0.5	N	1.054	3.084	0.114	1.887
		Y	0.133	2.85	0	NA
	1	N	1.082	2.915	0.112	1.807
		Y	0.108	0.685	0.1	0.874

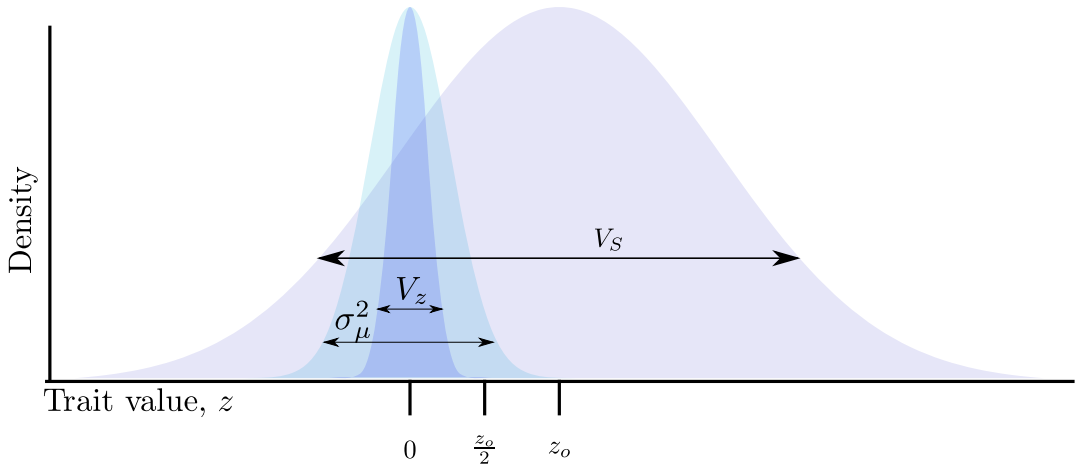


Figure 1: Schematic of the model. A Wright-Fisher population evolves to equilibrium around an optimum trait under Gaussian stabilizing selection with mean zero, where the parameter V_S represents the intensity of selection against extreme trait values ($w = e^{-z^2/2V_S}$). At equilibrium, the mean trait value is $\bar{z} \approx 0$ and the genetic variance, V_G equals the phenotypic variance V_z. Mutations arise at a constant rate with effect sizes, γ, drawn from a Gaussian distribution with mean zero and variance σ^2_μ. The optimum then shifts to $z_o > 0$, such that $w = e^{-(z-z_o)^2/2V_S}$. During adaptation, $\bar{z}$ approaches z_o due to allele frequency change and new mutations. At any point during adaptation, mutation with effect sizes $\gamma > (z_o - \bar{z})/2$ will overshoot the optimum if they reach high frequency or fix, and are therefore deleterious.

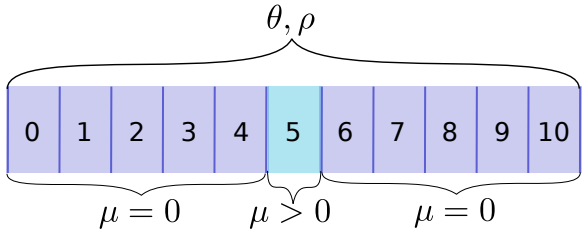
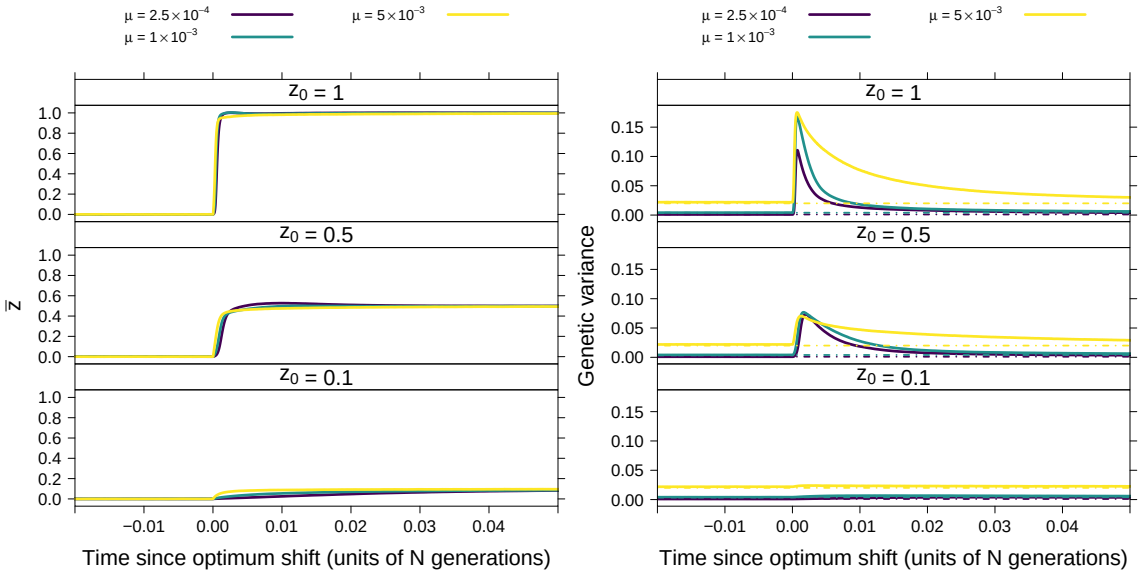


Figure 2: Schematic of a single locus for multi-locus simulations. The scaled neutral mutation and recombination rates, θ and ρ, respectively, apply to an entire locus. Neutral mutations and recombination events occur uniformly along a locus. For window-based analyses, locus is broken up into 11 equal-sized windows. Mutations affecting trait value only occur in the middle (sixth) window, at rate $\mu > 0$ per gamete, per generation. For simulations with multiple loci, the depicted layout is repeated multiple times and each locus is separated by 50 centiMorgans.



(a) Mean trait value.

(b) Mean genetic variance.

Figure 3: (3a) Mean trait value over time. (3b) Mean genetic variance over time. The dot-dashed lines correspond to $4\mu V_S$, which is the equilibrium genetic variance expected under the House-of-Cards model (Turelli, 1984).

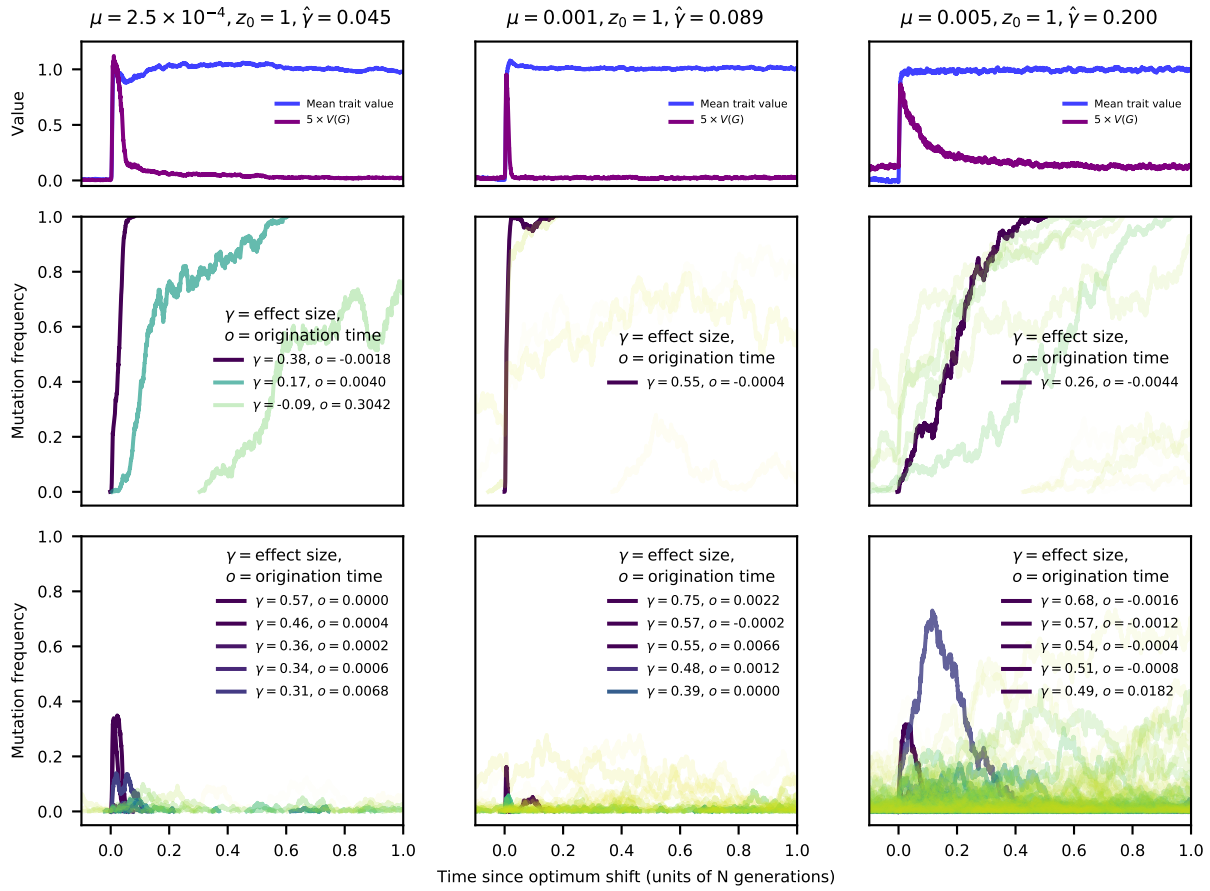


Figure 4: Trajectories of selected mutations. The three columns show results from a single simulation replicate for low, moderate, and high-mutation rate simulations, respectively. Parameter values are at the top of each column. The first row of plots shows the trait value and the genetic variance (multiplied by a constant for plotting purposes) over time, for up to N generations post optimum shift. As in Figure 3, the populations adapt quickly to the new optimum of $z_o = 1$. The middle row shows the frequency trajectories of fixations. Solid, darker (purple/blue) colors reflect larger effects on trait values, and more transparent colors in the green/yellow range reflect smaller effect sizes. Fixations of large effect on fitness ($|\gamma| > \hat{\gamma}$) are indicated in the legend. The bottom row shows the frequency trajectories of mutations that are eventually lost. The coloration is as for the fixations, and any mutations that did not reach a frequency of one percent are excluded. A maximum of five mutations, corresponding to the five largest $|\gamma|$ are included in the legend in the final row. Figure S1 shows the same data on a smaller time scale, showing the details on allele frequency change during the rapid adapation to the new optimum.

