

Dynamics of Neutral and Selected Alleles When the Offspring Distribution Is Skewed

Ricky Der,^{*1} Charles Epstein,[†] and Joshua B. Plotkin^{*}

^{*}Department of Biology and [†]Department of Mathematics, University of Pennsylvania, Philadelphia, Pennsylvania 19104

ABSTRACT We analyze the dynamics of two alternative alleles in a simple model of a population that allows for large family sizes in the distribution of offspring number. This population model was first introduced by Eldon and Wakeley, who described the backward-time genealogical relationships among sampled individuals, assuming neutrality. We study the corresponding forward-time dynamics of allele frequencies, with or without selection. We derive a continuum approximation, analogous to Kimura's diffusion approximation, and we describe three distinct regimes of behavior that correspond to distinct regimes in the coalescent processes of Eldon and Wakeley. We demonstrate that the effect of selection is strongly amplified in the Eldon–Wakeley model, compared to the Wright–Fisher model with the same variance effective population size. Remarkably, an advantageous allele can even be guaranteed to fix in the Eldon–Wakeley model, despite the presence of genetic drift. We compute the selection coefficient required for such behavior in populations of Pacific oysters, based on estimates of their family sizes. Our analysis underscores that populations with the same effective population size may nevertheless experience radically different forms of genetic drift, depending on the reproductive mechanism, with significant consequences for the resulting allele dynamics.

WHEREAS natural selection undoubtedly plays an important role in adaptation, genetic drift is recognized as an equally or even more important force in shaping the patterns of heritable variation in a population. At a qualitative level, genetic drift implies that variation will tend to be lost from a population over time, even in the absence of selection, and that, in the presence of selection, the fitter type in a population is not guaranteed to fix. Our understanding of genetic drift and its interplay with natural selection has been shaped primarily by the population models introduced by Wright and Fisher (Wright 1931; Fisher 1958) and Moran (Moran 1958). Most work in population genetics rests implicitly on the Wright–Fisher framework, including Kimura's work on fixation probabilities (Kimura 1955), Ewens' sampling formula (Ewens 1972; Lessard 2007), Kingman's coalescent (Kingman 1982), tests of neutrality (Hudson *et al.* 1987; Tajima 1989; McDonald and Kreitman 1991; Fu and Li 1992; Fay and Wu 2000), and techniques for inferring mutation rates and selection pres-

ures (Sawyer and Hartl 1992; Yang and Bielawski 2000; Bustamante *et al.* 2001; Desai and Plotkin 2008).

The behavior of the Wright–Fisher model is remarkably robust to modifications of the model's underlying assumptions. Indeed, many other population genetic models, including the Moran process (Moran 1958), and a number of Karlin–Taylor and Cannings processes (Karlin and McGregor 1964; Cannings 1974; Ewens 2004), share the same diffusion limit as the Wright–Fisher model (Möhle 2001) and therefore have very similar behavior in sufficiently large populations. Likewise, the Kingman coalescent—that is, the backward-time description of genealogical relationships among a sample of individuals drawn from a Wright–Fisher population—is also robust to deviations in the underlying model assumptions (Kingman 1982).

Nevertheless, the Wright–Fisher model and its diffusion limit are not appropriate in all circumstances. As a result, in some cases the standard model of genetic drift does not provide an accurate description of variation in a population or accurately predict how drift disrupts the force of natural selection. The Wright–Fisher model is founded on a set of assumptions about how organisms reproduce and the way in which the constraint of finite population size enforces dependencies among individuals' offspring. In particular, the Wright–Fisher model assumes that individuals each produce

a Poisson-distributed number of offspring each generation, subject to the constraint of a constant population size (Karlin and McGregor 1964). This formulation excludes the possibility of a highly skewed distribution of offspring numbers, which has been observed empirically in some species. Marine species in particular, as well as some plants and fungi, sometimes produce a very large number of offspring when faced with high mortality early in life (Hedgewood 1994).

