

Genetic Signatures of Evolutionary Rescue by a Selective Sweep

Matthew M. Osmond¹ and Graham Coop

Center for Population Biology and Department of Evolution and Ecology, University of California, Davis, California 95616

ORCID IDs: 0000-0001-6170-8182 (M.M.O.); 0000-0001-8431-0302 (G.C.)

ABSTRACT One of the most useful models in population genetics is that of a selective sweep and the consequent hitch-hiking of linked neutral alleles. While variations on this model typically assume constant population size, many instances of strong selection and rapid adaptation in nature may co-occur with complex demography. Here, we extend the hitch-hiking model to evolutionary rescue, where adaptation and demography not only co-occur but are intimately entwined. Our results show how this feedback between demography and evolution determines—and restricts—the genetic signatures of evolutionary rescue, and how these differ from the signatures of sweeps in populations of constant size. In particular, we find rescue to harden sweeps from standing variance or new mutation (but not from migration), reduce genetic diversity both at the selected site and genome-wide, and increase the range of observed Tajima's D values. For a given initial rate of population decline, the feedback between demography and evolution makes all of these differences more dramatic under weaker selection, where bottlenecks are prolonged. Nevertheless, it is likely difficult to infer the co-incident timing of the sweep and bottleneck from these simple signatures, never mind a feedback between them. Temporal samples spanning contemporary rescue events may offer one way forward.

KEYWORDS bottleneck; hitch-hiking; pairwise diversity; soft sweep; Tajima's D

THE simple models used to predict the genetic signatures of selective sweeps have been incredibly helpful in understanding and identifying population genetic signals of adaptation (e.g., Maynard Smith and Haigh 1974; Kaplan *et al.* 1989; reviewed in Stephan 2019). These models are usually based on constant-sized, Wright–Fisher populations. Meanwhile, many instances of adaptation—and thus selective sweeps—in nature will co-occur with complex demography. In fact, many of the most well-known examples of selective sweeps have arisen following a rather extreme and sudden change in the environment (e.g., after the application of insecticides, Sedghifar *et al.* 2016, or antimalarial drugs, Nair *et al.* 2003), which could have simultaneously imposed sharp demographic declines. Attempts to capture such complex demographic scenarios typically impose qualitatively appropriate changes in population size (e.g., Hermisson and Pennings

2005). Indeed, a number of studies have explored the genetic signatures of selective sweeps during demographic bottlenecks (e.g., Innan and Kim 2004; Teshima *et al.* 2006; Wilson *et al.* 2014). However, these demographies are nearly always chosen in the absence of an explicit population model and independently of evolution.

Here, we model selective sweeps in a scenario where demography and adaptive evolution are not independent. In particular, we model an instance of evolutionary rescue (Gomulkiewicz and Holt 1995; reviewed in Bell 2017), where a sudden environmental change causes population decline that is reverted by a selective sweep. Under this framework, rescue is a simultaneous demographic bottleneck and selective sweep, where each affects the other. First, because the mean absolute fitness of the population changes with the beneficial allele's frequency, the depth and duration of the bottleneck depends on the dynamics of the selective sweep, *i.e.*, evolution affects demography. Second, the probability that the beneficial allele establishes depends on its growth rate and thus the rate of population decline, *i.e.*, demography affects evolution. Together, this feedback between demography and evolution restricts the range of dynamics that are possible, and therefore also restricts the range of genetic

Copyright © 2020 by the Genetics Society of America

doi: <https://doi.org/10.1534/genetics.120.303173>

Manuscript received March 8, 2020; accepted for publication May 6, 2020; published Early Online December 5, 2020.

Supplemental material available at figshare: <https://doi.org/10.25386/genetics.12284324>.

¹Corresponding author: Center for Population Biology and Department of Evolution and Ecology, University of California, 1 Shields Ave., Davis, California 95616. E-mail: mmosmond@ucdavis.edu

signatures we should expect to observe. Our goal here is to describe the range of genetic signatures that (this model of) evolutionary rescue allows, to help elucidate how rescue may obscure inferences of past selection and demography and to identify patterns that could be used to infer rescue in nature.

Most theory on evolutionary rescue to date (reviewed in Alexander *et al.* 2014) has focused on the probability of rescue. Recently, however, some attention has been given to the dynamics of population size (Orr and Unckless 2014), the probability of soft sweeps (Wilson *et al.* 2017), and the genetic basis of adaptation (Osmond *et al.* 2020) given rescue in haploid or asexual populations. Here, we extend this line of thinking to three modes of rescue in diploid, sexual populations, and use coalescent theory and simulations of whole chromosomes to examine the genetic signatures at linked, neutral loci. Our focus is on three common genetic signatures: the number of unique lineages of the beneficial allele that establish (*i.e.*, the softness of the sweep), the pattern of nucleotide diversity around the selected site (*i.e.*, the dip in diversity), and the pattern of Tajima's D around the selected site (*i.e.*, skews in the site-frequency spectrum). We explore three modes of rescue, where the beneficial allele arises from either standing genetic variance, recurrent *de novo* mutation, or migration, and compare to selective sweeps arising in populations of constant size. Qualitatively, we find that for a given selection coefficient rescue causes faster, harder sweeps that produce wider, deeper dips in diversity, and more extreme values of Tajima's D . Due to the feedback between demography and evolution, the effect of rescue on the signatures of selective sweeps becomes more pronounced as the selection coefficient, and, thus, the probability of rescue, gets smaller.

Materials and Methods

Population dynamics

Deterministic trajectories: Consider a population of size $N(t)$ with a beneficial allele, A , at frequency $p(t)$ and an ancestral allele, a , at frequency $1-p(t)$. Assume nonoverlapping generations, and let W_{AA} , W_{Aa} , and W_{aa} be the absolute fitness (expected number of offspring) of each genotype. We are interested in the scenario where a population composed of primarily aa genotypes is declining at some rate d , *i.e.*, $W_{aa} = 1 - d$. The beneficial allele, A , is then assumed to act multiplicatively with the fitness of the ancestral background, such that $W_{Aa} = (1 - d)(1 + hs)$ and $W_{AA} = (1 - d)(1 + s)$. Throughout the text, we assume random mating, weak selection, $s \ll 1$, and additivity at the selected locus, $h = 1/2$, such that the allele frequency [cf. equation 5.3.12 in Crow and Kimura (1970)] and population dynamics [$N(t + 1) = \bar{W}(t)N(t)$, with $\bar{W}(t)$ the population mean fitness] can be approximated by

$$p(t) \approx \left(1 + \frac{1-p(0)}{p(0)} e^{-st/2}\right)^{-1} \quad (1)$$

$$N(t) \approx N(0)e^{[(1+s)(1-d)-1]t} \left(\frac{p(0)}{p(t)}\right)^{2(1-d)}.$$

Equation 1 allows for unbounded population growth; to prevent this we impose a simple carrying capacity at $N(0)$ by setting $N(t)$ to be the minimum of $N(0)$ and the value given in Equation 1.

Effective initial allele frequency: Equation 1 considers only the deterministic dynamics. We are, however, primarily concerned with a stochastic event: the establishment of a (potentially rare) beneficial allele. In rescue, we are particularly interested in only those instances where the beneficial allele establishes—otherwise the population is extinct. Conditioning on establishment creates a bias away from the deterministic trajectory. As shown in Maynard Smith (1971) [see Orr and Unckless (2014) for an application to evolutionary rescue], we can approximate this bias by assuming that a single copy of an allele that establishes with probability P_{est} instantaneously increases from an initial frequency of $1/(2N)$ to $1/(2NP_{\text{est}})$. In essence, dividing the initial allele frequency by the probability of establishment implies that an allele escaping random loss when rare will quickly increase to a larger frequency than otherwise expected before settling into its deterministic trajectory. As we will see below, we can combine this logic with the waiting time distribution for an establishing mutation to arrive to derive an effective initial allele frequency, p_0 . Using the deterministic prediction (Equation 1) with $p(0) = p_0$ then approximates the forward-time trajectory of an allele that is destined to fixation.

We derive the probability of establishment, P_{est} , assuming alleles do not interact (*i.e.*, a branching process). Under this assumption, a single copy of an allele that is expected to leave $1 + \epsilon$ copies in the next generation with a variance of v has an establishment probability of [Allen (2010), p. 172, see also Feller (1951), equation 5.7, for a derivation from a diffusion approximation]:

$$P_{\text{est}} = 1 - e^{-2\epsilon/v}. \quad (2)$$

When $d \leq 0$, we ignore density-dependence, and the growth rate of the establishing heterozygote then depends only on its absolute fitness, $\epsilon = W_{Aa} - 1 = (1 + sh)(1 - d) - 1 \approx sh - d$. The variance, v , depends on particulars of the lifecycle (see section *Genetic drift in the simulated lifecycle* in the Supplementary material for details). When $h = 1$ and $v = 1$, we retrieve the probability of establishment of a weakly beneficial mutation in an exponentially growing or declining haploid population with a Poisson offspring distribution, $P_{\text{est}} \approx 2(s - d)$ (Otto and Whitlock 1997).

Effective final allele frequency: Above, we have stitched together the initial stochastic phase of the allele frequency and population dynamics with the deterministic phase by using $p(0) = p_0$. We next incorporate the final phase, fixation, which is also stochastic. Incorporating this final phase is important here (as opposed to studies that predict only the forward-time dynamics; *e.g.*, Orr and Unckless 2014) because in looking at genetic signatures we are concerned primarily

with the dynamics backward in time from the point of fixation, and this third phase determines how fast fixation occurs.

As shown in Martin and Lambert (2015), we can treat the backward-in-time dynamics from fixation as we did the forward-in-time dynamics at establishment. In particular, the probability the heterozygote establishes (backward in time) in a population of mutant homozygotes is also given by Equation 2, but with the growth rate of a rare heterozygote in this population, ϵ , replaced by $-\epsilon$ (let's call this probability P_{est}^-). For instance, if the fitness of the heterozygote is $W_{Aa} = (1-d)(1+sh)$ and the fitness of the mutant homozygote is $W_{AA} = (1-d)(1+s)$, and the population is at carrying capacity, then the growth rate of the heterozygote depends only on its relative fitness, $-\epsilon = 1 - W_{Aa}/W_{AA} = 1 - (1+sh)/(1+s) \approx s(1-h)$, which is independent of demography, d .

Given the heterozygote establishes (backward-in-time), it will quickly increase to $1/P_{\text{est}}^-$ copies, implying that, in a population of size N , fixation effectively occurs when the deterministic trajectory reaches a frequency of $p_f = 1 - 1/(2NP_{\text{est}}^-)$. Below, we will explore scenarios where fixation is expected to occur while the population is at carrying capacity, $N(0)$. Thus, in all cases, we have the same effective final allele frequency,

$$p_f = 1 - \frac{1}{2N(0)P_{\text{est}}^-}, \quad (3)$$

where P_{est} is given by Equation 2, with $\epsilon = 1 - (1+sh)/(1+s)$, and ν is as given in the section *Genetic drift in the simulated lifecycle* in the Supplemental text.

Time to fixation: We can estimate the time to fixation in the additive model ($h = 1/2$) as the time it takes the deterministic approximation (Equation 1) to go from the effective initial frequency, p_0 , to the effective final frequency, p_f , which is [equation 5.3.13 in Crow and Kimura (1970)]

$$t_f = \frac{2}{s} \log \left[\frac{(1-p_0)p_f}{p_0(1-p_f)} \right]. \quad (4)$$

Backward-time dynamics: Setting $t = t_f - \tau$ in Equation 1 gives an approximation of the dynamics of allele frequency and population size backward in time,

$$p'(\tau) \approx \left(1 + \frac{1-p_f}{p_f} e^{s\tau/2} \right)^{-1} \quad (5)$$

$$N'(\tau) \approx N(0) e^{-d(t_f-\tau)} \left(\frac{1-p_0}{1-p'(\tau)} \right)^{2(1-d)},$$

from the point of fixation, $\tau = 0$, to the time of establishment, $\tau = t_f$. Note that the backward-time allele frequency dynamics do not depend on the initial frequency, p_0 , or demography, d , except in determining the maximum value of τ , t_f . Thus the backward-time allele frequency dynamics are expected to be the same in rescue as in populations of constant size with the

same carrying capacity (such that the final frequency, p_f , is the same). Meanwhile, the backward-time population size dynamics depend heavily on the initial frequency as the time at which the sweep establishes determines the depth and duration of the bottleneck.