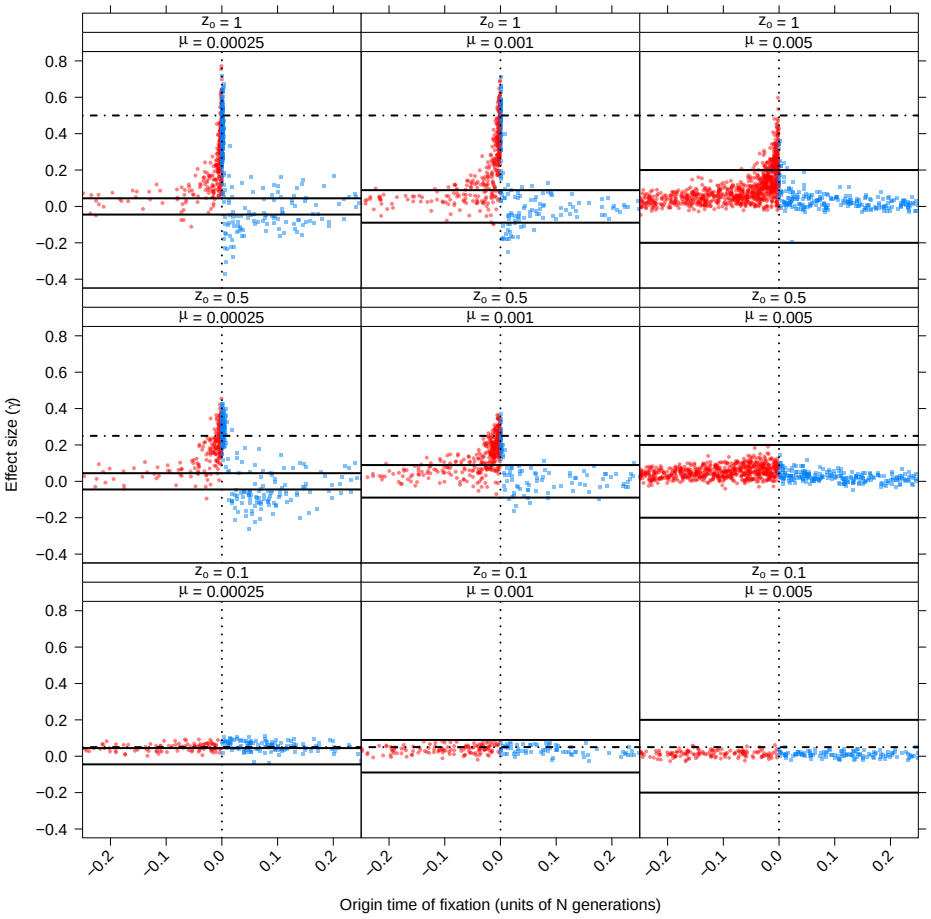


Figure 5: Fixations with large effects on trait values typically arise close to the time of the optimum shift. The x axis shows the origin time of each fixation and the y axis is that fixation's effect size, γ. Time is measured in units of N generations an scaled so that the optimum shift occurred at time zero. Red points are origin times prior to the optimum shift and are therefore fixations arising from standing variation. Blue points are fixations from new mutations. In each panel, the dot-dashed line is placed at $z_o/2$ and the solid lines are placed at $\pm\hat{\gamma}$ (de Vladar and Barton, 2014). Fixations of large effect on fitness ($|\gamma| > \hat{\gamma}$) typically arose near the time of the optimum shift and some populations fix mutations that over-shoot z_o, meaning that the effect size of the fixation is $\gamma > z_o/2$. Both phenomena are more pronounced as the optimum shift is more extreme (z_o larger) and the trait more mutation limited (μ smaller). For parameter combinations where over-shooting the optimum occurs, fixations from new mutations with $\gamma < 0$ are observed, many of which are of large effect when the trait is mutation-limited.

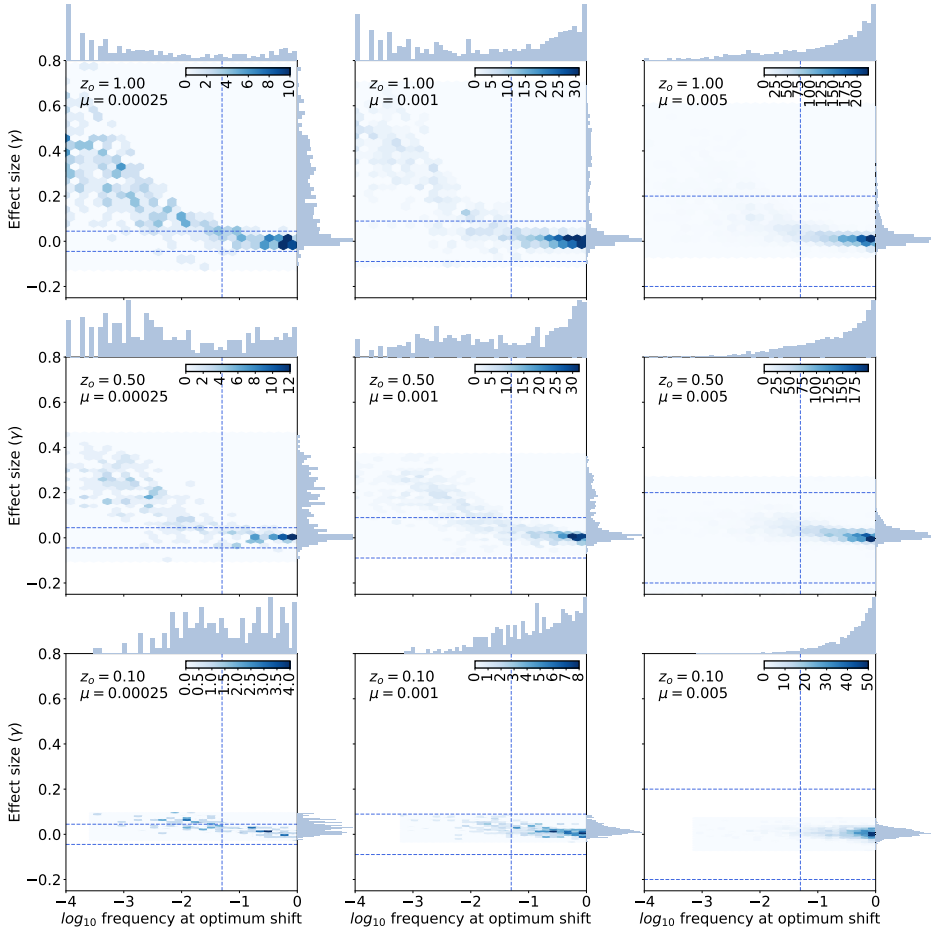


Figure 6: The joint distribution of frequency at the time of the optimum shift and effect size (γ) for fixations from standing variation. Each panel shows the joint density of the mutation frequency at the time of the optimum shift and effect size, using the raw data from Figure 5 The vertical dashed line represents a frequency of 0.05, which Przeworski *et al.* (2005) identified by simulation as a rough threshold below which sweeps from standing variation will show a reduction in levels of neutral sequence diversity at linked sites, provided that they are associated with large effect on fitness. The horizontal dashed lines are placed at $\pm\hat{\gamma}$ de Vladar and Barton (2014). As the trait is more mutation-limited (smaller μ) and/or the optimum shift more extreme (z_o larger), there is more of a tendency for sweeps from standing variation to be chosen from rarer mutations with larger effects on trait values. When sweeps from standing variation are from more common mutations, the effect sizes, are typically small ($|\gamma| < \hat{\gamma}$).

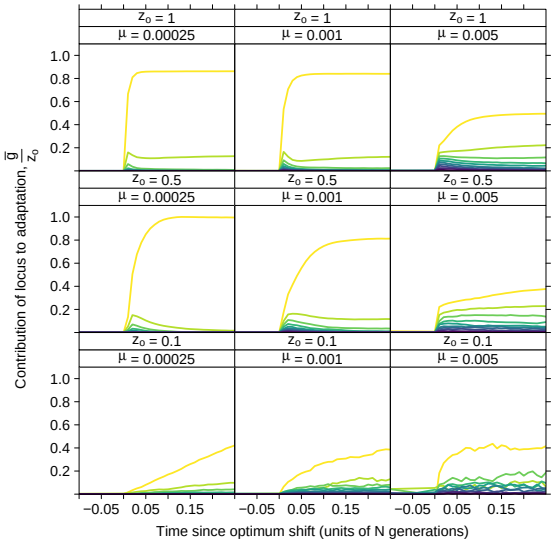


Figure 7: The contribution of individual loci to adaptation in a ten-locus system. For each locus, the average genetic value of an individual, due to mutations found at the locus, $\bar{g}$, is recorded and averaged across replicates before being divided by z_o, the optimum trait value. Each line is therefore the relative contribution of each locus towards adaptation, as a function of time since the optimum shift. The yellow-line represents the largest-effect locus, and colors darken through green to purple for loci with progressively smaller contributions to adaptation.