To accommodate species with the possibility of large family sizes, researchers have developed a form of coalescent theory that describes the genealogical relationships among samples from such a population, assuming neutrality (Pitman 1999; Sagitov 1999; Schweinsberg 2003; Eldon and Wakeley 2006, 2008, 2009; Sargsyan and Wakeley 2008). These generalized coalescents can differ dramatically from the standard Kingman coalescent, and they include the possibility of multiple mergers in the genealogy. As a result, for species with skewed offspring numbers the inferences reached by assuming the standard coalescent are likely to be incorrect (Eldon and Wakeley 2006). The forward-time dynamics of alleles in such populations have been defined in an abstract setting (Donnelly and Kurtz 1999; Birkner and Blath 2009), but their population-genetic consequences have yet to be worked out, especially those concerning the interaction between selection and genetic drift.

In this article, we study the forward-time population genetics of the population model introduced in Eldon and Wakeley (2006), and we describe their relationship to coalescent theory. Even though the backward-time perspective of such a population facilitates inferences from sampled genetic data, there are many merits to developing a corresponding forward-time perspective for such populations. Most importantly, a forward-time analysis allows us to incorporate the effects of natural selection, whose inclusion is difficult in a backward-time theory. Remarkably, we demonstrate that selection operates very differently in the Eldon–Wakeley model than it does in the standard Wright–Fisher or Moran model. In particular, the form of genetic drift that arises in populations with reproductive skew tends to amplify the effects of selection, relative to the standard form of Wright–Fisherian drift—even when both models are normalized to have the same variance-effective population size. The extent of selection amplification in the Eldon–Wakeley model can be so dramatic that advantageous alleles may even be guaranteed fixation in a population, despite the presence of genetic drift.

The Eldon–Wakeley model is an example of a generalized Wright–Fisher (GWF) process, a class of population processes first introduced and studied in Der *et al.* (2011). While the general methods and bounds of that earlier work apply to the Eldon–Wakeley process, the presentation in this article is considerably more refined and specific to the case of a skewed offspring distribution. In particular, we prove the convergence of the discrete chain to a particular Lambda population process, derive the appropriate scaling conditions on selection necessary for the existence of the contin-

uum limit, and present a new analytical formula to quantify the amount of selection amplification exhibited by the Eldon–Wakeley model, which may be generalized to the entire class of Lambda processes (see *Discussion*).

Our presentation is organized as follows. We start by describing the population model introduced by Eldon and Wakeley, and we specify it precisely in terms of a Markov transition matrix for the change in allele frequency between subsequent time steps. We next describe how to introduce selection into the model, by modifying the Markov transition matrix in a standard way. To analyze the forward-time behavior of the Eldon–Wakeley model, either with or without selection, we develop a continuum approximation, analogous to Kimura’s diffusion approximation of the Wright–Fisher model. We apply this continuum approximation to study the probability that a beneficial mutant will fix in the Eldon–Wakeley model, in three distinct regimes. We end by discussing how our results relate to empirical populations and to coalescent theory.

The Eldon–Wakeley Model

Eldon and Wakeley (2006) introduced a simple model of a population, based on the Moran process, that incorporates the possibility of large family sizes in the offspring distribution. They considered a population of N individuals, each of two types. At each time step, a single individual is chosen uniformly from the population and produces a random number U of offspring, drawn from a distribution of offspring numbers, P_U . The subsequent generation is then composed of the U offspring from the chosen individual supplemented by $N - U$ other individuals, randomly selected without replacement from the remainder of the population. To be clear, only a single individual contributes offspring in each reproduction event—the remaining individuals who neither contribute offspring nor die simply persist to the next time step. A schematic of this reproduction mechanism is shown in Figure 1.