The structured coalescent

Event rates: To explore the genetic patterns created by evolutionary rescue we next consider a random sample of chromosomes at the time the beneficial allele fixes. Focusing on a neutral locus that is recombination distance r from the selected site, we are interested in calculating the rate of coalescence, the rates of recombination and mutation off the selected background, and the rate of migration out of the population. If our sample of alleles has k distinct ancestors on the selected background τ generations before fixation, these rates are approximately [see Table 1 in Hudson and Kaplan (1988); equation 16 in Pennings and Hermisson (2006a); see section *Deriving the structured coalescent for derivations* in the Supplemental text]

$$p_{\text{coal}}(k, \tau) = \binom{k}{2} \frac{1}{2N_e'(\tau)p'(\tau)}$$

$$p_{\text{rec}}(k, \tau) = kr[1 - p'(\tau)]$$

$$p_{\text{mut}}(k, \tau) = k \frac{u[1 - p'(\tau)]}{p'(\tau)}$$

$$p_{\text{mig}}(k, \tau) = k \frac{m}{2N'(\tau)p'(\tau)}, \quad (6)$$

where $N_e'(\tau)$ is the effective population size τ generations before fixation (see section *Genetic drift in the simulated lifecycle* in the Supplemental text for details).

Event timing: We now use the instantaneous event rates (Equation 6) to calculate the probability that the most recent event is τ generations before fixation, and is either coalescence, recombination, mutation, or migration. Letting $i, j \in \{\text{coal}, \text{rec}, \text{mut}, \text{mig}\}$, the probability that i is the most recent event and occurs τ generations before fixation is (cf. equation 6 in Pennings and Hermisson (2006b))

$$P_i(k, \tau) = p_i(k, \tau) \left(\prod_{j \neq i} [1 - p_j(k, \tau)] \right) \prod_{l=0}^{\tau-1} \left(\prod_j [1 - p_j(k, l)] \right)$$

$$\approx p_i(k, \tau) \exp \left(- \sum_j \int_0^\tau p_j(k, l) dl \right), \quad (7)$$

i.e., the waiting time for an inhomogeneous exponential random variable. The approximation assumes the $p_i(k, \tau)$ are small enough such that, at most, one event happens each generation, with small probability, and the changes in the $p_i(k, \tau)$ from one generation to the next are small enough that we can approximate a sum across τ with an integral. As a technical aside, to speed computation we analytically solve

the integrals in Equation 7 under the assumption of unbounded population growth (Equation 1).

Genetic signatures at linked neutral loci

Pairwise diversity: One classic signature of a selective sweep is a dip in genetic diversity around the selected site (Maynard Smith and Haigh 1974; Kaplan *et al.* 1989). Here, we consider the average number of nucleotide differences between two randomly sampled sequences, π (Tajima 1983), focusing on sequences of length 1 (*i.e.*, heterozygosity).

We first consider our expectation for π at a site that is far enough away from the selected site to be unaffected by hitchhiking, which will provide us with an expectation for genome-wide average diversity. While not directly impacted by the sweep, these loci are indirectly impacted as the sweep dictates the severity of the population bottleneck. Our expectation for π at such a site in a population of constant effective size, N_e , is simply $\mathbb{E}[\pi|\text{unlinked}] = \theta \equiv 4N_eU$ (Watterson 1975), with U the per base per generation mutation rate at neutral loci. Ignoring neutral mutation input during the sweep, the π at a sufficiently loosely linked site in a population of changing size is this neutral expectation times the probability a sample of size two does not coalesce during the sweep (*cf.* equation 4 in Slatkin and Hudson (1991) and equation 7 in Griffiths and Tavaré (1994)),

$$\mathbb{E}[\pi|\text{unlinked}] \approx \theta \exp\left(-\int_0^{t_f} \frac{1}{2N_e(\tau)} d\tau\right). \quad (8)$$

We next consider sites that are close enough to the selected locus that they are directly affected by the selective sweep through hitchhiking. To keep the analysis simple, we assume that, if recombination or mutation moves one of the sampled alleles to the ancestral background before the two samples coalesce, then it is as if both samples were on the ancestral background from the start and therefore coalesce with each other as if they were at an unlinked locus (Equation 8). This assumption ignores the time it takes for the two samples to arrive on the ancestral background, and is therefore expected to underestimate diversity at moderately linked sites. There is also the possibility that no events occur in the history of the sample during the sweep. Then, if the sweep arose from mutation or migration, or from a single copy of the beneficial allele ($\kappa = 1$), the sample must coalesce. If we instead start with more than one beneficial copy, $\kappa > 1$, it is possible that the two samples had ancestors that were linked to distinct copies of the beneficial allele within the standing variation. While the coalescent naturally tells us the probability that two lineages do not coalesce during the sweep (since distinct copies of the beneficial allele are exchangeable; *e.g.*, Tavaré 1984), our continuous-time approximation of the coalescent (Equation 7) implicitly assumes a large number of copies of the beneficial allele at any one time. To account for the fact that κ is finite, and often small, we assume that, if, under the continuous-time approximation of the coalescent, there remain two lineages on the beneficial background at the

beginning of the sweep, then they coalesce with probability $1/\kappa$ [see the “instantaneous” approximation in Anderson and Slatkin (2007) for a similar approach]. Finally, we assume that each copy of the beneficial allele in the standing variance has a recent and unique mutational origin (implying the beneficial allele was sufficiently deleterious before the environmental change; *cf.* Prezeworski *et al.* 2005; Berg and Coop 2015), so that two distinct ancestral lineages are independent draws from a neutral population. Ignoring neutral mutations during the sweep, a simple approximation for diversity at any location in the genome is then

$$\mathbb{E}[\pi] \approx P_{\text{off}}(2, t_f)\mathbb{E}[\pi|\text{unlinked}] + P_{\emptyset}(2, t_f)(1 - 1/\kappa)\theta, \quad (9)$$

where $P_{\text{off}}(k, t_f) = \int_0^{t_f} [P_{\text{rec}}(k, \tau) + P_{\text{mut}}(k, \tau)]d\tau$ is the probability of recombination or mutation before coalescence during the sweep and $P_{\emptyset}(k, t_f) = \exp(-\sum_j \int_0^{t_f} p_j(k, \tau)d\tau)$ is the probability that no events have occurred in the history of the sample during the sweep ($j \in \{\text{rec, mut, coal}\}$). Equation 9 is a function of recombination distance, r , and approaches $\mathbb{E}[\pi|\text{unlinked}]$ as r gets large.

Tajima’s D : As a second genetic signature at linked neutral sites we consider Tajima’s D statistic (Tajima 1989), which measures the excess (positive D) or deficiency (negative D) of intermediate frequency polymorphisms relative to the standard neutral model. Quantitative predictions of Tajima’s D require one to consider samples of size greater than two, which quickly becomes complicated with selection and non-equilibrium demography. Instead, here we discuss the expected qualitative patterns based on intuition from the analysis presented above.

First, hard selective sweeps tend to produce star-like gene genealogies, with most samples coalescing near the beginning of the sweep, and recombination allowing a few samples to coalesce much further back in time (Kaplan *et al.* 1989). Hard sweeps therefore produce an excess of low and high frequency polymorphisms (Wakeley 2009, p. 120), leading to negative D (Braverman *et al.* 1995). Larger selection coefficients create more star-like genealogies (as there is then less time for recombination off the sweep), and, thus, more negative D when conditioned on a hard sweep. However, with sufficient standing genetic variation (SGV) or rates of recurrent mutation or migration, larger selection coefficients will tend to cause softer selective sweeps as they increase the probability any one copy establishes (Equation 2). Soft selective sweeps (by definition) allow some samples to coalesce further back in time, before the start of the sweep, even at the selected site. Soft sweeps therefore tend to have less effect on neutral genealogies, and, hence, on D , although sufficiently soft sweeps can actually cause positive D , by allowing intermediate-sized groups of samples to descend from different ancestors containing the beneficial allele (Pennings and Hermisson 2006b).

As linkage to the selected site decreases, so too does this skew in genealogies. In the case of a constant population size, D should asymptote to the neutral expectation of zero. In the case of rescue, however, the bottleneck will cause an excess of intermediate frequency polymorphisms (Wakeley 2009, p. 120), and, therefore, D should asymptote at some positive value (more positive with more severe bottlenecks).

Simulations

The simulation details are described in full in the section *Simulation details* in the Supplemental text. Briefly, the life-cycle described in *Simulated lifecycle* was simulated in SLiM 3 (Haller and Messer 2019) with tree-sequence recording (Haller *et al.* 2019). We simulated a 20 Mb segment of a chromosome, with all but one of the center loci neutral, with a per base recombination rate of $r_{bp} = 2 \times 10^{-8}$ and per base mutation rate at neutral loci of $U = 6 \times 10^{-9}$ (both inspired by *Drosophila* estimates; Mackay *et al.* 2012; Haag-Liautard *et al.* 2007). A simulation was considered successful and ended when the beneficial mutation was fixed and the population size had recovered to its initial size, $N(0)$. Successful simulations were recapitated with Hudson's coalescent algorithm (Hudson 1983, 2002), implemented in msprime (Kelleher *et al.* 2016), using $N_e \approx 4N(0)/7$ (see *Genetic drift in the simulated lifecycle*). Pairwise diversity, Tajima's D , site frequency spectra, and linkage disequilibrium were then calculated directly from the tree sequences using tskit (Kelleher *et al.* 2018), for a random sample of 100 contemporary chromosomes. We chose to simulate a population with initial census size $N(0) = 10^4$ declining at rate $d = 0.05$, where the beneficial allele had selection coefficient $s = 0.13$ or $s = 0.20$. This describes a relatively small population that is expected to go extinct in ≈ 200 generations in the absence of a sweep at a locus under strong selection (as might be expected following a severe environmental change).

Data availability statement

The authors state that all data necessary for confirming the conclusions presented in the article are represented fully within the article and Supplementary Material. Code used to derive and plot all results (Python scripts and Mathematica notebook) is available at <https://github.com/mmosmond/rescueCoalescent.git> and on figshare. File S1 refers to the Mathematica notebook, which is also provided in freely accessible forms (e.g., PDF). Supplemental material available at figshare: <https://doi.org/10.25386/genetics.12284324>.

Results

Rescue from standing genetic variation

We first consider rescue arising from genetic variation that is present at the time of the environmental change and ignore further mutations and migration. To avoid complicating the presentation, we assume the initial number of beneficial

alleles is given; treating the initial number as a random variable requires conditioning the distribution on a successful sweep (Hermisson and Pennings 2017).

The probability of rescue and soft selective sweeps: Given there are initially $\kappa \ll N(0)$ copies of the beneficial allele, the number that establish is roughly binomially distributed with κ trials and success probability P_{est} . The probability of rescue is the probability that at least one establishes (*cf.* equation 2 in Orr and Unckless 2014),

$$P_{\text{rescue}}^{\text{SGV}} = 1 - (1 - P_{\text{est}})^{\kappa}. \quad (10)$$

Conditioning on rescue, the expected number of establishing copies is

$$n_{\text{rescue}}^{\text{SGV}} = \frac{\kappa P_{\text{est}}}{P_{\text{rescue}}^{\text{SGV}}}. \quad (11)$$

Note that when the unconditioned expected number of establishing copies, κP_{est} , is small, the probability of rescue is roughly κP_{est} . Equation 11 then implies that we expect only one of the initial copies to establish, *i.e.*, we expect a hard sweep. For larger values of κP_{est} , rescue will often occur by a soft selective sweep (Hermisson and Pennings 2005), where multiple initial copies establish. The probability that multiple copies establish is the probability of rescue minus the probability of a hard sweep, $P_{\text{rescue}}^{\text{SGV}} - \kappa P_{\text{est}}(1 - P_{\text{est}})^{\kappa-1}$. Given rescue occurs, the probability it is due to a soft sweep can therefore be written as

$$P_{\text{soft|rescue}}^{\text{SGV}} = 1 - n_{\text{rescue}}^{\text{SGV}} \frac{1 - P_{\text{rescue}}^{\text{SGV}}}{1 - P_{\text{est}}}. \quad (12)$$

As expected, the probability of a soft sweep given rescue is 0 when we start with a single copy, $\kappa = 1$, and approaches 1 as the number of copies becomes large. Between these two extremes, we find that Equation 12 provides reasonable estimates for small d or large s but underestimates the probability of a soft sweep otherwise (Figure 1A), when beneficial alleles can persist at low numbers long enough to establish with some non-negligible probability of experiencing some selection as homozygotes (thus increasing their probability of establishment). Using the probability of establishment observed in the simulations corrects this error (see File S1). Because the expected number of copies that establishes given rescue (Equation 11) is nearly independent of P_{est} (as long as it is small), Equation 11 more closely matches simulations (Figure 1B).

