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

August 26, 2019

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, when phenotypes adapt to a new optimum trait value, the strength of selection on individual mutations decreases as the population adapts. Here, I use explicit forward simulations of a single trait with additive-effect mutations adapting to an optimum shift. Detectable "hitch-hiking" patterns are only apparent if i. the optimum shifts are large with respect to equilibrium variation for the trait, ii. traits are mutation-limited, and iii., large-effect mutations rapidly increase in frequency and eventually reach fixation, which typically occurs after the population reaches the new optimum. For the parameters simulated here,

partial sweeps do not appreciably affect patterns of linked variation, even when the mutations are strongly selected. For a given mutation rate, the number of sweeps from new mutations and from standing variation are approximately equal, even if the total mutational variance is high. Patterns of hitch-hiking are similar on average for the two classes of sweeps because sweeps standing variation involving large-effect mutations are rare when the optimum shifts. The distribution of effect sizes of new mutations has little effect on the time to reach the new optimum, but reducing the mutational variance increases the magnitude of hitch-hiking patterns. In general, populations reach the new optimum prior to the completion of any sweeps, and the times to fixation are longer for this model than for standard models of directional selection. The long fixation times are due to a combination of declining selection pressures during adaptation and the possibility of interference among weakly selected sites for traits with high mutation rates.

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 (Maynard-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 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).

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; Kimura, 1981). This model has been extended to multiple mutations in linkage equilibrium by several authors (Barton, 1986; de Vladar and Barton, 2014; Jain and Stephan, 2015, 2017b).

The equilibrium conditions of models of Gaussian 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). It is also common

to assume that the forwards and backwards mutation rates per locus are equal (Barton, 1986; de Vladar and Barton, 2014; Jain and Stephan, 2015, 2017b). Under these assumptions, and assuming distributions of mutational effects symmetric around zero, large-effect variants (*e.g.* those with effect sizes greater than $2\sqrt{2}\sqrt{V_S}$, where V_S is the variance of the Gaussian fitness function) will be near the boundaries while small-effect variants will be at frequencies near one-half (de Vladar and Barton, 2014; Jain and Stephan, 2017b). Here, "large" and "small" effect is with respect to the effect of a variant on the genetic load of a population (de Vladar and Barton, 2014).

The effect individual mutations on trait values affects their fixation probabilities and the time required to change the population's mean phenotype (Robertson, 1960). When most mutational effects are small and additive, fixations require on the order of the population size in generations, and most phenotypic change occurs through fixation (Robertson, 1960). Further, if the total intensity of selection on a trait is weak, then only large-effect mutations or mutations at high frequency would be expected to fix (Robertson, 1960). Recent theoretical work has attempted to clarify when sweeps should happen and when adaptation should proceed primarily via subtle allele frequency shifts. 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 contribute to adaptation. Under models of linkage equilibrium, additive mutational effects, and equal rates of forward and back mutation at a biallelic locus (Barton, 1986; de Vladar and Barton, 2014), polygenic traits adapt quickly to a sudden shift in the optimum via directional selection (Jain and Stephan, 2017b). In an infinitely large population, mutations that are rare at the time of the optimum shift may fix if their effect sizes are not overly large relative to the magnitude of the shift. The number of large-effect sweeps during adaptation depends on the magnitude of the shift and the average effect size of segregating

variants (Jain and Stephan, 2017b). After the directional phase, selection becomes disruptive and mutations affecting fitness are fixed or lost in order to reduce the genetic load of the population.

Under a model of a trait with a small number of phenotypic classes, Höllinger *et al.* (2019) 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 (defined as when mean fitness has recovered following the optimum shift) proceeds via frequency shifts at many loci.

While the work described above identifies the conditions where 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 among 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). Höllinger *et al.* (2019) were able to accommodate partial linkage by simplifying how mutations affect phenotype and focusing on the dynamics up until a particular mean trait value was first reached. 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. By restricting mutations affecting the trait to specific

"loci" (within which linkage is still relatively loose), and allowing neutral mutations to occur over much larger genomic intervals containing the loci, I describe the physical distances over which hitch-hiking during polygenic adaptation leaves detectable signatures. The simulations conducted here are therefore analogous to those used to study the spatial dynamics of linked selection via the structured coalescent (Kaplan *et al.*, 1988; Braverman *et al.*, 1995; Kim and Stephan, 2002; Przeworski, 2002). The key conceptual difference is that the model of adaptation is changed from constant directional selection to the sudden optimum shift models involving a continuous trait considered in de Vladar and Barton (2014); Jain and Stephan (2015, 2017b). I also investigate the effect of the recombination rate on the time to adaptation and the fixation time of beneficial mutations with respect to the mean time required to adapt to the new optimum.

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 haploid genome 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_γ. Mutations have additive effects on trait value and therefore an individual's genetic value, z, is the sum of all effect sizes in that individual.

Here, I use the term "locus" to refer to a continuous genomic region within which mutation and recombination events occur uniformly. Within a locus, mutations occur at positions according to a uniform contiguous distribution according to an infinitely many sites scheme. Thus, each mutation results in a biallelic variant and, in the case of trait-affecting mutations, the derived allele affects trait values. What I refer to here are mutations are typically referred

to as loci in much of the theoretical literature (Turelli, 1984; Barton, 1986; Robertson, 1956, 1960; de Vladar and Barton, 2014; Jain and Stephan, 2015, 2017b).

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. 160). 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 $z_o > 0$.

It is important to note that I am considering all of the heritable variation for the trait to be modeled in the genomic regions that are explicitly simulated. Thus, the approach is similar in spirit to that of de Vladar and Barton (2014), but with partial linkage. An alternative would be to allow for a genetic background that also evolves, for which we are not tracking mutation fates. Chevin and Hospital (2008) used the latter approach to mathematically model the dynamics of large-effect mutations in an infinitesimal background and Stetter *et al.* (2018) used a simple version of this method to simulate the dynamics of quantitative traits evolving under truncation selection.

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}, which is the total mutation rate per haploid genome. The total mutation rate per diploid, U, is 2μ. The three post-shift optima used were $z_o = 0.1$, 0.5, and 1. For all combinations of μ and z_o, $\sigma_\gamma = 0.25$. The mutational variance of the trait is therefore $2\mu\sigma_\gamma^2$, or 3.25×10^{-5}, 1.25×10^{-4}, or 6.25×10^{-4}, respectively, for each mutation rate. Assuming a trait with heritability equal to one-half, these parameters result in mutational variances of $O(10^{-2})$, which is the upper limit of the ranges reported based on experimental results (Lynch (1988), Falconer and Mackay (1996), p. 349). These mutation rates correspond to $\Theta = 4N\mu$ values of 5, 20, and 100, respectively, meaning we expect sweeps to high frequency, mixes of partial and complete sweeps, and adaptation primarily by allele frequency changes, respectively, as the population approaches the new optimum (Höllinger *et al.*, 2019).

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 at a locus and r is the mean number of recombination events per diploid at a locus. Mutation and recombination events occur uniformly along a locus and each locus is separated by 50 centiMorgans (cM). 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 and 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 ($\mu/10$).

At each locus, mutations affecting the trait occur only in the middle window (Figure 2). Therefore, the mean number of recombination events per diploid is ≈ 0.0045 in the middle window where trait-affecting variants arise. Similarly, the mean number of new mutations per diploid at a given locus affecting the trait is $\mu/5$. For the largest mutation rate used here ($\mu = 0.005$), the ratio of recombination events to new mutations affecting the trait in this window is nine to one. The entire genome consists of ten such loci, for a total mutation rate of μ and a total $\theta = 10^4$.

In a second set of simulations, also involving ten unlinked loci, I vary the distribution of effect sizes (DES) such that probability that $N\gamma^2 \geq 100$ is held constant. For Gaussian distributions of effect sizes, the mean γ is zero, as above, and σ_γ is found by numerical optimization using `scipy` (Jones *et al.*, 2001–) to give the desired $Pr(N\gamma^2 \geq 100)$. 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.

In a third set of simulations, I vary $\rho = 4Nr$, the recombination rate within each locus. I ran 256 replicates of these simulations using the tree sequence recording algorithm (Kelleher *et al.*, 2018) implemented in `fwdpy11` version 0.3.2. For these simulations, I recorded the entire population as nodes in the tree sequences for each of 200 generations after the

optimum shift. Recording nodes from these time points allows them to be analyzed after the simulation has completed. Each replicate was simulated twice. The first run simply output metadata about mutations that reached fixation. The second run was performed with the same random number seed as the first and used the metadata from the first run to track linkage disequilibrium around fixations over time, outputting those data along with the tree sequence for the simulation.

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. All statistics were obtained from 50 randomly chosen diploids.

z-scores for the nS_L statistic Individual values of the nS_L statistic (Ferrer-Admetlla *et al.*, 2014) from the first and last window of each locus were binned into intervals of size 0.1 based on derived frequency. These windows were used because they are the furthest from mutations affecting trait values, and thus the least affected by linked selection. The data from all loci were combined and the mean and standard deviation of each bin were used to obtain z-scores for markers from the remaining windows.

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.

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, 0.2.1, , 0.3.2, and 0.5.1 (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 (http://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 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.

Open source tools used Data processing and plotting relied heavily on the following open-source libraries for the `R` language (R Core Team, 2016) `dplyr` (Wickham and Grolemund, 2017), `ggplot2` (Wickham and Grolemund, 2017), `lattice` (Sarkar, 2008) as well as the following Python libraries: `pandas` (McKinney, 2017), `numpy` (VanderPlas, 2016) , `matplotlib` (Hunter, 2007; VanderPlas, 2016), and `seaborn` (http://seaborn.pydata.org). The `sqlite3` library (www.sqlite.org) facilitated data exchange between Python and `R` via the `pandas` and `dplyr` libraries, respectively.

Results

Simulation parameters

The results below are based on simulations using two different models of genetic linkage among variants affecting trait values. The first set of simulations consist of a single genomic region where the number of recombination events per diploid, r, are Poisson-distributed with a mean of one-half and recombination events occur uniformly along the "chromosome". Selected mutations also occur uniformly according to a Poisson process with mean μ per haploid genome (implying $U = 2\mu$ new mutations per diploid on average). Here, I use $\mu = 2.5\times10^{-4}$, 0.001, or 0.005. No neutral mutations occur and the simulations are only used to study the dynamics of selected mutations and the trait values over time. The mutation model is the infinitely many sites model (Kimura, 1969) and the distribution of effect sizes is symmetric around zero. For the majority of results, I use a Gaussian distribution.

In a second set of simulations, mutations affecting the trait occur at ten unlinked "loci". These layout of these loci are shown in Figure 2 and are analogous to those used in simulations of selection applying the structured coalescent (Hudson and Kaplan, 1995; Braverman *et al.*, 1995, and others). I am therefore using the term differently than its common use in quantitative genetics, where a locus and a biallelic mutation are equivalent (*e.g.* de Vladar and Barton, 2014). Here, a locus is a contiguous genomic segment within which mutation and recombination occur, giving rise to many possible alleles (combinations of mutations) with different effect sizes. Mutations affecting the trait only occur in the center of a locus at rate $\mu/10$, keeping the total mutation rate per individual the same as in the simulations described above. By restricting trait-affecting variants to the middle of each locus, we can examine both the temporal and the spatial dynamics of linked selection as a function of time since the optimum shift. The total recombination rate across an individual locus is given by the scaled recombination rate $\rho = 4Nr = 1,000$, giving $r = 0.05$, or about five

cM. The total genetic distance across all ten loci is therefore 500 cM, 450 of which are due to free recombination between loci, leaving 50cM to be explicitly tracked by the simulation. The details of the ratio of new mutations affecting the trait to recombination events within loci are in METHODS. Neutral mutations occur uniformly along the entire locus are rate $\theta = 4N\mu_n$, where μ_n is the mutation rate to neutral variants (per gamete, per locus).

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). For a model of a single trait under Gaussian stabilizing selection with a constant optimum, $s = \frac{\gamma^2}{2V_S}$ (SIMONS *et al.* (2018), see also KIMURA and CROW (1978)). Here, I fix $V_S = 1$, and therefore the relevant scaled strength of selection acting on a segregating variant is $N\gamma^2$. For many of the results presented below, it is helpful to treat the dynamics of strongly selected mutations separately. To do so, I define a "large-effect" variant as having $N\gamma^2 \geq 100$, meaning that the deterministic force of selection is much stronger than that of drift.

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 (see SIMULATION PARAMETERS above for details). I evolved populations for $10N$ generations to mutation-selection equilibrium around an optimum trait value of $z_o = 0$, at which point z_o is changed to 0.1, 0.5, or 1.0 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. They differ from the approach of HÖLLINGER *et al.* (2019) in that I simulate continuous

traits and do not stop evolution once a specific mean fitness is first reached.

The mean trait value, $\bar{z}$, rapidly approaches the new optimum, typically reaching the new optimum within one hundred generations (Figure 3a , see also de Vladar and Barton (2014); Jain and Stephan (2017b); Höllinger *et al.* (2019)). 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 ≈ 200 generations 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 $\bar{z}$, 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 sizes of variants where $N\gamma^2 \geq 100$. The legends also contain the origin times, o, of these large-effect mutations, measured as generations since the optimum shift.

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 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). Here, because of the one-way mutation model, such large-effect variants are at frequencies near zero.

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 $N\gamma^2 \geq 100$. Such large effect mutations only exist for a relatively brief period of time 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 mutations 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.

Figure 4 suggests that fixation times are rather long, on the order of N generations even for mutations with large $N\gamma^2$. These long fixation times are in fact typical, and mutations with $N\gamma^2 > 100$ typically fix in $N/2$ to N generations (Figure S2), which is long relative to the deterministic expectation for strongly selected sweeps from new mutations (Stephan *et al.*, 1992). Large-effect mutations that reach fixation arise close to the time of the optimum shift (Figure S3) and typically show shorter fixation times (Figure S4). In general, the number of sweeps from new mutations and from standing variants are similar, although fixations of smaller-effect standing variants are more common in simulations with higher μ (Figure S5). In Figure S5, a sweep from a new mutation is defined as a mutation arising withing 100 generations of the optimum shift and then reaching fixation. While somewhat arbitrary, this definition is justified by the rapid mean time to adaptation (Figure 3). In this model, large-effect standing variants that fix after the optimum shift were rare at the time of the shift (Figure S6). Small-effect mutations are also typically rare at mutation-selection balance, in

particular when the mutation rate is small (Figure S6).

Taken together, Figures S3, S4, and S6 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). Given that large-effect fixations from standing variation are typically rare at the onset of directional selection (Figure S6), we may also expect them to affect 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 S6), such mutations are unlikely to generate the patterns of haplotype diversity associated with "soft sweeps" because such patterns require strong selection on mutations at intermediate frequencies (Garud *et al.*, 2015).

