INVITED REVIEW

MOLECULAR ECOLOGY WILEY

Sweepstakes reproduction facilitates rapid adaptation in highly fecund populations

Bjarki Eldon[1] | Wolfgang Stephan[2]

[1]Institute of Evolution and Biodiversity Science, Natural History Museum Berlin, Berlin, Germany

[2]Department of Biology, University of Munich, Munich, Germany

Correspondence
Bjarki Eldon, Institute of Evolution and Biodiversity Science, Natural History Museum Berlin, Berlin, Germany.
Email: bjarki.eldon@mfn.berlin

Abstract

Adaptation enables natural populations to survive in a changing environment. Understanding the mechanics of adaptation is therefore crucial for learning about the evolution and ecology of natural populations. We focus on the impact of random sweepstakes on selection in highly fecund haploid and diploid populations partitioned into two genetic types, with one type conferring selective advantage. For the diploid populations, we incorporate various dominance mechanisms. We assume that the populations may experience recurrent bottlenecks. In random sweepstakes, the distribution of individual recruitment success is highly skewed, resulting in a huge variance in the number of offspring contributed by the individuals present in any given generation. Using computer simulations, we investigate the joint effects of random sweepstakes, recurrent bottlenecks and dominance mechanisms on selection. In our framework, bottlenecks allow random sweepstakes to have an effect on the time to fixation, and in diploid populations, the effect of random sweepstakes depends on the dominance mechanism. We describe selective sweepstakes that are approximated by recurrent sweeps of strongly beneficial allelic types arising by mutation. We demonstrate that both types of sweepstakes reproduction may facilitate rapid adaptation (as defined based on the average time to fixation of a type conferring selective advantage conditioned on fixation of the type). However, whether random sweepstakes cause rapid adaptation depends also on their interactions with bottlenecks and dominance mechanisms. Finally, we review a case study in which a model of recurrent sweeps is shown to essentially explain population genomic data from Atlantic cod.

KEYWORDS
adaptation, high fecundity, natural selection, offspring number distribution, recruitment dynamics, sweepstakes reproduction

1 | INTRODUCTION

Many instances from the natural world show that evolutionary adaptation may occur quite rapidly. Well-known examples of rapid

adaptation in response to environmental changes include colour variation in guppies (Reznick, 2011), field mice (Vignieri et al., 2010) and peppered moth (Cook et al., 2012), insecticide resistance in *Drosophila* (Daborn et al., 2002), beak size changes in Darwin's

finches (Grant & Grant, 2020) and limb development in Anolis lizards (Losos, 2009). The genetic architecture underlying these phenotypic traits ranges from a few genes of major effect as in the peppered moth (van't Hof et al., 2011) to highly polygenic systems of very small effects at individual sites such as human height (Turchin et al., 2012), although the effects of polygenic adaptation on human height appear to be overestimated (Berg et al., 2019; Sohail et al., 2019).

Mirroring this wide range of genetic architectures, the evolutionary genetic models that have been proposed describe adaptation at a single locus (or very few loci) to polygenic adaptation involving numerous sites. Best known are the models for single loci. Clearly, very strong positive directional selection at a single locus may explain fast adaptation, such as in the case of peppered moth (Haldane, 1924). Haldane's deterministic model has been extended in several directions to make it suitable for data analysis. The extension proposed by Maynard Smith and Haigh (1974) for studying genetic hitchhiking (selective sweeps) is most valuable.

On the other hand, polygenic adaptation caused by a large number of weakly selected loci of small effects is not nearly as well studied as the case of strong positive selection leading to selective sweeps (Stephan, 2019). Interest of population geneticists in this type of selection was only very recently evoked by Pritchard et al. (2010) and Pritchard and Di Rienzo (2010). These authors predicted that—in contrast to selective sweeps—allele frequencies may change by small amounts when a large number of genetic loci of minor effect sizes govern a phenotypic trait, but it was unclear whether such polygenic selection can explain rapid adaptation defined as follows.

How do we quantify 'rapid adaptation'? To define rapid adaptation, we assume that a population is at equilibrium when a sudden change of the fitness optimum occurs (e.g. due to a sudden shift in the environment) such that the optimum is placed to another value. If a trait is controlled by a single gene, the rate of response to selection may be quantified by the inverse of the time to fixation of a selected type given that it will go to fixation. Thus, in our simulations, we obtain the rate of adaptation from measuring the sojourn time of a selected allele going to fixation (given that it does so). For a diploid random-mating population of size S (assuming no dominance) evolving according to the Wright-Fisher model, the mean time to fixation (conditional on fixation), taking $s > 0$ as the selective advantage of the genotype homozygous for the beneficial type relative to the homozygous wild type, is approximately $4\ln(2Ns)/s$ if $2Ns$ is sufficiently large (van Herwaarden & van der Wal, 2002). For rapid adaptation to occur $2Ns$ must be at least 100 (i.e. of about the same order of magnitude as necessary for the detection of hard selective sweeps; Stephan, 2019).

In the case of polygenic evolution, rapid adaptation may be defined in a similar way. In this case, the response of a phenotypic trait to selection may be quantified by the rate at which the population mean of the trait reaches the new fitness optimum after a sudden environmental shift. Using deterministic population genetic models of polygenic adaptation that have been analysed in the past 10 years (in particular, the deterministic model of de Vladar & Barton, 2014),

this rate can be calculated (assuming linkage equilibrium). Jain and Stephan (2015, 2017a, 2017b) and Stephan (2016) developed formulas for the case when the effect sizes of the alleles at most loci are large relative to a scaled mutation rate such that directional selection at each of those loci is very strong, and for the opposite parameter range in which most loci involved have small effects and hence exert weak selection. While the first case is closely related to monogenic adaptation discussed above (Jain & Stephan, 2017b), in the second case the new optimum is approached exponentially (already derived by Lande, 1976), if the shift of the optimum is not too large relative to the equilibrium genetic variance. Importantly, if the effect sizes of the loci are exponentially distributed, the equilibrium variance is proportional to the number of loci (with small effects) governing the trait. Thus, if the number of loci is sufficiently large, rapid adaptation through polygenic selection may occur. Yet, populations of finite size simulations based on the classical Wright-Fisher model suggest that genetic drift slows down the speed of polygenic adaptation to some extent (John & Stephan, 2020).

Here, we extend a population genetic model of adaptation with Wright-Fisher drift to models of genetic reproduction that strongly deviate from the standard model of genetic drift. In particular, we consider the impact of random sweepstakes on adaptation in highly fecund populations. In random sweepstakes, the distribution of individual recruitment success is assumed to be highly skewed. In contrast to the Wright-Fisher model, this results in a huge variance in individual recruitment success (the number of offspring). We investigate here in which way and to what extent this type of genetic drift affects adaptation when positive selection acts at a single site.

Natural highly fecund populations are diverse and widely found (Eldon, 2020). By 'high fecundity', we refer to the ability of organisms to produce numbers of juveniles (potential offspring) at least on the order of the census population size. The evolution of a population over one generation is seen as each individual (or pair of individuals in diploid populations) independently producing a random number of juveniles according to a given model (i.e. probability distribution, see e.g. Equations 4 or 1); from the total pool of juveniles, a given number of them is sampled (uniformly at random and without replacement) to form a new set of reproducing individuals. These include broadcast spawners such as the Antarctic limpet (*N. concinna*), Atlantic cod, Japanese sardines, crop-infesting fungi and corals (Agrios, 2010; Árnason, 2004; Árnason & Halldórsdóttir, 2015; Barfield et al., 2022; Niwa et al., 2016; Vendrami et al., 2021). Even viruses may be classified as being highly fecund (Irwin et al., 2016; Timm & Yin, 2012). Broadcast spawners produce huge numbers of juveniles to counter the high mortality among them (i.e. Type III survivorship). A central question regarding highly fecund populations is if the recruitment dynamics, or the distribution of individual recruitment success (the offspring number distribution) may be characterized by 'sweepstakes reproduction', or a highly skewed offspring number distribution. In this context, it is important to understand how sweepstakes reproduction comes about. The Wright-Fisher model may be viewed as a model of high fecundity, where *every* individual

produces a huge number of potential offspring (e.g. gametes), or at least an order of magnitude larger than the population size (see e.g. Der et al., 2011). Then, assuming that the population size is much smaller than the number of juveniles produced by each individual, sampling juveniles uniformly and without replacement to form a new generation of individuals is well approximated by the surviving offspring sampling a parent with replacement, the sampling mechanism of the Wright-Fisher model. This mechanism does not lead to sweepstakes reproduction, since the number of surviving offspring from any given individual will be negligible relative to a large total population size. High fecundity cannot on its own produce sweepstakes reproduction. The key ingredient for sweepstakes reproduction is a mechanism that turns high fecundity into a skewed offspring number distribution, or a skewed individual recruitment success. An important open question in evolution and ecology is to infer sweepstakes reproduction in highly fecund populations, and to identify the actual mechanism of sweepstakes reproduction given the evidence of it.

Since broadcast spawners characterized by Type III survivorship are highly fecund, it is plausible that the reproductive output among broadcast spawning individuals may be skewed, in the sense that occasionally, at any given time, a significant proportion of surviving offspring come from a few parents. The skew in reproductive output generated in this way without involving selection has been named 'sweepstakes reproduction' and has been claimed to 'play a major role in shaping marine biodiversity' (Hedgecock, 1994; Hedgecock & Pudovkin, 2011). We will refer to this type of sweepstakes reproduction in a highly fecund population as 'random sweepstakes'(Árnason et al., 2023). Broadcast spawning and Type III survivorship combine to generate the possibility for a few lucky individuals to produce a significant number of surviving juveniles through chance matching of reproduction with favourable environmental conditions, thus being a mechanism turning high fecundity into sweepstakes reproduction.

Sweepstakes reproduction may also be generated through natural selection. In this mechanism, juveniles produced at any given time are seen as having to pass through independent selective filters during development from earliest juvenile stage to reproductive age, with the result that the genetic constitution of the surviving juveniles is, on average, different from that of non-surviving juveniles (Williams, 1975). New recombined genotypes are continuously generated and carried to high frequency in a population chasing an ever-changing optimum. We will refer to this type of sweepstakes reproduction as 'selective sweepstakes', in which natural selection acts as the mechanism turning high fecundity into sweepstakes reproduction (Árnason et al., 2023).

Recruitment dynamics of natural populations, in particular the distribution of individual recruitment success, are central to the mechanisms shaping genetic diversity. Improved understanding of recruitment dynamics is therefore required for illuminating the ecology, population connectivity, local adaptation and resilience of natural populations, for better conservation and management of fish stocks, and for much needed further development of population and evolutionary genetic theory (Botsford et al., 2001;

Cowen & Sponaugle, 2009; Eldon, 2020; Fu & Li, 1999; Gagnaire et al., 2015; Grant et al., 2016; Selkoe et al., 2016; Wakeley, 2004).