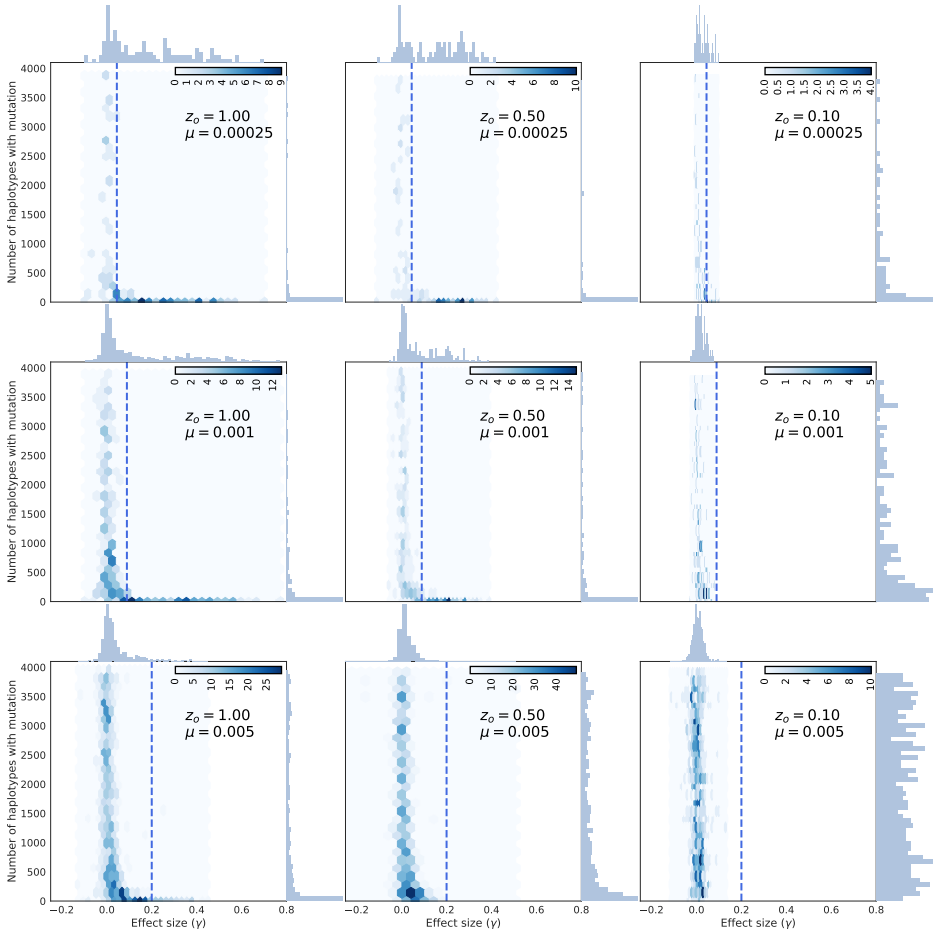
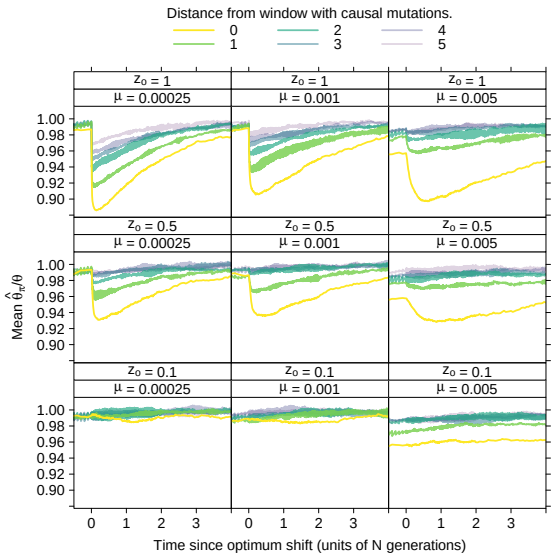
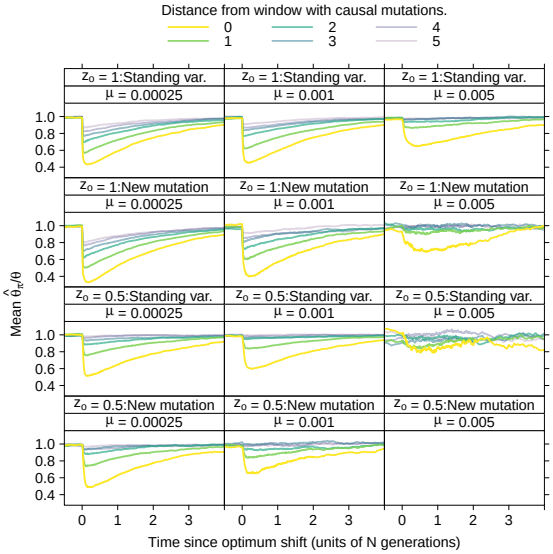


Figure 8: The number of haplotypes associated with fixations from standing variation of different effect sizes. Each panel shows the effect size of a fixation from standing variation (x axis) and the number of unique haplotypes in the entire population containing that mutation. The number of haplotypes for each mutation is taken immediately prior to the optimum shift and excludes any mutations that arose that generation. Thus, all mutations found on a single haplotype are more than one generation old. In each panel, the dashed vertical blue line corresponds to an effect size of $\hat{\gamma} = 2\sqrt{2}\sqrt{\mu \times V_S}$, above which a mutation has a large effect on fitness (de Vladar and Barton, 2014).

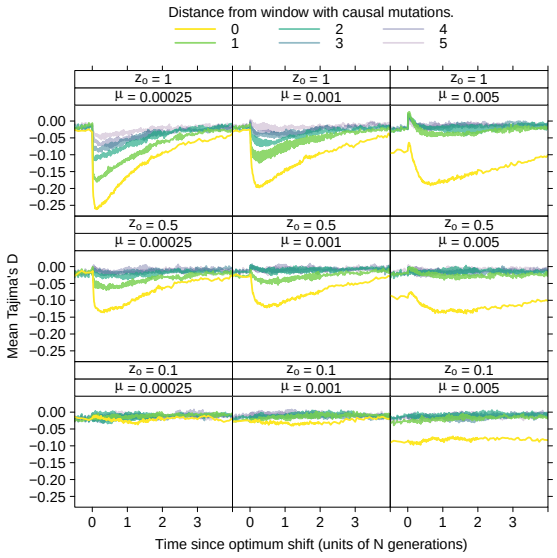


(a) Mean reduction in diversity over time.

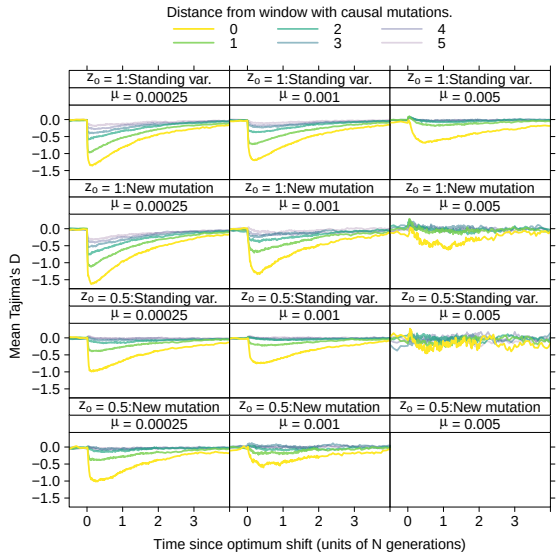


(b) Mean reduction in diversity over time, conditioning on large-effect substitutions.

Figure 9: The average reduction in genetic diversity due to the optimum shift. Here, diversity is measured as the mean number of pairwise differences between sequences (Tajima, 1983) divided by $\theta = 4N_e\mu$. 9a shows overall reduction in diversity, averaging across all ten loci and across all replicates. The mean value is shown separately for windows of increasing distance from the central window containing selected mutations (Figure 2). A distance of zero refers to the central window. There are an equal number of flanking windows on either side of the central window (Figure 2) and therefore a distance of one averages the results for the windows immediately to the left and right of the central window and a distance of two is averaging the next window on either side, etc.. 9b shows the mean reduction in diversity at loci experiencing a sweep due to a mutation of large effect, *i.e.* $\gamma \geq \hat{\gamma}$ (de Vladar and Barton, 2014). Sweeps from standing variants (those arising prior to the optimum shift) and new mutations arising after the shift are shown separately. There are no lines for the lower right panel because no large-effect fixations occur in that part of the parameter space. No panels are shown for an optimum trait value of $z_o = 0.1$ for the same reason.

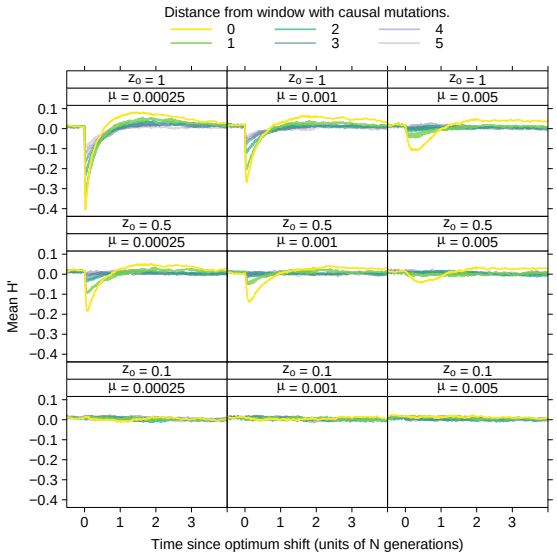


(a) Mean Tajima's D over time

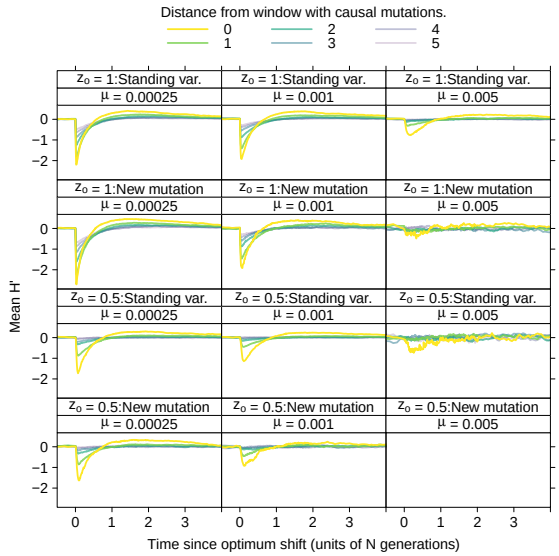


(b) Mean Tajima's D over time, conditioning on large-effect substitutions.

Figure 10: Tajima's (1989) D statistic over time. 10a shows the average value of Tajima's (1989) D over time. The data are shown separately for windows of different distances from the central window where mutations affecting the trait arise (Figure 2) 10b shows the mean value of D conditioning on a locus fixing a mutation with effect size $\gamma \geq \hat{\gamma}$ (de Vladar and Barton, 2014). These loci are separated by whether the fixation was from standing variation, meaning a mutation predating the optimum shift, or from a new mutation arising after the shift. Panels with no data imply that no fixations of large effet occurred.