Effective initial allele frequency and the backward-time dynamics: As each of the $n_{\text{rescue}}^{\text{SGV}}$ expected establishing copies given rescue is expected to rapidly reach $1/P_{\text{est}}$ copies (see section *Effective initial allele frequency*), it is as if the sweep began deterministically (Equation 1) with initial frequency (*cf.* equation S1.4 in Orr and Unckless 2014)

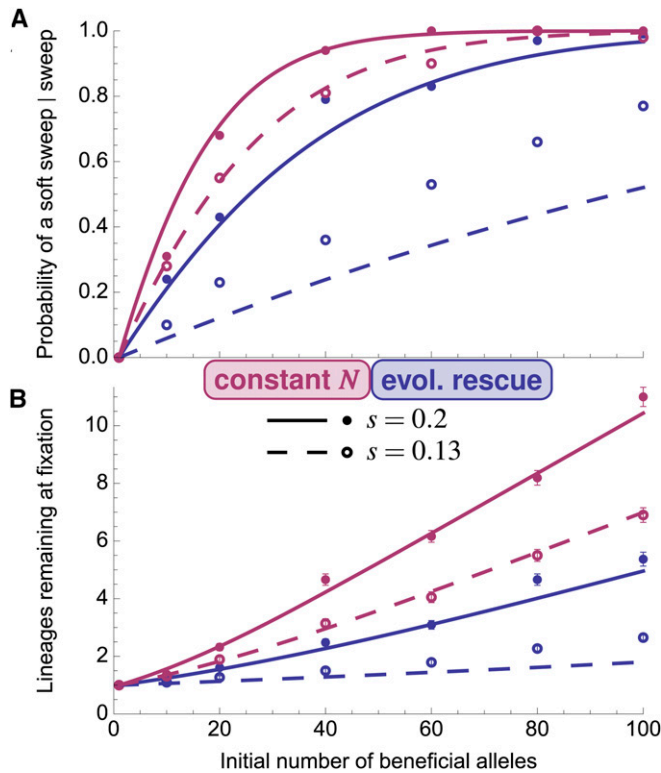


Figure 1 Soft sweeps from standing genetic variation. (A) The probability that more than one initial copy of the beneficial allele has descendants at the time of fixation in rescue (blue; $d = 0.05$) or in a population of roughly constant size (red; $d = 0$) as a function of the initial number of copies of the beneficial allele, κ . (B) The expected number of initial copies of the beneficial allele that have descendants at the time of fixation. The curves are Equation 12 (A) and Equation 11 (B). Each point is based on 100 replicate simulations where a sweep (and population recovery) was observed. Error bars in (B) are standard errors. Parameters: $N(0) = 10^4$.

$$p_{0|\text{rescue}}^{\text{SGV}} = \frac{n_{\text{rescue}}^{\text{SGV}}}{2N(0)P_{\text{est}}} = \frac{\kappa}{2N(0)} \frac{1}{p_{\text{rescue}}^{\text{SGV}}}. \quad (13)$$

This same conditioning applies in a population of constant size (*i.e.*, $d = 0$). As the decline rate, d , increases, the probability a copy establishes, P_{est} , and, hence, the probability of rescue, $p_{\text{rescue}}^{\text{SGV}}$, declines, making the conditioning stronger (*i.e.*, increasing the difference between the effective and true initial allele frequencies). This causes selective sweeps to get started faster as d increases, implying that, all else equal, rescue sweeps are expected to be shorter than those in populations of constant size.

Using $p_{0|\text{rescue}}^{\text{SGV}}$ as p_0 in Equations 4 and 5 gives our semideterministic prediction of the backward-time dynamics. As shown in Figure 2, this does a reasonably good job of describing the simulation results. Looking closer, we find that our predictions do an excellent job of approximating the mean allele frequency as it departs from 1, but typically begin to underestimate the mean allele frequency as frequencies drop lower (see File S1 for more detail). Consequently, we tend to underestimate the mean time to fixation (compare

arrows and stars in Figure 2) and overestimate mean population size. This error arises even under a true branching process (*e.g.*, under clonal reproduction) in a population of constant size, and becomes worse as selection becomes weaker (see File S1). The reason for this error is that one minus the effective final allele frequency, $1 - p_f = q_f$ is exponentially distributed with expectation $1/(2N(0)P_{\text{est}}^-)$ (Martin and Lambert 2015), where P_{est}^- is the backward-time establishment probability of the heterozygote in a population of mutant homozygotes and we have assumed fixation occurs when the population size is $N(0)$ (which is expected to be true in all of our numerical examples). The variance in the effective final allele frequency therefore increases like $1/(P_{\text{est}}^-)^2 \sim 1/s^2$ as s declines, meaning that, under weaker selection, there is much more variance in the time spent in the stochastic phase near fixation. Combining this distribution of waiting times to leave the stochastic phase with the nonlinearities of the subsequent deterministic phase causes our semideterministic approximation to generally underestimate allele frequency. Despite this error, we forge ahead and use our simple semideterministic approximations in the structured coalescent (Equations 6 and 7).

Before moving on to the coalescent, however, a number of biological insights can be gleaned from Figure 2 that will help explain downstream results: (1) the backward-time allele frequency dynamics are the same in rescue as they are in populations of constant size, since they depend only on the effective final allele frequency, p_f , which does not depend on d (Equation 3), (2) fixation times are shorter in rescue as, forward-in-time, those sweeps tend to get started faster (*i.e.*, the effective initial frequency increases with d ; Equation 13), (3) weaker selection causes larger population bottlenecks as the deterministic portion of the sweep is then slower and a given frequency change has less impact on population mean fitness (Equation 1), and (4) both the allele frequency and population size dynamics depend only weakly on the initial number of mutants, despite large differences in the probability of rescue, because sweeps that are less likely to occur get started faster (Equation 13).

The structured coalescent: For a linked neutral allele in a sample of size 2, Figure 3 compares the predicted probability of recombination off the selective sweep (opaque dashed curves) and coalescence on the beneficial background (opaque solid curves) with the mean probabilities given the allele frequency and population size dynamics observed in simulations (transparent curves). We find that our approximations do relatively well near fixation, where our predictions of allele frequency are better (Figure 2), but that our underestimates of allele frequency at later times cause us to generally overestimate rates of recombination and coalescence on the beneficial background near the beginning of the sweep. Despite this error, we capture the qualitative dynamics and can draw out a number of interesting biological consequences: (1) the timing of the bottleneck during rescue pushes coalescent times toward the present, so that the distributions of

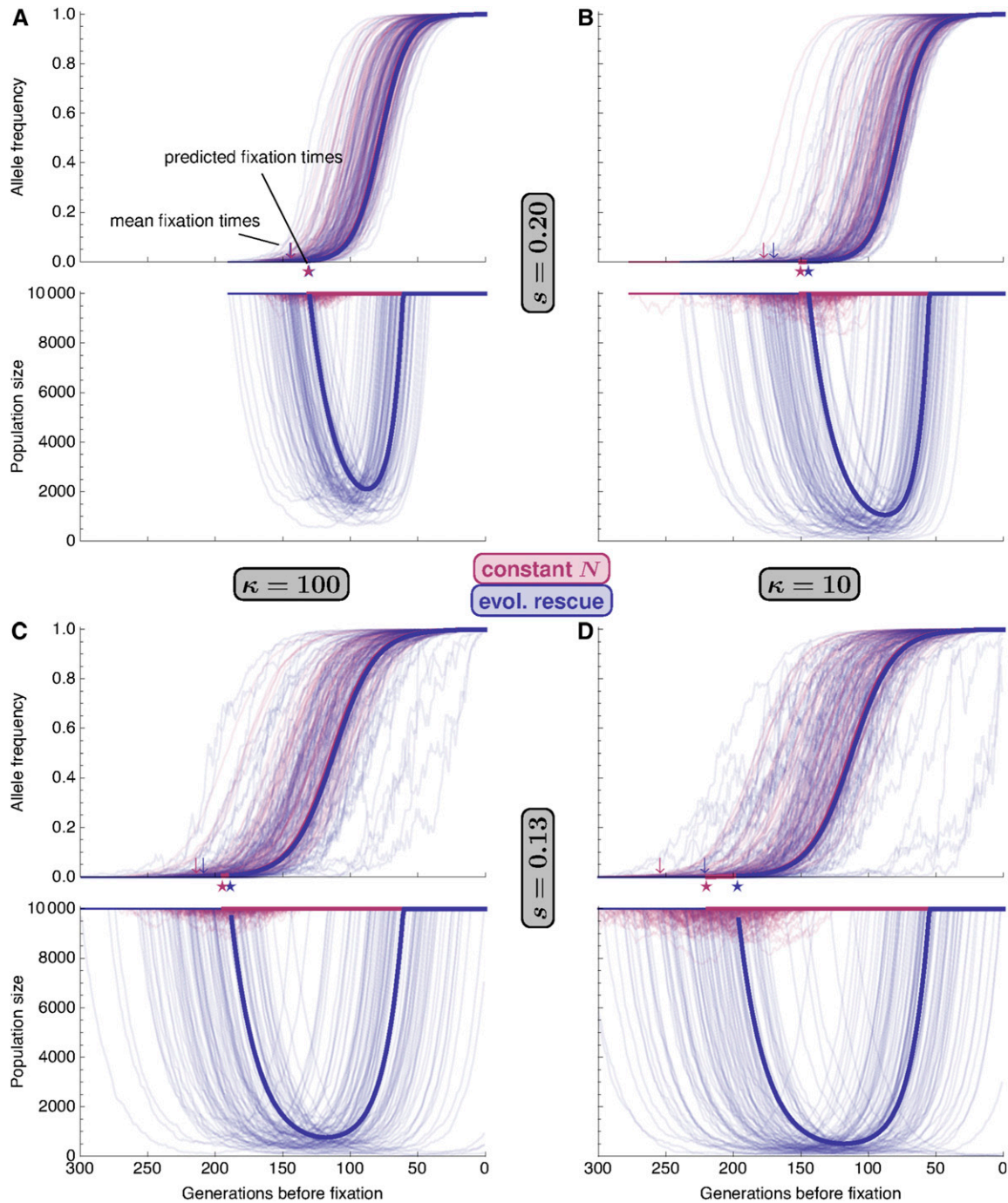


Figure 2 Allele frequency, $p'(\tau)$, and population size, $N'(\tau)$ at generation τ before fixation during a selective sweep from standing genetic variation in evolutionary rescue (blue; $d = 0.05$) and in a population of roughly constant size (red; $d = 0$). The thick solid curves are analytic approximations (Equation 5), using $p_0 = p_{0|rescue}^{SGV}$ (Equation 13) as the initial frequency and p_f (Equation 3) as the final allele frequency. The thin curves are 100 replicate simulations where a sweep (and population recovery) was observed. The stars show the predicted time to fixation (Equation 4). The arrows show the mean time to fixation observed in simulations.

coalescence and recombination times overlap more than in populations of constant size, (2) the bottleneck also increases the overall probability of coalescence in rescue, which reduces the probability of recombination off the sweep (compare areas under dashed curves), and (3) the difference in the structured coalescent between rescue and a population of constant size is larger under weaker selection. This latter

point nicely illustrates the coupling between demography and evolution in rescue; while weaker selection creates a slower sweep, and, hence, more time for recombination off the beneficial background, it also slows population recovery, leading to longer and deeper bottlenecks that increase coalescence enough to counteract the additional time provided for recombination.