Fitness effects of mutations during adaptation

In this section, I explore in more detail the strength of selection on individual mutations during the directional phase of selection. These dynamics are relevant to the long fixation times noted in the previous section and also to the extent to which hitch-hiking will affect patterns of linked variation, which is the topic of the next section. As the focus of the remaining sections will be on patterns of variation during adaptation, we switch from simulating a single large region to simulating ten unlinked regions. The only difference between these simulations and those described above is the genetic map and the position of mutations affecting trait values (see Methods). The total scaled recombination rate of $\rho = 4Nr = 10^4$ is divided equally into ten loci with $\rho = 10^3$, and each locus is separated by 50cM. Within each locus, trait-affecting mutations are restricted to a central "window" of each locus (Figure 2). The length of this central window is ≈ 0.05 cM and the mutation rate

at each locus is equal ($\mu/10$). Neutral mutations will occur uniformly along each locus at the scaled rate $\theta = 4N\mu_n = 1,000$. This layout for an individual locus (Figure 2) is the same as in coalescent simulations of directional selection (Kaplan *et al.*, 1988; Braverman *et al.*, 1995; Przeworski, 2002; Kim and Stephan, 2002) and allows the properties of neutral variation to be efficiently measured at fixed distances from regions with mutations affecting fitness.

Figure 5 plots the dynamics of mutations in a ten locus system for one replicate of each of the three mutation rates used here. In each column, the gray vertical line is the time the population first reaches a mean trait value of $0.9z_o$, which corresponds to a mean fitness of at least 0.9 for each replicate. For simplicity, we will call this the time of adaptation. The top row of Figure 5a shows the frequency trajectories of mutations that eventually fixed. These replicates were chosen because each had one fixation of a strongly selected mutation with a similar effect size. As the mutation rate increases, the genetic background of these fixing variants becomes more polygenic. As a result, the initial rate of frequency change of the fixation lessens because other mutations are involved in the response to the optimum shift, some of which may contribute to adaptation but not fix in the long term. For all replicates, the fixations are at different loci (separated by at least 50cM) with one exception. For the high mutation rate case, there is one fixation of a mutation with small γ at the same locus as the fixation with $N\gamma^2 \geq 100$.

Figure 5b shows the frequency dynamics of mutations arising prior to adaptation that were eventually lost. As the mutation rate increases, there are more mutations with $N\gamma^2 \geq 100$ increasing in frequency during adaptation. For the lowest mutation rate simulated here, two such mutations are decreasing in frequency prior to adaptation. At $\mu = 0.001$, four strongly selected mutations sweep to frequencies > 0.10 and are later lost. For $\mu = 0.005$, several large-effect mutations experience more modest increases in frequency during adaptation. These results are well-predicted by the theory presented in Höllinger *et al.*

(2019) under a simpler model of a polygenic trait. From left to right, the columns of Figure 5a and 5b correspond to the domains where complete sweeps, a mixture of complete and partial sweeps, and subtle allele frequency shifts, respectively, are the expected outcomes on the time scale of the population reaching its new optimum (Höllinger *et al.*, 2019).

A direct comparison to the quantitative predictions from Jain and Stephan (2017b) is difficult because their expressions depend on the assumption of equal forwards and backwards mutation rates at each position. However, several qualitative comparisons can be made. First, the simulations presented here are comparable to the "most effects are large" case from Jain and Stephan (2017b) because the trait variance *increases* during adaptation (also see de Vladar and Barton (2014)). In this domain, large allele frequency shifts are expected when the effect size of a mutation is larger than the mean. The second row in Figure 5 shows the mean deviation of a genotype with a given mutation standardized by the standard deviation in trait values (the "z"-score). The mutations that fix (Figure 5a) are all initially found in heterozygous genotypes with trait values multiple standard deviations above the mean. Such mutations are not necessarily the largest-effect variants present at the time of the optimum shift, which is seen for the two higher mutation rates in Figure 5. The mutations that do eventually fix were initially at higher frequency and/or associated with higher-fitness genotypes than large-effect mutations that were eventually lost.

As the population adapts, the deviation in trait value (from the population mean) for a mutation with a given effect size decreases. These z scores decrease because the genetic variance transiently increases following the optimum shift (Figure 3b, de Vladar and Barton (2014); Jain and Stephan (2017b)) because mutations are increasing in frequency and the variance is a function of allele frequency times the squared effect size. Mutations causing larger deviations are expected to become lost, as seen most clearly in the first column of Figure 5b–the blue and green mutations over- and under- shoot the optimum, respectively. At low mutation rates, there is a tendency to slightly over shoot on average

(Figure 3) because such mutations will have larger initial increases in allele frequency than smaller-effect variants.

In summary, the dynamics in Figure 5 are qualitatively compatible with the predictions make by Jain and Stephan (2017b) for the conditions for a mutation to sweep in the "most effects are large" regime. Variants that sweep have large initial effects on trait values relative to standing variation and will not overshoot the optimum by large amounts. It is useful to note here that regimes where the genetic variance decreases during adaptation are not observed for any of the simulations presented here. The decrease in variance is seen in the "most effects are small" domain where the equilibrium frequency of variants prior to the optimum shift is one-half, which maximizes the variance (de Vladar and Barton, 2014; Jain and Stephan, 2017b). Adaptation to the new optimum displaces allele frequencies, reducing the variance from its maximum value (see, for example Figure 9 of de Vladar and Barton (2014)).

Finally, we can turn to the long fixation times. These are, in part, due to the decreasing strength of selection on individual mutations during time period where directional selection occurs. The final row of Figure 5 shows the relative deviation in due to genotypes carrying each mutation over time. As expected, genotypes with fitness above the mean increase in frequency, and these genotypes are associated with trait values multiple standard deviations closer to the new optimum. As the mean trait value approaches the new optimum, the relative excess fitness of these genotypes declines, approaching the reciprocal of the population size. Once the population has adapted, these mutations have small effects on phenotypic variation and their long-term dynamics are governed by underdominant selection against phenotypic variance (Robertson, 1956; Kimura, 1981). Note that all of the sweeping alleles in Figure 5 are from standing variation (origin times less than zero) and are rare at the onset of directional selection (also see Figure S6).

The displacement of allele frequencies during the directional phase moves trait-affecting

alleles to intermediate frequencies (Figures 4, S1, and 5). Mutations with frequencies greater than one-half will be weakly favored and expected to fix, while those with frequencies less than one-half will most likely removed from the population (Robertson, 1956; Kimura, 1981). In the absence of further mutation during adaptation, Robertson (1960) showed that the total "advance" of a trait under weak selection (the time required to move the mean trait value to its maximum by the fixation of all favorable alleles) takes on the order of N generations, and the observations in Figures 4 and 5 are consistent with this conjecture.

Finally, traits with higher mutation rates have larger numbers of small-effect mutations segregating prior to adaptation (Figure 5). Once the population is adapted, the deviations from mean fitness tend to be small for most genotypes, and the large-effect mutants are not yet fixed, implying that interference (Hill and Robertson, 1966) may also increase fixation times when the mutation rate is higher. We will return to the role of interference below. The observation in Figures 4 and 5 of mutations not reaching fixation by the time the new optimum is hit consistent with previous results from other authors (Höllinger *et al.*, 2019; Chevin and Hospital, 2008; Jain and Stephan, 2017b).

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 and the parameters defining a locus are detailed above in Simulation parameters). The trait parameters are the same as in the previous section. Figure S7 shows the mean number of mutations affecting the trait over time in these simulations.

Figure 6 summarizes patterns of variation in the central window (Figure 2) of each locus where large-effect mutations segregate during adaptation to the new optimum. The figure is based on the same data as Figure 5. The first two rows plot the frequency trajectories of eventual fixations and losses, and the next three rows summarize patterns of variation

calculated from a random sample of individuals. These summaries of variation only show deviations from equilibrium values consistent with directional selection at loci where large-effect fixations occur. Further, the deviations are more pronounced when the mutation rate is smaller. The partial sweeps occurring at intermediate mutation rates (middle column of Figure 6) are not associated with strong signals of hitch-hiking, at least when the sample size is relatively small as is the case here. The time when a given statistic shows its maximum departure from equilibrium values differs for each statistic and, for the replicate with $\mu = 0.001$, the maximum departure may occur ≈ 100 generations after the time to adaptation. Visually, however, one could argue that haplotype diversity tends to minimize closer to the time to adaptation than the summaries of the site frequency spectrum.

Figure 7 shows patterns of variation along each of the ten loci from an additional simulated replicate for each of the parameters shown in Figure 5 and Figure 6. Each line corresponds to a different time point in the approach to the new optimum, showing data for the first time the population mean trait value crosses the thresholds of $\bar{z} \geq 0.1$, ≥ 0.5, and ≥ 0.9. While the values are noisy along a genome, it is apparent that directional selection is affecting patterns of variation at linked sites in the replicates with smaller mutation rates. In the leftmost column, where $\mu = 2.5 \times 10^{-4}$, an excess of high-frequency derived variants is seen at locus four, along with a reduction in haplotype diversity. A standing variant with $N\gamma^2 \geq 100$ swept to high frequency at this locus during adaptation. In the middle column ($\mu = 10^{-3}$), one sees a less dramatic reduction in haplotype diversity at locus ten, where a strongly selected standing variant reached high frequency. For these two replicates, there is some evidence of reduced haplotype diversity at loci eight and five, respectively, that is not associated with any fixations. The final column, where $\mu = 5 \times 10^{-3}$, there is no obvious temporal nor spatial pattern to variation in diversity levels, and the largest deviations from the background are not associated with the fixation of beneficial mutations.

Overall, Figure 6 and Figure 7 suggest that patterns of strong hitch-hiking are more likely

at loci where mutations with large values of $N\gamma^2$ fix. Moreover, such mutations must arise on average before the mean time to adaptation. Below, when looking at average patterns of variation over time and along genomes, we will distinguish patterns of variation where fixations meeting these conditions occur from the mean pattern expected from a randomly chosen locus.

The site frequency spectrum over time

The expected histogram of mutation frequencies in a sample (the "site frequency spectrum") is a geometrically decreasing function of increasing mutation frequency under the standard neutral model (Wakeley, 2008). Departures from this expectation are often summarized as single numbers whose expectations are ≈ 0 under this null model. In this section, I describe the average dynamics of two widely used statistics (Tajima, 1989; Zeng *et al.*, 2006) as a function of both time since the optimum shift and of distance from trait-affecting mutations.

Figure 8 shows the average behavior of Tajima's (1989) D over time. Figure 8a 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.

Averaging over loci experiencing large-effect fixations, again defined as having scaled effect sizes $N\gamma^2 \geq 100$, Figure 8b shows a stronger hitch-hiking 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. Qualitatively similar patterns hold for the overall reduction

in diversity (Figure S8) and the H' statistic (Figure S9). The latter statistic returns to equilibrium rather rapidly, consistent with previous results (Przeworski, 2002).

It is of particular interest to compare the patterns of linked selection at loci where large-effect variations fixed from standing variation to those where new mutations of large effect fixed. The latter case is typically referred to as a "hard" selective sweep, and gives rise to a reduction in diversity and a skew in the site frequency spectrum over physical distances depending on the ratio of the strength of selection to the recombination rate (Maynard-Smith and Haigh, 1974; Fay and Wu, 2000; Przeworski, 2002; Kaplan *et al.*, 1989; Durrett and Schweinsberg, 2004). When large-effect sweeps occur from standing variation, their effects on linked neutral variation become less dramatic as the frequency at the onset of positive directional selection increases (Innan and Kim, 2004; Przeworski *et al.*, 2005; Berg and Coop, 2015), thus "softening" the signal of hitch-hiking.

Here, large-effect fixations from new mutations and from standing variants have similar average effects on statistics like D and H' (Figures 8 and S9). Figure 9 shows the number of haplotypes at a locus associated with sweeps from standing variation as a function of the effect size of the variant. Here, a haplotype is defined as a unique genotype at a locus, including all neutral and non-neutral variants. Large-effect sweeps from standing variation are either extremely rare (at high μ) or are rare at the time of the optimum shift when μ is small, and therefore associated with few haplotypes.

Power to reject the null model using the site-frequency spectrum

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 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 S10a 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 S10b), of which there are relatively few per replicate (Figure S11). For a large optimum shift, the most mutation-limited trait simulated here shows over 50% of "significant" D values are at loci 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.

The behavior of H' is similar to that of D, but power decreases more rapidly with time since the optimum shift (Figure S12a, (also see Przeworski, 2002)). The composite likelihood ratio (CLR) test of (Nielsen *et al.*, 2005) also uses high-frequency derived variants to detect strong sweeps from new mutations. The power of this test, estimated using `SweeD` (Pavlidis *et al.*, 2013), is low overall and decays rapidly with time (Figure S13a), much like the H' statistic (Figure S12). 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 S13b). For smaller optimum shifts and more polygenic traits, at most one-half of "significant" loci experienced large-effect substitutions, and a "significant" test was almost equally likely to be associated with a sweep from standing variation or from a new mutation (Figure S13b).

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). Summaries of haplotype

homozygosity are widely used to detect recent selection (Voight *et al.*, 2006; Ferrer-Admetlla *et al.*, 2014) and are indirect summaries of the underlying linkage disequilibrium in the data (Sabatti and Risch, 2002).

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. A negative value of the statistic implies longer runs of homozygosity around the derived allele. Figure S14 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 S14b).

Shortly after the optimum shift, the mean z-score becomes positive (Figure S14). 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 S15) Thus, the positive z-scores in Figure S14 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 9). This difficulty in interpretation is a general issue arising from the fact that patterns of variation due to strong sweeps from standing variation overlap considerably with those of older sweeps from new mutations (Schrider *et al.*, 2015).

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 S16, S17, and S18).

Robustness to variation in the recombination rate

In this section, I explore the effect of varying the scaled recombination rate within a locus, ρ. At higher mutation rates, longer fixation times are more likely as ρ decreases (Figure 10). In individual replicates, there is a tendency towards negative disequilibrium among beneficial mutations ($\gamma > 0$, Figure S19), suggesting a role for interference among selected sites affecting times to fixation (Hill and Robertson, 1966; Felsenstein, 1974). In the previous sections, the ratio of ρ to θ within loci is one, which is roughly "human"-like (REFS). For species like *Drosophila melanogaster*, where $\rho \gg \theta$ (Haddrill *et al.*, 2005), fixation times will be much shorter on average (Figure 10). Note that the effect of recombination rate on fixation time is most dramatic for $\mu = 0.005$, which is also the part of the parameter space explored here where fixations of larger-effect ($N\gamma^2 \geq 1,000$) are rare.

The within-locus recombination rate has no discernible effect on $\bar{z}$ nor on V_G (Figure S20). The differences in the height of the "spike" in V_G when $\mu = 2.5 \times 10^{-4}$ show no clear pattern with ρ and are thus attributable to Monte Carlo error in estimating a second-order statistic from 256 replicates.