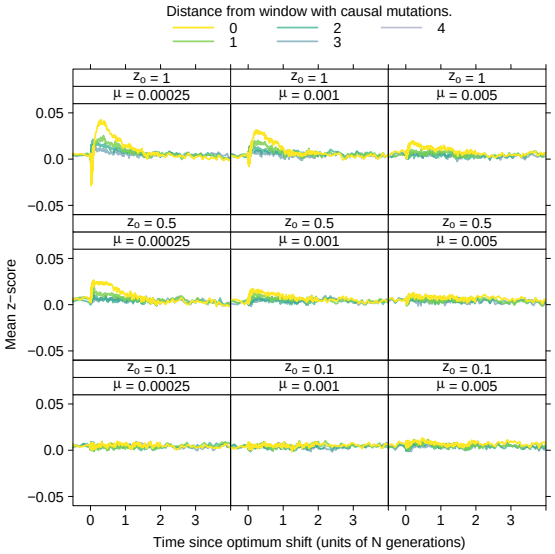


(a) Mean H' over time

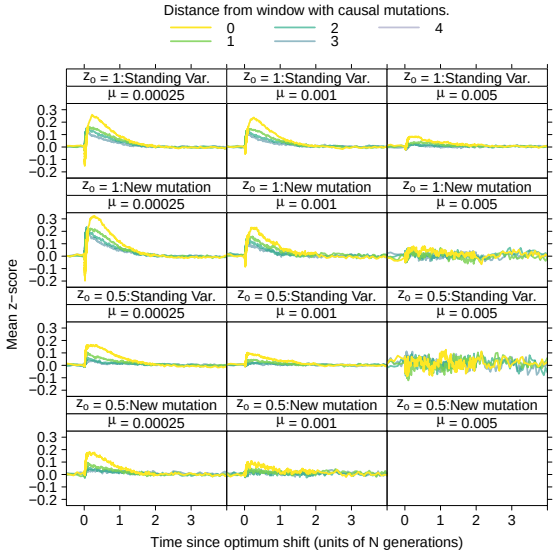


(b) Mean H' over time, conditioning on large-effect substitutions.

Figure 11: Average behavior of H' (Zeng *et al.*, 2006) over time. 11a shows the mean H' over time, averaging over both loci and over replicates. The data are plotted as a function of increasing window distance from where mutations affecting the trait occur (Figure 2). 11b shows the mean of H' over time, conditioning on the fixation of a large-effect mutation either from standing variation or from a new mutation arising after the optimum shift. Panels with no data imply that no fixations of large-effect variants occurred.

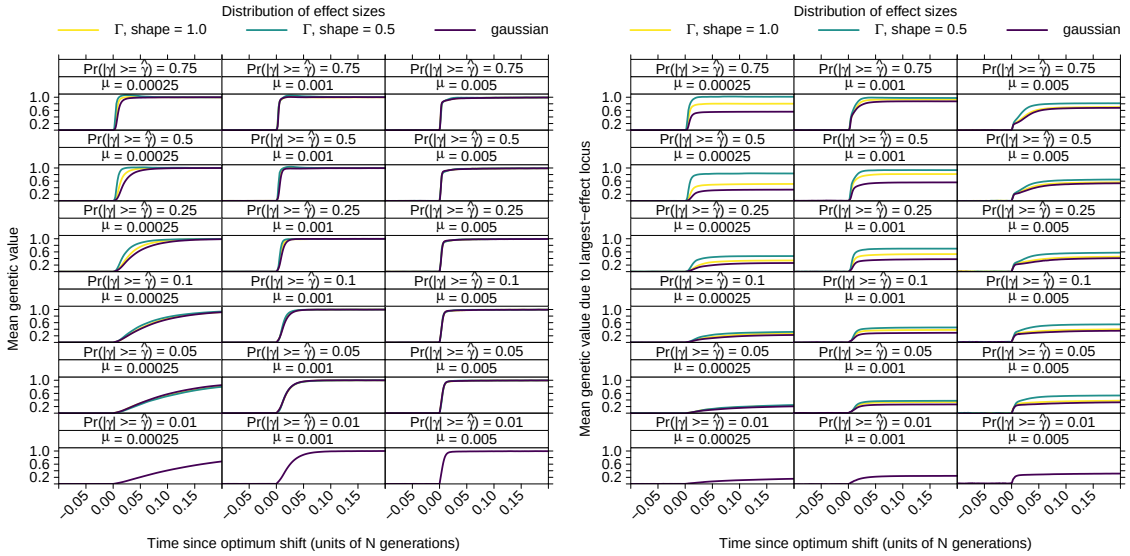


(a) Mean z-score over time



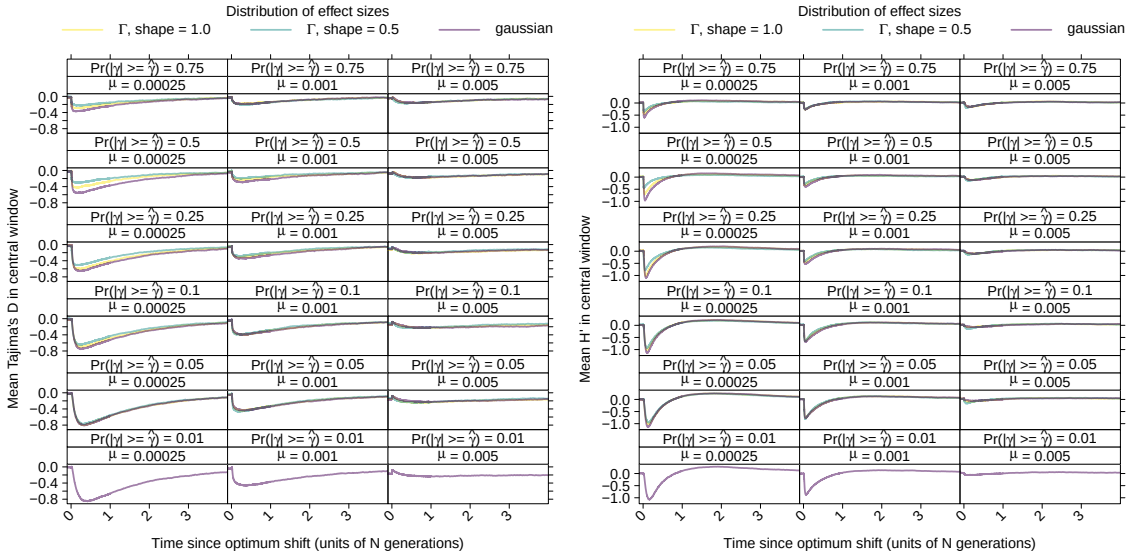
(b) Mean z-score over time, conditioning on large-effect substitutions.

Figure 12: The mean z-score corresponding to the nS_L statistic (Ferrer-Admetlla *et al.*, 2014) over time. The data are plotted for different windows according to the scheme shown in Figure 2. 12a shows the average value over time and 12b shows the average value conditioning on fixation of large effect from standing variation or from a new mutation. Panels with no data imply that no fixations from large-effect mutations occurred.



(a) (b)

Figure 13: The average dynamics of adaptation as a function of the fraction of newly-arising mutations of large effect. For a given mutation rate, μ, the probability that a new mutation is of large effect, $Pr(|\gamma| \geq \hat{\gamma})$, is varied for three different distributions of effect sizes (DES), the Gaussian distribution and Gamma distributions with two difference shape parameters. 13a The mean trait value over time. 13b The mean genetic value at the largest-effect locus.



(a) (b)

Figure 14: The average deviation from equilibrium patterns of variation as a function of the probability that a new mutation is of large effect. As in Figure 13, the mutation rate is held constant for three different distributions as $Pr(|\gamma| \geq \hat{\gamma})$ varies. 14a The mean of Tajima's (1989) D over time in the central window of a locus, which is where mutations affecting the trait arise (Figure 2). 14b The mean H' (Zeng *et al.*, 2006) over time in the central window of a locus.

Supporting Information

			Large effect ($\hat{\gamma} \leq \gamma < z_o/2$)			Small effect ($0 \leq \gamma < \hat{\gamma}$)		
μ	$\hat{\gamma}$	z_o	Proportion	$\bar{\gamma}$	Mean $2Ns_0$	Proportion	$\bar{\gamma}$	Mean $2Ns_0$
2.5×10^{-4}	0.045	0.1	0.008	0.047	36.08	0.071	0.022	19.79
		0.5	0.27	0.139	581.81			108.43
		1.0	0.41	0.208	1702.71			218.11
1×10^{-3}	0.089	0.1	0	NA	NA	0.140	0.044	34.39
		0.5	0.201	0.164	662.45			209.22
		1.0	0.337	0.237	1886.47			423.44
5×10^{-3}	0.20	0.1	0	NA	NA	0.29	0.095	49.72
		0.5	0.053	0.224	832.97			419.92
		1.0	0.189	0.311	2312.84			863.37

Table S1: Simulation parameters and relationship to proportions of large- and small- effect mutations and initial strengths of selection.