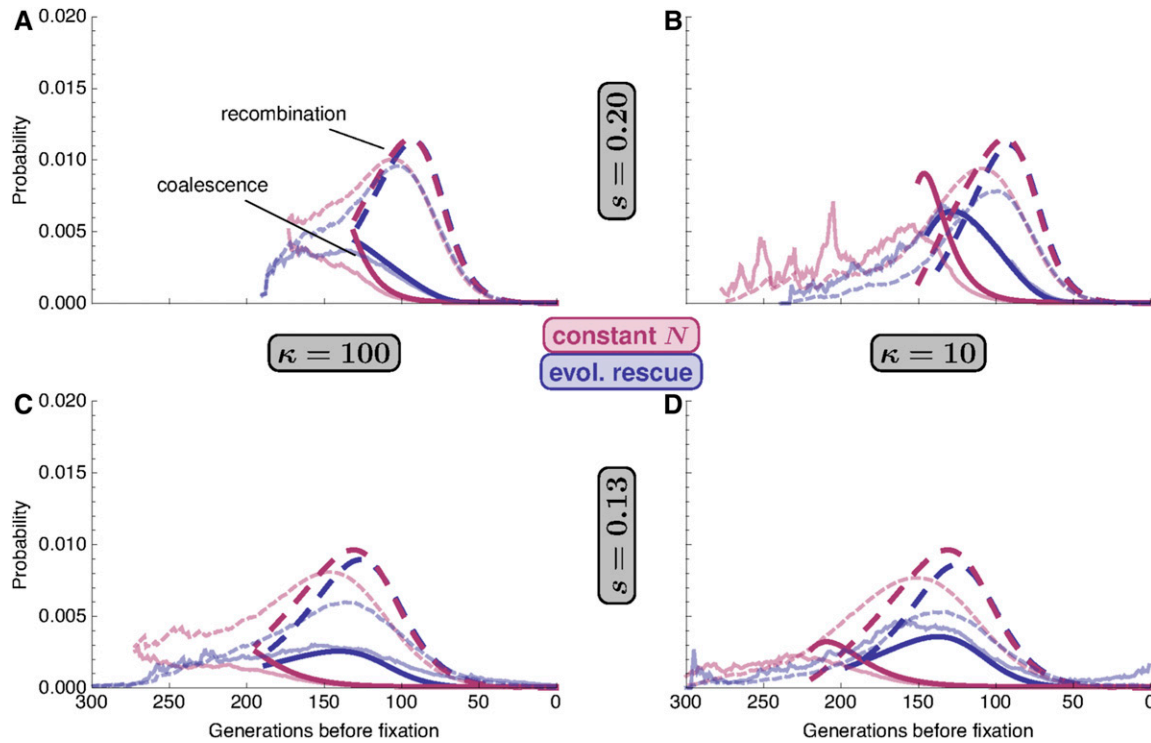


Figure 3 The timing of events in the structured coalescent (Equation 7) for a sample of size $k = 2$ at a linked neutral locus ($r = 0.01$) during a selective sweep from standing genetic variation in evolutionary rescue (blue; $d = 0.05$) or in a population of roughly constant size (red; $d = 0$). The opaque curves use the analytic expressions for allele frequency and population size (Equation 5) while the transparent curves show the mean probabilities given the observed allele frequency and population size dynamics in 100 replicate simulations (these become more variable back in time as less replicates remain polymorphic).

Genetic diversity: Figure 4 compares our predictions of relative pairwise diversity, $\mathbb{E}[\pi]/\mathbb{E}[\pi|\text{unlinked}]$ (Equations 8 and 9), against simulations, $\pi/\bar{\pi}$. Despite the errors in the predictions of the dynamics above (Figure 2 and Figure 3), our approximations capture the relative rates of recombination and coalescence. Two main conclusions emerge: (1) rescue tends to deepen dips in diversity when soft sweeps are possible, *i.e.*, it hardens soft sweeps because the bottleneck increases the probability of coalescence [decreasing $P_{\emptyset}(2, T)$ at the selected site, $r = 0$; Equation 9], and (2) rescue generally produces wider dips in diversity due to excess, and earlier, coalescence as well shorter sweeps (the magnitude and timing of the bottleneck, as well as lower establishment probabilities, decreases $P_{\text{off}}(2, T)$ at a given $r > 0$; Equation 9). Note that when the probability of rescue is very small (Figure 4D), we predict the diversity at the selected site to actually be higher under rescue than in a population of constant size. This is driven by our prediction of fixation time; low probabilities of rescue imply high effective initial allele frequencies, leaving less time for coalescence to occur (compare solid curves in Figure 3D). This prediction is not, however, borne out in the observed simulations, as we tend to overestimate coalescence in populations of constant size and underestimate it in rescue (Figure 3D).

If the population was sampled both before and after the selective sweep (or we have very good estimates of its

mutation rate and long-term effective population size), the absolute amount of pairwise diversity contains additional information. In the Supplemental Material, Figure S1 compares our predictions of $\mathbb{E}[\pi]$ (Equation 9) against simulations, showing that the bottleneck decreases background levels of diversity as well. This decrease is more evident under weaker selection, where bottlenecks are more pronounced. While our predictions qualitatively match the mean simulation results, our tendency to underestimate allele frequencies, and, thus, overestimate harmonic mean population sizes during rescue (Figure 2) causes us to generally overestimate background diversity in these cases. To correct for this, we can replace our prediction for background diversity, $\mathbb{E}[\pi|\text{unlinked}]$ (Equation 8), in the rescue scenario with the observed genome-wide average diversity (dashed curves in Figure S1). A very similar result is achieved by using effective population sizes observed in simulations in Equation 8 (results not shown). Using the observed mean diversity level may be justified by the fact that genome-wide diversity can be measured directly from data and is highly variable across populations (Tajima 1983).

Tajima's D : Figure 5 shows the Tajima's D values observed in simulations. There are two main take-aways: (1) the bottleneck during rescue causes positive background D and (2) rescue sweeps tend to be harder and thus cause a greater

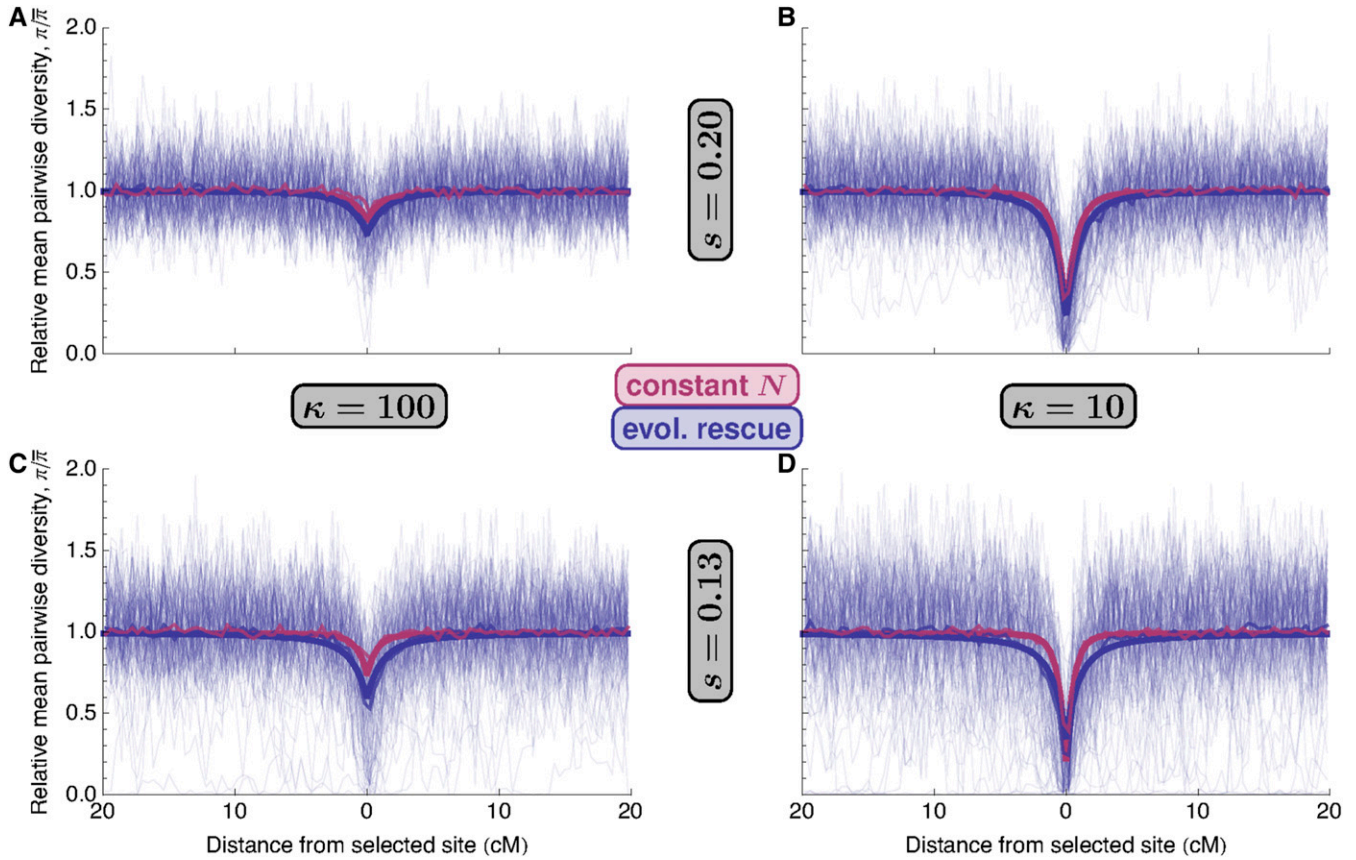


Figure 4 Relative mean pairwise diversity, $\pi/\bar{\pi}$, after a selective sweep from standing genetic variation in evolutionary rescue (blue; $d = 0.05$) or in a population of roughly constant size (red; $d = 0$). The thickest curves are $\mathbb{E}[\pi]/\mathbb{E}[\pi|\text{unlinked}]$ (Equation 9). The thinnest curves are 100 replicate simulations (rescue only for clarity) and the slightly thicker curves are simulation means (often obscured by prediction). Parameters: $N(0) = 10^4$.

decrease (or smaller increase) of D at the selected site. Together, these patterns cause rescue to “stretch out” (e.g., panel D) or even invert (e.g., panel C) the pattern of D observed in populations of constant size.

A simultaneous, but independent, bottleneck and sweep: It will of course be much harder to differentiate rescue from a simultaneous, but independent, bottleneck and sweep. In *A simultaneous, but independent, bottleneck and sweep from SGV* (Supplemental Material) we compare rescue, with expected effective population size N_r (from Equation 1), to sweeps that occur in populations suddenly bottlenecked to size N_r . Because rescue has higher population sizes at the beginning and end of the sweep the sweeps take longer, leaving more time for coalescence, which leads to harder sweeps and lower genome-wide diversity. To see if the slightly different timings of recombination relative to coalescence are detectable, we examine the full site frequency spectrum as well as linkage disequilibrium. There is perhaps a slightly more uniform spectrum across intermediate frequency mutations at tightly linked sites under rescue, as expected given coalescence then overlaps more with recombination (similar to the effect of recurrent mutation, Pennings and Hermisson 2006b). Linkage disequilibrium

is elevated under rescue, largely mirroring the patterns we see in absolute diversity.