Unlike the mean trait value and variance, the mean temporal dynamics of summaries of variation data are strongly affected by ρ (Figure S21). Figure S22 shows how the within-locus recombination rate affects patterns of haplotype diversity in a ten locus system with $\sigma_\gamma = 0.25$. When ρ is small, the impact of linked selection is much more apparent. These effects of the local recombination rate on patterns of hitch-hiking are expected from standard

theory of directional selection because both the magnitude and extent along the genome of linked selection depends on the ratio of the recombination rate to the selection coefficient (Kaplan *et al.*, 1989; Durrett and Schweinsberg, 2004; Nielsen *et al.*, 2005).

Varying the distribution of effect sizes

The results described in the previous sections are based on a Gaussian distribution of effect sizes whose standard deviation is held constant. In this section, I vary the distribution of effect sizes (DES) such that the fraction of mutations with $N\gamma^2 \geq 100$ 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. 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. I compare the results of Gaussian distributions of effect sizes to two gamma distributions with shape parameters of one and one-half. The gamma distributions with a shape parameter of one results in a mutational variance similar to that of the Gaussian distribution, but the gamma is skewed towards higher effect sizes. The gamma with a shape parameter of one-half increased the mutational variance by approximately one order of magnitude due to increased skew towards larger effect sizes. For a trait with a heritability of one-half, the Gaussian distribution and the gamma with a shape of one give a ratio of the mutational variance to the environmental variance of 2×10^{-3} to 3×10^{-3} when the proportion of new mutations with $N\gamma^2 \geq 100$ is 0.1. These values are close to the mean of $\approx 10^{-3}$ reported for a variety of traits (Lynch (1988), Falconer and Mackay (1996), p. 349)

Varying the fraction of large-effect mutations has a weak effect on the mean time to reach the new optimum, with mutation-limited traits adapting more slowly on average when the majority of variants are of small effect (Figure S23a). This observation should be unsurprising as the population must wait longer for a strongly selected mutation in this case.

Patterns of variation expected due to hitch-hiking are more extreme when $Pr(N\gamma^2 \leq 100)$ is small, as the trait is effectively more mutation-limited with respect to strongly selected variants (Figure S24). 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 both the mutation rate and $Pr(N\gamma^2 \geq 100)$ are small.

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).

The simulations performed here have several important differences from recent theoretical treatments of adaptation to sudden optimum shifts (see below). However, the conditions for a selective sweep are consistent with predictions made using theoretical results from from Jain and Stephan (2017b) and Höllinger *et al.* (2019). Mutations with large effects on trait values at the time of the optimum shift are most likely to rise in frequency (Figure 4 and 5). When several such mutations co-segregate, those with the highest initial frequencies tend to reach fixation. If initial frequencies are similar, the variant with the highest initial fitness typically fixes. For a given distribution of effect sizes, faster sweeps are more likely

686 at lower mutation rates.

687 When considering the pattern of hitch-hiking at a locus, the presence or absence of
688 a large-effect fixation at a locus is a reliable predictor of the magnitude of hitch-hiking
689 patterns. As expected, such fixations are more common when the trait is mutation-limited
690 (Höllinger *et al.*, 2019) and thus strong departures from equilibrium patterns of variation
691 are not expected for more polygenic traits (Figure 8). For the optimum shift model considered
692 here, the strength of selection is not constant over time (Figure 5, see also Kimura (1981)).
693 Thus, genotypes containing variants that were initially strongly favored by selection are
694 much more weakly selected by the time the population has reached the new optimum. This
695 weakening of selection increases fixation times to the order of the population size (Figure S2),
696 which is much longer than the times $\ll N$ generations expected for directional selection in
697 large populations (Stephan *et al.*, 1992).

698 Here, the number of selected mutations segregating over time ranged from dozens to
699 several hundred, as a function of the underlying mutation rate (Figure S7). At high mutation
700 rates, the number of segregating loci are roughly the same as some of the results presented in
701 de Vladar and Barton (2014). Here, however, the partial linkage among sites in this work
702 leads to some negative linkage disequilibrium (Figure S19), which is a signal of interference
703 (Hill and Robertson, 1966; Felsenstein, 1974). This interference has little effect on
704 the mean time to adaptation, but fixation times are increased. The lack of effect on time to
705 adaptation is driven by initial large fitness differences among genotypes (Figure 5, also see
706 Höllinger *et al.* (2019)). Once the population is close to the new optimum, selection on
707 individual genotypes is much weaker 5, setting up the conditions for interference to affect
708 fixation times (Hill and Robertson, 1966).

709 The distribution of effect sizes has different effects on properties of the trait than on
710 patterns of hitch-hiking. The mutation rates used here span the parameter space from
711 partial and complete sweeps being most common to the optimum being reached via allele

frequency shifts of many mutations (Figure 4, Höllinger *et al.* (2019)). In general, the mean time to adapt is not strongly affected by the DES if the fraction of new mutations of large effect is constant (Figure S23a). For a given mutation rate, lowering the mutational variance lowers the probability of a strongly selected mutation, resulting in stronger signals of hitch-hiking (Figure S24). Effectively, the trait becomes more mutation limited when the mutational variance is lower, although the effect es most pronounced when the mutation rate itself is small (Figure S24). 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 S24).

The genetic model assumed here does not lead to sweeps of large-effect mutations from common variants (frequencies greater than, say, five percent). Rather, the stabilizing selection around the initial optimum keeps large-effect mutations rare, such that sweeps from such standing variations start at low frequencies. Importantly, it is not possible to tune the model parameters to obtain sweeps from large-effect, but common, variants with high probability. Changing the strength of stabilizing selection (V_S) preserves the rank orders of fitness 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 *detectable* "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 important to note a key methodological difference between this work and that of other authors. Höllinger *et al.* (2019) 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 Höllinger *et al.* (2019)). Here, we observe fixations of large effect for the mutation rates corresponding to $\Theta = 4N\mu = 100$ (Figure 4), which Höllinger *et al.* (2019) show is the parameter range where adaptation occurs primarily by changes in allele frequency. These results are consistent with the theoretical predictions from Höllinger *et al.* (2019), as the fixations in the simulations described here take place on time scales longer than the mean time to reach the new optimum. In the rightmost column of Figure 4, the population has adapted quickly, with the fixations occurring over a much longer 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 Höllinger *et al.* (2019).

This work (and that of Höllinger *et al.* (2019)) differs from the analytical and numerical work of de Vladar and Barton (2014); Jain and Stephan (2015, 2017b) in several key aspects. First, we consider irreversible mutation here (the infinitely many sites model of Kimura (1969)), while de Vladar and Barton (2014) assumed equal rates of forward and reverse mutation (see also Barton (1986); Jain and Stephan (2015, 2017b)). The infinitely many sites model used here was chosen because it is the most commonly used mutational model for investigating the effects of linked selection during adaptation (*e.g.* Braverman *et al.*, 1995; Przeworski, 2002; Przeworski *et al.*, 2005). I also allow for partial linkage among sites, which is a key difference from the work based on the Barton (1986) framework, which assumes free recombination. As noted above, partial linkage affects the long-term dynamics of selected mutations (Figure 10).

I have focused 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 predictions 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 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 S6, also see Berg and Coop (2015); Przeworski *et al.* (2005); Orr and Betancourt (2001); Hermisson and Pennings (2005)). The only test statistic based on patterns of SNP variation for detecting polygenic adaptation that I am aware of 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) and explicit modeling of trait architectures at nor near the infinitesimal limit Robertson (1970, 1977). 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 was redone using effect sizes from the UK Biobank data (Berg *et al.*, 2019; Sohail *et al.*, 2019).

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 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 simulations here also model the entirety of heritable variation for the trait. An alternative approach would be to allow for a unlinked additive genetic background with its own mutational variance. Such an approach would be straightforward assuming an infinitesimal model for the background, as has been done previously (Chevin and Hospital, 2008; Stetter *et al.*, 2018). Stetter *et al.* (2018) simulated "domestication" traits evolving to a new optimum via truncation selection and a heritable background affecting the focal trait. They concluded that the contribution of genetic background to several outcomes of interest (speed of adaptation, fixations of beneficial mutations, etc.) was of overall less importance to the dynamics than the variance in mutational effect sizes, σ_γ. Clearly, however, the details will depend on the specifics of the model, with Chevin and Hospital (2008) at one extreme and the current work at perhaps the other. Here, the simulations with high mutation rates imply that any single segregating variant finds itself in a mutation-rich genetic background of up to several hundred segregating variants, the majority of which have small fitness effects (Figure S7). Another appealing alternative would be to simulate entire genomes using an adaptation of Robertson's (1977) method to incorporate tree sequence recording (Kelleher *et al.*, 2018) and large-effect mutations occurring at some rate. Such a scheme would generate large-effect genomic regions through two different mechanisms: the

occasional large-effect mutation as well as via large-effect haplotypes arising from stochastic recombination events (Sachdeva and Barton, 2018).

It may also be 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 observations of allelic heterogeneity underlying variation in complex traits (Clark, 1998; 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).

Acknowledgments

The author thanks Jaleal Sanjak, David Lawrie, Jeffrey Ross-Ibarra, Markus Stetter, Tony Long, Joachim Hermisson, Nick Barton, Kavita Jain, and two anonymous reviewers for valuable discussions and feedback on the manuscript. 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.

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.*, 2019 Reduced signal for polygenic adaptation of height in UK biobank. Elife **8**.

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. bioRxiv : 419275.

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.

Clark, A. G., 1998 Mutation-selection balance with multiple alleles. Genetica **102-103**: 41–47.

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.

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

Falconer, D. S. and T. F. C. Mackay, 1996 *Introduction to Quantitative Genetics (4th Edition)*. Pearson, 4 edition edition.

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

Felsenstein, J., 1974 The evolutionary advantage of recombination. Genetics **78**: 737–756.

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.

Haddrill, P. R., K. R. Thornton, B. Charlesworth and P. Andolfatto, 2005 Multilocus patterns of nucleotide variability and the demographic and selection history of drosophila melanogaster populations. Genome Res. **15**: 790–799.

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

Hill, W. G. and A. Robertson, 1966 The effect of linkage on limits to artificial selection. Genet. Res. **8**: 269–294.

Höllinger, I., P. S. Pennings and J. Hermisson, 2019 Polygenic adaptation: From sweeps to subtle frequency shifts. PLoS Genet. **15**: e1008035.

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

Hunter, J. D., 2007 Matplotlib: A 2D graphics environment. Comput. Sci. Eng. **9**: 90–95.

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.

Kim, Y. and W. Stephan, 2002 Detecting a local signature of genetic hitchhiking along a recombining chromosome. Genetics **160**: 765–777.

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

Kimura, M., 1981 Possibility of extensive neutral evolution under stabilizing selection with special reference to nonrandom usage of synonymous codons. Proc. Natl. Acad. Sci. U. S. A. **78**: 5773–5777.

Kimura, M. and J. F. Crow, 1978 Effect of overall phenotypic selection on genetic change at individual loci. Proc. Natl. Acad. Sci. U. S. A. **75**: 6168–6171.

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.

Lynch, M., 1988 The rate of polygenic mutation. Genet. Res. **51**: 137–148.

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

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.

Robertson, A., 1960 A theory of limits in artificial selection. Proceedings of the Royal Society of London B: Biological Sciences **153**: 234–249.

Robertson, A., 1970 A theory of limits in artificial selection with many linked loci. In K.-I. Kojima, editor, *Mathematical Topics in Population Genetics*. Springer Berlin Heidelberg, Berlin, Heidelberg, 246–288.

Robertson, A., 1977 Artificial selection with a large number of linked loci. In *Proceeding of the international conference on quantitative genetics*. 307–322.

Sabatti, C. and N. Risch, 2002 Homozygosity and linkage disequilibrium. Genetics **160**: 1707–1719.

Sachdeva, H. and N. H. Barton, 2018 Replicability of introgression under linked, polygenic selection. Genetics **210**: 1411–1427.

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.

Sohail, M., R. M. Maier, A. Ganna, A. Bloemendal, A. R. Martin *et al.*, 2019 Polygenic adaptation on height is overestimated due to uncorrected stratification in genome-wide association studies. Elife **8**.

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.

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.

VanderPlas, J., 2016 *Python Data Science Handbook: Essential Tools for Working with Data*. O'Reilly Media, 1 edition edition.

Voight, B. F., S. Kudaravalli, X. Wen and J. K. Pritchard, 2006 A map of recent positive selection in the human genome. PLoS Biol. **4**: e72.

Wakeley, J., 2008 *Coalescent Theory: An Introduction*. W. H. Freeman.

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.

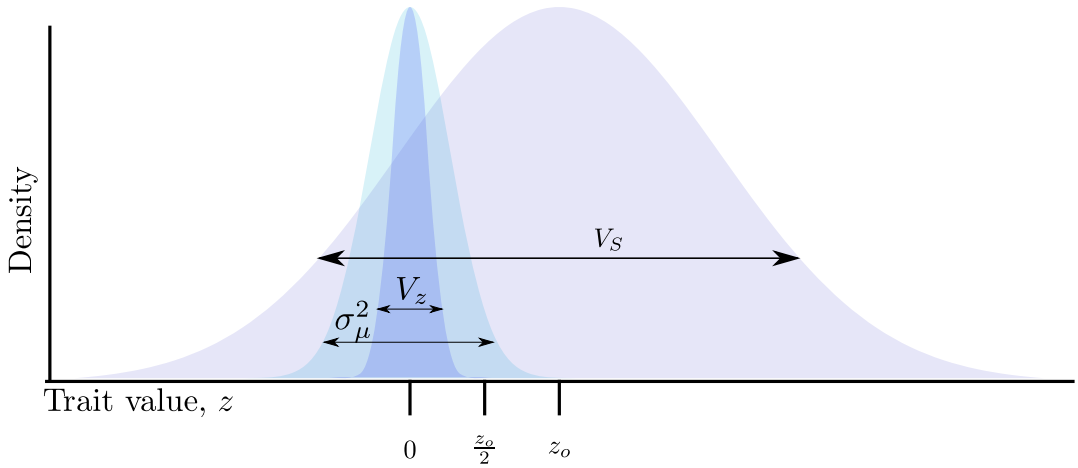


Figure 1: Schematic of the model. A Wright-Fisher population evolves to equilibrium around an optimum trait under Gaussian stabilizing selection with mean zero, where the parameter V_S represents the intensity of selection against extreme trait values ($w = e^{-z^2/2V_S}$). At equilibrium, the mean trait 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.