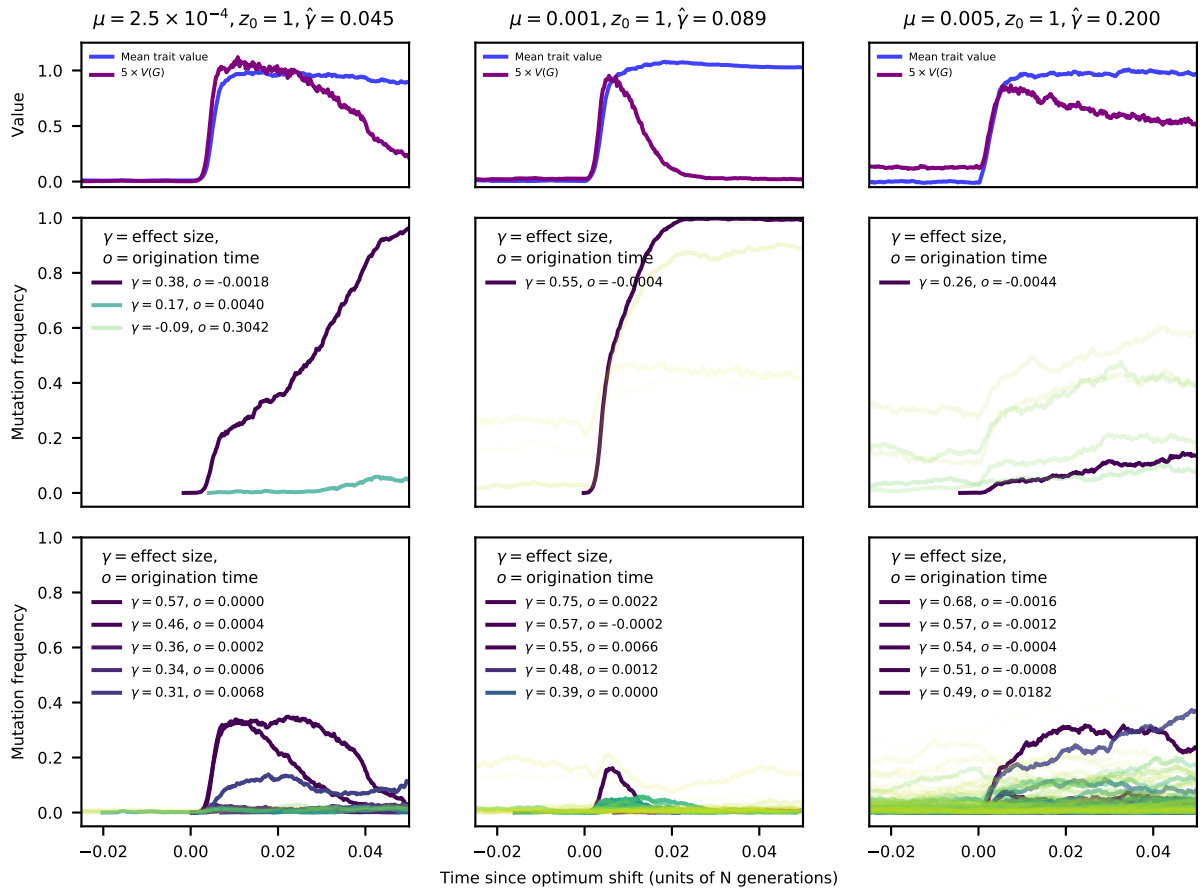


Figure S1: Trajectories of selected mutations, shown over a short time scale. The data and colors are identical to Figure 4.

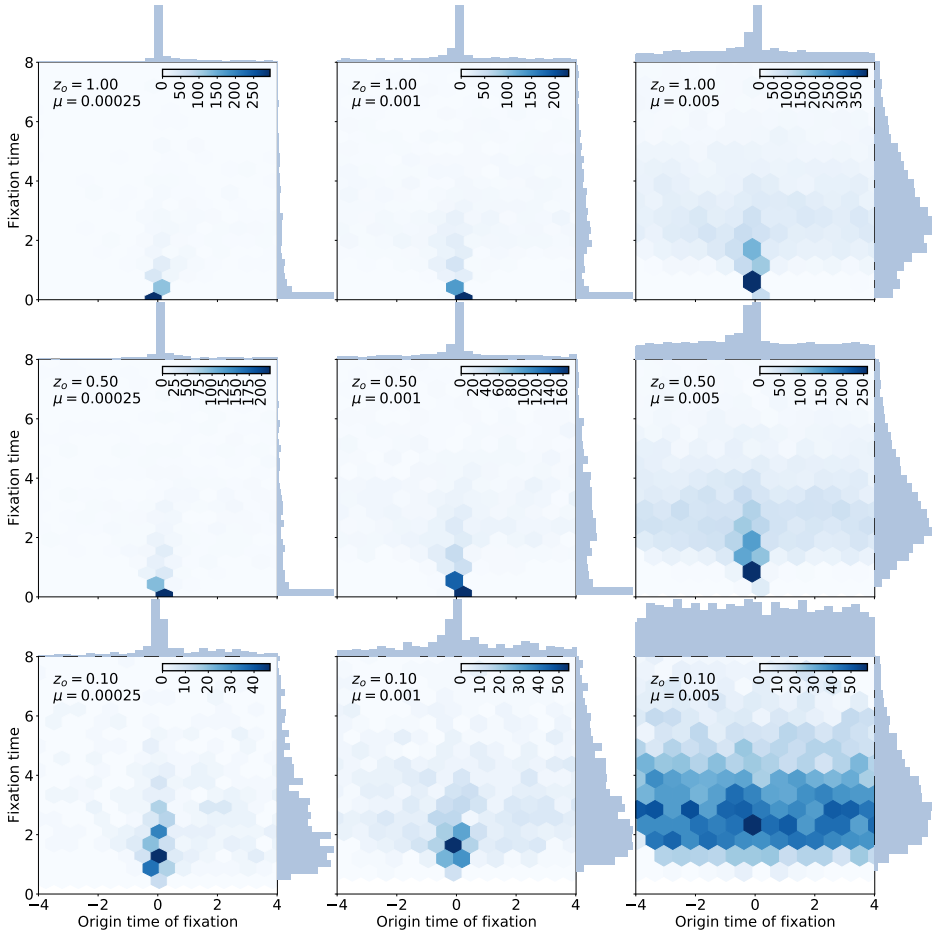


Figure S2: The time to fixation as a function of origin time. The x axis shows the origin time of each fixation in units of N generations and scaled such that the optimum shift is at time zero. The y axis is the fixation time, again in units of N generations. Each panel shows the joint density of origin and sojourn time as well as the marginal distributions, and the raw data are the same as in Figure 5. As the trait is more mutation-limited (smaller μ) and/or the optimum shift more extreme (larger z_o), the fixation time decreases because the effect sizes increase (Figure 5).

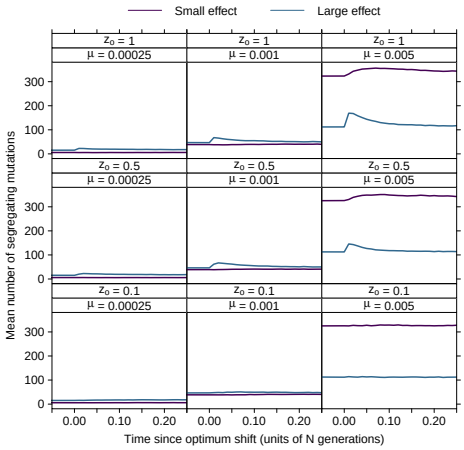
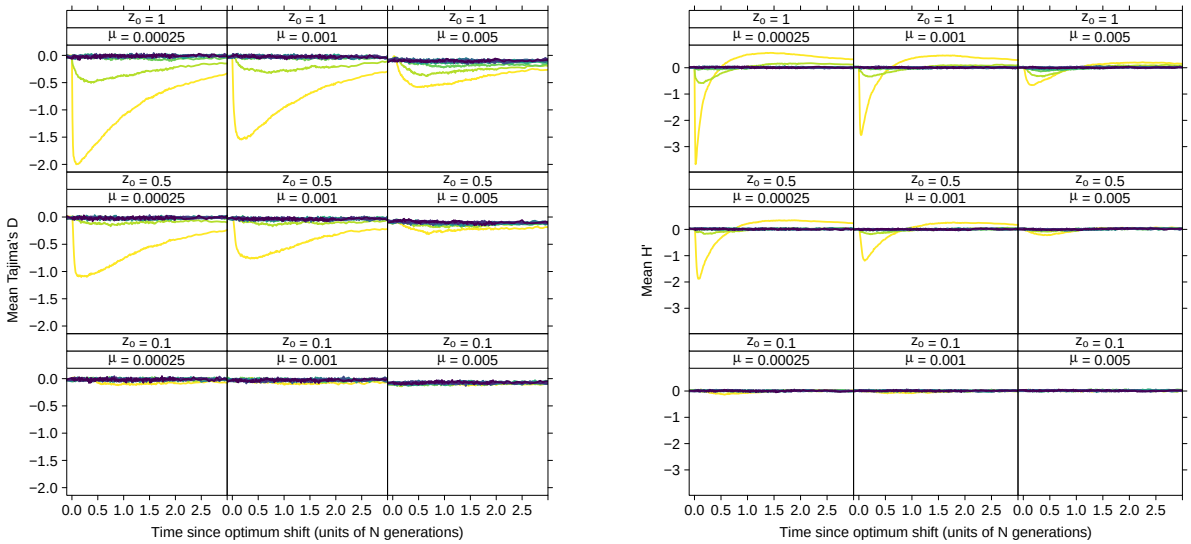


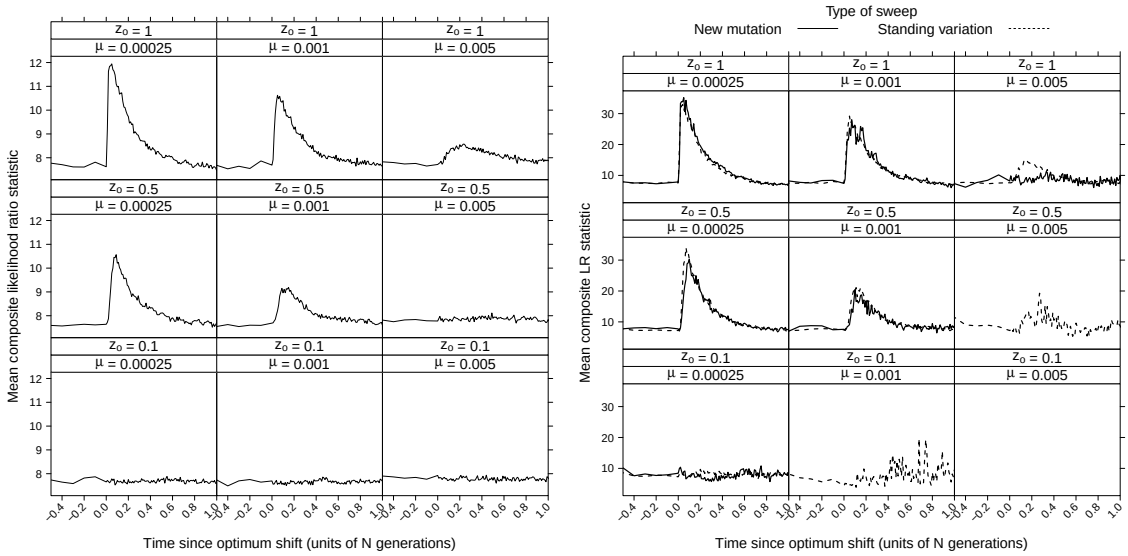
Figure S3: The mean number of segregating mutations affecting the trait over time. The different lines are for mutations of small effect ($|\gamma| < \hat{\gamma}$) and large-effect variants ($|\gamma| > \hat{\gamma}$)