Rescue by *de novo* mutation

The probability of rescue and soft selective sweeps: When there are few copies of the beneficial allele at the time of environmental change, rescue may depend on mutations arising *de novo* at the selected site during population decline. To predict the allele frequency and population size dynamics in this scenario, we then need to derive the waiting time until a rescue mutation successfully establishes. Ignoring unsuccessful mutations, the first successful rescue mutation arrives according to a time-inhomogeneous Poisson process with rate, $\lambda(t) = 2N(t)uP_{\text{est}}$, where $2N(t) = 2N(0)e^{-dt}$ describes the decline in the number of copies of the ancestral allele and u is rate at which ancestral alleles mutate to rescue alleles. Thus, the probability that a rescue mutation has established by time T is $F(T) = 1 - e^{-\int_0^T \lambda(t)dt}$. Taking the limit as time goes to infinity then gives the probability of rescue [cf. equation 10 in Orr and Unckless (2008)]

$$P_{\text{rescue}}^{\text{DNM}} = 1 - e^{-2N(0)uP_{\text{est}}/d}. \quad (14)$$

We can also calculate the probability of a soft sweep during rescue from recurrent mutation. Taking into account the

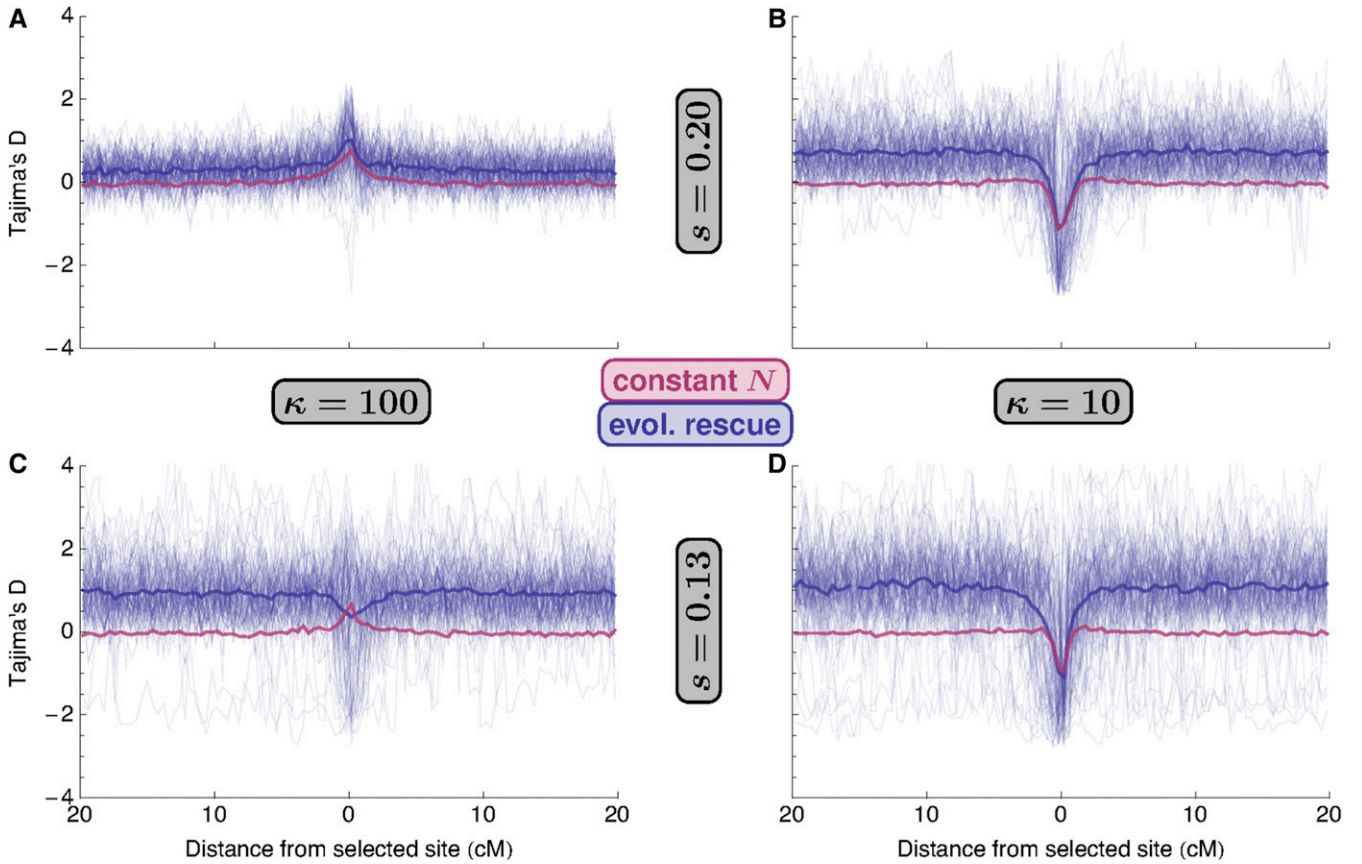


Figure 5 Tajima's D after a selective sweep from standing genetic variation in evolutionary rescue (blue; $d = 0.05$) or in a population of roughly constant size (red; $d = 0$). Thin curves show 100 replicate simulations (rescue only for clarity) and thicker curves show simulation means. Parameters: $N(0) = 10^4$.

beneficial alleles present at time t , the rate at which additional copies arise and establish is $\lambda(t) = 2N(t)q(t)uP_{\text{est}}(t)$, where $2N(t)q(t)$ is the number of ancestral alleles at time t , and $P_{\text{est}}(t)$ is the probability of establishment at time t , which changes with allele frequency, and thus time, because allele frequency influences the genotypes the rescue allele experiences (*i.e.*, its marginal fitness). Thus the number of mutations that arise and fix is Poisson with rate $\int_0^\infty \lambda(t)dt$. To gain intuition, we make a very rough approximation, assuming $q(t) \approx 1$ while mutations are arriving [*i.e.*, while $N(t)$ is still large], so that $P_{\text{est}}(t) \approx P_{\text{est}}$ and we get the same Poisson rate we derived above for the first successful mutation, $\int_0^\infty \lambda(t)dt \approx 2N(0)uP_{\text{est}}/d$. The resulting probability distribution for the number of mutations that establish is analogous to the result in a model with haploid selection [*cf.* equation 7 in Wilson *et al.* (2017)]. Dividing the expected number of establishing mutations by the probability of rescue, the expected number that establish given rescue is

$$n_{\text{rescue}}^{\text{DNM}} = \frac{-\log(1 - p_{\text{rescue}}^{\text{DNM}})}{p_{\text{rescue}}^{\text{DNM}}}. \quad (15)$$

With these same approximations, the probability of a soft selective sweep given rescue (*i.e.*, the probability that more than one copy establishes) is

$$P_{\text{soft|rescue}}^{\text{DNM}} = 1 - n_{\text{rescue}}^{\text{DNM}} (1 - p_{\text{rescue}}^{\text{DNM}}), \quad (16)$$

as in a haploid population [equation 8 in Wilson *et al.* (2017)].

Both of these approximations tends to be underestimates (Figure 6), especially when selection is weak. This is partly because we also tend to underestimate the probability of rescue, which is elevated by increases in the frequency of the beneficial allele, which increases the marginal fitness of the ancestral allele whenever $h > 0$ (prolonging its persistence and, therefore, creating more opportunity for mutation) and increases the marginal fitness of the mutant allele (increasing the establishment probability). However, using the observed probability of rescue in Equations 15 and 16 still results in underestimates (see File S1) because ignoring increases in the beneficial allele frequency becomes less reasonable as more copies are established.

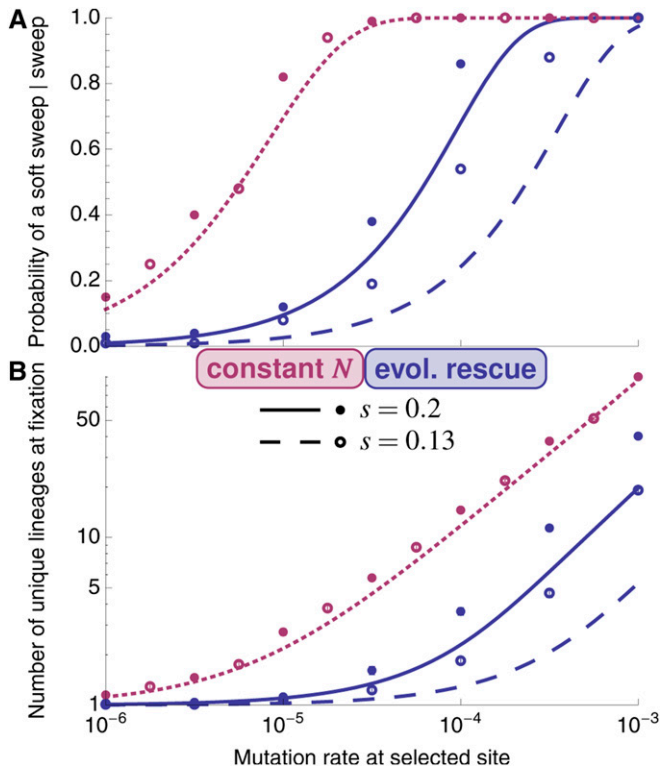


Figure 6 Soft sweeps from *de novo* mutation. (A) The probability that more than one mutant lineage has descendants at the time of fixation in rescue (blue; $d = 0.05$) or in a population of roughly constant size (red; $d = 0$) as a function of the mutation rate, u . (B) The expected number of mutant lineages that have descendants at the time of fixation. The blue curves are Equation 16 (A) and Equation 15 (B). The red curves are from Ewens' sampling formula [equation 11 (A) and equation 12 (B) in Pennings and Hermisson (2006a), with $\theta = 2N_e(0)u = 8N(0)u/7$ and $n = 2N(0)$]. Each point is based on 100 replicate simulations where a sweep (and population recovery) was observed. Error bars in (B) are standard errors. Parameters: $N(0) = 10^4$.

Effective initial allele frequency and the backward-time dynamics:

Following Orr and Unckless (2014) (see their Text S2 and our File S1 for more detail), taking the derivative of the cumulative distribution of waiting times for the first successful mutation, $F(T)$, and dividing by the probability of rescue gives the probability distribution function for the arrival time of the first establishing rescue mutation given rescue, $f(t)$. As discussed above, while the first establishing mutation is still rare it will grow exponentially at rate $\epsilon = (1 + sh)(1 - d) - 1$, and conditioned on its establishment, will very quickly reach $1/P_{\text{est}}$ copies. Integrating over the distribution of arrival times then gives the expected number of copies of this successful mutation at time t since the environmental change given rescue, $\int_0^\infty (e^{(t-\tau)\epsilon}/P_{\text{est}})f(\tau)d\tau$. Solving this integral, evaluating at $t = 0$, and dividing by the total number of alleles at the time of environmental change, $2N(0)$, it is therefore as if the successful mutation was present at the time of environmental change, with frequency

$$p_{0|\text{rescue}}^{\text{DNM}} = \frac{1}{2N(0)} \frac{1}{P_{\text{est}}} \times \frac{(1 - P_{\text{rescue}}^{\text{DNM}})[\Gamma(1 + \epsilon/d, \log(1 - P_{\text{rescue}}^{\text{DNM}})) - \Gamma(1 + \epsilon/d)]}{P_{\text{rescue}}^{\text{DNM}} \log(1 - P_{\text{rescue}}^{\text{DNM}})^{\epsilon/d}}, \quad (17)$$

where $\Gamma(z)$ is the gamma function [Equation 6.1.1 in Abramowitz and Stegun (1972)] and $\Gamma(a, x)$ is the incomplete gamma function [equation 6.5.3 in Abramowitz and Stegun (1972)]. The factor $1/P_{\text{est}} > 1$ increases the effective initial frequency, because we have conditioned on establishment, while the last factor decreases the effective initial frequency, because we must wait for the mutation to arise. When the unconditioned expected number of rescue mutations, $-\log(1 - P_{\text{rescue}}^{\text{DNM}}) = 2uN(0)P_{\text{est}}/d$, is small, this cancels out and the last factor in Equation 17 becomes approximately $d/(d + \epsilon)$, which is independent of the mutation rate (as in the haploid case; Orr and Unckless 2014). That is, conditioning on unlikely rescue, rescue mutations arise earlier in populations that decline faster.