Here, we use simulations to investigate how random sweepstakes and recurrent bottlenecks (in a highly fecund population) affect positive selection when acting at a single site. In Section 2, population models of random sweepstakes are introduced. In Section 3, the simulation results are presented and discussed. A case study of population genomic data of Atlantic cod (Árnason et al., 2023), which found a good fit of the data to a population model of recurrent sweeps (Durrett & Schweinsberg, 2005), seen as approximating selective sweepstakes, is reviewed in Section 4. A brief conclusion listing the main results from the simulations and important remaining follow-up projects is presented in Section 5. In the Supporting Information, we give a brief summary of the mathematical formulation of random sweepstakes, and some further examples of the effect of random sweepstakes, recurrent bottlenecks and dominance mechanisms on selection.

2 | MODELLING RANDOM SWEEPSTAKES

A natural model of random sweepstakes describes the probability distribution of the random number of juveniles, or potential offspring, contributed by each individual (see Equation 1) in any given generation. From the pool of juveniles, a given number is sampled without replacement to form a new set of reproducing individuals (Schweinsberg, 2003). The probability distribution for the random number of juveniles produced by a given individual in the Schweinsberg (2003) model is given by

$$\lim_{x \to \infty} Cx^{\alpha} \mathbb{P}(X_1 \geq x) = 1, \tag{1}$$

where $C > 0$ is a constant of proportionality (i.e. ensuring that the limit in Equation (1) is one) and $\alpha > 0$ is a constant determining the skewness of the distribution (Schweinsberg, 2003). The formulation in Equation (1) should be understood as specifying how the probability of producing at least x juveniles behaves for very large x (at least on the order of the population size). In a population evolving according to Equation (1), most individuals will produce a small (relative to the population size) number of juveniles and so a few surviving offspring; however, occasionally a single individual (the probability of two or more individuals each producing a large number of juveniles will be negligible in a large population) will produce a large (relative to the population size) number of juveniles and so a significant number of surviving offspring (illustrated in Figure 1b; in Figure 1a, we illustrate for comparison the evolution of a haploid population over one generation according to the Wright-Fisher model). The model described in Equation (1) therefore corresponds well to random sweepstakes occurring in a broadcast spawner evolving according to Type III survivorship; individuals must have the capacity to produce huge numbers of juveniles to counter the high mortality among the juveniles, and once in a while a lucky individual matches reproduction with favourable environmental conditions so a significant fraction of the lucky
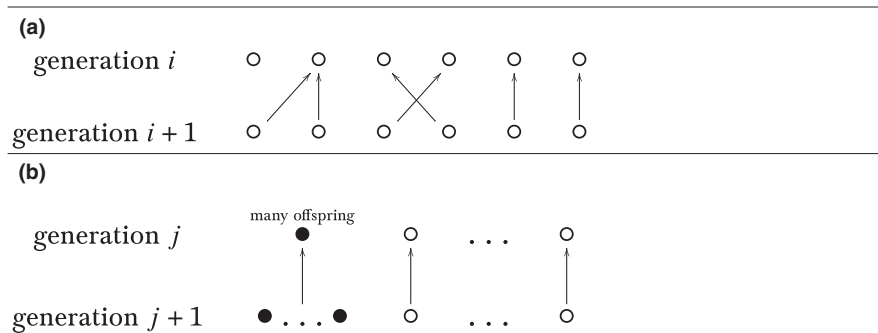
FIGURE 1 A simple illustration of the difference between the Wright-Fisher model (a) and a model of random sweepstakes (b), with the arrows indicating the parent of the surviving offspring. In (a), a haploid population evolves according to the Wright-Fisher model, where each individual (gene copy) produces at most a small (relative to the population size) number of surviving offspring; in (b), a single randomly picked individual contributes a large (relative to the population size) number of surviving offspring (indicated with filled circles), while all other individuals contribute the remaining number of offspring according to the Wright-Fisher model.
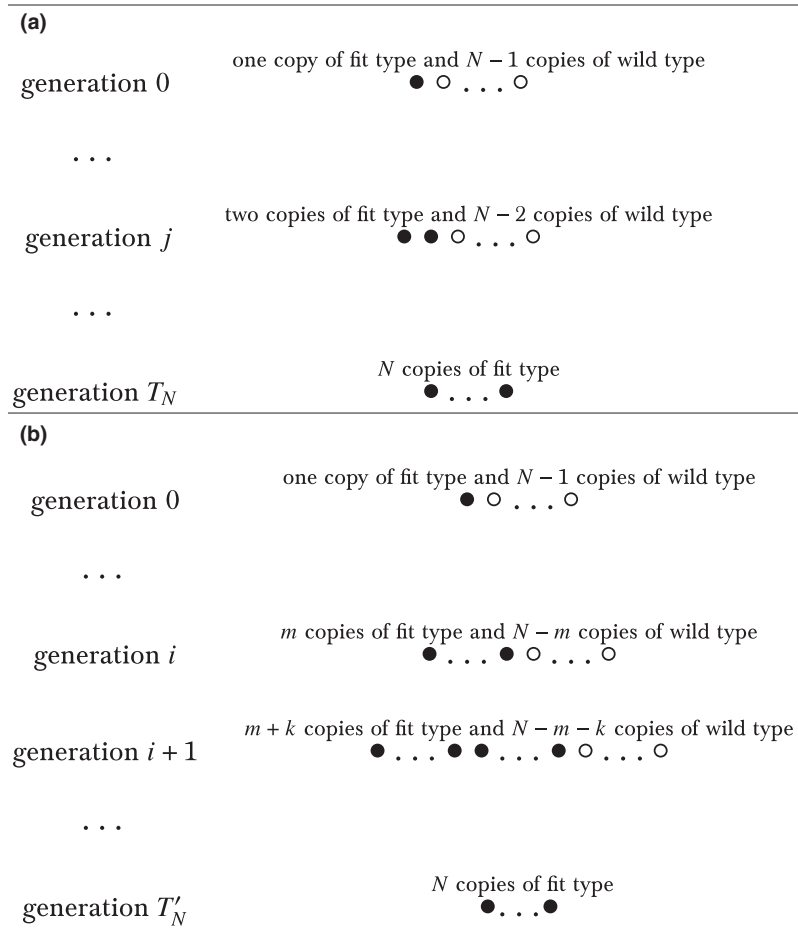
individual's juveniles survive. How often such large families occur, and how large they will be, depends on the value of α in Equation (1) or (3) determining the skewness of the probability distribution for the number of juveniles produced by any given individual; the smaller α is the higher the chance of producing many juveniles. From the total pool of juveniles, assuming there are enough of them (at least N of them in a haploid population of size N), which is almost guaranteed in a large population, provided each individual produces more than one juvenile on average (Schweinsberg, 2003), we then sample N juveniles uniformly at random and without replacement to form a new set of reproducing individuals. The parameter α in Equation (1) is the quantity determining how quickly the probability of producing at least x juveniles decays as x increases. In the case $0 < \alpha < 2$, large families occur often enough and are large enough to affect the evolution of the population. That means that the ancestral process (the process tracking the random ancestral relations of sampled gene copies) is in the domain of attraction of a particular example of a multiple-merger coalescent (which, in contrast to the Kingman-coalescent, admits mergers of at least three ancestral lineages) generally referred to as the Beta$(2 - \alpha, \alpha)$-coalescent (Schweinsberg, 2003).

Population genetic models incorporating random sweepstakes are in fundamental ways different from the classical Wright-Fisher model. For example, the process tracking the frequency of a genetic type segregating in a population evolving according to random sweepstakes is in the domain of attraction of a jump diffusion where the process, in addition to evolving according to the well-known Wright-Fisher diffusion, admits discontinuous jumps (Birkner & Blath, 2009). The jumps correspond to the occurrence of large families involving a number of copies of the type being tracked (a brief overview of the mathematical formulation is given in Section S1). Correspondingly, when looking at genealogies of samples, models of random sweepstakes are in the domain of attraction of multiple-merger coalescent models (a coalescent is a probabilistic description of the random ancestral relations of a sample of gene copies from an arbitrarily large hypothetical population evolving according to a given model; Berestycki, 2009; Eldon & Wakeley, 2006; Huillet & Möhle, 2011, 2013; Möhle, 2011; Pitman, 1999; Sagitov, 1999;

Sargsyan & Wakeley, 2008; Schweinsberg, 2003), where a random number of ancestral lineages of a sample of gene copies merges at any given time (Donnelly & Kurtz, 1999; Pitman, 1999; Sagitov, 1999; Schweinsberg, 2000). Coalescent theory provides a framework for the development of powerful inference methods (Wakeley, 2007). Multiple-merger coalescent processes predict patterns of population genetic data that are different from predictions of the classical Wright-Fisher model and similar models (Birkner, Blath, & Eldon, 2013b; Blath et al., 2016). To identify random sweepstakes in natural populations, one can therefore apply coalescent-based inference to population genetic data (Birkner et al., 2011; Birkner & Blath, 2008; Birkner, Blath, & Steinrücken, 2013; Eldon, 2011, 2016; Eldon et al., 2015; Freund & Siri-Jégousse, 2021; Koskela, 2018; Koskela & Berenguer, 2019). In contrast to multiple merger coalescents, the gene genealogy of a sample from a population evolving according to the Wright-Fisher model, or similar model (where 'similar' refers to certain conditions on the offspring number distribution; Möhle & Sagitov, 2001; Sagitov, 1999) is described by the Kingman-coalescent, in which no more than two ancestral lineages merge each time (Berestycki, 2009; Kingman, 1982a, 1982b, 1982c; Tajima, 1983). Intuitively one would expect a multiple-merger coalescent to describe gene genealogies under random sweepstakes, since whenever a large family occurs, it will involve a number of the ancestral lineages with non-negligible probability in a large population. Correspondingly, looking forward in time, a number of copies of the genetic type being tracked will be involved in the large family event, leading to a jump in the type frequency process. Figure 1 records a simple illustration of the difference between the Wright-Fisher model and random sweepstakes for haploid populations. The occasional occurrence of a large family as shown in Figure 1b can induce jumps in the frequency of the type being tracked. Our aim is to investigate, given that a population evolves according to random sweepstakes, how does that affect the fate of advantageous mutations, and by implication adaptation (Figure 2).

Even though the model in Equation (1) seems natural, it has two main drawbacks. One is that individuals are assumed to be able to produce arbitrarily many juveniles. Although some organisms are extremely fecund, the assumption of unbounded fecundity is unrealistic.

FIGURE 2 Random sweepstakes and fixation of an advantageous type starting in one copy. In (a) the population is assumed to evolve according to the Wright-Fisher (or similar) model, so that the number of copies of the fit type increases in small (relative to the population size) amount between generations. In (b), the population evolves according to random sweepstakes, and the number of copies of the fit type can increase by a significant amount between generations when an individual carrying the fit type produces many juveniles. On average, one would then expect, with T'_N denoting the time to fixation (conditional on fixation) under random sweepstakes and T_N under the Wright-Fisher model, $T'_N < T_N$.