(a) Mean of Tajima's D over time

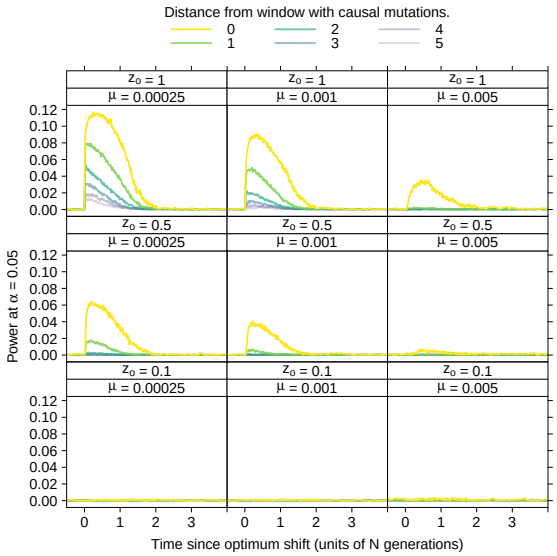
(b) Mean H' over time

Figure S4: Average patterns of hitch-hiking as a function of a locus' contribution to adaptation. The average values of Tajima's (Tajima, 1989) D (S4a) and the H' statistic (Zeng *et al.* (2006), S4b) are shown for the central window (where mutations affecting trait values occur) of each locus, and loci are ranked by their contribution to adaptation over time (see Figure 7). As in Figure 7, brighter colors label the loci contributing the most to adapation and colors darken as loci make less of a contribution. The pattern of hitch-hiking is strongest for the loci contributing the most to adaptation (yellow) and departures from the null model will be generally undetectable for the second-largest-effect locus (green). No data are shown for the case of $z_o = 0.1$ because there is no discernable pattern of hitch-hiking for any parameter combination (as in Figures 10-12).

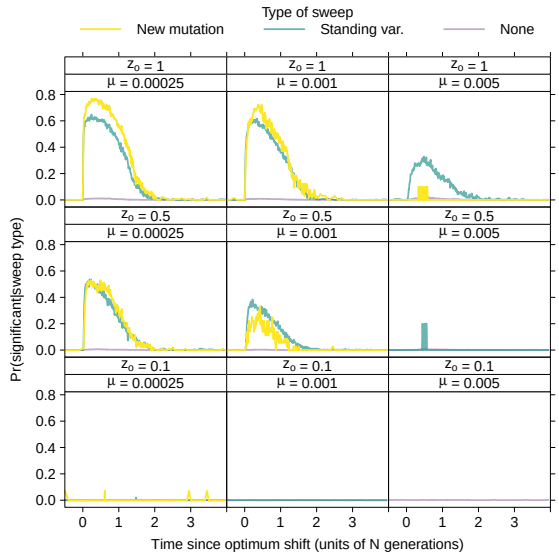


(a) (b)

Figure S5: Temporal behavior of the composite likelihood ratio (CLR) statistic (Nielsen *et al.*, 2005) over time. All results were obtained using SweeD (Pavlidis *et al.*, 2013). S5a The average composite likelihood ratio statistic over time. S5b The average CLR over time, conditional on a fixation of large effect. The mean CLR is shown separately for loci with large-effect fixations from new mutations and from standing variation.

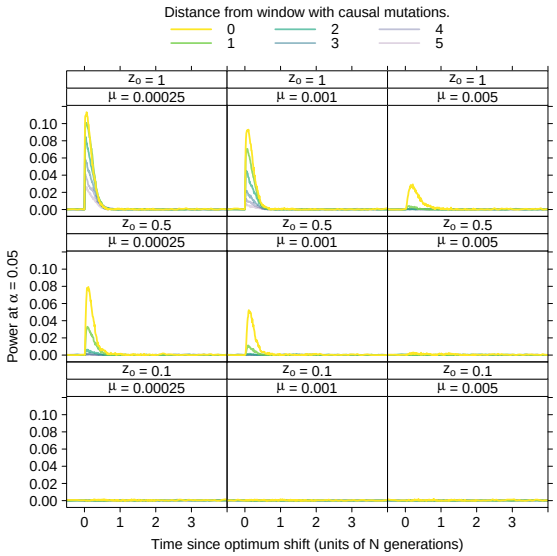


(a) Power of Tajima's D

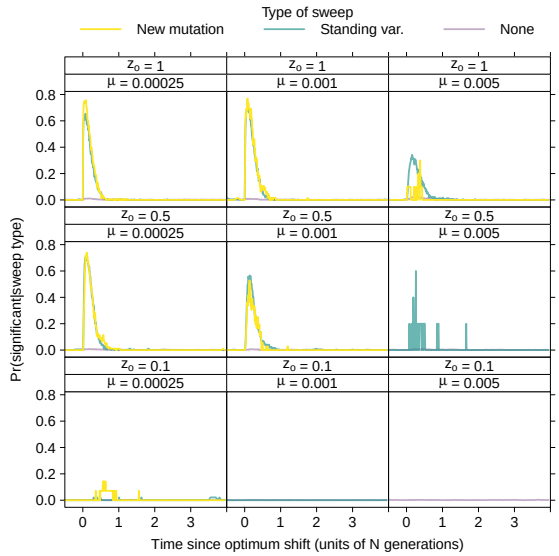


(b) Fraction of significant loci with large-effect fixations.

Figure S6: Power of Tajima's (1989) to reject the standard neutral model. S6a The power to reject the null model in windows of increasing distance from selected site as a function of time since the optimum shift, corrected so that the per-window rejection rate under the null model is 0.05. The windows are labelled using the same scheme as Figure 10. S6b The probability of a "significant" window at a locus, conditional on the fixation of a large-effect mutation at the locus.

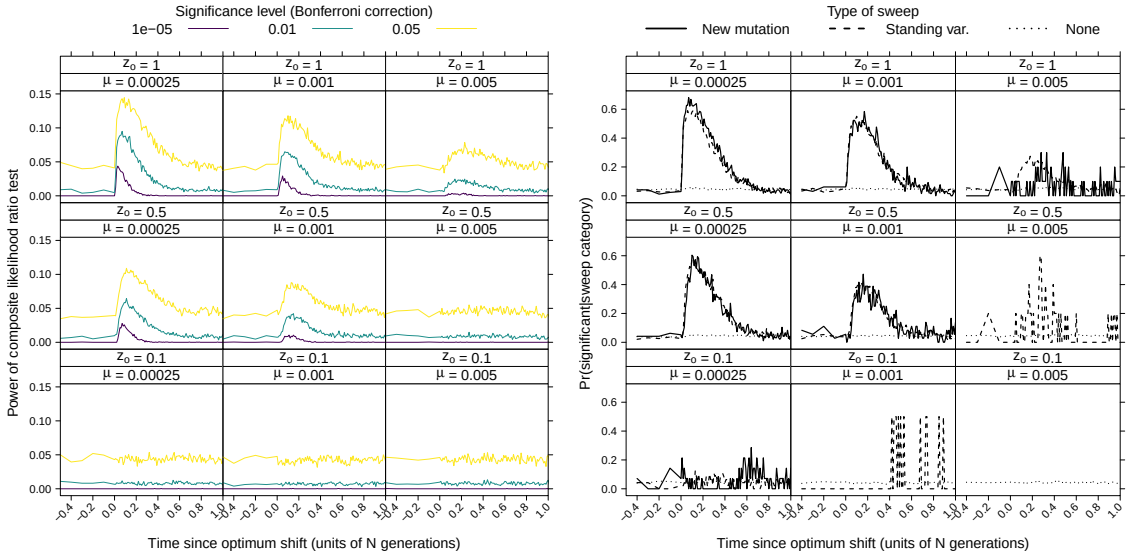


(a) Power of H'



(b) Fraction of significant loci with large-effect fixations.

Figure S7: Power of H' (Zeng *et al.*, 2006) to reject the standard neutral model. S7a The power to reject the null model in windows of increasing distance from selected site as a function of time since the optimum shift, corrected so that the per-window rejection rate under the null model is 0.05. The windows are labelled using the same scheme as Figure 11. S7b The probability of a "significant" window at a locus, conditional on the fixation of a large-effect mutation at the locus.



(a) (b)

Figure S8: Power of the composite likelihood ratio test (CLRT) to reject the null model. All results are based on 10^6 coalescent simulations using `msprime` (Kelleher *et al.*, 2016) and analyzed using SweeD (Pavlidis *et al.*, 2013). S8a The power to reject the null model at three different significance levels, after application of a Bonferonni correction. S8b The probability of a significant CLRT, conditional on the type of sweep.