With constant population size, the time of establishment of the first successful mutation is irrelevant for the backward-in-time dynamics, and, therefore, for the resulting genetic signatures (in contrast, when $d > 0$ we need to predict the time of establishment to predict the population size dynamics; Equation 5). For populations of constant size, we need only determine the effective allele frequency at the time the sweep begins, which is simply $1/(2N(0)P_{\text{est}})$, independent of the mutation rate.

Figure S2 compares our analytical approximations of the dynamics (using Equations 5 and 17) against individual-based simulations. As with rescue from standing genetic variance (Figure 2), the nonlinearities of the deterministic phase cause our semideterministic approximation to underestimate mean allele frequencies and overestimate mean population sizes as the allele frequency declines from fixation. We again see that rescue sweeps tend to be shorter than those in populations of constant size (due to lower establishment probabilities) and that the bottleneck sizes and sweep times given rescue depend little on the probability of rescue.

The structured coalescent: Figure S3 shows the timing of coalescence, recombination, and mutation for a sample of size two at a linked neutral locus. The patterns of coalescence and recombination are essentially the same as observed during sweeps from standing genetic variation (Figure 3). The timing of mutation is qualitatively similar to that of coalescence, peaking near establishment (since both depend on the inverse of allele frequency). The main effect of the excess coalescence during rescue is a reduced probability of mutating off the sweep (compare areas under dotted curves), which is exacerbated by rescue sweeps being shorter.

Genetic signatures at linked neutral loci: Figure S4 shows relative pairwise diversity around the selected site. The patterns here are nearly identical to those found after sweeps from standing genetic variation (Figure 4), although the

difference between rescue and the constant population size case are larger here, perhaps due partly to the reduction in rates of mutation off the selected sweep (and also because our parameter choice causes the bottlenecks to be deeper when rescue occurs by mutation). Figure S5 shows Tajima's D values around the selected site, which are nearly identical to those observed after sweeps from standing genetic variation (Figure 5).

On a related note, Wilson *et al.* (2017) reasoned that, because soft sweeps from recurrent mutation are expected when rescue is likely while hard sweeps are expected when rescue is rare (Equation 16), population bottlenecks will tend to be more extreme when rescue occurs by a hard selective sweep. From this they argued it might actually be easier to detect soft sweeps from patterns at linked neutral loci, as bottlenecks are expected to obscure the signal. Here, we show the importance of conditioning on rescue, which roughly equalizes bottleneck sizes across scenarios with very different probabilities of rescue (*e.g.*, due to differences in mutation rate; Figure S2), potentially making harder sweeps easier to detect due to their greater effect on local gene genealogies (Figures S3–S5).

Rescue by migrant alleles

The probability of rescue and soft selective sweeps: Rescue can also arise from beneficial alleles that arrive via migration. Assuming that the number of migrant alleles that replace a resident allele each generation is Poisson with mean m , the waiting time until the first successful migrant is exponential with rate $\lambda = mP_{\text{est}}$. Given the population is expected to persist in the absence of beneficial alleles for $\log(N(0))/d$ generations, the probability of rescue is therefore roughly the probability the first successful migrant arrives by then,

$$P_{\text{rescue}}^{\text{MIG}} = 1 - N(0)^{-mP_{\text{est}}/d}. \quad (18)$$

Under rescue from standing variation or new mutation we derived the probability and extent of soft selective sweeps from the forward-time process. In contrast, under rescue from migration we use the coalescent. In particular, Equation 6 shows that the per generation probability of migration and coalescence depend on population size and beneficial allele frequency in the same way. This similar form implies that the relative rates at which lineages coalesce and migrate at the selected site does not depend on the population size and allele frequency. Pennings and Hermisson (2006a) used this fact to show that, in an ideal population of constant size, the number of unique migrant haplotypes contributing to a present-day sample, as well as their proportions, is described by Ewens' sampling formula (pp. 334ff in Ewens 2004) when we replace θ with $2m$. Powerfully, this results holds even in nonideal populations of changing size (as briefly noted by Pennings and Hermisson 2006a, pp. 1081–1082)—including during evolutionary rescue—as long as the relationship between the effective and census population sizes remains the same [*i.e.*, if $N_e(t)/N(t)$ is a constant; in which case we

now replace θ with $2mN_e/N$ in Ewens' sampling formula]. Thus the softness of a sweep from migration depends only on the migration rate and variance in gamete numbers (σ^2 , which determines N_e/N ; see section *Genetic drift in the simulated lifecycle* in the Supplemental text), and is the same during rescue as it is in a population of constant size (*i.e.*, it is independent of d). The analogous result for rescue by *de novo* mutation does not hold (as it does for a population of constant size, Pennings and Hermisson 2006a), since the rate of mutation is not inversely proportional to population size (Equation 6).

Here, we use just two properties of Ewens' sampling formula, the expected number of unique migrants among a sample of size n (p. 336 in Ewens 2004),

$$n_{\text{rescue}}^{\text{MIG}} = \sum_{j=1}^n \frac{2mN_e/N}{j-1+2mN_e/N}, \quad (19)$$

and the probability that this is >2 (equation 10.9 in Ewens 2004),

$$P_{\text{soft|rescue}}^{\text{MIG}} = 1 - \prod_{j=1}^{n-1} \frac{j}{j+2mN_e/N}. \quad (20)$$

Figure 7 shows that these formulas perform very well, even when we sample the entire population [$n = 2N(0)$].

Population dynamics and the coalescent: In the section *Population dynamics and the coalescent under rescue by migration* in the Supplemental text, we derive the expected backward-time dynamics and structured coalescent under rescue by migration. The results are closely analogous to rescue from mutation. We do not explore genetic signatures at linked neutral sites in this case as these depend on the demographic history of the metapopulation. Previous work has explored some potential signatures of migrant sweeps in populations of constant size (*e.g.*, Setter *et al.* 2019).

Discussion

Here, we have explored genetic signatures of evolutionary rescue by a selective sweep. By allowing demography to depend on the absolute fitness of the genotypes that comprise the population, we explicitly invoke a feedback between demography and evolution. This feedback restricts the range of dynamics, and, thus, the signatures, that one should expect to observe. We find that, because the probability an allele with a given selective advantage establishes is reduced in declining populations (Equation 2; see also Otto and Whitlock 1997), selective sweeps causing rescue are expected to be harder than those in populations of constant size when sweeps arise from standing genetic variance or recurrent mutation (Figures 1 and 6; consistent with Wilson *et al.* 2014, 2017). Further from the selected locus, the demographic bottleneck experienced during rescue increases the rate of coalescence

relative to mutation and recombination (Figure 3 and Figure S3), creating wider dips in relative diversity (Figure 4 and Figure S4) and lower absolute diversity genome-wide (Figure S1; consistent with Innan and Kim 2004). Like absolute diversity, Tajima's D captures both the hardening of the sweep and the demographic bottleneck, causing D to often reach both higher and lower values under rescue (Figure 5 and Figure S5). These differences between evolutionary rescue and standard sweeps all become larger under weaker selection (where the heterozygote has a smaller growth rate, $s/2 - d$) as the slower sweeps that result imply deeper, longer bottlenecks during rescue. If, instead, we compare rescue to populations bottlenecked to the same effective population size during the sweep, we find that the remaining signatures of rescue are harder sweeps that take longer to complete, allowing more time for coalescence, and, hence, lower genome-wide diversity and elevated linkage disequilibrium (Figure A3). The subtle difference in the timing of recombination between the two scenarios, with recombination overlapping more with coalescence during rescue, may also increase variation in the number of derived mutations at a site (as expected following soft sweeps from recurrent mutation in populations of constant size; Pennings and Hermisson 2006b), giving rise to site frequency spectra with perhaps slightly "flatter bottoms" (Figure A2).

In contrast to standing variance or mutation, when sweeps arise from a constant rate of migration demography has no affect on the number of beneficial alleles that establish (as briefly noted by Pennings and Hermisson 2006a), and, thus, rescue has no affect on the hardness of the sweep (Figure 7). Further, because the rates of coalescence and migration are both inversely proportional to the number of beneficial alleles [$N(t)p(t)$, Equation 6, Figure A5], the distribution of the number and frequency of migrant haplotypes spanning the selected site is given by Ewens' sampling formula (Ewens 1972), with mutation replaced by migration (Pennings and Hermisson 2006a). The signatures of rescue by migration at linked neutral loci depend on the demographic and selective history of the metapopulation, which we do not explore here.

Evolutionary rescue has been explored theoretically (e.g., Gomulkiewicz and Holt 1995; Uecker and Hermisson 2016; Anciaux *et al.* 2018) and observed repeatedly in both experiments (e.g., Bell and Gonzalez 2009; Lindsey *et al.* 2013; Ramsayer *et al.* 2013) and in host–pathogen systems in nature (e.g., Wei *et al.* 1995; Feder *et al.* 2016). More recently, a number of studies have examined genetic data following purported evolutionary rescue in the wild, including bats altering hibernation to survive white-nose syndrome (Gignoux-Wolfsohn *et al.* 2018), killifish deleting receptors to tolerate pollution (Oziolor *et al.* 2019), and tall waterhemp evolving herbicide resistance (Kreiner *et al.* 2019). In the latter two cases, there is strong evidence of a recent selective sweep by a very beneficial allele (in one of these cases the evidence includes reduced nucleotide diversity at the selected site; Kreiner *et al.* 2019). Genetic evidence for a putatively recent demographic bottleneck was presented in only one of these

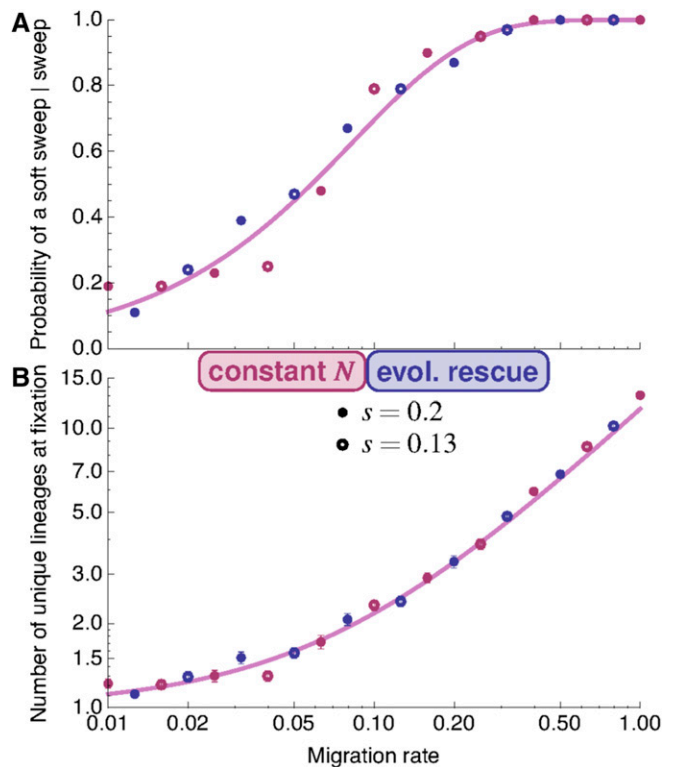


Figure 7 Soft sweeps from migrant alleles. (A) The probability that more than one migrant lineage has descendants at the time of fixation in rescue (blue; $d = 0.05$) or in a population of roughly constant size (red; $d = 0$) as a function of the migration rate, m . (B) The expected number of migrant lineages that have descendants at the time of fixation. The curves are Equation 20 (A) and Equation 19 (B) using $n = 2N(0)$ (i.e., sampling the entire population). Each point is based on 100 replicate simulations where a sweep (and population recovery) was observed. Error bars in (B) are standard errors. Parameters: $N(0) = 10^4$.

studies (Oziolor *et al.* 2019), being detected from genome-wide reductions in nucleotide diversity and increases in Tajima's D relative to populations that did not experience the same environmental change. Developing a statistical model to fit our theory to such data in a more quantitative way could lead to parameter estimates, which could then inform how likely rescue may have been (or if persistence was assured regardless, $d \approx 0$). To do so one would also have to incorporate the time since fixation, as the site frequency spectra and linkage disequilibrium will begin a return to their neutral expectations following the sweep and bottleneck (e.g., the sweep signatures in Tajima's D and linkage disequilibrium decay within $\sim 2N_e$ generations; Przeworski 2002). Ignoring the time since fixation would therefore cause underestimates of s and d , making rescue look more like a weaker sweep in a constant population. In contrast, the deeper and wider dips in genetic diversity observed under rescue would likely inflate estimates of selection if demography was ignored.