(a)

generation 0 — one copy of fit type and $N-1$ copies of wild type

. . .

generation j — two copies of fit type and $N-2$ copies of wild type

. . .

generation T_N — N copies of fit type

(b)

generation 0 — one copy of fit type and $N-1$ copies of wild type

. . .

generation i — m copies of fit type and $N-m$ copies of wild type

generation $i+1$ — $m+k$ copies of fit type and $N-m-k$ copies of wild type

. . .

generation T'_N — N copies of fit type

The second assumption involves the scaling of time used to pass to a continuous-time limit process as population size tends to infinity. In passing to a coalescent limit (in probability theory this means convergence in a certain sense, in our case in terms of finite-dimensional distributions), one scales time with a quantity usually denoted c_N, which is the probability that two distinct gene copies sampled at the same time derive from the same parental copy in the previous generation. For a haploid population of size N evolving according to the Wright-Fisher model, $c_N = 1/N$ (so that N generations correspond to one coalescent time unit). In the random sweepstakes model given in Equation (1) with $1 < \alpha < 2$, we have $c_N \propto N^{1-\alpha}$. This means that, for estimates of α close to one (e.g. Árnason & Halldórsdóttir, 2015), an unrealistically high population size is required to recover the observed genetic variation (Eldon, 2020). To see this, we consider the expected number of segregating sites (see Equation S13 in Section S5) of a sample of size n of a non-recombining contiguous chromosome segment of length L: $\mathbb{E}[B(n)]$ denotes the expected tree size (i.e. the expected size of the random gene genealogy connecting the n sampled gene copies under a given population model) with time measured in coalescent time units, and μ the per site per generation mutation rate. We can safely take $\mathbb{E}[B(n)] = \mathcal{O}(1)$, and then we obtain (see Section S5), assuming we observe m segregating sites:

$$N \approx \left(\frac{1}{\mu} \frac{m}{L} \right)^{\frac{1}{\alpha-1}}. \qquad (2)$$

Assuming around one percent of sites are segregating, that is, $m/L \approx 0.01$, and μ of order $\mathcal{O}\left(10^{-8}\right)$ as for Pacific cod (Canino et al., 2010), we see that for estimates of α close to one, it would be difficult to recover the observed amount of genetic variation without requiring an unrealistically large population size.

To address the assumption of the Schweinsberg (2003) model (see Equation 1) regarding unbounded fecundity, one can apply an upper bound on the number of juveniles produced by an individual (Eldon & Stephan, 2018). Suppose X_1 denotes the random number of juveniles produced by a given individual, consider the mass function (i.e. the probability of producing x juveniles)

$$\mathbb{P}(X_1 = x) = \mathbb{1}_{\{1 \le x \le u(N)\}} \left(\frac{1}{x^\alpha} - \frac{1}{(1+x)^\alpha} \right) \frac{(1+u(N))^\alpha}{(1+u(N))^\alpha - 1}, \qquad (3)$$

where $u(N)$ is an increasing positive function of N and representing an upper bound on the number of juveniles a given individual can produce. The parameter $\alpha > 0$ determines how the probability of producing x juveniles decreases with increasing x. The smaller α is the higher the probability of producing many juveniles, and $\mathbb{1}_{\{A\}} := 1$ if A holds, and zero otherwise. The model in Equation (3) is a variant of the model in Equation (1). The behaviour of the model with respect to varying α and $u(N)$ is being rigorously investigated for a planned future publication. The model in Equation (3), or some variant of it, in addition to modelling random sweepstakes (as α decreases the probability of producing

many juveniles increases) can also serve as a natural, realistic and mathematically tractable alternative to the Wright-Fisher model (as α increases the probability of producing many juveniles decreases).

In order to address the second drawback of the Schweinsberg (2003) model regarding the scaling of time, we consider a simple variant of the model in Equation (3). A similar approach is also adopted in the works of Eldon & Wakeley, 2006, and Huillet & Möhle, 2013, where convergence to coalescents are obtained with less problematic scalings of time. Suppose with probability ε_N all current N individuals produce juveniles according to Equation (3) with with α 'small', or an increased probability of producing many juveniles, with probability $1 - \varepsilon_N$ we take α 'large' (representing a decreased probability of producing many juveniles); in both cases, we suppose that the cutoff $u(N)$ is proportional to the population size as population size N increases, for example, taking $u(N) = KN$ for some finite constant $K > 0$. The advantages of this approach are twofold. One is that in this model ordinary reproduction (in which each individual produces a small number, relative to the population size, of juveniles with high probability) occurs most of the time. Occasionally (i.e. with probability ε_N), however, reproduction matches favourable environmental conditions, and each individual produces juveniles according to Equation (3) with smaller α or with an increased probability of producing many juveniles. In this way, random sweepstakes can be modelled to be strong enough to impact the evolution of the population without being the overwhelming force. We also claim that the second advantage of the mixture model regards the scaling of time (recall Equation 2) required to pass to a coalescent limit (this is also being investigated in a planned future work).

Now, we describe the mixture model we use for the simulations. Let $L(\alpha, u(N))$ denote the law (probability distribution) of the number of juveniles with mass function as in Equation (3), and take $\alpha_1 \in (0, 2)$ and $\alpha_2 \geq 2$ as fixed. The quantity α_1 represents an increased probability of producing many juveniles, and α_2 a decreased probability of doing so. Assuming the cutoff $u(N)$ is at most of order N, that is, we assume the fecundity is high enough to impact the evolution of the population, but not necessarily an unbounded fecundity, we will write the mixture-model as, where $X_1, \ldots, X_N$ denote the random number of juveniles produced by the current N individuals,

$$X_1, \ldots, X_N \sim \begin{cases} L(\alpha_2, u(N)) & \text{with probability } 1 - \varepsilon_N \\ L(\alpha_1, u(N)) & \text{with probability } \varepsilon_N \end{cases}, \quad (4)$$

that is, the $X_1, \ldots, X_N$ are independent and identically distributed as specified in Equation (4). The meaning of Equation (4) is as just described, that with probability $1 - \varepsilon_N$, the number of juveniles is distributed according to Equation (3) with $\alpha = \alpha_2$ (a decreased probability of producing many juveniles), and with probability ε_N, we take $\alpha = \alpha_1$ (i.e. reproduction matches favourable environmental conditions and every individual produces juveniles with an increased probability of producing many juveniles). Similarly one can keep α fixed between one and two and randomize the cutoff $u(N)$; however, we will restrict ourselves to the model in Equation (4). The model in Equation (4) is a natural way

of formulating random sweepstakes in a broadcast spawner; most of the time individuals produce a small (relative to the population size) number of juveniles $(\alpha = \alpha_2 \geq 2)$; occasionally (with probability ε_N) favourable environmental conditions match reproduction so that there is a higher chance of producing a larger number of juveniles $(\alpha = \alpha_1 \in (0, 2))$.

Models corresponding to Equation (4) for diploid (or polyploid) populations would necessarily involve simultaneous multiple-merger coalescent processes where at least two distinct groups of ancestral lineages could merge at a time (Birkner et al., 2018; Birkner, Blath, & Eldon, 2013a; Blath et al., 2016; Koskela & Berenguer, 2019; Möhle & Sagitov, 2003; Sagitov, 2003; Schweinsberg, 2000). The Beta$(2 - \alpha, \alpha)$-coalescent based on the original population model of random sweepstakes (see Equation 1; Schweinsberg, 2003) for both haploid and diploid (Birkner et al., 2018) populations has been implemented in the state-of-the-art simulation package MSPRIME (Baumdicker et al., 2021).

3 | THE IMPACT OF RANDOM SWEEPSTAKES ON SELECTION

We measure the impact of selection on the evolution of a population through the time to fixation of a selectively advantageous type conditional on fixation of the type. We denote this quantity by $\tau_N(y)$ (see Section S2), where we start with y copies of the type. We see the probability (denoted $p_N(y)$, see Equation (S10) in Section S2 in Supporting Information) of fixation of a selectively advantageous type when starting with y copies of the type more as an auxiliary quantity. The quantity $p_N(1)$ is the probability of fixation of the beneficial type when starting from one copy of the type, and $\tau_N(1)$ is the expected time to fixation of the beneficial type conditional on fixation of the type and starting with one copy of the beneficial type. We are interested in investigating the effect of random sweepstakes on $\tau_N(1)$ under a simple model of viability selection. The fixation probability can inform about adaptation, including the occurrence of resistance to antibiotics, and about loss of genetic variation (Patwa & Wahl, 2008). The expected time $\tau_N(1)$ is well defined since we consider a finite population, so that the boundaries (0, N) will be reached in finite time almost surely. The expected fixation time can inform about the rate of adaptation given that it will occur. In our simulations, we focus mainly on the time to fixation to demonstrate that sweepstakes reproduction may facilitate rapid adaptation.

As we have discussed (Section 2), models of random sweepstakes admit jumps in the genetic type frequency. This means that classical diffusion techniques (Feller, 1951; Kimura, 1957) are not applicable. Nevertheless, some mathematical results have been obtained on the impact of random sweepstakes on selection (Der et al., 2011, 2012; Foucart, 2013). One way to approach this problem might be to identify the limiting generators (as population size tends to infinity) of the forward-in-time process and then work with the generators as done by Der et al. (2011) who introduce generalized Wright-Fisher models (see Section S1). Consider the following simple model of

random sweepstakes. Suppose a population evolves according to a discrete-time Moran model (i.e. one randomly picked individual produces offspring, and an equivalent amount of individuals perish to keep the population size constant). The distribution of the random number of offspring (V) produced by the parent at any given time is

$$\mathbb{P}(V = v) = \mathbb{1}_{\{v=1\}}(1 - \varepsilon_N) + \mathbb{1}_{\{v=\lfloor \psi N \rfloor\}}\varepsilon_N. \qquad (5)$$

In Equation (5) (as in Equation 4), ε_N can be understood as the probability of matching reproduction with favourable environmental conditions, and when that happens the parent produces $\lfloor \psi N \rfloor$ surviving offspring. Each time as many individuals as produced perish to keep the population size N constant, and $0 < \psi, \varepsilon_N < 1$ (Eldon & Wakeley, 2006). Viewing time in units of $\mathcal{O}(N^\gamma)$ generations for some $\gamma \in (0, 2]$, it can be shown that the model is in the domain of attraction of a Λ-coalescent, indeed for $\gamma = 2$ the limiting coalescent is a mixture of the Kingman-coalescent and a multiple-merger coalescent (Eldon & Wakeley, 2006). The model in Equation (5) is unrealistic in assuming that exactly the same fraction (ψ) of the population is replaced in each sweepstakes event, but it is among the simplest models of random sweepstakes, and its simplicity does facilitate some mathematical results to be obtained (Der et al., 2012; Eldon & Freund, 2018; Matuszewski et al., 2018). In particular, fixation probabilities under selection are studied for generalized Wright-Fisher processes (Der et al., 2011). Assuming random sweepstakes according to Equation (5) in a framework involving selection, conditions on the strength of selection can be identified under which fixation of the fitter of two genetic types is assured (Der et al., 2012). The results on assured fixation are limit results as population size $N \to \infty$ (Der et al., 2012). The results of a simulation study of the effect of random sweepstakes on selection in a finite haploid population, based on the model in Equation (3) and not on the mixture model in Equation (4), indicate that fixation is anything but given in a finite population evolving according to random sweepstakes, that is, the fixation probability is clearly decreased as α tends to one, the distribution of the number of juveniles is highly skewed, and individuals are allowed to produce numbers of juveniles an order of magnitude larger than the population size (Eldon & Stephan, 2018).