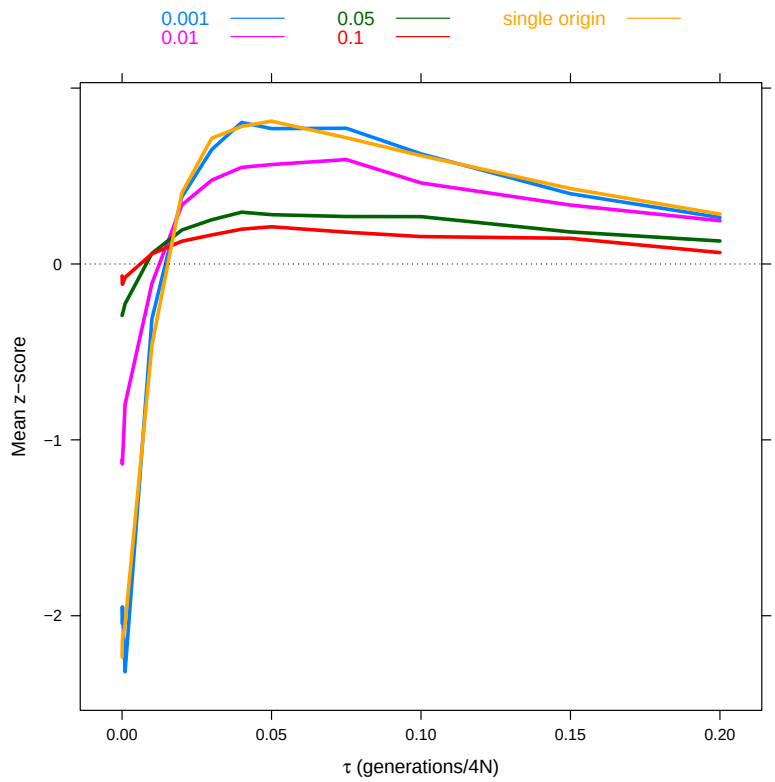
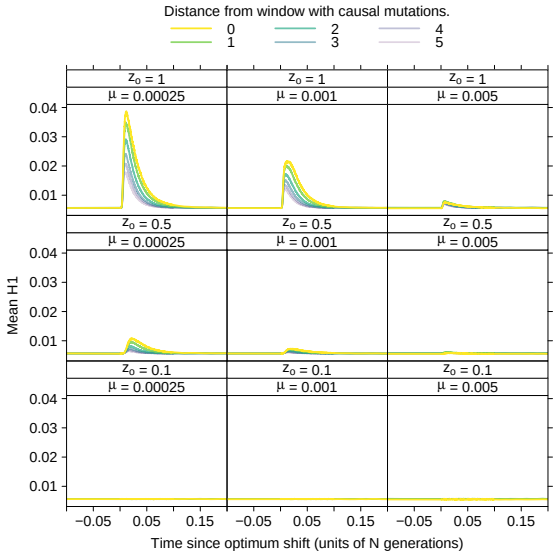
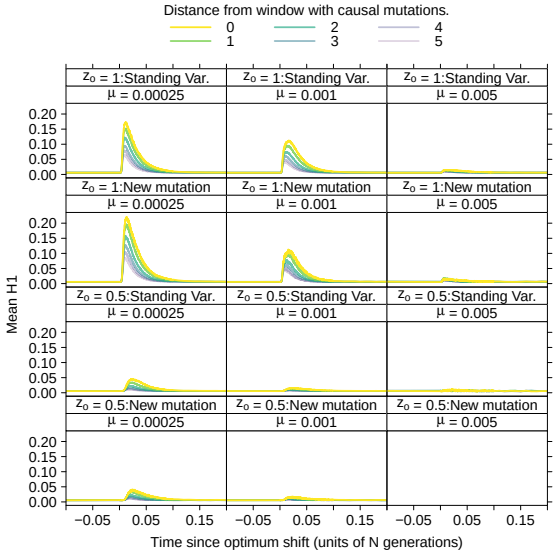


Figure S9: The average z-score associated with the nS_L statistic (Ferrer-Admetlla *et al.*, 2014) as a function of the time since the fixation of a strongly-selected mutation. All lines are based on 100 replicates of a coalescent simulation (Kern and Schrider, 2016) of a sample size of 100 haplotypes and a region of size $\theta = \rho = 100$. A single sweep with strength of selection $\alpha = 2Ns$ occurs in the middle of the region. The time since the fixation (τ) and the initial starting frequency of the mutation vary. A sweep from a single origin is a "hard sweep" with initial frequency $1/(2N_e)$.

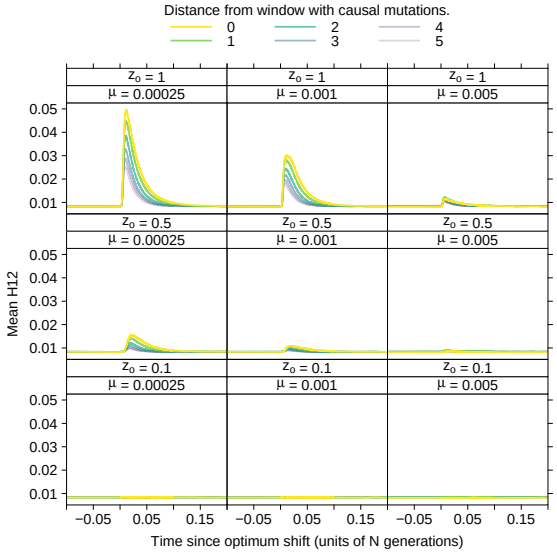


(a) Mean $H1$ over time

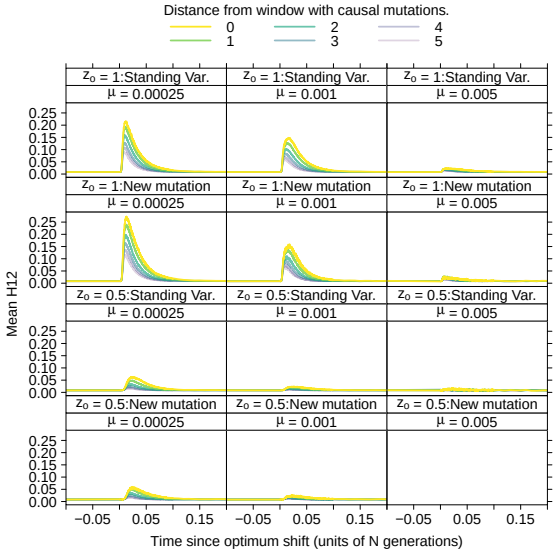


(b) Mean $H1$ over time, conditioning on large-effect substitutions.

Figure S10: The mean $H1$ statistic (Garud *et al.*, 2015) over time. S10a shows the average value over time and S10b shows the average value conditioning on a fixation of large effect either from standing variation or from a new mutation.

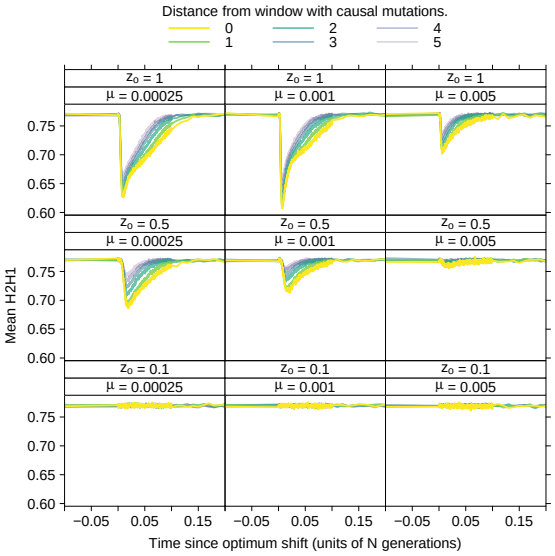


(a) Mean $H12$ over time

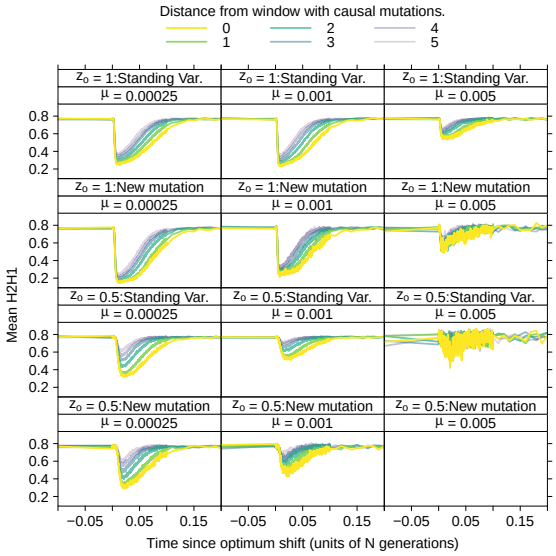


(b) Mean $H12$ over time, conditioning on large-effect substitutions.

Figure S11: The mean $H12$ statistic (Garud *et al.*, 2015) over time. S11a shows the average value over time and S11b shows the average value conditioning on a fixation of large effect either from standing variation or from a new mutation.



(a) Mean $H2/H1$ over time



(b) Mean $H2/H1$ over time, conditioning on large-effect substitutions.

Figure S12: The mean $H2/H1$ statistic (Garud *et al.*, 2015) over time. S12a shows the average value over time and S12b shows the average value conditioning on a fixation of large effect either from standing variation or from a new mutation.