A lingering question that helped motivate this study is whether one could ever infer evolutionary rescue from genetic data alone. Such a possibility would greatly help assess the

relevance of rescue in nature. In the empirical examples just discussed, we implied that evidence for a sweep and a bottleneck is consistent with rescue. Yet stronger support for rescue would come from inferring the coincident timing of the sweep and bottleneck. Our analysis suggests that detecting coincident sweeps and bottlenecks from site frequency spectra and linkage disequilibrium will be difficult as the relative timing of coalescence and recombination leaves only subtle signatures. Consistent with this, explicit estimates of the timing of a sweep (e.g., Ormond *et al.* 2016) or a bottleneck (reviewed in Beichman *et al.* 2018) given a genetic sample collected from a single time point come with relatively large amounts of uncertainty. Sampling before and after the potential rescue event may provide one way forward, as it would help tighten the bounds on the timing of the bottleneck and sweep. However, it should be noted that, even if the sweep and bottleneck appear to have co-occurred, this correlation in timing does not imply it was caused by a feedback between demography and evolution. It is, of course, difficult to say in any case—without observing replicate populations go extinct or performing experiments—whether extinction would have occurred (or will occur) without adaptive evolution, as required by the strict definition of evolutionary rescue. More experiments (such as in Rêgo *et al.* 2019) that explore the genetic consequences of verified rescue may help develop a robust genetic signal of the feedback between demography and evolution during rescue, or at least determine if such a feedback can be detected in real genomes.

A strength of the above analysis is that we have explicitly modeled a feedback between demography and evolution, restricting the range of genetic signatures we consequently expect to observe. To take a recent example, Harris *et al.* (2018) have claimed that the lower reductions in genetic diversity within HIV populations adapting to less efficient drugs (as observed by Feder *et al.* 2016) could be due to weaker bottlenecks or slower sweeps rather than sweeps being softer, *i.e.*, arising from multiple mutations. Fortunately, in this case genetic time-series data were available to show that the ability of HIV to reliably adapt on a short time-scale necessitates mutation rates and selection coefficients that imply adaptation by soft sweeps is likely (Feder *et al.* 2018). Formally modeling a feedback between demography and evolution also helps narrow the relevant parameter range. For example, under a haploid version of the model explored here (as is applicable to HIV) the minimum population size during rescue by new mutations is $N(0)s(2N(0)s)^{-d/s}/(s-d)$ [equation 22 in Orr and Unckless (2014)]. Thus, for a given initial population size, $N(0)$, and selection coefficient, s , the smallest minimum population size is $e \ln S / (2s)$, where $S = 2N(0)s$, implying that the smallest minimum population size consistent with the model is roughly proportional to $1/s$. The imposed feedback between demography and evolution therefore precludes simultaneously slow sweeps and large bottlenecks [for example, negating the two smallest bottleneck sizes in Figure 3A of Harris *et al.* (2018) and constraining one to the upper right

portion of Figure 3A,B in Feder *et al.* (2018)]. While it is very likely in this case that soft sweeps are indeed the cause of the pattern (Feder *et al.* 2018), incorporating an explicit model of how demography and evolution interact could help focus future debates.

The model presented here is but one model of evolutionary rescue, which involved a number of assumptions. One of these is that the beneficial allele acts multiplicatively with the ancestral background, so that the absolute fitness its carriers, $(1+sh)(1-d)$ and $(1+s)(1-d)$ —and thus the probability it establishes—are affected by the decline rate of the ancestral genotype, d . Meanwhile, this same assumption caused the relative fitness of the mutants, $1+sh$ and $1+s$ —and thus the allele frequency dynamics during the sweep, once started—to be independent of the initial decline rate (Equation 1). If, instead, we made the absolute fitness of the heterozygote and mutant homozygote independent of the initial decline rate, say $1+sh$ and $1+s$, then the reverse would be true; the relative fitness of the mutant would depend on the initial rate of population decline, d , while its absolute fitness while establishing would not. We expect these effects would, however, largely cancel out (higher establishment probabilities will cause longer sweeps and more coalescence, but slower changes in allele frequency will allow more recombination off the sweep). In any case, because our results depend primarily on the absolute and relative fitness of the heterozygote, the alternative model just described may closely match the model analyzed in detail here when sh is replaced by $(sh+d)/(1-d)$.

We have also assumed that the beneficial allele acts additively with the ancestral allele at that locus ($h = 1/2$). This assumption has been made for convenience; it yields a simple explicit approximation of the allele frequency and population dynamics (Equation 5) that allows closed-form solutions of the structured coalescent (Equation 7), greatly accelerating our computation of the coalescent and pairwise diversity (Equation 9). Alternative forms of dominance are, however, likely and will impact our results. At one extreme, a completely recessive beneficial allele ($h = 0$) is much less likely to establish when compared to additivity, even in a constant population [compare Equation 2 with $v = 1$ and $\epsilon = s/2$ to equation 15 in Kimura (1962)], making rescue nearly impossible in outcrossing populations (Uecker 2017). In fact, any $h < d/s$ will cause the heterozygote to have a negative growth rate (*i.e.*, be subcritical), meaning that establishment will rely on the stochastic persistence of subcritical lineages until they create a supercritical mutant homozygote, analogous to the fixation of underdominant alleles or chromosomal rearrangements. As it is very unlikely that such alleles will establish in time to rescue the population (compared to supercritical heterozygotes), we have neglected this parameter range in order to focus on parameter values that are more likely to be reflected in empirical systems where rescue has occurred. It may, however, be possible to combine our approach with that taken for sweeps with arbitrary dominance in populations of constant size (Ewing *et al.* 2011) and the probability of

establishment of underdominant alleles [e.g., equation 3 in Lande (1979)]. At the other extreme, complete dominance ($h = 1$) will greatly increase the probability of establishment and rescue (Uecker 2017), as well as population mean fitness and, thus, population size. All else equal, we therefore expect rescue to have less effect on the signatures of selective sweeps relative to those in populations of constant size when the rescuing allele is more dominant. Given that the marginal fitness of the beneficial allele will not depend on allele frequency under complete dominance, we expect dynamics much like the haploid model (Orr and Unckless 2014), where simple predictions of allele frequency and population size are more accurate.

Finally, it is of course possible to model rescue under much more complex lifecycles and population structure (e.g., as expected for the evolution of malarial drug resistance; Kim *et al.* 2014), at least using simulations. More complex lifecycles, such as those of parasites like *Plasmodium* and HIV, could cause the bottleneck to have additional impacts on the resulting genetic signature. For example, our populations are obligate sexual outcrossers, meaning that the probability of recombination does not depend on the population size (*cf.* Equation 6) as everyone must mate with someone. However, with selfing and/or facultative sex (genetic exchange), rates of recombination could be lower at lower population densities (as expected for HIV), which would increase the impact of bottlenecks on resulting genetic signatures.

Evolutionary rescue is only one example of myriad processes where demography and evolution feedback on one another. Our approach—combining forward-time eco-evolutionary models with coalescent theory to predict genetic signatures—could be used in many other scenarios. For instance, adaptive colonization of new habitat (a.k.a., adaptive niche expansion) is a closely related process for which a similar approach has already been taken (Kim and Gulisija 2010). As in the case of rescue, explicitly modeling the feedback between demography and evolution in adaptive niche expansion changes the expected signatures left behind by selective sweeps as compared to (bottlenecked) Wright-Fisher populations. Such an approach is interesting from a conceptual point of view, improving our understanding of how eco-evolutionary dynamics affect genetic signatures. But further, given the computational power and simulation platforms available today, it is no longer necessary to restrict oneself to Wright-Fisher populations—researchers may now simulate under much more ecologically realistic models. As our models become more complex, developing simple approximations will become increasingly important for intuition.

Acknowledgments

We thank Pleuni Pennings and Alison Feder for helpful discussions, Ben Haller for comments on our SLiM code, and Joachim Hermisson and the anonymous reviewers for insightful comments. Financial support was provided by the Center for Population Biology at the University of

California - Davis (fellowship to M.M.O.), Banting (Canada; fellowship to M.M.O.), and the National Institute of General Medical Sciences of the National Institutes of Health (NIH R01 GM108779, awarded to G.C.).

Literature Cited

- Abramowitz, M., and I. A. Stegun (Editors), 1972 *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, National Bureau of Standards, United States Department of Commerce, Gaithersburg, MD.
- Alexander, H. K., G. Martin, O. Y. Martin, and S. Bonhoeffer, 2014 Evolutionary rescue: linking theory for conservation and medicine. *Evol. Appl.* 7: 1161–1179. <https://doi.org/10.1111/eva.12221>
- Allen, L. J., 2010 *An Introduction to Stochastic Processes with Applications to Biology*, Chapman and Hall/CRC, Boca Raton, FL. <https://doi.org/10.1201/b12537>
- Anciaux, Y., L.-M. Chevin, O. Ronce, and G. Martin, 2018 Evolutionary rescue over a fitness landscape. *Genetics* 209: 265–279. <https://doi.org/10.1534/genetics.118.300908>
- Anderson, E. C., and M. Slatkin, 2007 Estimation of the number of individuals founding colonized populations. *Evolution* 61: 972–983. <https://doi.org/10.1111/j.1558-5646.2007.00080.x>
- Beichman, A. C., E. Huerta-Sanchez, and K. E. Lohmueller, 2018 Using genomic data to infer historic population dynamics of nonmodel organisms. *Annu. Rev. Ecol. Syst.* 49: 433–456. <https://doi.org/10.1146/annurev-ecolsys-110617-062431>
- Bell, G., 2017 Evolutionary rescue. *Annu. Rev. Ecol. Syst.* 48: 605–627. <https://doi.org/10.1146/annurev-ecolsys-110316-023011>
- Bell, G., and A. Gonzalez, 2009 Evolutionary rescue can prevent extinction following environmental change. *Ecol. Lett.* 12: 942–948. <https://doi.org/10.1111/j.1461-0248.2009.01350.x>
- Berg, J. J., and G. Coop, 2015 A coalescent model for a sweep of a unique standing variant. *Genetics* 201: 707–725. <https://doi.org/10.1534/genetics.115.178962>
- 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.
- Crow, J. F., and M. Kimura, 1970 *An Introduction To Population Genetics Theory*, Harper & Row, New York.
- Ewens, W. J., 1972 The sampling theory of selectively neutral alleles. *Theor. Popul. Biol.* 3: 87–112. [https://doi.org/10.1016/0040-5809\(72\)90035-4](https://doi.org/10.1016/0040-5809(72)90035-4)
- Ewens, W. J., 2004 *Mathematical population genetics*, Springer-Verlag, New York. <https://doi.org/10.1007/978-0-387-21822-9>
- Ewing, G., J. Hermisson, P. Pfaffelhuber, and J. Rudolf, 2011 Selective sweeps for recessive alleles and for other modes of dominance. *J. Math. Biol.* 63: 399–431. <https://doi.org/10.1007/s00285-010-0382-4>
- Feder, A. F., S.-Y. Rhee, S. P. Holmes, R. W. Shafer, D. A. Petrov *et al.*, 2016 More effective drugs lead to harder selective sweeps in the evolution of drug resistance in HIV-1. *eLife* 5: e10670. <https://doi.org/10.7554/eLife.10670>
- Feder, A. F., P. S. Pennings, and D. A. Petrov, 2018 The clarifying role of time series data in the population genetics of HIV. *bioRxiv* 495275 (Preprint posted on December 29, 2018) <https://doi.org/10.1101/495275>
- Feller, W., 1951 Diffusion processes in genetics. In *Proceedings of the second Berkeley symposium on mathematical statistics and probability*. The Regents of the University of California, Oakland, CA.
- Gignoux-Wolfsohn, S. A., M. Pinsky, K. Kerwin, C. Herzog, M. Hall *et al.*, 2018 Genomic signatures of evolutionary rescue in bats