3.1 | Haploid populations

Assuming a haploid population of constant size N evolving by random sweepstakes according to Equation (3) under a simple model of viability selection, simulation results indicate that both $p_N(1)$ and $\tau_N(1)$ are much smaller under random sweepstakes than under ordinary (i.e. individuals produce small number, relative to the population size, of juveniles with high probability) reproduction (Eldon & Stephan, 2018). That is, the chance of fixation is significantly smaller under random sweepstakes, and the expected time to fixation when fixation happens is shorter under random sweepstakes. One must view these results in the context of the expected number of beneficial mutations in the whole population over a period of time.
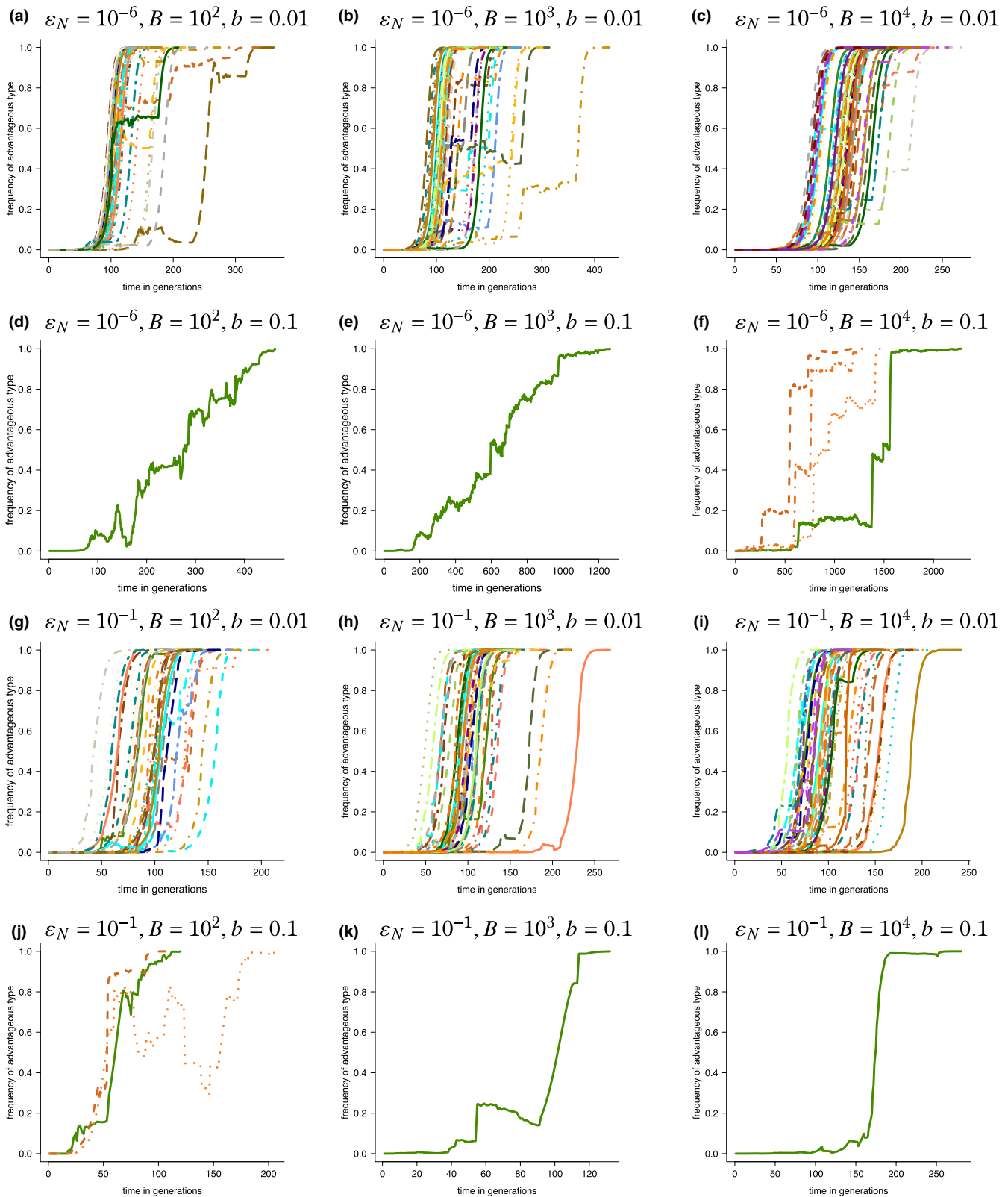
The mixture model in Equation (4) may moderate the reduction in the probability of fixation due to random sweepstakes relative to Equation (3), the model in the study of Eldon & Stephan, 2018. Thus, if the number of beneficial variants that occur in the population before the most recent common ancestor of the whole population is reached is sufficiently large that a somewhat smaller chance of fixation of any one of them due to random sweepstakes will not significantly alter their overall effect, then one would expect that random sweepstakes would affect the evolution of the population when new beneficial mutations arise. Figure 2b records a simple illustration of the effect of random sweepstakes on selection through the fixation of an advantageous type (for comparison Figure 2a illustrates fixation in a haploid population evolving according to the Wright-Fisher model).

To investigate the impact of random sweepstakes acting through Equation (4) on selection, we model selection as follows. Suppose a haploid population of constant size N evolves in discrete generations. In every generation, all N current individuals independently contribute juveniles according to Equation (4). The population is partitioned into two genetic types, one conferring viability weight one, and the other viability weight e^{-s}. Throughout, we start with the fitter type in one copy. Each juvenile inherits the genetic type of its parent; we exclude mutation. Given a pool of S_N juveniles, we sample independent exponentials each with the rate of the viability weight of a given juvenile. That is, if there are S_y juveniles with viability weight one, and S_{N-y} juveniles ($S_y + S_{N-y} = S_N$) with viability weight e^{-s}, we sample S_y independent exponentials with rate one each, and S_{N-y} independent exponentials with rate e^{-s}. The N juveniles with the smallest exponentials survive to form the next generation of reproducing individuals. This is a way to let selection influence the viability of each juvenile, and has been applied in a previous investigation on the effects of random sweepstakes on selection (Eldon & Stephan, 2018). In the present work, we model random sweepstakes based on Equation (4), whereas earlier (Eldon & Stephan, 2018), we kept both α and the cutoff $u(N)$ fixed. In Figures 3 and 4 (see also Figures S1–S5), we show examples of excursions to fixation for several scenarios. We do not aim for precise estimates of $\tau_N(1)$ or $p_N(1)$ (Equation S10), but rather to see main trends in how random sweepstakes modelled through Equation (4) affect $\tau_N(1)$ and $p_N(1)$. In the Figures the scale of the abscissa (horizontal axis) may vary between subplots. In each panel, the colours and line types are only meant to distinguish between the trajectories; for any given panel, the trajectories shown were all obtained under identical conditions. The C++ code written for the simulations is freely available on github (see Section 6.1).

The size of natural populations changes in time; in particular, there may be randomly occurring bottlenecks (sharp reduction in population size). For example, bottlenecks have been suggested to be an important factor in the evolution of resistance of pathogenic bacteria to antibiotics (Mahrt et al., 2021). We model bottlenecks as follows. Suppose there is a fixed upper bound on the population size, that is, the total number of individuals in the population cannot be more than some fixed number N. We can think of this number as the

carrying capacity of the environment. In any given generation, a bottleneck occurs with a fixed probability. Should a bottleneck occur, we sample a fixed number (denoted N_b) of individuals uniformly at random that will survive the bottleneck (the remaining individuals, the ones not surviving the bottleneck, are discarded). We then check

if any of the surviving individuals contain the beneficial type; if not, we stop since the beneficial type is then lost from the population. On the other hand, if the beneficial type is fixed among the surviving individuals, we stop and record a fixation of the type. If the beneficial type is present among the surviving individuals but has not fixed, the

(a) $\varepsilon_N = 10^{-6}, B = 10^2, b = 0.01$

(b) $\varepsilon_N = 10^{-6}, B = 10^3, b = 0.01$

(c) $\varepsilon_N = 10^{-6}, B = 10^4, b = 0.01$

(d) $\varepsilon_N = 10^{-6}, B = 10^2, b = 0.1$

(e) $\varepsilon_N = 10^{-6}, B = 10^3, b = 0.1$

(f) $\varepsilon_N = 10^{-6}, B = 10^4, b = 0.1$

(g) $\varepsilon_N = 10^{-1}, B = 10^2, b = 0.01$

(h) $\varepsilon_N = 10^{-1}, B = 10^3, b = 0.01$

(i) $\varepsilon_N = 10^{-1}, B = 10^4, b = 0.01$

(j) $\varepsilon_N = 10^{-1}, B = 10^2, b = 0.1$

(k) $\varepsilon_N = 10^{-1}, B = 10^3, b = 0.1$

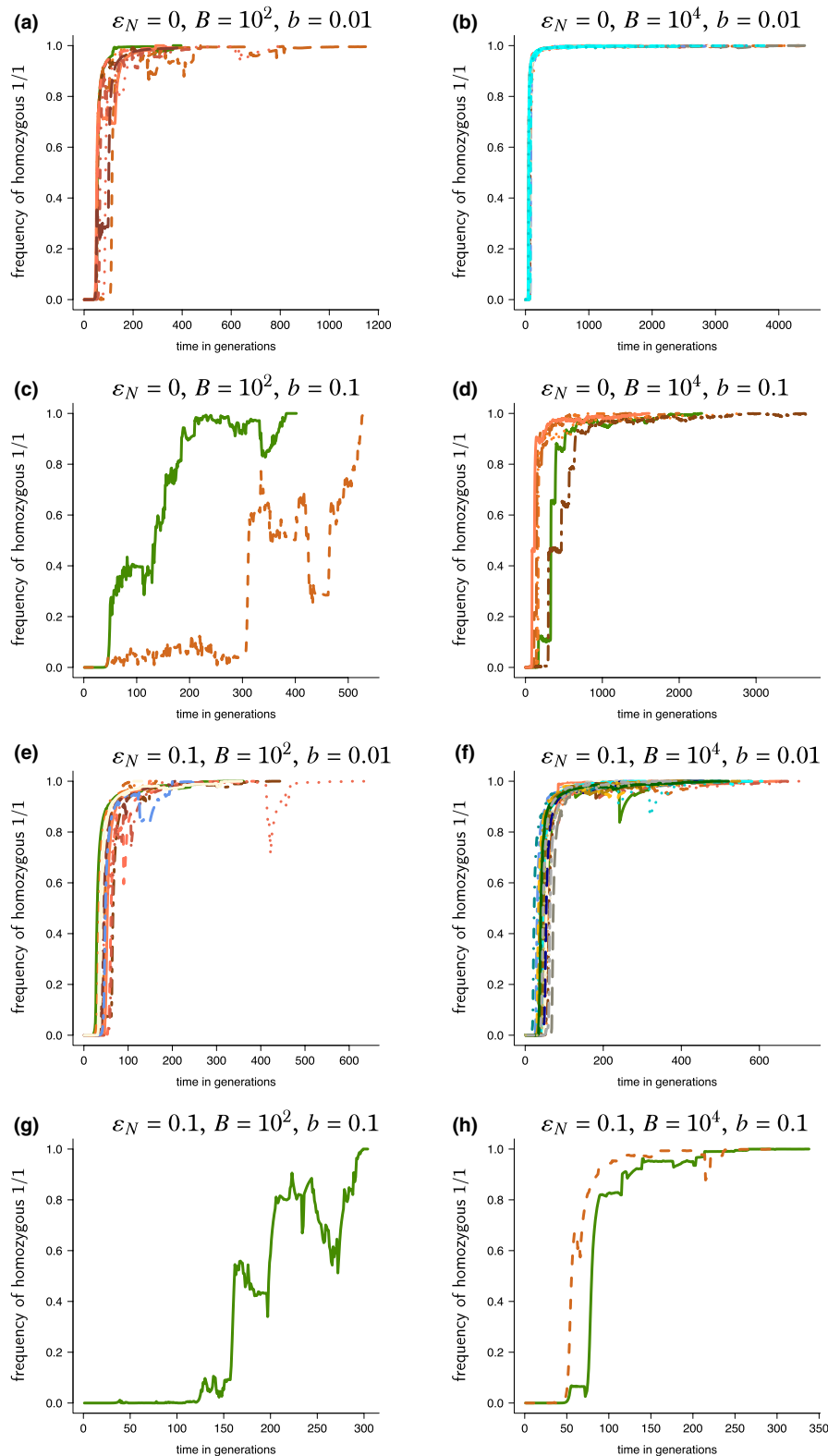(l) $\varepsilon_N = 10^{-1}, B = 10^4, b = 0.1$