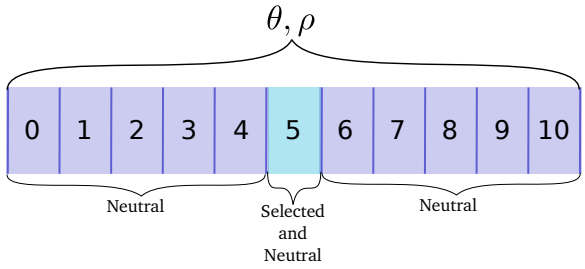
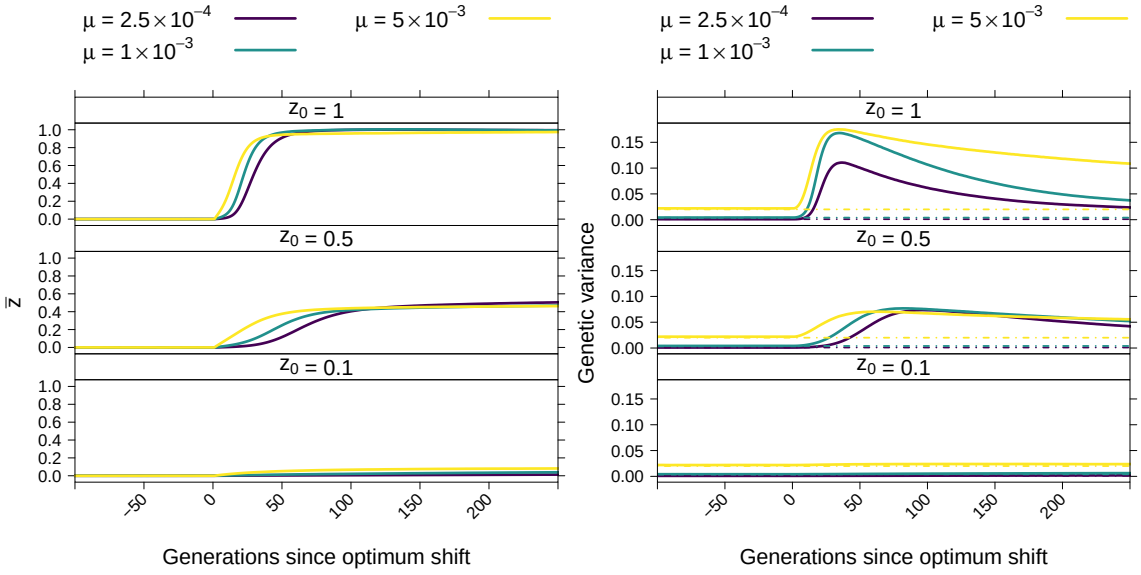


Figure 2: Schematic of a single locus for multi-locus simulations. The scaled neutral mutation and recombination rates, θ and ρ, respectively, are modeled as uniform processes across a locus. A locus is divided into 11 windows of equal size. Mutations affecting the trait only occur in the central window, shown in pale blue. Multiple loci are 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).

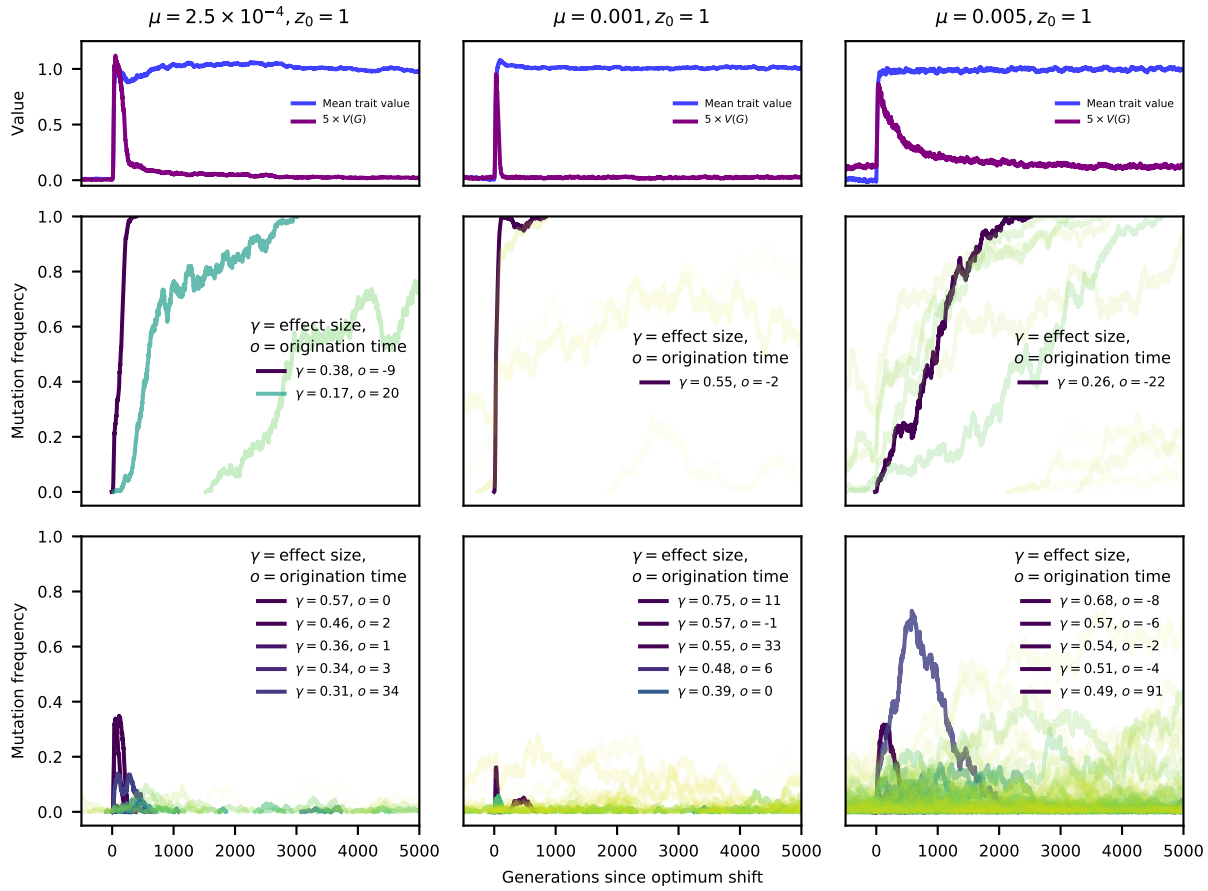
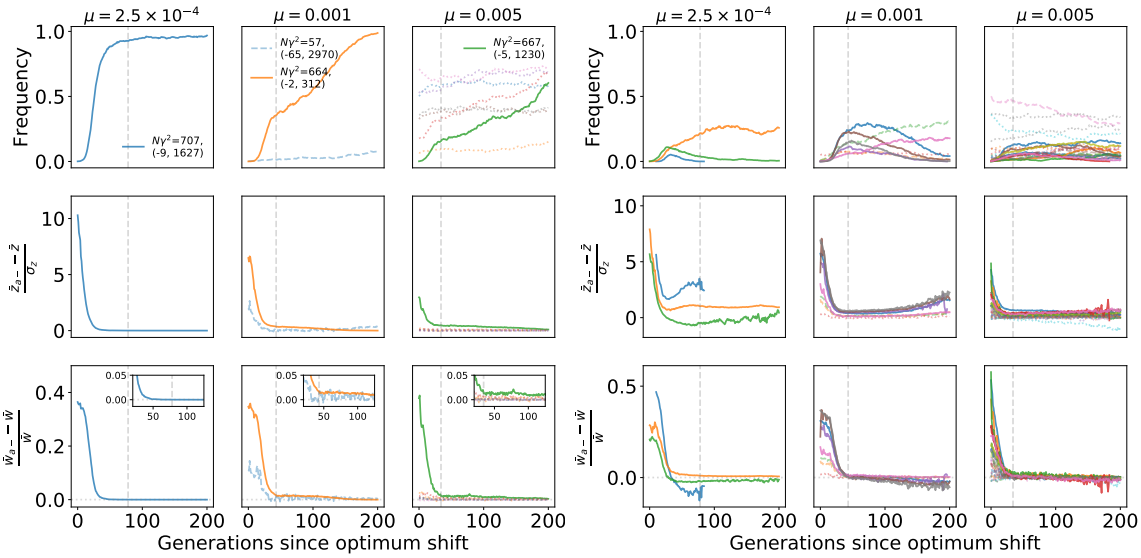


Figure 4: Trajectories of selected mutations. The three columns show results from a single simulation replicate for low, moderate, and high-mutation rate simulations, respectively. Parameter values are at the top of each column. The first row of plots shows the trait value and the genetic variance (multiplied 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 with effect sizes $N\gamma^2 \geq 100$ 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 adaptation to the new optimum.



(a) Eventual fixations

(b) Eventual losses

Figure 5: Phenotypic and fitness effects of fixations and losses. The data shown are for a single simulation replicate with $\sigma_\gamma = 0.25$, $z_o = 1$ and the mutation rate μ shown atop each column. The mutation rate shown is the sum over loci and individual loci mutate at equal rates ($\mu/10$). In all panels, solid lines refer to $N\gamma^2 \geq 100$, dashed lines $10 \leq N\gamma^2 < 100$ and dotted lines $1 \leq N\gamma^2 < 10$. The vertical line is the generation when the mean trait value first crossed 90% of the new optimal value ($0.9z_o$). (a) The dynamics of fixations. The top row shows the frequency trajectory of mutations that eventually reached fixations. For mutations with $N\gamma^2 \geq 10$, the legend shows $N\gamma^2$ and the mutation's origin and fixation times in parentheses, scaled so that zero is the time of the optimum shift. Defining an $a-$ genotype to be any genotype containing at least one copy of these "focal" mutations, the second row shows the mean deviation from the mean trait value for the focal genotypes, standardized by the phenotypic standard deviation. The final row shows the mean relative deviation in fitness for $a-$ genotypes. The horizontal line in the last row is placed at the reciprocal of the population size ($1/N$). (b) The dynamics of losses. Plotting is identical to (a), but the data are filtered to only include mutations arising prior to the population first crossing $0.9z_o$ and eventually reaching a frequency of $q \geq 0.05$.

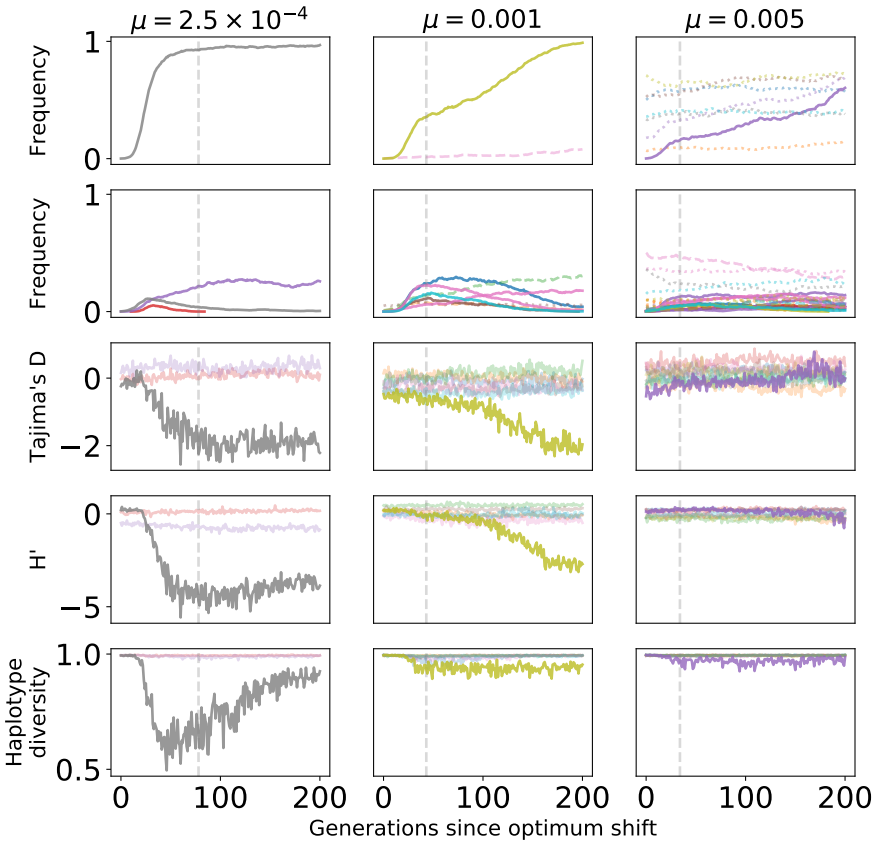


Figure 6: Signals of directional selection in single replicates of a ten-locus system. The data shown are based on the same simulations as in Figure 5. The first two rows show frequency trajectories for fixations and losses, respectively, with the colors indicating the locus where the mutation is found. The vertical gray line is the generation when the mean trait value first crosses 90% of the optimal trait value. The remaining rows show Tajima's (1989) D, H' (Zeng *et al.*, 2006), and haplotype diversity in a random sample of 50 diploids, calculated using genotypes taken from the central "window" of a locus where causal mutations are occurring (Figure 2).

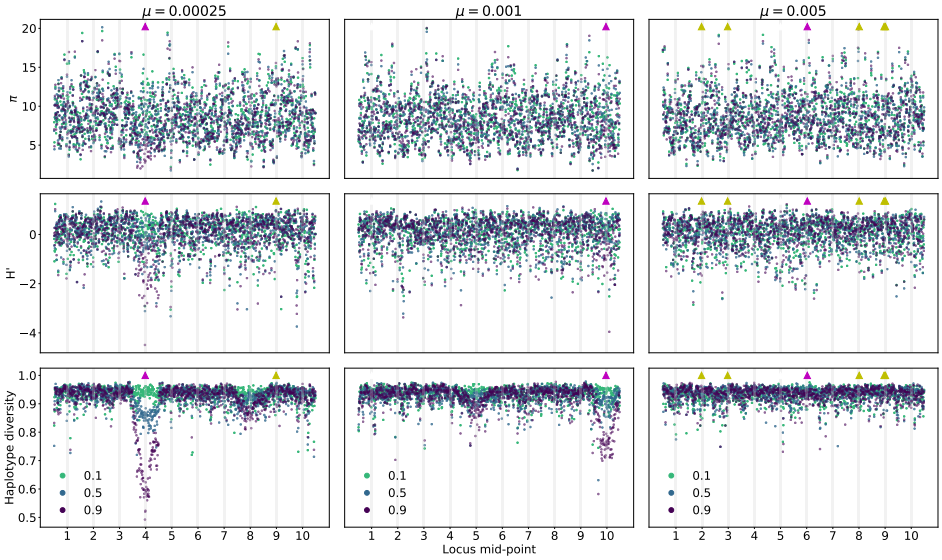
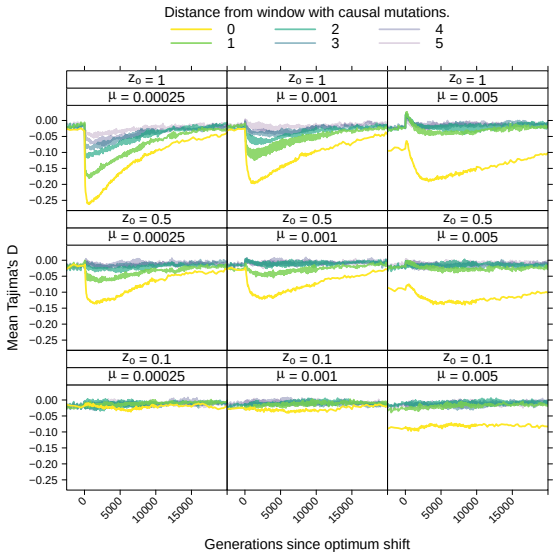
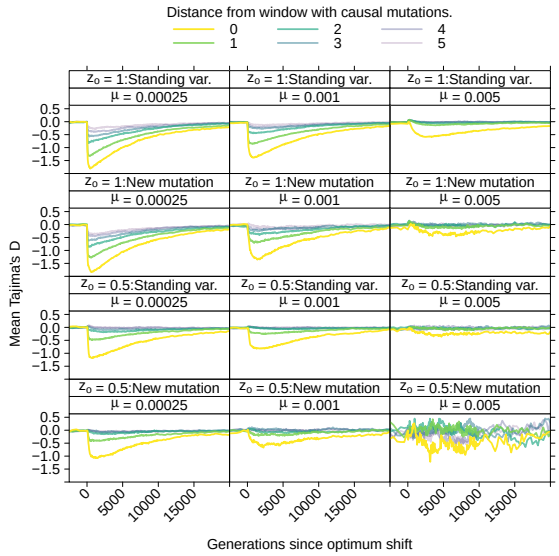


Figure 7: Patterns of genetic variation along genomes in a ten locus system during adaptation to an optimum value of $z_o = 1$ and $\sigma_\gamma = 0.25$. The mutation rate shown is the sum over loci and individual loci mutate at equal rates ($\mu/10$). Each column corresponds to a single simulated replicate with the mutation rate given at the top. The three rows correspond to nucleotide diversity (π), H' (Zeng *et al.*, 2006), and haplotype diversity, respectively. The three point colors refer to statistics calculated from 50 randomly chosen diploids in the first generation that the population mean trait value first crossed values of at least 0.1, 0.5, or 0.9. The gray shades refer to the locations within each locus where mutations affecting trait values occur (Figure 2). The triangles along the top of each panel show where fixations occurred. Triangles pointing up are fixations from standing variation. Magenta refers to fixations with scaled effect sizes $N\gamma^2 \geq 100$ and yellow refers to $1 \leq N\gamma^2 \leq 10$.