This model coincides with the classical Moran process when the distribution P_U is concentrated at two individuals. Eldon and Wakeley (2006) considered the following distribution of offspring numbers,

$$\begin{aligned} \mathbb{P}(U = 2) &= 1 - N^{-\gamma} \\ \mathbb{P}(U = \lambda N) &= N^{-\gamma}, \end{aligned} \quad (1)$$

where $0 \leq \lambda \leq 1$ and $\gamma > 0$ are parameters of the model. The Eldon–Wakeley process is thus a mixture of two mechanisms: a typical Moran step in which one individual replicates and displaces another, which occurs with probability $1 - N^{-\gamma}$, and a rarer event in which a fixed fraction λ of the population is replaced by the offspring from a single individual, which occurs with probability $N^{-\gamma}$. The rare large-family-size events can alternatively be thought of as rapid bottlenecks, in which a fraction λ of the population is removed and immediately replaced by the offspring of a single

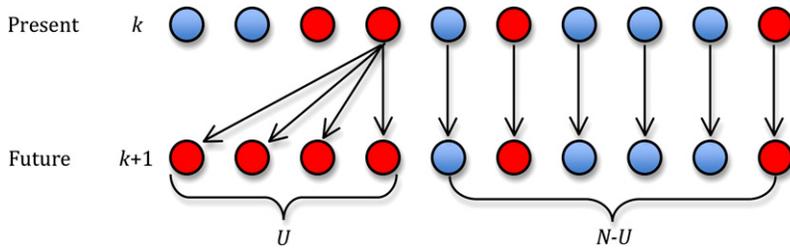


Figure 1 Schematic diagram of the Eldon–Wakeley population model.

individual. The parameter γ controls the frequency of such “bottleneck” events.

Our analysis in this article focuses on the Eldon–Wakeley offspring distribution given by Equation 1, but many of our results can be generalized to a large class of offspring distributions P_U (see *Discussion*).

Eldon and Wakeley (2006) analyzed the genealogical relationships in a sample from the population, assuming neutrality. We introduce selection into the Eldon–Wakeley model and study the forward-time behavior of the population. We are especially concerned with understanding the fixation probability of an adaptive mutant and how this probability differs from that of the standard Moran model.

The forward-time transition matrix

Here we study the forward-time evolution of a population under the Eldon–Wakeley model with selection, in the absence of mutation. We keep track of the number of individuals in generation k of type 1, denoted by X_k . The process X_k is a Markov chain on the states $\{0, \dots, N\}$, with some transition matrix \mathbf{P}_{ij} , $0 \leq i, j \leq N$. We decompose the transition matrix \mathbf{P} into a product of two matrices: one transition matrix due to selection alone and another due to the (neutral) reproduction process alone. In other words, we introduce selection by first altering the relative frequencies of the two types and then proceeding as in the neutral Eldon–Wakeley case. This form of viability selection coincides with the standard form of selection in the Moran model. (Other forms of selection, such as fecundity selection, generally coincide as well in the limit of large population size; see *Discussion*.)

More specifically, we define the transition matrix \mathbf{P} as

$$\mathbf{P} = \mathbf{S}\mathbf{Q}. \quad (2)$$

The selection matrix \mathbf{S} can be thought as the transition distribution for the standard logistic growth process, which, originating in state i , moves to state $[(1 + s/N)i \cdot N / ((1 + s/N)i + (N - i))]$ with probability 1. We use the notation $[x]$ to denote the integer part of the real number x . Under neutrality, $\mathbf{S} = \mathbf{I}$, the identity matrix. The matrix \mathbf{Q} describes the neutral genetic drift of the Eldon–Wakeley model; it may be written as the mixture

$$\mathbf{Q} = (1 - N^{-\gamma})\mathbf{M} + N^{-\gamma}\mathbf{J}, \quad (3)$$

where \mathbf{M} is the standard transition matrix for the two-type neutral Moran process, and \mathbf{J} is a “jump” matrix modeling

the rare bottleneck events. The rows of \mathbf{J} are in general mixtures of hypergeometric distributions (see *Appendix*); to a high degree of approximation, \mathbf{J} is a matrix of zeroes except for the index pairs $(i, i + [\lambda(N - i)])$ and $(i, i - [\lambda i])$, whose entries are respectively i/N and $1 - i/N$, denoting jumps in the allele frequency caused by the bottleneck events.