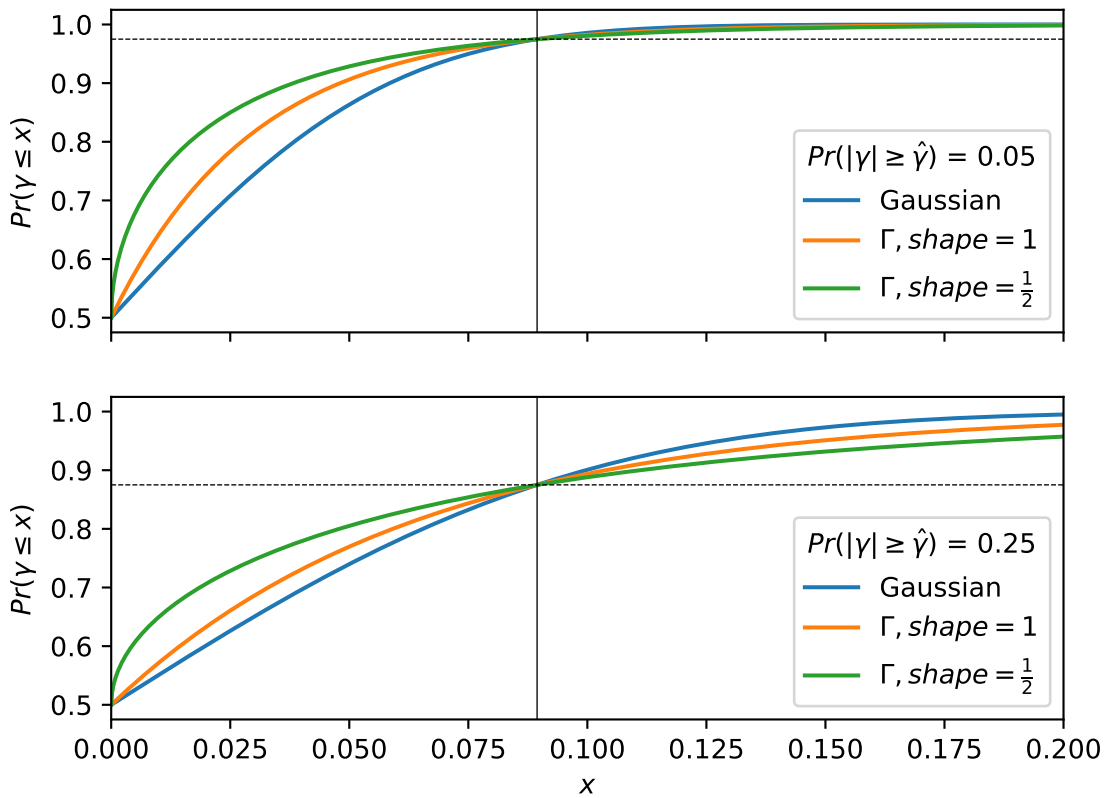
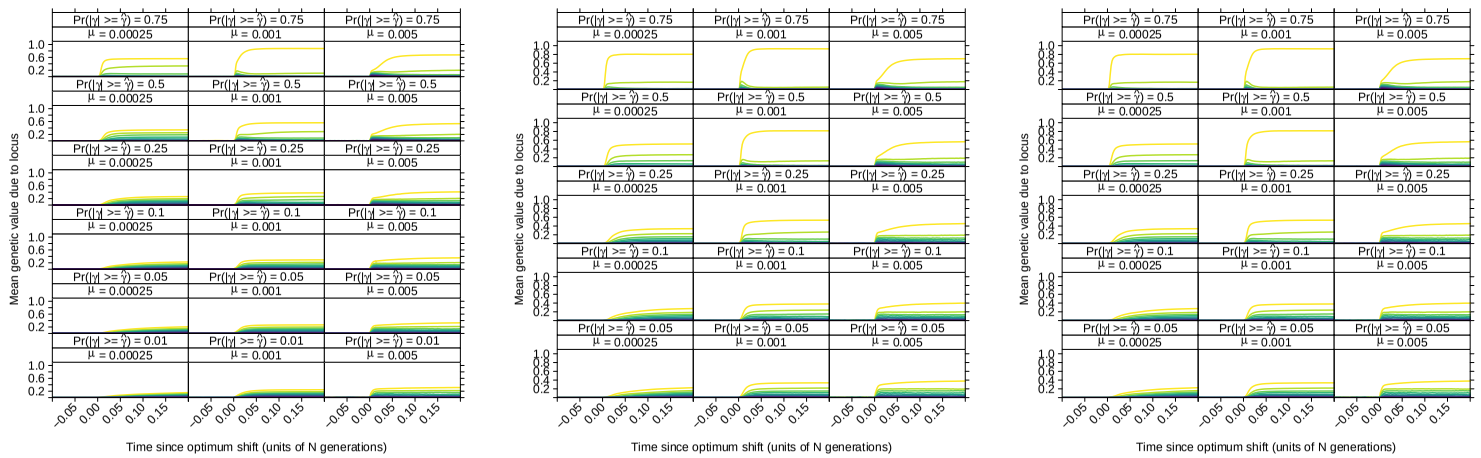
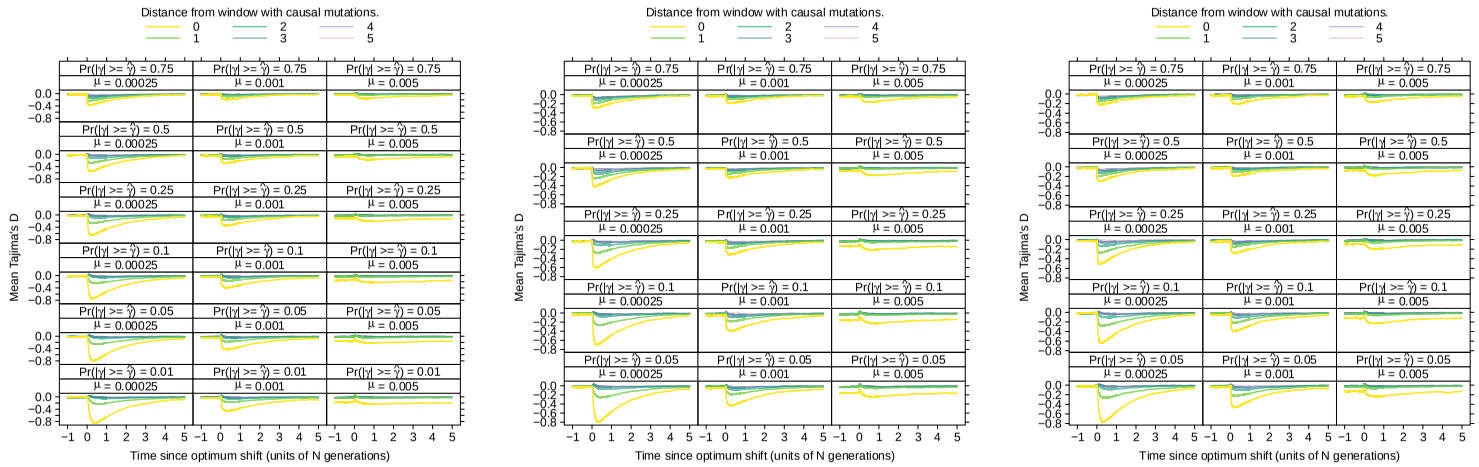


Figure S13: Examples of distribution of fitness effects (DES) when the probability that a new mutation is of large effect, $Pr(|\gamma| \geq \hat{\gamma})$, is held constant. The two panels show the cumulative density functions of three different DES for two different probabilities of large-effect variants, 0.05 and 0.25. The vertical line corresponds to $\hat{\gamma}$ (de Vladar and Barton, 2014), assuming $\mu = 10^{-3}$ and $V_S = 1$. The dashed horizontal line is placed at $1 - Pr(|\gamma| \geq \hat{\gamma})/2$. For simulations, I use DES that are symmetric around a value of $x = 0$ such that the probability that $Pr(|\gamma| \geq \hat{\gamma})$ is fixed at a specific value.



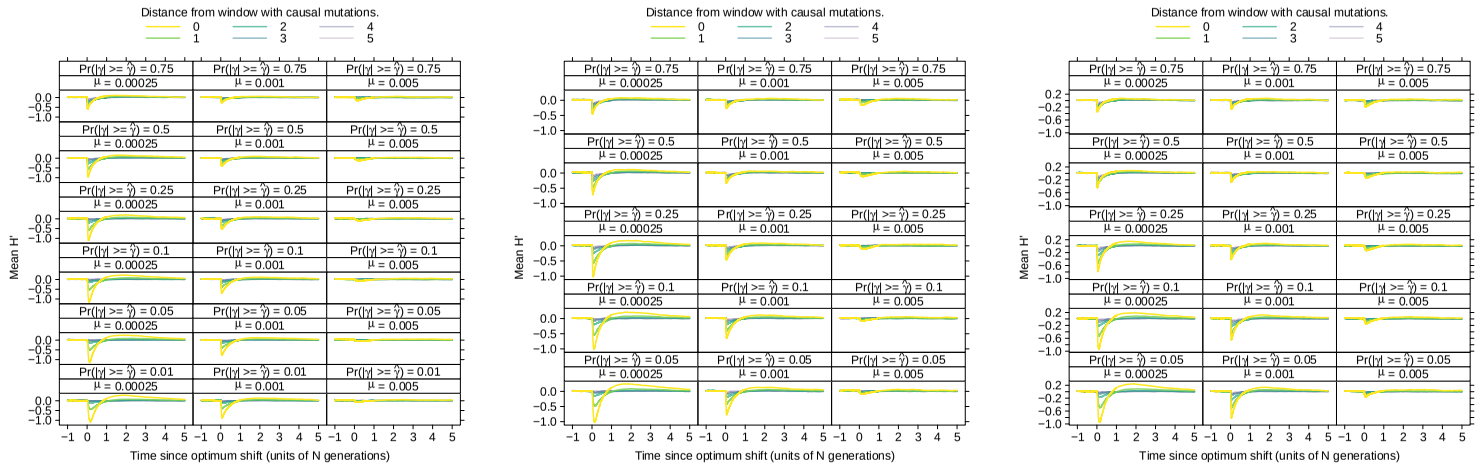
(a) (b) (c)

Figure S14: The contribution of individual loci to adaptation in a ten-locus system, shown for different distributions of effect sizes and different probabilities that a newly-arising mutation has a large effect. The layout of the figure is the same as in Figure 7 and the data are the same as shown in Figure 13a. The difference from Figure 13a is that all loci are now shown, with the yellow line representing the largest-effect locus and colors darkening from green to purple for loci with smaller contributions to adaptation. S14a Gaussian DES. S14b Gamma DES with shape parameter equal to one. S14c Gamma DES with shape parameter equal to one-half.



(a) (b) (c)

Figure S15: The mean Tajima's (1989) D over time, shown for different distributions of effect sizes and different probabilities that a newly-arising mutation has a large effect. S15a Gaussian distribution of effect sizes. S15b Gamma distribution of effect sizes with shape parameter equal to one. S15c Gamma distribution of effect sizes with shape parameter equal to one-half.



(a) (b) (c)

Figure S16: The mean H' (Zeng *et al.*, 2006) over time, shown for different distributions of effect sizes and different probabilities that a newly-arising mutation has a large effect. S16a Gaussian distribution of effect sizes. S16b Gamma distribution of effect sizes with shape parameter equal to one. S16c Gamma distribution of effect sizes with shape parameter equal to one-half.

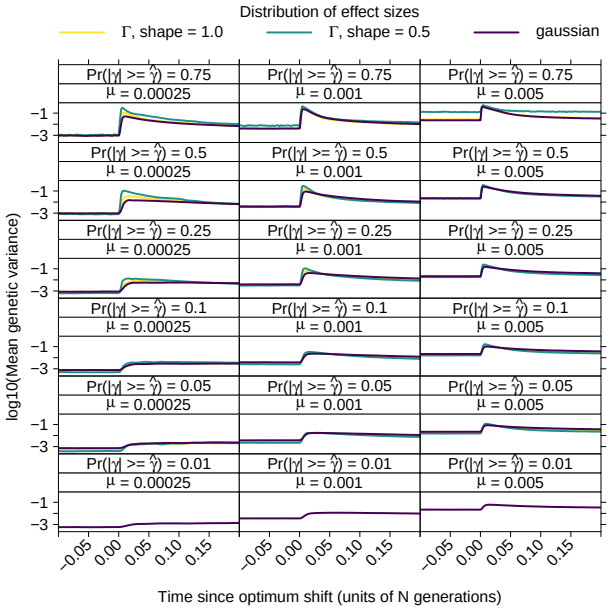


Figure S17: . The average genetic variation over time as a function of the probability that a new mutations is of large effect.