(a) Mean Tajima's D over time



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

Figure 8: Tajima's (1989) D statistic over time. 8a 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) 8b shows the mean value of D conditioning on a locus fixing a mutation with effect size $N\gamma^2 \geq 100$. 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.

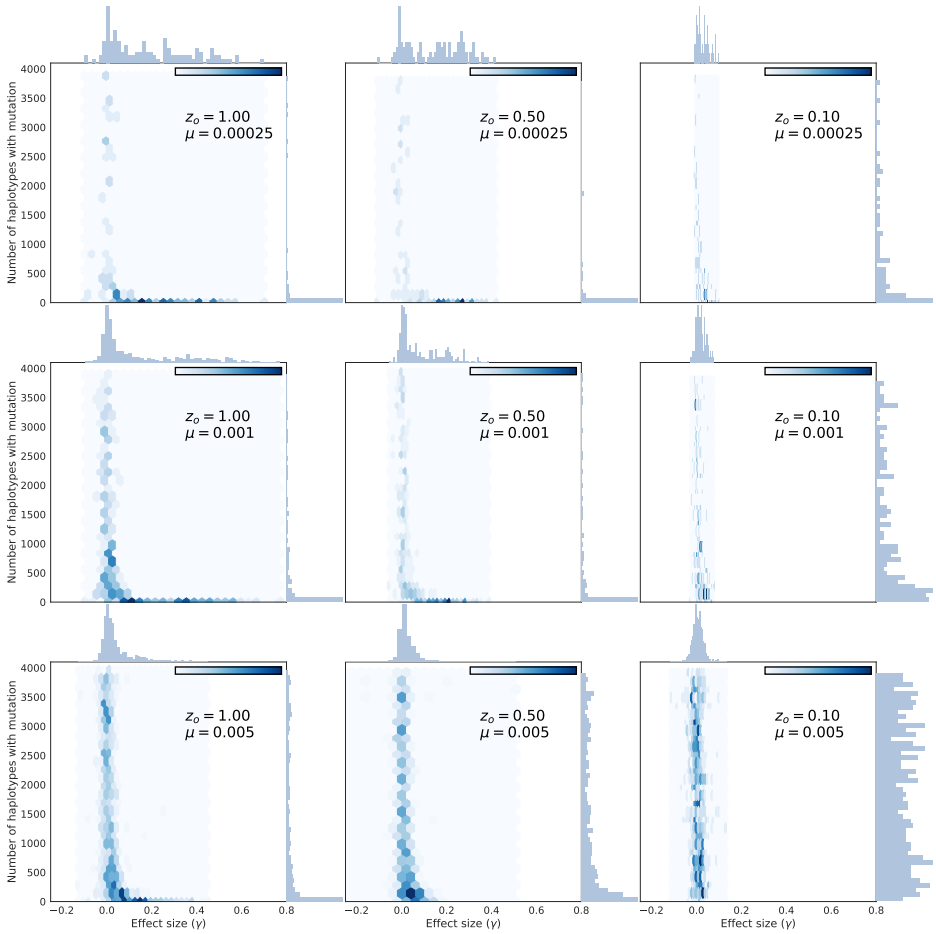


Figure 9: The number of haplotypes associated with fixations from standing variation of different effect sizes. Each panel shows the effect size of a fixation from standing variation (x axis) and the number of unique haplotypes in the entire population containing that mutation. The number of haplotypes 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.

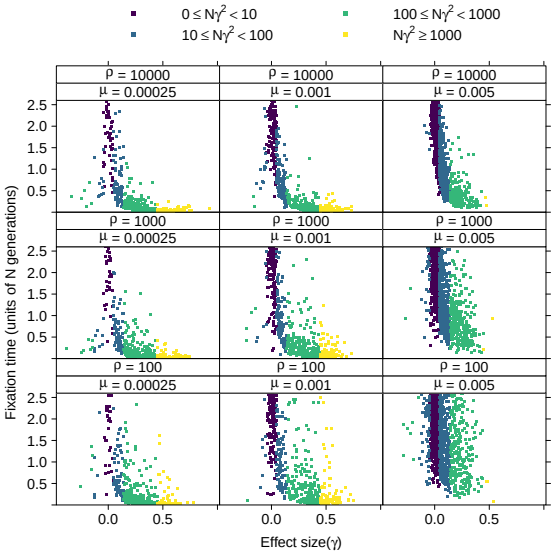


Figure 10: The sojourn times of fixations in a ten locus system with varying recombination rates within loci. The x axis is the effect size of a fixation, and the y axis is its time to fixation scaled by the population size. The expected fixation time due to drift alone is 4 and the distribution of fixation times under neutrality has a long tail including large values. The points are collected from 256 replicates for each parameter combination and colored by order-of-magnitude ranges of their scaled strength of selection ($N\gamma^2$).

Supporting Information

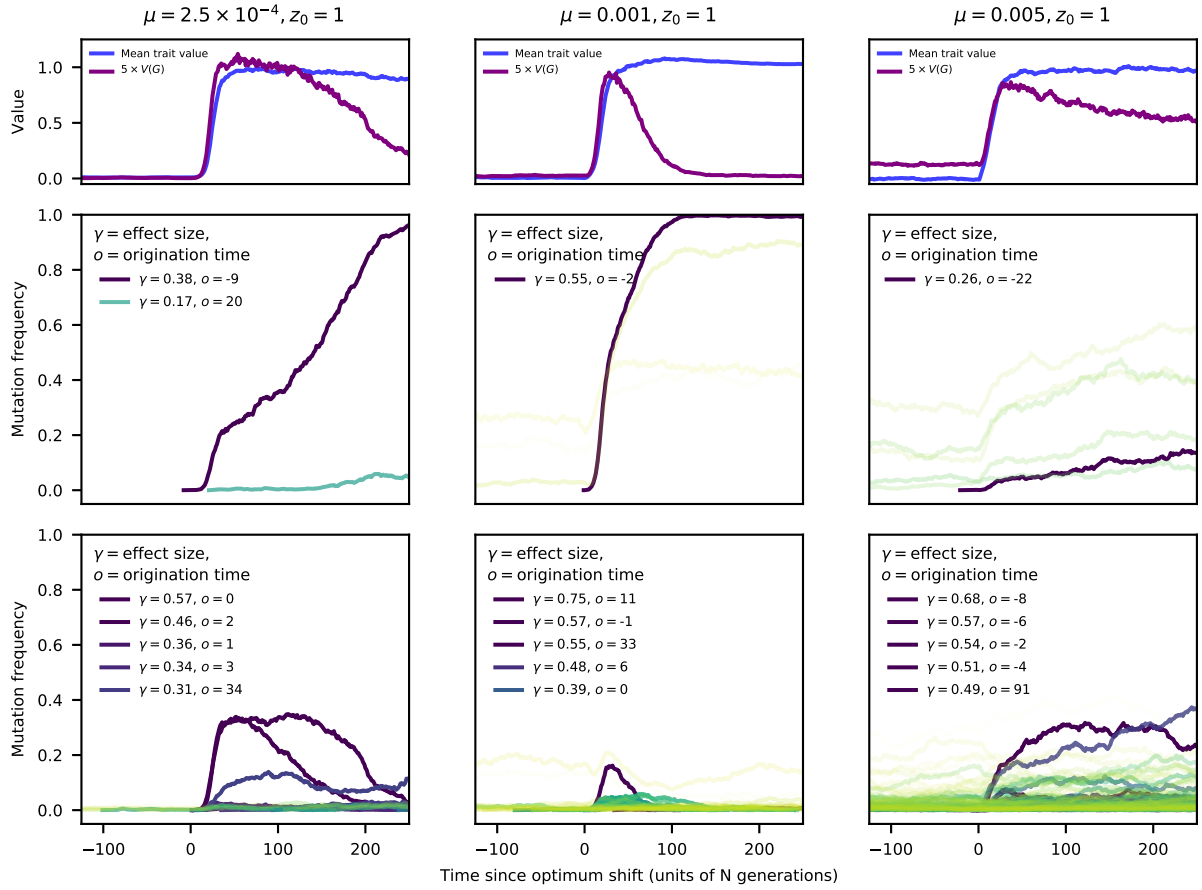


Figure S1: Trajectories of selected mutations, shown over a short time scale. The data and colors are identical to Figure 4.

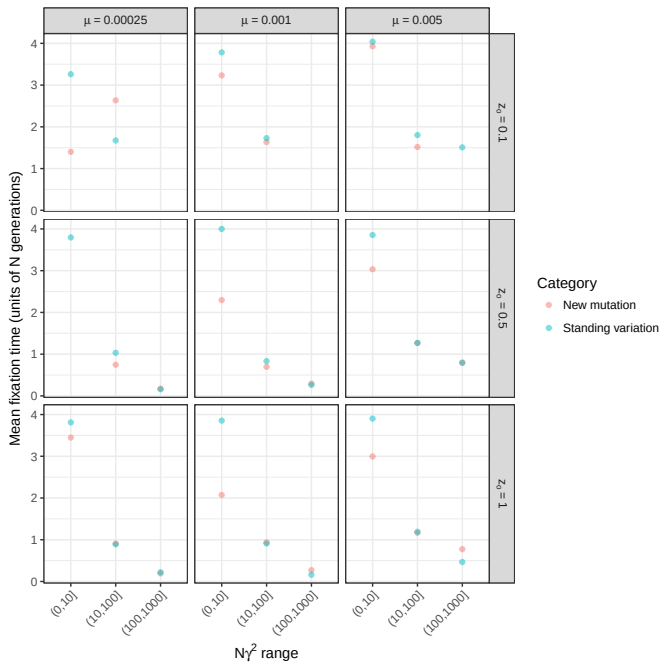


Figure S2: The mean time to fixation of a selected mutation in a single-region simulation. The mean times are scaled by population size (N) to facilitate comparison to the expectation in the absence of selection, which is 4. The mean values are separated by both scaled strength of selection ($N\gamma^2$) and by fixations from standing variation versus new mutations arising in the first one hundred generations after the optimum shift.

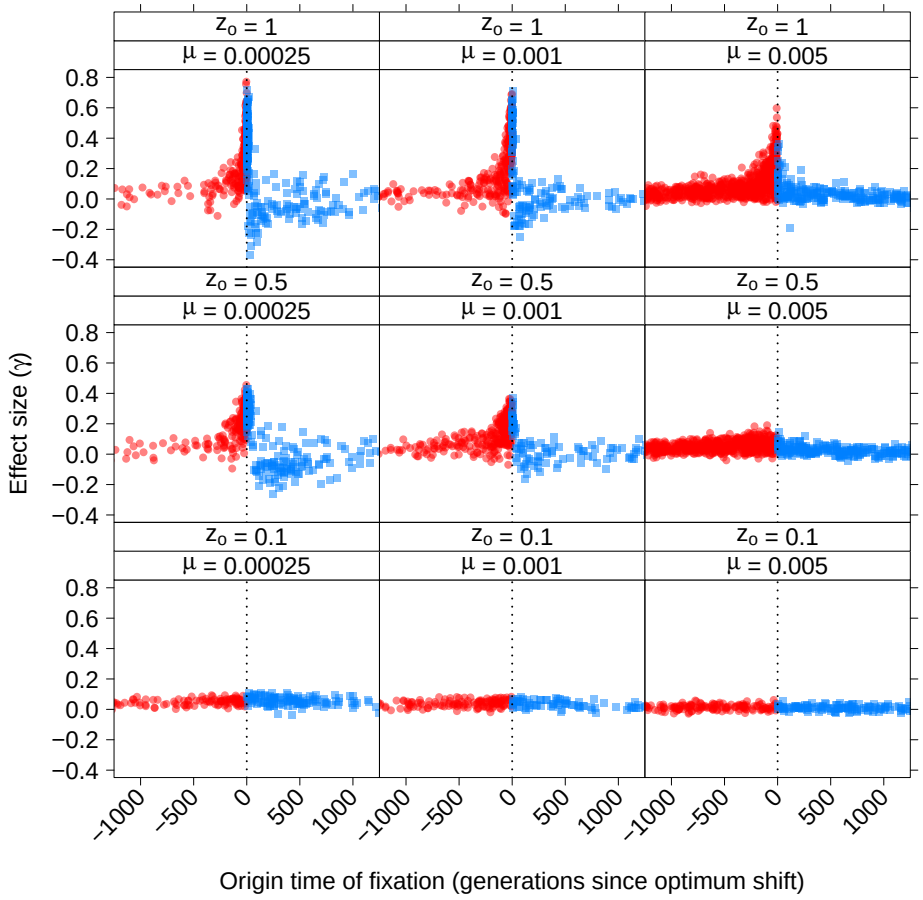


Figure S3: Fixations with large effects on trait values typically arise close to the time of the optimum shift. The x axis shows the origin time of each fixation and the y axis is that fixation's effect size, γ. Time is measured in units of N generations an scaled so that the optimum shift occurred at 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.