FIGURE 3   Random sweepstakes and recurrent bottlenecks in a haploid population. Examples of excursions to fixation of a type conferring a selective advantage in a haploid population of maximum size $N = 10^6$ evolving according to random sweepstakes as in Equation (4). The mechanism described in Equation (4) is that occasionally, or with probability ε_N, reproduction matches favourable conditions so that individuals have a higher chance of producing many juveniles. Most of the time however, or with probability $1 - \varepsilon_N$, individuals have a lower chance of producing many juveniles. Here, we take $\alpha_1 = 0.75$ representing increased probability of producing many juveniles, $\alpha_2 = 3$ representing decreased probability of producing many juveniles. Furthermore, we take the cutoff $u(N) = N$ (each individual can produce at most $u(N)$ juveniles), strength of selection $s = 0.5$ throughout, and with ε_N the probability of having a higher chance of producing many juveniles (i.e. the probability of $\alpha = \alpha_1$), the number B of individuals surviving a bottleneck, and the probability b a bottleneck occurs in any given generation as shown. In any given generation the current individuals produce juveniles according to Equation (3) with $\alpha = \alpha_1$ with probability ε_N, and with $\alpha = \alpha_2$ with probability $1 - \varepsilon_N$, see Equation (4). Results from 10^2 (a, b, c), 16394 (j), otherwise from 10^3 experiments. The scale of the time (horizontal) axis may differ between the subplots. In each panel, the trajectories, shown as the frequency of the fit type as a function of time, were obtained under identical conditions.

surviving individuals produce juveniles according to the given model (Equation 4). If the total number of juveniles is less than N, all the juveniles are assumed to survive; in this way, we allow the population to recover from the bottleneck. Otherwise, we assign a viability weight (weight one to the fit type and weight e^{-s} to the wild type for haploid populations; for diploid populations see Equation S11) to each juvenile and sample N of them as described above. We write B for the event a bottleneck occurs in a given generation and B^c for the event a bottleneck does not occur in a given generation, and we define $\mathbb{1}_{\{B\}} := 1$ if B occurs and zero otherwise. Denoting N_t as the population size at time t, we take $N_{t+1} = \min(S_M, N)$, where S_M denotes the total number of juveniles produced by $M = \mathbb{1}_{\{B\}} N_b + \mathbb{1}_{\{B^c\}} N_t$ individuals. Thus, viability selection comes into effect only when the total number of juveniles exceeds N. This way of modelling bottlenecks is somewhat similar to the model in Equation (5) (see Eldon & Wakeley, 2006). The model in Equation (5) can be seen as a model of an instantaneous bottleneck followed by immediate recovery of the population. A bottleneck can happen in any generation. At the start of a generation, we toss a coin with probability b of a bottleneck, and if a bottleneck happens, we sample a fixed number B of individuals to survive the bottleneck and produce juveniles. Thus, our model corresponds to the model in Equation (5) (Eldon & Wakeley, 2006), which can be seen as a model of randomly occurring bottlenecks with immediate recovery; in our framework, how quickly a population recovers from a bottleneck depends on the reproduction parameters α and the upper bound $u(N)$ on the number of juveniles an individual (or a parent pair in diploid populations) can produce (see Equation 3).

In Figure 3, we investigate the effects of randomly occurring bottlenecks on the evolution of a haploid population evolving according to Equation (4). To recall the mechanism described in Equation (4), individuals produce a small number (relative to the population size) of juveniles most of the time (with probability $1 - \varepsilon_N$) but with a chance (represented by ε_N), when reproduction matches favourable conditions, of producing an increased number of juveniles, with carrying capacity $N = 10^6$, $\alpha_1 = 0.75$ (representing an increased probability of producing many juveniles according to Equation 3), $\alpha_2 = 3$ (representing a decreased probability of producing many juveniles), cutoff $u(N) = N$ (each individual can produce at most $u(N)$ juveniles), and selection strength $s = 0.5$ throughout. The parameters that vary as shown in Figure 3 are ε_N (the probability of matching reproduction

with favourable environmental conditions, so that individuals have a higher chance of producing many juveniles, i.e. $\alpha = \alpha_1$ in Equation 3), bottleneck size B (the number of individuals surviving a bottleneck), and b the probability of a bottleneck in a given generation. Bottlenecks clearly affect the probability of fixation; the top row shows the excursions to fixation for 10^2 experiments, but the remaining panels all show excursions for 10^3 experiments except for panel (j), where there are over 16,000 experiments. We emphasize that we are interested in uncovering broad trends in how random sweepstakes and bottlenecks affect selection, and we are not aiming for precise estimates of $\tau_N(1)$ and $p_N(1)$ (Equation S10). Figure 3 shows that if bottlenecks on average occur frequently (high probability of a bottleneck), the probability of fixation of the advantageous type ($p_N(1)$, Equation S10) is reduced. We claim that a lower $p_N(1)$ would be expected from a high frequency of bottlenecks, since then a bottleneck will occur with high probability while the beneficial type is still in low frequency (we always start with the beneficial type in one copy), and so there is a good chance of losing the type through a bottleneck. Furthermore, if bottlenecks occur frequently but are not too severe, the time to fixation is increased if the random sweepstakes are not too severe (top two rows); if random sweepstakes occur with high probability ($\varepsilon_N = 0.1$, bottom two rows) they clearly cancel out the effect of bottlenecks on the time to fixation. We estimate that strong random sweepstakes as given by the scenario in Figure 3g–l reduce the probability of fixation roughly tenfold compared to the case of weak sweepstakes Figure 3a–f.

The fixation trajectory (or a part of it) of a beneficial type in a haploid population has been described with a logistic differential equation (LDE; Kaplan et al., 1989; Stephan et al., 1992; see Equation S14 in Section S6). See the work of Schweinsberg & Durrett, 2005, for another approach to describe fixation trajectories. Figure S1 in Section S2 gives examples of trajectories well approximated by a LDE even in the presence of (moderated with $u(N) = N$) random sweepstakes (Figure S1b). Frequent recurrent bottlenecks (Figure 3d–f, j–l) clearly cause significant deviations from the logistic curve; even less frequently occurring bottlenecks (Figure 3a–c, g–i) generate notable deviations from the LDE. The trajectory a given mutation travels towards fixation may inform about the shape of the site-frequency spectrum of a sample. There is pervasive U-shape of the site-frequency spectrum in genomic data from Atlantic cod (Árnason et al., 2023), and the mutations at observed

(a) $\varepsilon_N = 0$, $B = 10^2$, $b = 0.01$
(b) $\varepsilon_N = 0$, $B = 10^4$, $b = 0.01$
(c) $\varepsilon_N = 0$, $B = 10^2$, $b = 0.1$
(d) $\varepsilon_N = 0$, $B = 10^4$, $b = 0.1$
(e) $\varepsilon_N = 0.1$, $B = 10^2$, $b = 0.01$
(f) $\varepsilon_N = 0.1$, $B = 10^4$, $b = 0.01$
(g) $\varepsilon_N = 0.1$, $B = 10^2$, $b = 0.1$
(h) $\varepsilon_N = 0.1$, $B = 10^4$, $b = 0.1$

FIGURE 4 Complete dominance and fixation trajectories. Examples of excursions to fixation of the genotype (1 / 1) homozygous for the type at a single locus conferring selective advantage in a diploid population evolving according to Equation (4). The mechanism described by Equation (4) is that occasionally (with probability ε_N) individuals have a higher chance of producing many juveniles, but most of the time (with probability $1 - \varepsilon_N$) individuals produce a small number (relative to the population size) number of juveniles with high probability, and experiencing randomly occurring bottlenecks, with carrying capacity $2N = 10^6$, $\alpha_1 = 0.75$ representing a higher chance of producing many juveniles, $\alpha_2 = 3$ representing a lower chance of producing many juveniles, cutoff $u(N) = 2N$ meaning that each parent pair produces at most $u(N)$ juveniles, strength of selection $s = 0.5$ throughout, with ε_N, the number B of individuals surviving a bottleneck, and the probability b of a bottleneck in any given generation as shown. Here, we consider the case of complete dominance of the beneficial type with the heterozygote as fit as the homozygote for the fit type with weight one, and the homozygote for the wild type least fit with weight e^{-4s} (see Equation S11) in Section S3. Results shown from 40 experiments (a, b, e) and otherwise from 10^2 experiments. The scale of the time (horizontal) axis may differ between the subplots. In each panel, the trajectories shown were obtained under identical conditions. The excursions are shown as $n_2(t) / n(t)$, where $n_2(t)$ is the number of copies of the homozygous 1 / 1 type at time t, and $n(t)$ is the total number of gene copies in the population at time t.

segregating sites may represent mutations travelling along a fixation trajectory on their way to fixation. The trajectories of a completely dominant fit type (the heterozygote is as fit as the homozygote for the fit type, see Equation (S12) in Section S3) are characterized by an extended time with the fit type at high frequency (Figure 4). This type of a trajectory indicates that mutations that we pick up in a

sample showing a U-shaped site-frequency spectrum may either be mutations under positive selection travelling on this kind of a trajectory, or other mutations hitchhiking with a positive mutation: this may explain the excess (relative to predictions of the Kingman-coalescent) of mutations in high frequency in the right tail of a U-shaped site-frequency spectrum. Forward-in-time simulations (using

SLiM Haller & Messer, 2016) of strongly beneficial dominant or semi-dominant positive mutations arising in a population evolving according to the Wright-Fisher model yielded site-frequency spectrum matching quite well the observed spectra from Atlantic cod (Árnason et al., 2023). Taken together, we believe that these results show that knowledge of the shape of fixation trajectories may inform about the footprint of selection in data.