The “offspring distribution” of the process is the distribution represented by the row $\mathbf{Q}_{1,j}$, for $j = 0, 1, 2, \dots$. Its variance, given by (4), is an important quantity that will determine the appropriate time scaling of the continuum limit. For large N the offspring variance is

$$\sigma_N^2 = 2N^{-1} + \lambda^2 N^{-\gamma+1}, \quad N \rightarrow \infty. \quad (4)$$

From the above, σ_N^2 can be seen as a mixture of two components—a term $2/N$ arising from the Moran part of the process and a second term arising from the bottleneck component. Note that when $\gamma < 2$, *i.e.*, when the bottleneck events are sufficiently common, the offspring variance is dominated by the contribution from the jump process, whereas for $\gamma > 2$, it is dominated by the Moran component. This transition point in the offspring variance has important consequences for the nature of the continuum limit, discussed below.

A Continuum Approximation of the Eldon–Wakeley Process

Analysis of the Moran (or Wright–Fisher) model is typically facilitated by a continuum limit, which becomes accurate in the limit of large population size $N \rightarrow \infty$ (Kimura 1955; Ewens 2004). For the Moran model, the continuum limit is a diffusion—*i.e.*, it has continuous sample paths and it can be described by the second-order partial differential equations introduced by Kimura and others. By contrast, the continuum approximation of the Eldon–Wakeley model is complicated by the fact that no diffusion process (with continuous sample paths) is an appropriate limit point, due to the occasional large family sizes. These extreme reproduction events produce jumps in the allele frequencies, and hence discontinuous sample paths appear in the limit. Nonetheless, the techniques described in Der (2010; Der *et al.* 2011) can be applied to the Eldon–Wakeley model, with or without selection, to derive a continuous-time/space-limiting process that accurately approximates the discrete Eldon–Wakeley process.

State, time, and selection scalings

To derive a continuum approximation of a discrete process X_k we must choose how to scale time and space, as $N \rightarrow \infty$. Such scalings involve replacing the number i of individuals with the frequency $x = i/b_N$ and the generation number k by the time $t = kc_N$, for some choice of sequences $\{b_N\}$, $\{c_N\}$. The continuum limit is then the process

$$\tilde{X}_t = \lim_{N \rightarrow \infty} \frac{1}{b_N} X_{[t/c_N]}. \quad (5)$$

In the classical Moran model, for example, $b_N = N$ and $c_N = N^{-2}$ and the resulting continuum limit \tilde{X}_t is Kimura's classical diffusion process. It can be shown for a large class of neutral population models (Der 2010; Der *et al.* 2011) that the choice of state-space scaling dictates the choice of time scaling—up to a constant factor—in the sense that no other scaling leads to a nontrivial limiting process. Since we are interested in studying fixation events, we choose to normalize the state space by the population size, so that $b_N = N$; the general theory then indicates that the time scaling must be proportional to

$$c_N = \frac{\sigma_N^2}{N}, \quad (6)$$

where σ_N^2 is the offspring variance, so that t units of continuous time correspond to t/c_N discrete generations. Using (4), we see that in the Eldon–Wakeley model, when $0 < \gamma < 2$, time must be scaled proportional to $N^{-\gamma}$ and, when $\gamma \geq 2$, proportional to N^{-2} , which coincides with the usual Moran scaling. In the *Appendix*, we give an independent derivation of these constraints on c_N .

The choice of time scaling then dictates the appropriate scaling for the selection coefficient s , to produce a nontrivial balance between selection and drift. In the *Appendix*, we show that this scaling must satisfy

$$\begin{aligned} s &= O(N^{1-\gamma}), & \gamma \leq 2 \\ s &= O(N^{-1}), & \gamma > 2. \end{aligned} \quad (7)$$

Note that this scaling for the selection coefficient in the Eldon–Wakeley model coincides with the mutation scaling described in the corresponding continuum coalescent (Eldon and Wakeley 2006).