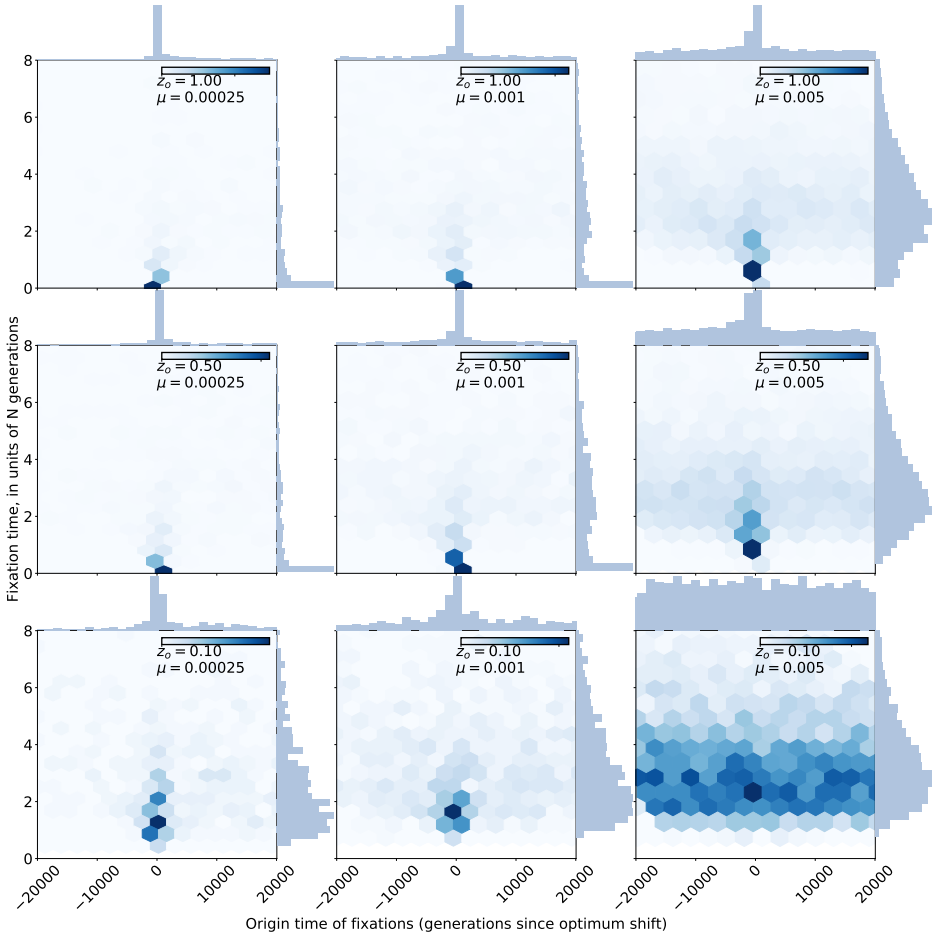


Figure S4: The time to fixation as a function of origin time. The x axis shows the origin time of each fixation in units of N generations and scaled such that the optimum shift is at time zero. The y axis is the fixation time, again in units of N generations. Each panel shows the joint density of origin and sojourn time as well as the marginal distributions, and the raw data are the same as in Figure S3. 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 S3).

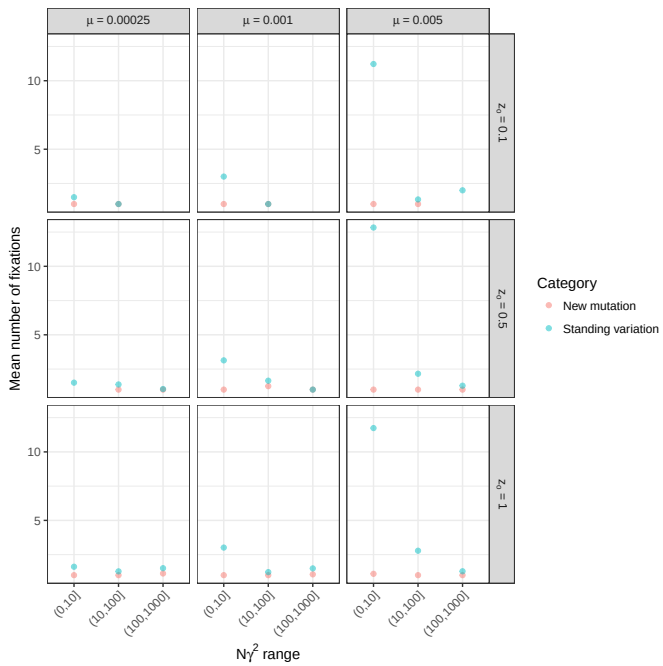


Figure S5: The mean number to fixations of selected mutations in a single-region simulation. The mean values are separated by both scaled strength of selection ($N\gamma^2$) and by fixations from standing variation versus new mutations arising in the first one hundred generations after the optimum shift.

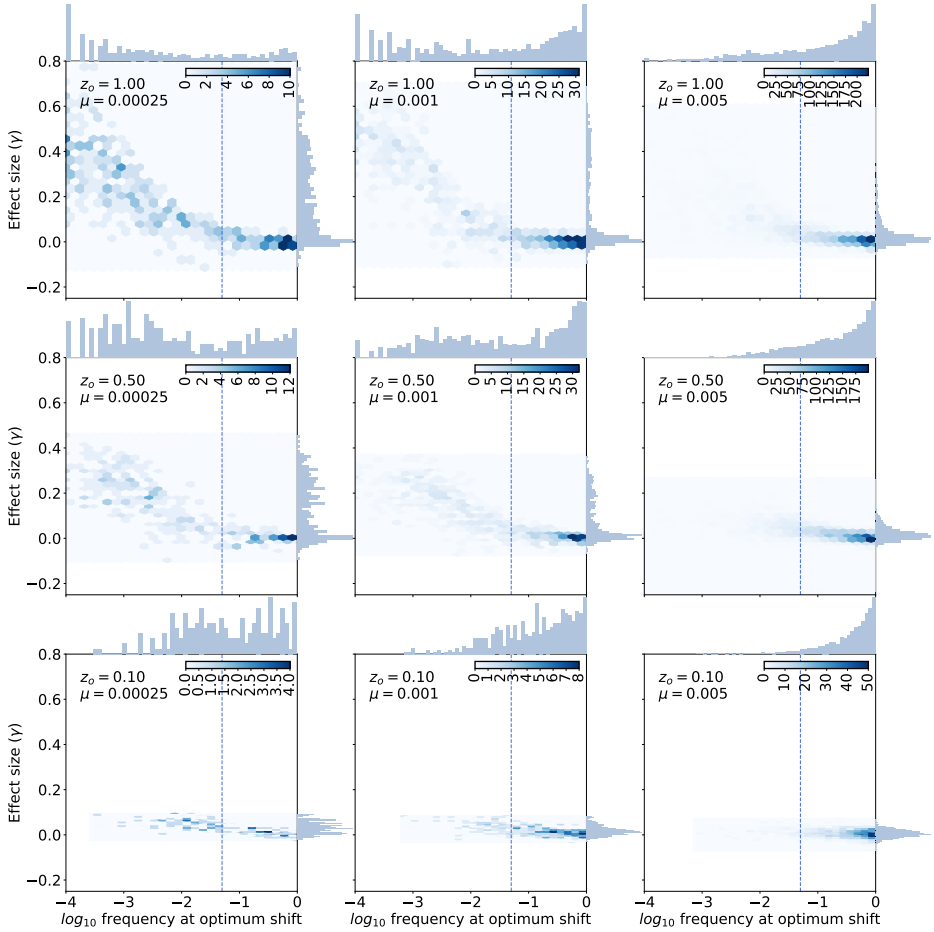


Figure S6: The joint distribution of frequency at the time of the optimum shift and effect size (γ) for fixations from standing variation. Each panel shows the joint density of the mutation frequency at the time of the optimum shift and effect size, using the raw data from Figure S3 The vertical dashed 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. 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. Sweeps from common variants are associated with small effect sizes.

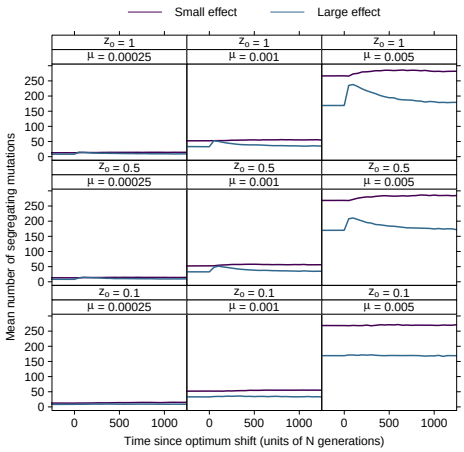
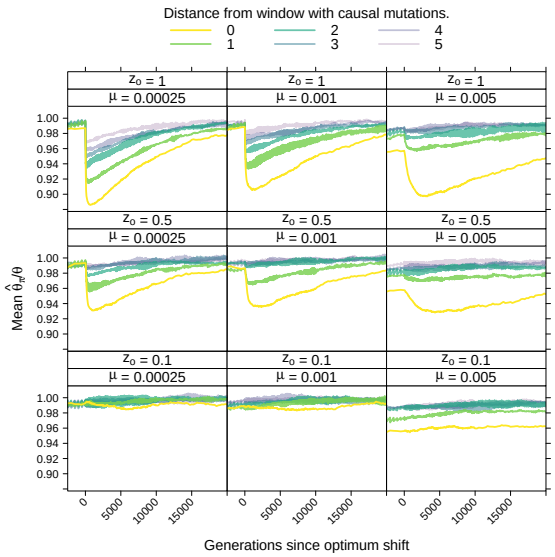
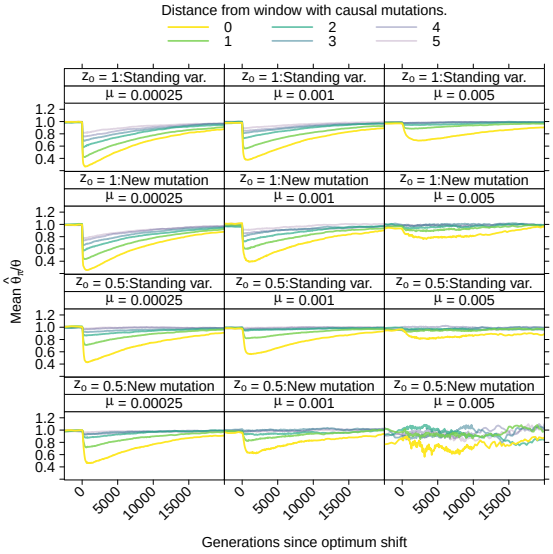


Figure S7: The mean number of segregating mutations affecting the trait over time in ten locus simulations. The different lines are for mutations of small effect ($N\gamma^2 < 100$) and large-effect variants ($N\gamma^2 \geq 100$).

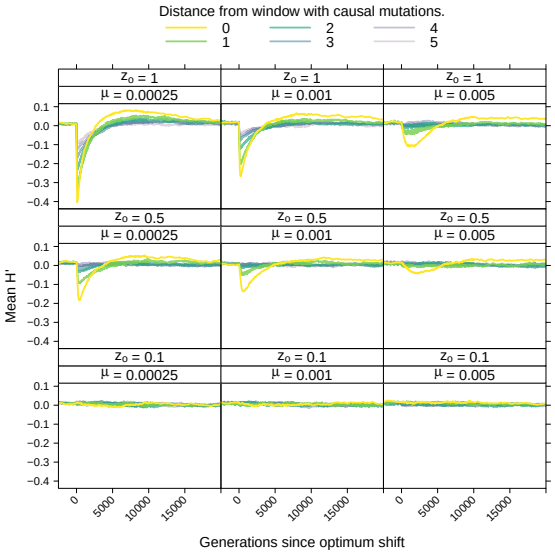


(a) Mean reduction in diversity over time.

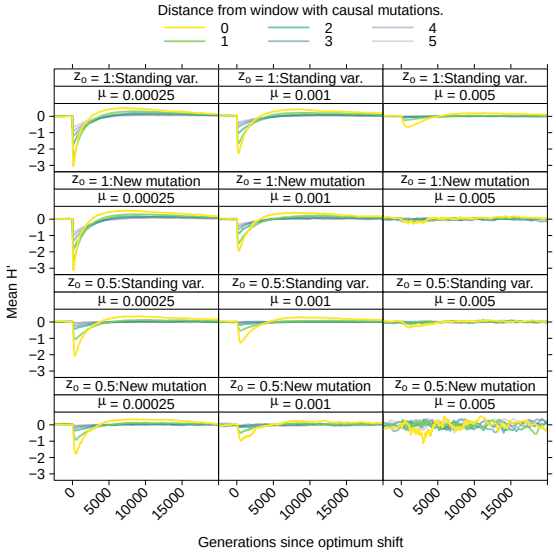


(b) Mean reduction in diversity over time, conditioning on large-effect substitutions.

Figure S8: 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 $\theta = 4N_e\mu$. S8a 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.. S8b shows the mean reduction in diversity at loci experiencing a sweep due to a mutation with effect size $N\gamma^2 \geq 100$. 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.

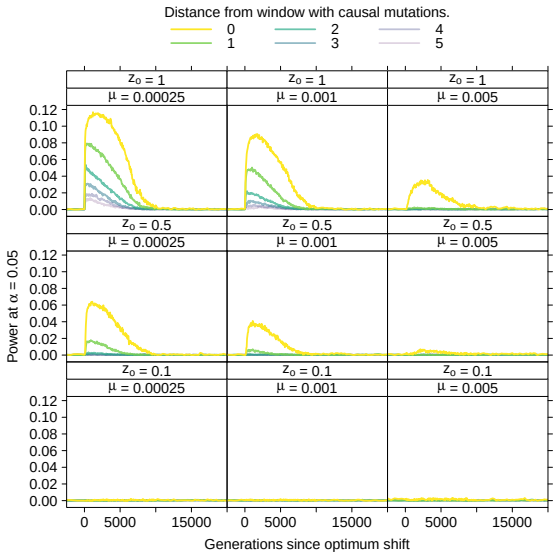


(a) Mean H' over time

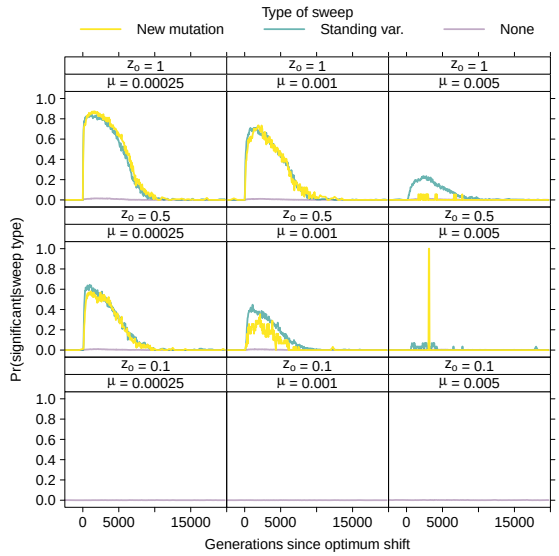


(b) Mean H' over time, conditioning on large-effect substitutions.