To summarize the simulation results for haploid populations and focusing on the time to fixation ($\tau_N(1)$ in Equation S10) frequently occurring bottlenecks increase $\tau_N(1)$ relative to infrequently occurring bottlenecks when random sweepstakes occur infrequently (Figure 3a–f). Increasing the frequency of random sweepstakes (Figure 3g–l) counteracts the effect of bottlenecks on $\tau_N(1)$. In the absence of bottlenecks, random sweepstakes as modelled in Equation (4) with a bound on the number of juveniles have little effect on $\tau_N(1)$ (Figure S1). In the complete absence of random sweepstakes (Figure S2), increasing the frequency of bottlenecks (Figure S2d–f) increases $\tau_N(1)$; however, this depends on the severity of the bottleneck. Bottlenecks thus allow random sweepstakes to have an effect on $\tau_N(1)$. The upper bound ($u(N)$ in Equation 3) for the number of juveniles produced by any individual remains fixed at the carrying capacity. This means that when a bottleneck occurs, the population size becomes smaller than the cutoff, thus increasing the chance for individuals to produce a large number of juveniles relative to the population size. This is consistent with previous simulation results for haploid populations of fixed size, where it was seen that taking the cutoff larger than the population size shortened $\tau_N(1)$ (Eldon & Stephan, 2018).

3.2 | Diploid populations

In addition to haploid populations, we consider the effect of random sweepstakes and randomly occurring bottlenecks on selection in diploid populations. To this end, we consider a diploid population of maximum size $2N$ diploid individuals (the carrying capacity). In any given generation, the current diploid individuals arbitrarily form pairs, and the pairs then independently produce juveniles according to Equation (4). The juveniles are assigned gene copies following Mendel's laws, that is, each diploid juvenile receives one gene copy from each diploid parent. The genotype of each given juvenile then determines the viability weight as described in Equation S11 in Section S3. We then proceed as previously described for haploid populations. Diploidy gives us an opportunity to investigate the joint effects of dominance mechanisms, random sweepstakes and randomly occurring bottlenecks on $p_N(1)$ and $\tau_N(1)$ (Equation S10). We will consider complete dominance and incomplete dominance of the fit type as well as the case with the fit type being recessive (see Equation (S12) in Section S3). In all cases, the optimal genotype is the homozygous $1/1$ type. In the complete dominance case, it makes sense to consider $1/1$ as the optimal type, since heterozygotes contain the 0 type, so while there are heterozygotes in the population there is always a chance of a $0/0$ type.

In Figure 4, we compare the effects of random sweepstakes and randomly occurring bottlenecks on $p_N(1)$ and $\tau_N(1)$ defined in Equation (S10) in the case of complete dominance of the beneficial type (Equation S12). Random sweepstakes ($\varepsilon_N = 0.1$, bottom two rows) reduce $p_N(1)$ (Equation S10) only slightly. The effect of random sweepstakes on $\tau_N(1)$ is particularly noticeable in the case of a 'weak' bottleneck (bottleneck size 10^4, right column). The case of the beneficial type showing incomplete dominance as defined in Equation (S12) is investigated in Section S3.2: see Figure S4.

In the case of the beneficial type being recessive (Figure S5 in Section S3.3) without random sweepstakes ($\varepsilon_N = 0$ in Equation 4), the time for fixation is longer if the bottlenecks are weak (more individuals surviving a bottleneck) (Figure S5b,d). The main effect of random sweepstakes is to shorten $\tau_N(1)$, while bottlenecks tend to increase $\tau_N(1)$. Furthermore, there is a clear qualitative difference in the excursions to fixation depending on the dominance mechanism (Figures S3–S5).

To summarize the simulation results for diploid populations and focusing on $\tau_N(1)$, the effect of random sweepstakes depends on the dominance mechanism. In the absence of bottlenecks (Figure S3), random sweepstakes have negligible effect on $\tau_N(1)$ when the fit type is incompletely dominant (Figure S3a,b), but clearly reduce $\tau_N(1)$ when the fit type is either completely dominant (Figure S3c,d), or recessive (Figure S3e,f). When the fit type is dominant (Figure 4), increasing the frequency of bottlenecks increases $\tau_N(1)$ in the absence of random sweepstakes (Figure 4a–d); introducing random sweepstakes (Figure 4e–h) largely negates the effect of bottlenecks on $\tau_N(1)$. Introducing frequently occurring bottlenecks to a population with an incompletely dominant fit type (Figure S4) increases $\tau_N(1)$ in the absence of random sweepstakes (Figure S4c,d) relative to less frequently occuring bottlenecks (Figure S4a,b); again, introducing random sweepstakes (Figure S4e–h) largely nullifies the effect of bottlenecks on $\tau_N(1)$. Similarly, one can compare the effects of random sweepstakes and bottlenecks on $\tau_N(1)$ when the fit type is recessive (Figure S5). As in the case of haploid populations, bottlenecks allow random sweepstakes to have an effect on $\tau_N(1)$ since the cutoff ($u(N)$ in Equation 3) remains fixed at the carrying capacity, so is larger than the population size when a bottleneck occurs.

4 | SELECTIVE SWEEPSTAKES

We have discussed the effect of random sweepstakes on selection. Now, we turn our focus on selection in the absence of random sweepstakes and discuss possible footprint of selection in data. For example, models of selection in haploid populations, where positive mutations have accumulated, and incorporating a form of clonal, or Hill-Robertson interference (Hill & Robertson, 1966), can lead to multiple-merger gene genealogies (Desai et al., 2013; Neher & Hallatschek, 2013; Schweinsberg, 2017). In contrast, underdominance of the fit type in a diploid population with two types, where the heterozygote is less fit than both homozygotes, results in the Kingman-coalescent describing the gene genealogy of a sample of

the fit type (Etheridge & Penington, 2022). Here, we do not focus on these models. We will be concerned with a model of selection that focuses on one selectively advantageous mutation at a time corresponding to our simulations of the evolution of a selectively advantageous type always starting in one copy.

Selective sweepstakes are a form of sweepstakes reproduction in which, in contrast to random sweepstakes, natural selection plays a key role. In selective sweepstakes, juveniles are seen as having to pass through independent selective filters as they go through the different developmental stages on their way to reproductive age (Williams, 1975). We are not aware of a mathematical model precisely for selective sweepstakes. One may, however, view selective sweepstakes as being well approximated by models of recurrent selective sweeps (Coop & Ralph, 2012; Durrett & Schweinsberg, 2004, 2005).

The Durrett-Schweinsberg model (Durrett & Schweinsberg, 2004, 2005) of recurrent selective sweeps (see Figure S6), each time from a new and strongly beneficial mutation, has been shown to explain population genomic data of Atlantic cod (Árnason et al., 2023). Furthermore, current models of random sweepstakes based on Equation (1) (Birkner et al., 2018; Schweinsberg, 2003), in addition to the Kingman coalescent incorporating complex demography and background selection, do not explain the Atlantic cod data (Árnason et al., 2023). To better understand the importance of this result we now briefly describe the Durrett-Schweinsberg model (Durrett & Schweinsberg, 2004, 2005). Consider a haploid population of constant size $2N$ evolving according to the continuous-time Moran model, that is, at exponentially distributed times a single individual contributes a single offspring, and another individual is removed to keep the population size constant. A new beneficial mutation occurs on a chromosome with rate proportional to $1/N$, where with probability $S > 0$ (independent of N) the new mutation will sweep to fixation in $\log(N)$ time units on average. Viewing time in N time units the gene genealogy of a neutral site on a chromosome converges (as $N \to \infty$) to a coalescent, which is a mixture of the Kingman coalescent and a multiple-merger coalescent (see Equation S6).

However, even though the Durrett-Schweinsberg model (see Section S4) essentially explains the U-shape (Figure 5) of the site-frequency spectrum of Atlantic cod, there are certain limitations to the model. It is essentially a haploid model, where selection acts directly on individual chromosomes rather than pairs of chromosomes in diploid individuals. Furthermore, in order to obtain a non-trivial coalescent (see Equation S6 in Section S1), the advantage of the beneficial mutation must be of order $\mathcal{O}(1)$ to lead to $\log(N)$ as the order of the average time it takes to sweep to fixation. Viewing time on the scale of N time units a sweep then generates instantaneous (multiple) mergers in the genealogy. However, to avoid an instantaneous merger of all the lineages, some lineages must be allowed to escape a sweep through recombination. This is a key element of the model. The short duration (on average) relative to the coalescent timescale of a sweep means that one must assume very high recombination rates in order to allow an escape during a sweep. Despite these limitations, the exceptionally good fit of the Durrett-Schweinsberg model to the Atlantic cod data shows that models of recurrent and
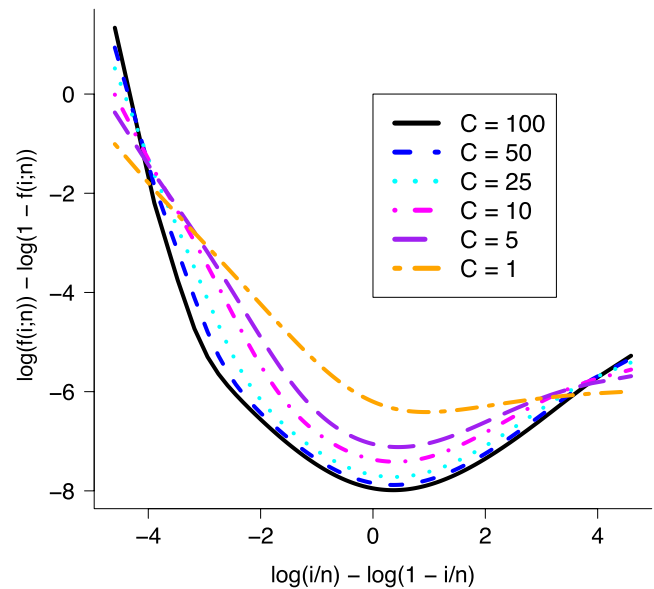
FIGURE 5 Recurrent sweeps predict U-shaped site-frequency spectrum. Logits (i.e. $\log(x) - \log(1 - x)$ for $0 < x < 1$) of the exact normalized expected branch length spectrum as a function of the logits of allele frequency, where $f(i; n) := \mathbb{E}\left[B_i(n)\right] / \mathbb{E}\left[B(n)\right]$, the normalized expected branch lengths supporting $i \in \{1, \ldots, n-1\}$ leaves corresponding to the size of a derived mutation (the number of times a derived mutation is observed in the sample), with sample size $n = 100$ and the parameter $\mathbb{C}$ from Equation (S6), the multiple-merger coalescent derived from the Durrett-Schweinsberg model of recurrent sweeps (Durrett & Schweinsberg, 2005a) as shown. The abscissa corresponds to a derived allele frequency (relative size of a mutation, logit scale), and the ordinate (vertical axis) corresponds to the expected number of derived mutations of a given size relative to the expected number of segregating sites (also on logit scale). The expected values were computed exactly using recursions (Birkner, Blath, & Eldon, 2013b).

pervasive selective sweeps, and multiple-merger coalescent models, are relevant for explaining genetic diversity in highly fecund natural populations. Indeed, related models of incomplete sweeps also lead to multiple-merger coalescents, with similar predictions of genetic diversity (Coop & Ralph, 2012).