- surviving white-nose syndrome. *bioRxiv*. 470294 (Preprint posted on November 15, 2018). <https://doi.org/10.1101/470294>
- Gomulkiewicz, R., and R. D. Holt, 1995 When does evolution by natural selection prevent extinction? *Evolution* 49: 201–207. <https://doi.org/10.1111/j.1558-5646.1995.tb05971.x>
- Griffiths, R. C., and S. Tavaré, 1994 Sampling theory for neutral alleles in a varying environment. *Philosophical Transactions of the Royal Society Series B: Biological Sciences* 344: 403–410. <https://doi.org/10.1098/rstb.1994.0079>
- Haag-Liautard, C., M. Dorris, X. Maside, S. Macaskill, D. L. Halligan *et al.*, 2007 Direct estimation of per nucleotide and genomic deleterious mutation rates in *Drosophila*. *Nature* 445: 82–85 [corrigenda: *Nature* 453: 128 (2008)]. <https://doi.org/10.1038/nature05388>
- Haller, B. C., and P. W. Messer, 2019 SLiM 3: forward genetic simulations beyond the Wright–Fisher model. *Mol. Biol. Evol.* 36: 632–637. <https://doi.org/10.1093/molbev/msy228>
- Haller, B. C., J. Galloway, J. Kelleher, P. W. Messer, and P. L. Ralph, 2019 Tree-sequence recording in SLiM opens new horizons for forward-time simulation of whole genomes. *Mol. Ecol. Resour.* 19: 552–566. <https://doi.org/10.1111/1755-0998.12968>
- Harris, R. B., A. Sackman, and J. D. Jensen, 2018 On the unfounded enthusiasm for soft selective sweeps II: examining recent evidence from humans, flies, and viruses. *PLoS Genet.* 14: e1007859. <https://doi.org/10.1371/journal.pgen.1007859>
- Hermisson, J., and P. S. Pennings, 2005 Soft sweeps: molecular population genetics of adaptation from standing genetic variation. *Genetics* 169: 2335–2352. <https://doi.org/10.1534/genetics.104.036947>
- Hermisson, J., and P. S. Pennings, 2017 Soft sweeps and beyond: understanding the patterns and probabilities of selection footprints under rapid adaptation. *Methods Ecol. Evol.* 8: 700–716. <https://doi.org/10.1111/2041-210X.12808>
- Hudson, R. R., 1983 Properties of a neutral allele model with intragenic recombination. *Theor. Popul. Biol.* 23: 183–201. [https://doi.org/10.1016/0040-5809\(83\)90013-8](https://doi.org/10.1016/0040-5809(83)90013-8)
- Hudson, R. R., 2002 Generating samples under a wright–Fisher neutral model of genetic variation. *Bioinformatics* 18: 337–338. <https://doi.org/10.1093/bioinformatics/18.2.337>
- Hudson, R. R., and N. L. Kaplan, 1988 The coalescent process in models with selection and recombination. *Genetics* 120: 831–840.
- Innan, H., and Y. Kim, 2004 Pattern of polymorphism after strong artificial selection in a domestication event. *Proc. Natl. Acad. Sci. USA* 101: 10667–10672. <https://doi.org/10.1073/pnas.0401720101>
- Kaplan, N. L., R. Hudson, and C. 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. <https://doi.org/10.1371/journal.pcbi.1004842>
- 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. <https://doi.org/10.1371/journal.pcbi.1006581>
- Kim, Y., A. A. Escalante, and K. A. Schneider, 2014 A population genetic model for the initial spread of partially resistant malaria parasites under anti-malarial combination therapy and weak intrahost competition. *PLoS One* 9: e101601. <https://doi.org/10.1371/journal.pone.0101601>
- Kim, Y., and R. Nielsen, 2004 Linkage disequilibrium as a signature of selective sweeps. *Genetics* 167: 1513–1524. <https://doi.org/10.1534/genetics.103.025387>
- Kim, Y., and D. Gulisija, 2010 Signatures of recent directional selection under different models of population expansion during colonization of new selective environments. *Genetics* 184: 571–585. <https://doi.org/10.1534/genetics.109.109447>
- Kimura, M., 1962 On the probability of fixation of mutant genes in a population. *Genetics* 47: 713.
- Kreiner, J. M., D. A. Giacomini, F. Bemm, B. Waithaka, J. Regalado *et al.*, 2019 Multiple modes of convergent adaptation in the spread of glyphosate-resistant *Amaranthus tuberculatus*. *Proc. Natl. Acad. Sci. USA* 116: 21076–21084 (erratum: *Proc. Natl. Acad. Sci.* 116: 23363). <https://doi.org/10.1073/pnas.1900870116>
- Lande, R., 1979 Effective deme sizes during long-term evolution estimated from rates of chromosomal rearrangement. *Evolution* 33: 234–251. <https://doi.org/10.1111/j.1558-5646.1979.tb04678.x>
- Lindsey, H. A., J. Gallie, S. Taylor, and B. Kerr, 2013 Evolutionary rescue from extinction is contingent on a lower rate of environmental change. *Nature* 494: 463–467. <https://doi.org/10.1038/nature11879>
- Mackay, T. F., S. Richards, E. A. Stone, A. Barbadilla, J. F. Ayroles *et al.*, 2012 The *Drosophila melanogaster* genetic reference panel. *Nature* 482: 173–178. <https://doi.org/10.1038/nature10811>
- Martin, G., and A. Lambert, 2015 A simple, semi-deterministic approximation to the distribution of selective sweeps in large populations. *Theor. Popul. Biol.* 101: 40–46. <https://doi.org/10.1016/j.tpb.2015.01.004>
- McVean, G. A., 2002 A genealogical interpretation of linkage disequilibrium. *Genetics* 162: 987–991.
- McVean, G., 2007 The structure of linkage disequilibrium around a selective sweep. *Genetics* 175: 1395–1406. <https://doi.org/10.1534/genetics.106.062828>
- Nair, S., J. T. Williams, A. Brockman, L. Paiphun, M. Mayxay *et al.*, 2003 A selective sweep driven by pyrimethamine treatment in southeast Asian malaria parasites. *Mol. Biol. Evol.* 20: 1526–1536. <https://doi.org/10.1093/molbev/msg162>
- Ormond, L., M. Foll, G. B. Ewing, S. P. Pfeifer, and J. D. Jensen, 2016 Inferring the age of a fixed beneficial allele. *Mol. Ecol.* 25: 157–169. <https://doi.org/10.1111/mec.13478>
- Orr, H. A., and R. L. Unckless, 2008 Population extinction and the genetics of adaptation. *Am. Nat.* 172: 160–169. <https://doi.org/10.1086/589460>
- Orr, H. A., and R. L. Unckless, 2014 The population genetics of evolutionary rescue. *PLoS Genet.* 10: e1004551. <https://doi.org/10.1371/journal.pgen.1004551>
- Osmond, M. M., S. P. Otto, and G. Martin, 2020 Genetic paths to evolutionary rescue and the distribution of fitness effects along them. *Genetics* 214: 493–510. <https://doi.org/10.1534/genetics.119.302890>
- Otto, S. P., and M. C. Whitlock, 1997 The probability of fixation in populations of changing size. *Genetics* 146: 723–733.
- Oziolor, E. M., N. M. Reid, S. Yair, K. M. Lee, S. G. VerPloeg *et al.*, 2019 Adaptive introgression enables evolutionary rescue from extreme environmental pollution. *Science* 364: 455–457. <https://doi.org/10.1126/science.aav4155>
- Pennings, P. S., and J. Hermisson, 2006a Soft sweeps II: molecular population genetics of adaptation from recurrent mutation or migration. *Mol. Biol. Evol.* 23: 1076–1084. <https://doi.org/10.1093/molbev/msj117>
- Pennings, P. S., and J. Hermisson, 2006b Soft sweeps III: the signature of positive selection from recurrent mutation. *PLoS Genet.* 2: e186. <https://doi.org/10.1371/journal.pgen.0020186>
- 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. <https://doi.org/10.1111/j.0014-3820.2005.tb00941.x>

- Ramsayer, J., O. Kaltz, and M. E. Hochberg, 2013 Evolutionary rescue in populations of *Pseudomonas fluorescens* across an antibiotic gradient. *Evol. Appl.* 6: 608–616. <https://doi.org/10.1111/eva.12046>
- Rêgo, A., F. J. Messina, and Z. Gompert, 2019 Dynamics of genomic change during evolutionary rescue in the seed beetle *Callosobruchus maculatus*. *Mol. Ecol.* 28: 2136–2154. <https://doi.org/10.1111/mec.15085>
- Sedghifar, A., P. Saelao, and D. J. Begun, 2016 Genomic patterns of geographic differentiation in *Drosophila simulans*. *Genetics* 202: 1229–1240. <https://doi.org/10.1534/genetics.115.185496>
- Setter, D., S. Mousset, X. Cheng, R. Nielsen, M. DeGiorgio *et al.*, 2019 VolcanoFinder: genomic scans for adaptive introgression. *bioRxiv* (Preprint posted on July 11, 2019)<https://doi.org/10.1101/697987>
- Slatkin, M., and R. R. Hudson, 1991 Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics* 129: 555–562.
- Smith, J., 1971 What use is sex? *J. Theor. Biol.* 30: 319–335. [https://doi.org/10.1016/0022-5193\(71\)90058-0](https://doi.org/10.1016/0022-5193(71)90058-0)
- Smith, J., and J. Haigh, 1974 The hitch-hiking effect of a favourable gene. *Genet. Res.* 23: 23–35. <https://doi.org/10.1017/S0016672300014634>
- Stephan, W., 2019 Selective sweeps. *Genetics* 211: 5–13. <https://doi.org/10.1534/genetics.118.301319>
- 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.
- Tavaré, S., 1984 Line-of-descent and genealogical processes, and their applications in population genetics models. *Theor. Popul. Biol.* 26: 119–164. [https://doi.org/10.1016/0040-5809\(84\)90027-3](https://doi.org/10.1016/0040-5809(84)90027-3)
- Teshima, K. M., G. Coop, and M. Przeworski, 2006 How reliable are empirical genomic scans for selective sweeps? *Genome Res.* 16: 702–712. <https://doi.org/10.1101/gr.5105206>
- Uecker, H., 2017 Evolutionary rescue in randomly mating, selfing, and clonal populations. *Evolution* 71: 845–858. <https://doi.org/10.1111/evo.13191>
- Uecker, H., and J. Hermisson, 2016 The role of recombination in evolutionary rescue. *Genetics* 202: 721–732. <https://doi.org/10.1534/genetics.115.180299>
- Wakeley, J., 2009 *Coalescent theory*, Roberts & Company, Winterville, NC.
- Watterson, G., 1975 On the number of segregating sites in genetic models without recombination. *Theor. Popul. Biol.* 7: 256–276. [https://doi.org/10.1016/0040-5809\(75\)90020-9](https://doi.org/10.1016/0040-5809(75)90020-9)
- Wei, X., S. K. Ghosh, M. E. Taylor, V. A. Johnson, E. A. Emini *et al.*, 1995 Viral dynamics in human immunodeficiency virus type 1 infection. *Nature* 373: 117–122. <https://doi.org/10.1038/373117a0>
- Wilson, B. A., D. A. Petrov, and P. W. Messer, 2014 Soft selective sweeps in complex demographic scenarios. *Genetics* 198: 669–684. <https://doi.org/10.1534/genetics.114.165571>
- Wilson, B. A., P. S. Pennings, and D. A. Petrov, 2017 Soft selective sweeps in evolutionary rescue. *Genetics* 205: 1573–1586. <https://doi.org/10.1534/genetics.116.191478>

Communicating editor: J. Hermisson