Figure S9: Average behavior of H' (Zeng *et al.*, 2006) over time. S9a 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). S9b 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.



(a) Power of Tajima's D



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

Figure S10: Power of Tajima's (1989) to reject the standard neutral model. S10a 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 labeled using the same scheme as Figure 8. S10b The probability of a "significant" window at a locus, conditional on the fixation of a large-effect mutation at the locus.

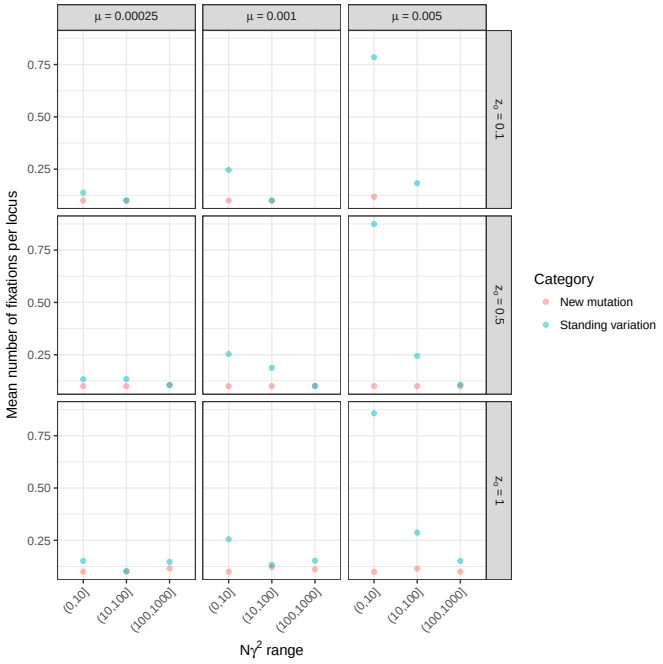
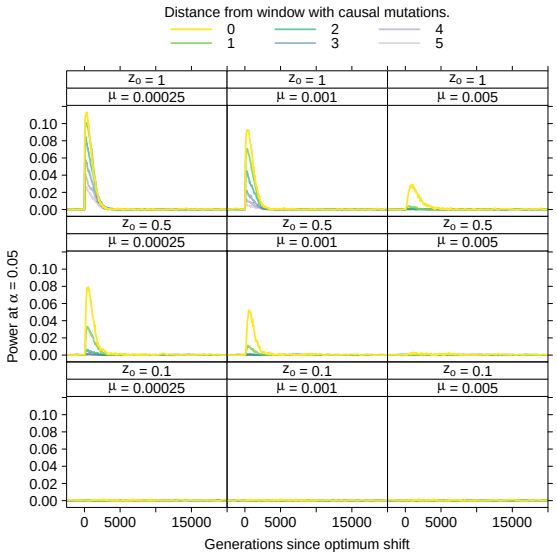
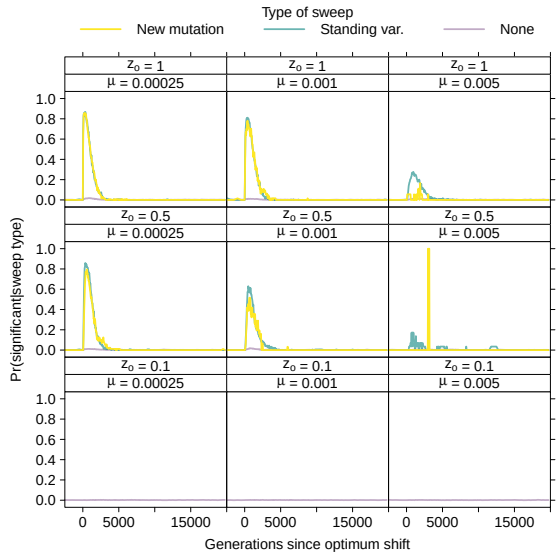


Figure S11: The mean number of fixations per locus. The means are taken for different ranges of $N\gamma^2$. Missing points mean no fixations in the relevant range.

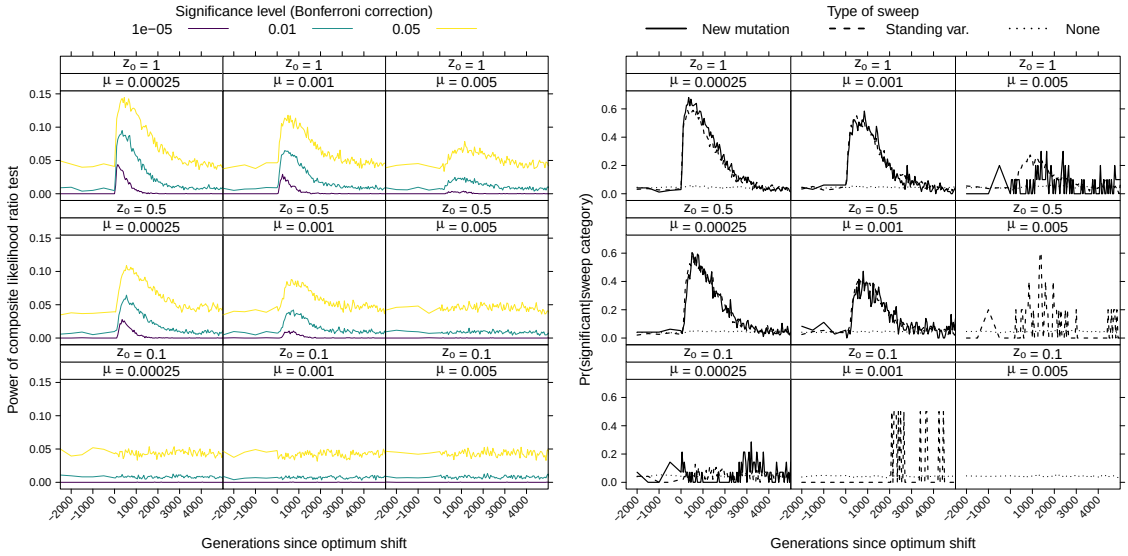


(a) Power of H'



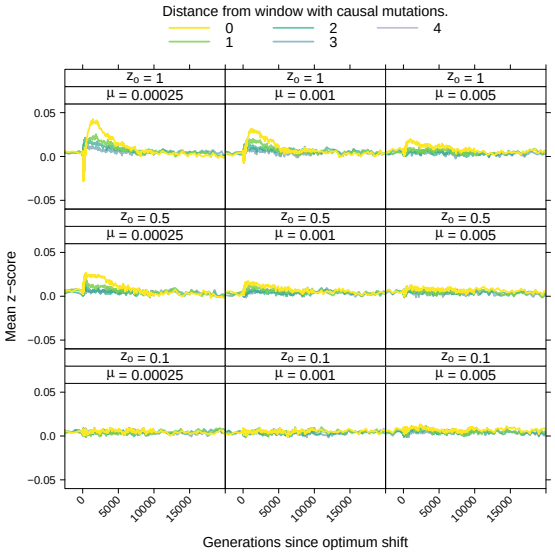
(b) Fraction of significant loci with large-effect fixations.

Figure S12: Power of H' (Zeng *et al.*, 2006) to reject the standard neutral model. S12a 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 labeled using the same scheme as Figure S9. S12b The probability of a "significant" window at a locus, conditional on the fixation of a large-effect mutation at the locus.

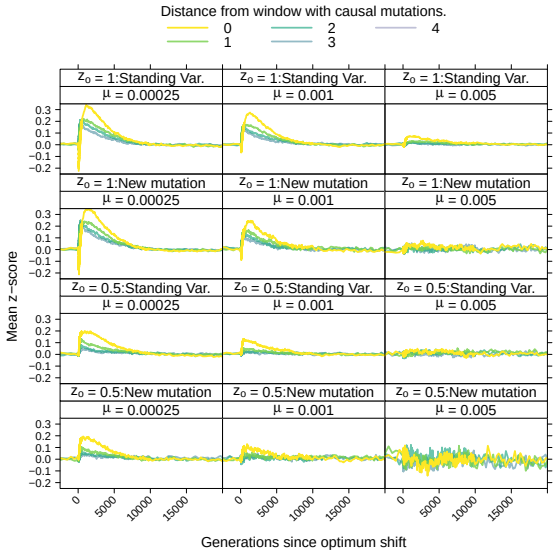


(a) (b)

Figure S13: 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). S13a The power to reject the null model at three different significance levels, after application of a Bonferonni correction. S13b The probability of a significant CLRT, conditional on the type of sweep.



(a) Mean z-score over time



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

Figure S14: 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. S14a shows the average value over time and S14b shows the average value conditioning on fixation of large effect from standing variation or from a new mutation.

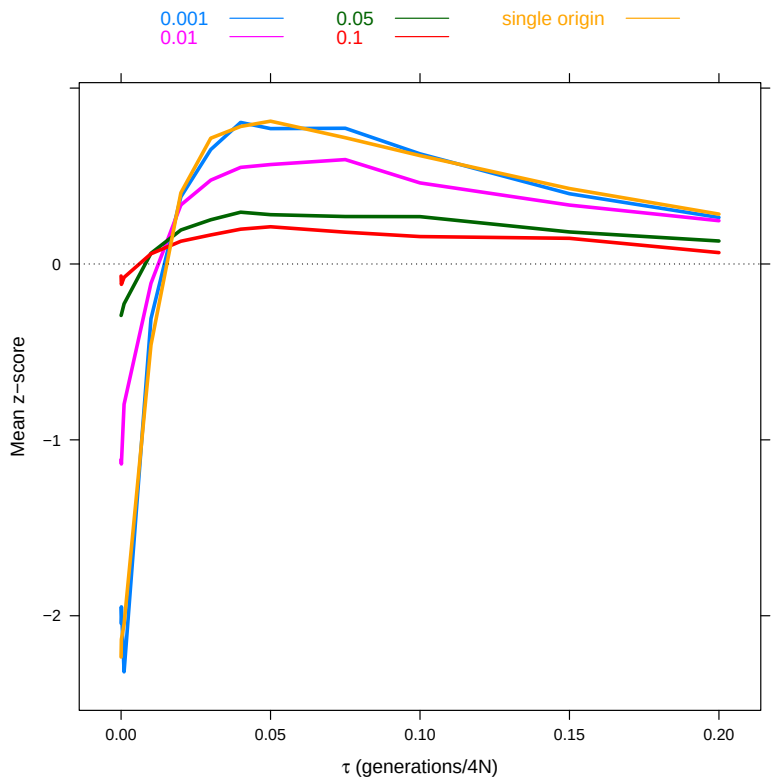
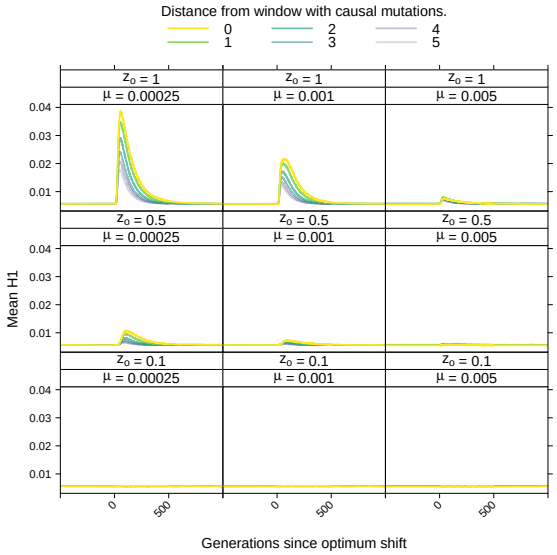
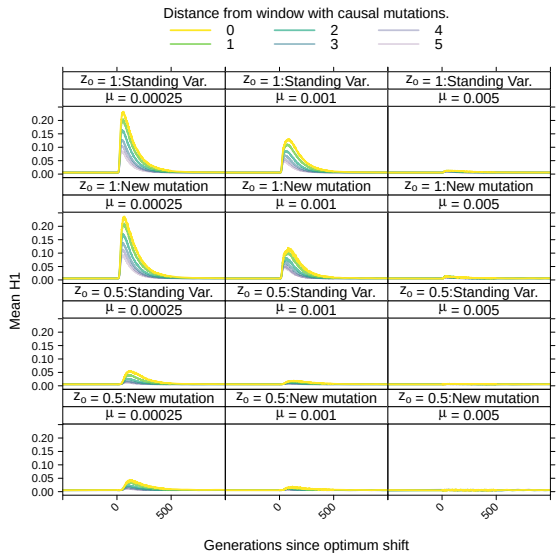


Figure S15: The average z-score associated with the nS_L statistic (Ferrer-Admetlla *et al.*, 2014) as a function of the time since the fixation of a strongly selected mutation. All lines are based on 100 replicates of a coalescent simulation (Kern and Schrider, 2016) of a sample size of 100 haplotypes 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 $1/(2N_e)$.

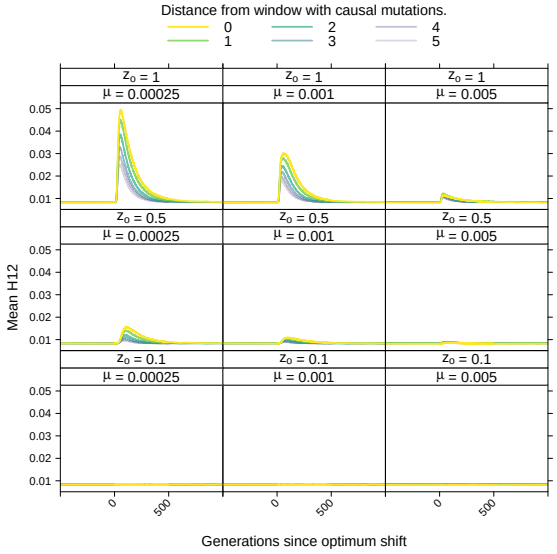


(a) Mean $H1$ over time

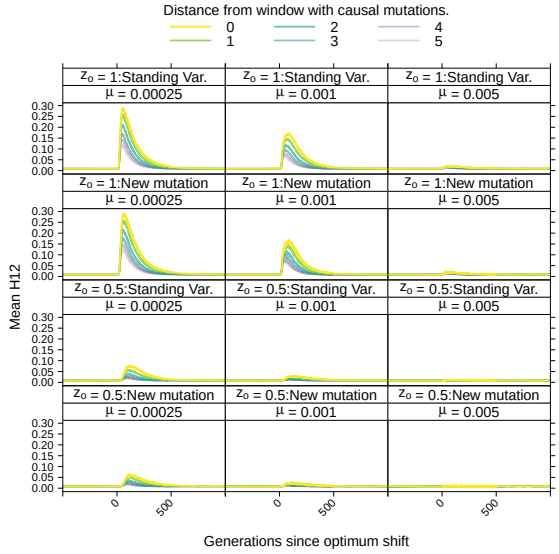


(b) Mean $H1$ over time, conditioning on large-effect substitutions.