5 | CONCLUSION

We have discussed adaptation in natural populations with a focus on highly fecund populations evolving according to sweepstakes reproduction. We have discussed two quite different mechanisms that turn high fecundity into skewed individual recruitment success. We have suggested new models of random sweepstakes that address the assumption of unbounded fecundity and the time scale issue and the related problem of recovering observed amount of genetic variation discussed above. Finally, we have used simulations to identify the main trends in how random sweepstakes affect fixation of a beneficial type under a simple model of viability selection with randomly occurring bottlenecks. The main impact of random sweepstakes is to reduce the time to fixation conditional on fixation (compared to

ordinary reproduction). This suggests that random sweepstakes facilitate rapid adaptation. Yet, as in the case of polygenic selection mentioned in Introduction, our one-locus model predicts that bottlenecks may increase the time to fixation (conditional on fixation) and thus limit the speed of adaptation.

We give examples of fixation trajectories to understand how a mutation sweeps to fixation, to try to understand how selection in diploid populations affects genetic diversity, and to learn about likely dominance mechanisms of new mutations in natural populations (Nanjundiah, 1993; Orr, 2010).

The fact that a model seen as approximating selective sweepstakes (see Section 4) essentially explains population genomic data of a highly fecund population indicates that sweepstakes reproduction also facilitates rapid adaptation through selective sweepstakes. Under the Durrett-Schweinsberg model a beneficial mutation sweeps to fixation with a probability of order $\mathcal{O}(1)$ independent of the population size, and the duration of a sweep is (on average) only of order $\log(N)$ time units (Durrett & Schweinsberg, 2005). The variants of the Schweinsberg model (Schweinsberg, 2003) we study here, see Equations 3 and 4, have yet to be compared to data; they may be shown to give as good or even better fit than the Durrett-Schweinsberg model. Extending these models of random and selective sweepstakes to the genomic scale, that is, to several chromosomes, remains an important future task, not least since U-shaped site-frequency spectra as predicted by multiple-merger coalescents and not by the Kingman-coalescent are observed across domains of life (Freund et al., 2022). Note that we see the Durrett-Schweinsberg model as a first approximation of selective sweepstakes. To the best of our knowledge, it remains to model recurrent sweeps in a population evolving according to random sweepstakes. A rigorous mathematical verification of our simulation results is also an important follow-up project. However, our answer to our main question is that sweepstakes reproduction facilitates rapid adaptation where we have defined 'rapid adaptation' based on the time to fixation of a beneficial type given that it does so.

AUTHOR CONTRIBUTIONS

Designed research: BE and WS; Performed research: BE and WS; Contributed new software for simulations and analysis: BE; Analyzed data: BE and WS; wrote the paper: BE and WS.

ACKNOWLEDGMENTS

OPEN RESEARCH BADGES

This article has earned an Open Data badge for making publicly available the digitally-shareable data necessary to reproduce the reported results. The data is available at 10.5061/dryad.xwdbrv1gf.

DATA AVAILABILITY STATEMENT

The C++ code written for the simulations is freely available at https://github.com/eldonb/sweepstakes_reproduction_facilitates_rapid_adaptation.

The code is also available at Dryad with the DOI 10.5061/dryad.xwdbrv1gf. The code behind Figure 5 is freely available at https://github.com/eldonb/Durrett_Schweinsberg_Expected_SFS.

BENEFIT-SHARING STATEMENT

The C++ code written for the simulations is being made freely available. The paper will be published Open-Access and therefore be made freely available.

ORCID

Bjarki Eldon https://orcid.org/0000-0001-9354-2391

REFERENCES

Agrios, G. N. (2010). *Introduction to plant pathology*. Elsevier.

Árnason, E. (2004). Mitochondrial cytochrome *b* variation in the high-fecundity Atlantic cod: Trans-Atlantic clines and shallow gene genealogy. *Genetics*, *166*, 1871–1885.

Árnason, E., & Halldórsdóttir, K. (2015). Nucleotide variation and balancing selection at the *Ckma* gene in Atlantic cod: Analysis with multiple merger coalescent models. *PeerJ*, *3*, e786.

Árnason, E., Koskela, J., Halldórsdóttir, K., & Eldon, B. (2023). Sweepstakes reproductive success via pervasive and recurrent selective sweeps. *eLife*, *12*, e80781.

Barfield, S., Davies, S. W., & Matz, M. V. (2022). Evidence of sweepstakes reproductive success in a broadcast-spawning coral and its implications for coral metapopulation persistence. *Molecular Ecology*, *32*, 696–702.

Baumdicker, F., Bisschop, G., Goldstein, D., Gower, G., Ragsdale, A. P., Tsambos, G., Zhu, S., Eldon, B., Ellerman, E. C., Galloway, J. G., Gladstein, A. L., Gorjanc, G., Guo, B., Jeffery, B., Kretzschmar, W. W., Lohse, K., Matschiner, M., Nelson, D., Pope, N. S., … Kelleher, J. (2021). Efficient ancestry and mutation simulation with msprime 1.0. *Genetics*, *220*, iyab229.

Berestycki, N. (2009). Recent progress in coalescent theory. *Ensaios Mathématicos*, *16*, 1–193.

Berg, J. J., Harpak, A., Sinnott-Armstrong, N., Joergensen, A. M., Mostafavi, H., Field, Y., Boyle, E. A., Zhang, X., Racimo, F., Pritchard, J. K., & Coop, G. (2019). Reduced signal for polygenic adaptation of height in UK biobank. *eLife*, *8*, e39725.

Birkner, M., & Blath, J. (2008). Computing likelihoods for coalescents with multiple collisions in the infinitely many sites model. *Journal of Mathematical Biology*, *57*, 435–465.

Birkner, M., & Blath, J. (2009). Measure-valued diffusions, general coalescents and population genetic inference. In J. Blath, P. Mörters, & M. Scheutzow (Eds.), *Trends in stochastic analysis* (pp. 329–363). Cambridge University Press.

Birkner, M., Blath, J., & Eldon, B. (2013a). An ancestral recombination graph for diploid populations with skewed offspring distribution. *Genetics*, *193*, 255–290.

Birkner, M., Blath, J., & Eldon, B. (2013b). Statistical properties of the site-frequency spectrum associated with Λ-coalescents. *Genetics*, *195*, 1037–1053.

Birkner, M., Blath, J., & Steinrücken, M. (2011). Importance sampling for lambda-coalescents in the infinitely many sites model. *Theoretical Population Biology*, *79*, 155–173.

Birkner, M., Blath, J., & Steinrücken, M. (2013). Analysis of DNA sequence variation within marine species using Beta-coalescents. *Theoretical Population Biology*, *87*, 15–24.

Birkner, M., Liu, H., & Sturm, A. (2018). Coalescent results for diploid exchangeable population models. *Electronic Journal of Probability*, *23*, 1–44.

Blath, J., Cronjäger, M. C., Eldon, B., & Hammer, M. (2016). The site-frequency spectrum associated with Ξ-coalescents. *Theoretical Population Biology*, *110*, 36–50.

Botsford, L. W., Hastings, A., & Gaines, S. D. (2001). Dependence of sustainability on the configuration of marine reserves and larval dispersal distance. *Ecology Letters*, *4*(2), 144–150.

Canino, M. F., Spies, I. B., Cunningham, K. M., Hauser, L., & Grant, W. S. (2010). Multiple ice-age refugia in Pacific cod, *Gadus macrocephalus*. *Molecular Ecology*, *19*(19), 4339–4351.

Cook, L. M., Grant, B. S., Saccheri, I. J., & Mallet, J. (2012). Selective bird predation on the peppered moth: The last experiment of Michael Majerus. *Biology Letters*, *8*(4), 609–612.

Coop, G., & Ralph, P. (2012). Patterns of neutral diversity under general models of selective sweeps. *Genetics*, *192*(1), 205–224.

Cowen, R. K., & Sponaugle, S. (2009). Larval dispersal and marine population connectivity. *Annual Review of Marine Science*, *1*, 443–466.

Daborn, P., Yen, J., Bogwitz, M., Le Goff, G., Feil, E., Jeffers, S., Tijet, N., Perry, T., Heckel, D., Batterham, P., Feyereisen, R., Wilson, T. G., & ffrench-Constant, R. (2002). A single P450 allele associated with insecticide resistance in *Drosophila*. *Science*, *297*(5590), 2253–2256.

de Vladar, H. P., & Barton, N. (2014). Stability and response of polygenic traits to stabilizing selection and mutation. *Genetics*, *197*(2), 749–767.

Der, R., Epstein, C., & Plotkin, J. B. (2012). Dynamics of neutral and selected alleles when the offspring distribution is skewed. *Genetics*, *191*(4), 1331–1344.

Der, R., Epstein, C. L., & Plotkin, J. B. (2011). Generalized population models and the nature of genetic drift. *Theoretical Population Biology*, *80*(2), 80–99.

Desai, M. M., Walczak, A. M., & Fisher, D. S. (2013). Genetic diversity and the structure of genealogies in rapidly adapting populations. *Genetics*, *193*(2), 565–585.

Donnelly, P., & Kurtz, T. G. (1999). Particle representations for measure-valued population models. *Annals of Probability*, *27*, 166–205.

Durrett, R., & Schweinsberg, J. (2004). Approximating selective sweeps. *Theoretical Population Biology*, *66*, 129–138.

Durrett, R., & Schweinsberg, J. (2005). A coalescent model for the effect of advantageous mutations on the genealogy of a population. *Stochastic Processes and their Applications*, *115*(10), 1628–1657.

Eldon, B. (2011). Estimation of parameters in large offspring number models and ratios of coalescence times. *Theoretical Population Biology*, *80*, 16–28.

Eldon, B. (2016). Inference methods for multiple merger coalescents. In P. Pontarotti (Ed.), *Evolutionary biology: Convergent evolution, evolution of complex traits, concepts and methods* (pp. 347–371). Springer.

Eldon, B. (2020). Evolutionary genomics of high fecundity. *Annual Review of Genetics*, *54*(1), 213–236.

Eldon, B., Birkner, M., Blath, J., & Freund, F. (2015). Can the site-frequency spectrum distinguish exponential population growth from multiple-merger coalescents? *Genetics*, *199*, 841–856.

Eldon, B., & Freund, F. (2018). Genealogical properties of subsamples in highly fecund populations. *Journal of Statistical Physics*, *172*(1), 175–207.

Eldon, B., & Stephan, W. (2018). Evolution of highly fecund haploid populations. *Theoretical Population Biology*, *119*, 48–56.

Eldon, B., & Wakeley, J. (2006). Coalescent processes when the distribution of offspring number among individuals is highly skewed. *Genetics*, *172*, 2621–2633.

Etheridge, A., & Penington, S. (2022). Genealogies in bistable waves. *Electronic Journal of Probability*, *27*, 1–99.

Feller, W. (1951). Diffusion processes in genetics. In *Proceedings of the second Berkeley symposium on mathematical statistics and probability* (pp. 227–246). University of California Press.

Foucart, C. (2013). The impact of selection in the Λ-Wright-Fisher model. *Electronic Communications in Probability*, *18*, 1–10.