Figure S16: The mean $H1$ statistic (Garud *et al.*, 2015) over time. S16a shows the average value over time and S16b shows the average value conditioning on a fixation of large effect either from standing variation or from a new mutation.

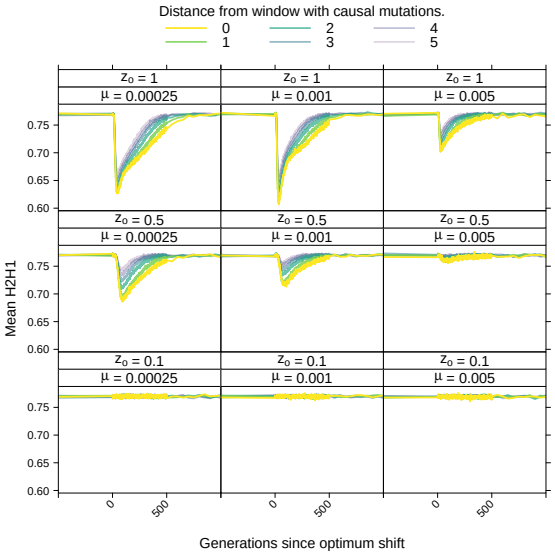


(a) Mean $H12$ over time

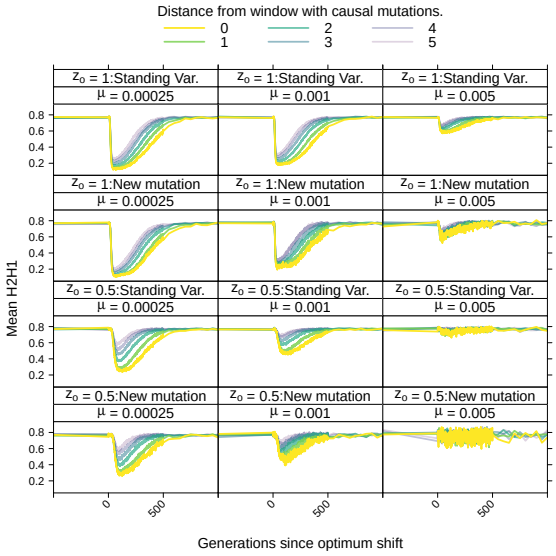


(b) Mean $H12$ over time, conditioning on large-effect substitutions.

Figure S17: The mean $H12$ statistic (Garud *et al.*, 2015) over time. S17a shows the average value over time and S17b 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 S18: The mean $H2/H1$ statistic (Garud *et al.*, 2015) over time. S18a shows the average value over time and S18b shows the average value conditioning on a fixation of large effect either from standing variation or from a new mutation.

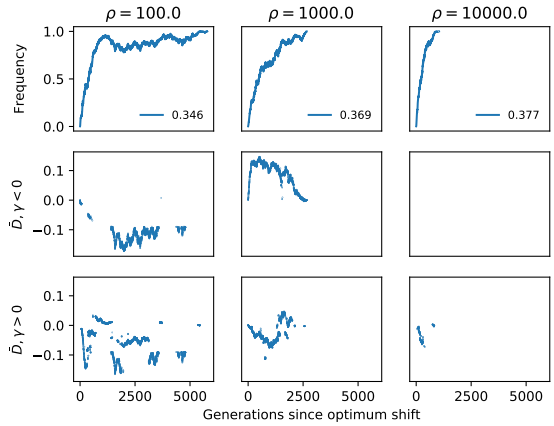


Figure S19: Within-locus linkage disequilibrium with fixations. Each column represents a single simulation replicate of a ten-locus system with $\mu = 5 \times 10^{-3}$ and a different value of $\rho = 4Nr$ per locus (column labels). Here, each simulation happens to have one fixation from a new mutation, whose frequency trajectory is shown in the top row and the effect sizes are shown in the panel legends. The second and third rows show the mean value of $D = p_{AB} - p_A p_B$, where p_A is the frequency of the sweeping variant and p_B is the frequency of another variant affecting trait values at the same locus. D is calculated after filtering variants with minor allele frequencies less than 0.1. The middle row shows $\bar{D}$ with respect to loci with effect sizes $\gamma > 0$ and the third row is with respect to $\gamma > 0$. Gaps in the $\bar{D}$ data mean that no mutations passing the frequency filter were present in that generation.

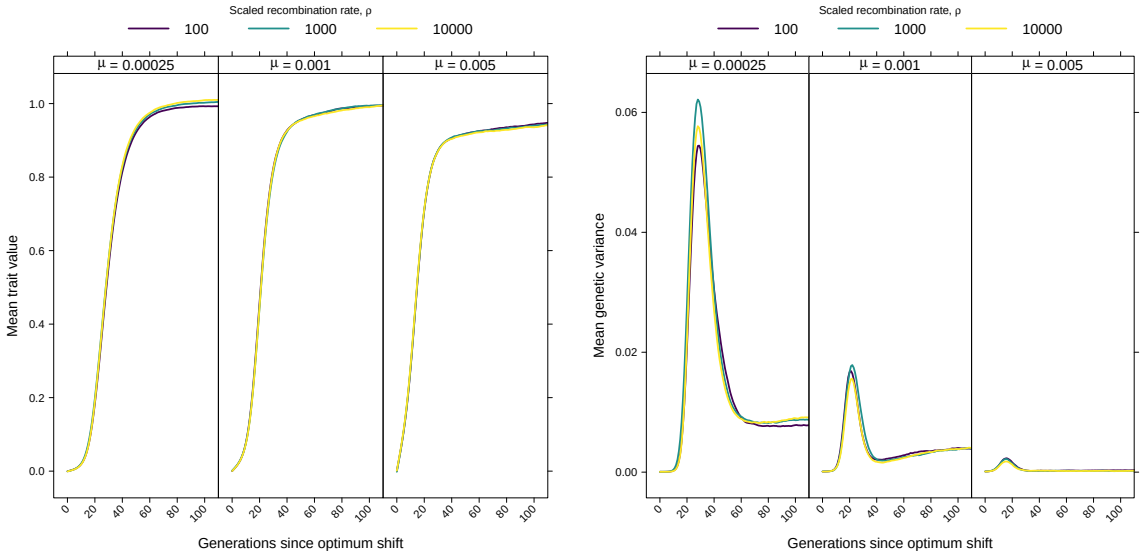
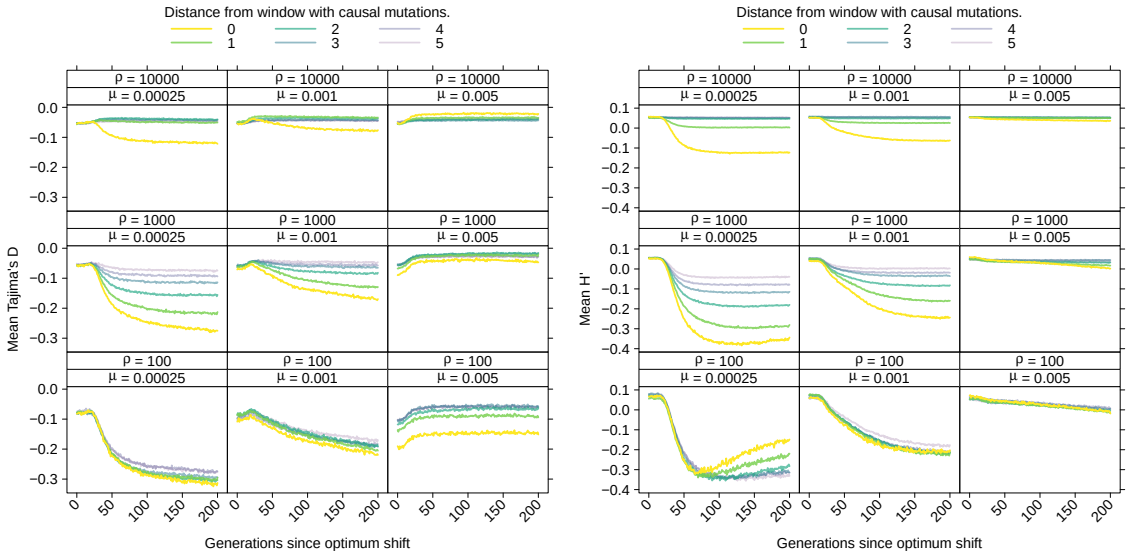


Figure S20: The mean genetic value and variance over time in a ten-locus system as the within-locus recombination rate $\rho = 4Nr$ varies.



(a) (b)

Figure S21: The mean behavior of Tajima's (1989) D (a) and H' (Zeng *et al.*, 2006) (b) over time as the recombination rate within a locus, $\rho = 4Nr$, is varied. The layout of each panel is as in Figure 8, averaging across loci with the same distance from the central window of each locus (Figure 2) each generation.

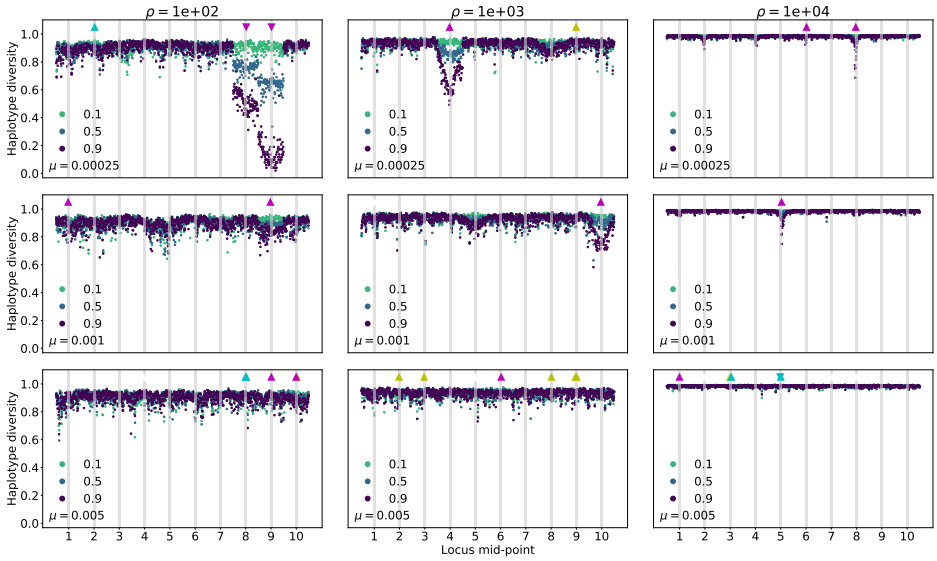
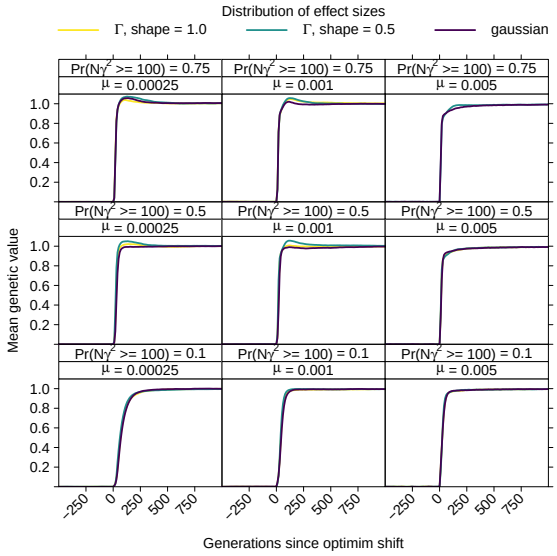
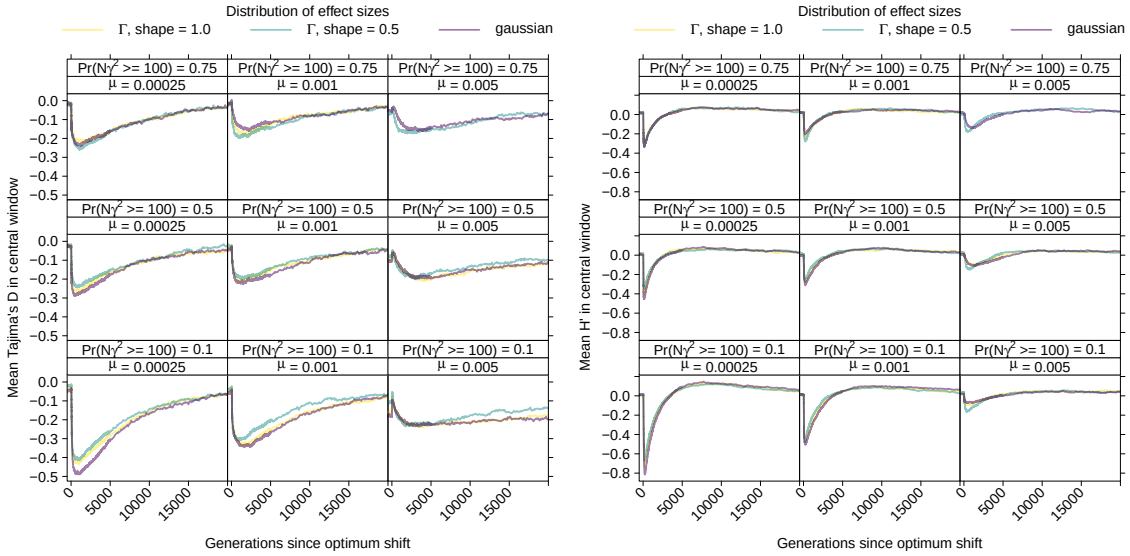


Figure S22: Haplotype diversity along the genome in a ten locus system as the within-locus recombination rate varies. Haplotype diversity is shown for three replicate simulations adapting to an optimum trait value of $z_o = 1$. The center column shows the same data as the final row of Figure 7. The left and right column decrease and increase ρ to 100 and 10,000, respectively. The triangles along the top of each panel show where fixations occurred. Triangles pointing up are fixations from standing variation, and those pointing downwards are fixations from new mutations arising within 100 generations of the optimum shift. Magenta refers to fixations with scaled effect sizes $N\gamma^2 > 100$, cyan indicates $10 < N\gamma^2 \leq 100$, and yellow refers to $1 \leq N\gamma^2 \leq 10$.



(a)

Figure S23: 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 has an effect size $N\gamma^2 \geq 100$ is varied for three different distributions of effect sizes (DES), the Gaussian distribution and Gamma distributions with two difference shape parameters. S23a The mean trait value over time.



(a) (b)

Figure S24: 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 S23, the mutation rate is held constant for three different distributions as $Pr(N\gamma^2 \geq 100)$ varies. S24a 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). S24b The mean H' (Zeng *et al.*, 2006) over time in the central window of a locus.