Freund, F., Kerdoncuff, E., Matuszewski, S., Lapierre, M., Hildebrandt, M., Jensen, J. D., Ferretti, L., Lambert, A., Sackton, T. B., & Achaz, G. (2022). Interpreting the pervasive observation of U-shaped site frequency spectra. *bioRxiv*. https://doi.org/10.1101/2022.04.12.488084

Freund, F., & Siri-Jégousse, A. (2021). The impact of genetic diversity statistics on model selection between coalescents. *Computational Statistics & Data Analysis*, *156*, 107055.

Fu, Y.-X., & Li, W.-H. (1999). Coalescing into the 21st century: An overview and prospects of coalescent theory. *Theoretical Population Biology*, *56*(1), 1–10.

Gagnaire, P.-A., Broquet, T., Aurelle, D., Viard, F., Souissi, A., Bonhomme, F., Arnaud-Haond, S., & Bierne, N. (2015). Using neutral, selected, and hitchhiker loci to assess connectivity of marine populations in the genomic era. *Evolutionary Applications*, *8*(8), 769–786.

Grant, P. R., & Grant, B. R. (2020). *How and why species multiply*. Princeton University Press.

Grant, W. S., Árnason, E., & Eldon, B. (2016). New DNA coalescent models and old population genetics software. *ICES Journal of Marine Science*, *73*(9), 2178–2180.

Haldane, J. (1924). A mathematical theory of natural and artificial selection. *Transactions of the Cambridge Philosophical Society*, *23*, 19–41.

Haller, B. C., & Messer, P. W. (2016). SLiM 2: Flexible, interactive forward genetic simulations. *Molecular Biology and Evolution*, *34*(1), 230–240.

Hedgecock, D. (1994). Does variance in reproductive success limit effective population sizes of marine organisms? In A. R. Beaumont (Ed.), *Genetics and evolution of marine organisms* (pp. 122–134). Chapman and Hall.

Hedgecock, D., & Pudovkin, A. I. (2011). Sweepstakes reproductive success in highly fecund marine fish and shellfish: A review and commentary. *Bulletin of Marine Science*, *87*, 971–1002.

Hill, W. G., & Robertson, A. (1966). The effect of linkage on limits to artificial selection. *Genetical Research*, *8*(3), 269–294.

Huillet, T., & Möhle, M. (2011). Population genetics models with skewed fertilities: Forward and backward analysis. *Stochastic Models*, *27*, 521–554.

Huillet, T., & Möhle, M. (2013). On the extended Moran model and its relation to coalescents with multiple collisions. *Theoretical Population Biology*, *87*, 5–14.

Irwin, K. K., Laurent, S., Matuszewski, S., Vuilleumier, S., Ormond, L., Shim, H., Bank, C., & Jensen, J. D. (2016). On the importance of skewed offspring distributions and background selection in virus population genetics. *Heredity*, *117*, 393–399.

Jain, K., & Stephan, W. (2015). Response of polygenic traits under stabilizing selection and mutation when loci have unequal effects. *G3: Genes, Genomes, Genetics*, *5*(6), 1065–1074.

Jain, K., & Stephan, W. (2017a). Modes of rapid polygenic adaptation. *Molecular Biology and Evolution*, *34*(12), 3169–3175.

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

John, S., & Stephan, W. (2020). Important role of genetic drift in rapid polygenic adaptation. *Ecology and Evolution*, *10*(3), 1278–1287.

Kaplan, N. L., Hudson, R. R., & Langley, C. H. (1989). The "hitchhiking effect" revisited. *Genetics*, *123*(4), 887–899.

Kimura, M. (1957). Some problems of stochastic processes in genetics. *The Annals of Mathematical Statistics*, *28*(4), 882–901.

Kingman, J. F. C. (1982a). The coalescent. *Stochastic Processes and their Applications*, 13, 235–248.

Kingman, J. F. C. (1982b). Exchangeability and the evolution of large populations. In G. Koch & F. Spizzichino (Eds.), *Exchangeability in probability and statistics* (pp. 97–112). North-Holland.

Kingman, J. F. C. (1982c). On the genealogy of large populations. *Journal of Applied Probability*, 19A, 27–43.

Koskela, J. (2018). Multi-locus data distinguishes between population growth and multiple merger coalescents. *Statistical Applications in Genetics and Molecular Biology*, 17(3), 20170011.

Koskela, J., & Berenguer, M. W. (2019). Robust model selection between population growth and multiple merger coalescents. *Mathematical Biosciences*, 311, 1–12.

Lande, R. (1976). Natural selection and random genetic drift in phenotypic evolution. *Evolution*, 30(2), 314.

Losos, J. (2009). *Lizards in an evolutionary tree*. University of California Press.

Mahrt, N., Tietze, A., Künzel, S., Franzenburg, S., Barbosa, C., Jansen, G., & Schulenburg, H. (2021). Bottleneck size and selection level reproducibly impact evolution of antibiotic resistance. *Nature Ecology & Evolution*, 5(9), 1233–1242.

Matuszewski, S., Hildebrandt, M. E., Achaz, G., & Jensen, J. D. (2018). Coalescent processes with skewed offspring distributions and nonequilibrium demography. *Genetics*, 208(1), 323–338.

Maynard Smith, J., & Haigh, J. (1974). The hitch-hiking effect of a favourable gene. *Genetical Research*, 23(1), 23–35.

Möhle, M. (2011). Coalescent processes derived from some compound Poisson population models. *Electronic Communications in Probability*, 16, 567–582.

Möhle, M., & Sagitov, S. (2001). A classification of coalescent processes for haploid exchangeable population models. *Annals of Probability*, 29, 1547–1562.

Möhle, M., & Sagitov, S. (2003). Coalescent patterns in diploid exchangeable population models. *Journal of Mathematical Biology*, 47, 337–352.

Nanjundiah, V. (1993). Why are most mutations recessive? *Journal of Genetics*, 72(2–3), 85–97.

Neher, R. A., & Hallatschek, O. (2013). Genealogies of rapidly adapting populations. *Proceedings of the National Academy of Sciences of the United States of America*, 110(2), 437–442.

Niwa, H.-S., Nashida, K., & Yanagimoto, T. (2016). Reproductive skew in Japanese sardine inferred from DNA sequences. *ICES Journal of Marine Science*, 73(9), 2181–2189.

Orr, H. A. (2010). The population genetics of beneficial mutations. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, 365(1544), 1195–1201.

Patwa, Z., & Wahl, L. M. (2008). The fixation probability of beneficial mutations. *Journal of the Royal Society Interface*, 5(28), 1279–1289.

Pitman, J. (1999). Coalescents with multiple collisions. *Annals of Probability*, 27, 1870–1902.

Pritchard, J. K., & Di Rienzo, A. (2010). Adaptation–not by sweeps alone. *Nature Reviews Genetics*, 11(10), 665–667.

Pritchard, J. K., Pickrell, J. K., & Coop, G. (2010). The genetics of human adaptation: Hard sweeps, soft sweeps, and polygenic adaptation. *Current Biology*, 20(4), R208–R215.

Reznick, D. N. (2011). *The origin then and now*. Princeton University Press.

Sagitov, S. (1999). The general coalescent with asynchronous mergers of ancestral lines. *Journal of Applied Probability*, 36, 1116–1125.

Sagitov, S. (2003). Convergence to the coalescent with simultaneous mergers. *Journal of Applied Probability*, 40, 839–854.

Sargsyan, O., & Wakeley, J. (2008). A coalescent process with simultaneous multiple mergers for approximating the gene genealogies of many marine organisms. *Theoretical Population Biology*, 74, 104–114.

Schweinsberg, J. (2000). Coalescents with simultaneous multiple collisions. *Electronic Journal of Probability*, 5, 1–50.

Schweinsberg, J. (2003). Coalescent processes obtained from supercritical Galton-Watson processes. *Stochastic Processes and their Applications*, 106, 107–139.

Schweinsberg, J. (2017). Rigorous results for a population model with selection II: Genealogy of the population. *Electronic Journal of Probability*, 22, 1–54.

Schweinsberg, J., & Durrett, R. (2005). Random partitions approximating the coalescence of lineages during a selective sweep. *The Annals of Applied Probability*, 15, 1591–1651.

Selkoe, K. A., Aloia, C. C., Crandall, E. D., Iacchei, M., Liggins, L., Puritz, J. B., von der Heyden, S., & Toonen, R. J. (2016). A decade of seascape genetics: Contributions to basic and applied marine connectivity. *Marine Ecology Progress Series*, 554, 1–19.

Sohail, M., Maier, R. M., Ganna, A., Bloemendal, A., Martin, A. R., Turchin, M. C., Chiang, C. W., Hirschhorn, J., Daly, M. J., Patterson, N., Neale, B., Mathieson, I., Reich, D., & Sunyaev, S. R. (2019). Polygenic adaptation on height is overestimated due to uncorrected stratification in genome-wide association studies. *eLife*, 8, e39702.

Stephan, W. (2016). Signatures of positive selection: From selective sweeps at individual loci to subtle allele frequency changes in polygenic adaptation. *Molecular Ecology*, 25(1), 79–88.

Stephan, W. (2019). Selective sweeps. *Genetics*, 211(1), 5–13.

Stephan, W., Wiehe, T. H., & Lenz, M. W. (1992). The effect of strongly selected substitutions on neutral polymorphism: Analytical results based on diffusion theory. *Theoretical Population Biology*, 41(2), 237–254.

Tajima, F. (1983). Evolutionary relationships of DNA sequences in finite populations. *Genetics*, 105, 437–460.

Timm, A., & Yin, J. (2012). Kinetics of virus production from single cells. *Virology*, 424(1), 11–17.

Turchin, M. C., Chiang, C. W., Palmer, C. D., Sankararaman, S., Reich, D., & Hirschhorn, J. N. (2012). Evidence of widespread selection on standing variation in Europe at height-associated SNPs. *Nature Genetics*, 44(9), 1015–1019.

van Herwaarden, O. A., & van der Wal, N. J. (2002). Extinction time and age of an allele in a large finite population. *Theoretical Population Biology*, 61(3), 311–318.

van't Hof, A. E., Edmonds, N., Dalíková, M., Marec, F., & Saccheri, I. J. (2011). Industrial melanism in British peppered moths has a singular and recent mutational origin. *Science*, 332(6032), 958–960.

Vendrami, D. L. J., Peck, L. S., Clark, M. S., Eldon, B., Meredith, M., & Hoffman, J. I. (2021). Sweepstake reproductive success and collective dispersal produce chaotic genetic patchiness in a broadcast spawner. *Science Advances*, 7(37), eabj4713.

Vignieri, S. N., Larson, J. G., & Hoekstra, H. E. (2010). The selective advantage of crypsis in mice. *Evolution: International Journal of Organic Evolution*, 64(7), 2153–2158.

Wakeley, J. (2004). Recent trends in population genetics: More data! More math! Simple models? *Journal of Heredity*, 95(5), 397–405.

Wakeley, J. (2007). *Coalescent theory*. Roberts & Co.

Williams, G. (1975). *Sex and evolution. Monographs in population biology*. Princeton University Press.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.