The limiting generator

In the *Appendix* we derive the continuum limit of the Eldon–Wakeley process. Such limits are characterized by an operator G , the infinitesimal generator for the continuum process, and an associated Kolmogorov backward equation, analogous to the diffusion equation of Kimura. As discussed above, we scale time by $c_N = N^{-\gamma}$ for $\gamma < 2$ and by $c_N = N^{-2}$ for $\gamma \geq 2$. In accordance with (7), we define the corresponding scaled selection coefficients as $\beta = \lim_{N \rightarrow \infty} sN$ for $\gamma \geq 2$ and $\beta = \lim_{N \rightarrow \infty} sN^{\gamma-1}$ when $\gamma < 2$. Let $\tilde{X}_t^{(N)} = (1/N)X_{[t/c_N]}$ be the rescaled process and

$u_N(x, t) = \mathbb{E}[\tilde{X}_t^{(N)} | \tilde{X}_0 = x]$ be its expected value at time t . Then $\tilde{X}_t^{(N)}$ converges to a limiting process \tilde{X}_t , and $u_N(x, t)$ converges to $u(x, t) = \mathbb{E}[\tilde{X}_t | \tilde{X}_0 = x]$, which satisfies the backward equation

$$\frac{\partial u(x, t)}{\partial t} = \beta x(1-x) \frac{\partial u(x, t)}{\partial x} + G_D u(x, t), \quad \text{with } u(x, 0) = x, \quad (8)$$

where

$$G_D u(x, t) = \begin{cases} x(1-x) \frac{\partial^2 u(x, t)}{\partial x^2}, & \gamma > 2, \\ x(1-x) \frac{\partial^2 u(x, t)}{\partial x^2} + xu(x + (1-x)\lambda, t) - u(x, t) \\ \quad + (1-x)u(x - \lambda x, t), & \gamma = 2, \\ xu(x + (1-x)\lambda, t) - u(x, t) + (1-x)u(x - \lambda x, t), & \gamma < 2. \end{cases} \quad (9)$$

As with the original discrete process, the limiting continuum process splits into three regimes, depending on the jump frequency γ . In all three regimes the operator on the right-hand side of (8) decomposes into two terms: a portion $\beta x(1-x)(\partial/\partial x)$ that describes selection and a portion G_D that describes genetic drift. Note that the term describing selection is the standard first-order advection term in Kimura's diffusion equation. The portion describing genetic drift, however, may differ from the standard form of Wright–Fisher drift.

We first study the behavior of the Eldon–Wakeley model without selection, meaning the continuum process with G_D alone.

For $\gamma > 2$, the discrete Eldon–Wakeley model is well approximated by the standard Kimura diffusion, as $N \rightarrow \infty$. In other words, when large family sizes are sufficiently rare, then the continuum limit corresponds to the standard diffusion used in population genetics. In this regime, the Eldon–Wakeley process behaves much like the standard Moran model.

When $\gamma < 2$, the generator G_D corresponds to a neutral two-type model known as the Λ -process (Pitman 1999; Möhle 2001; Birkner and Blath 2009) with Λ a Dirac measure at the point λ . In this process, allele frequencies exhibit a “jump and hold” behavior: jumps that increase the allele frequency occur with probability x and have size $(1-x)\lambda$, and jumps that reduce the allele frequency occur with probability $(1-x)$ and have size λx . In between these jumps the allele frequency remains fixed, for an exponentially distributed amount of time whose mean does not depend on the state x .

At the critical point $\gamma = 2$, the continuum limit is a mixture of the Λ -process and Kimura's diffusion: jumps still occur as in the case $\gamma < 2$, but the allele frequency follows classical genetic drift between the jump events.

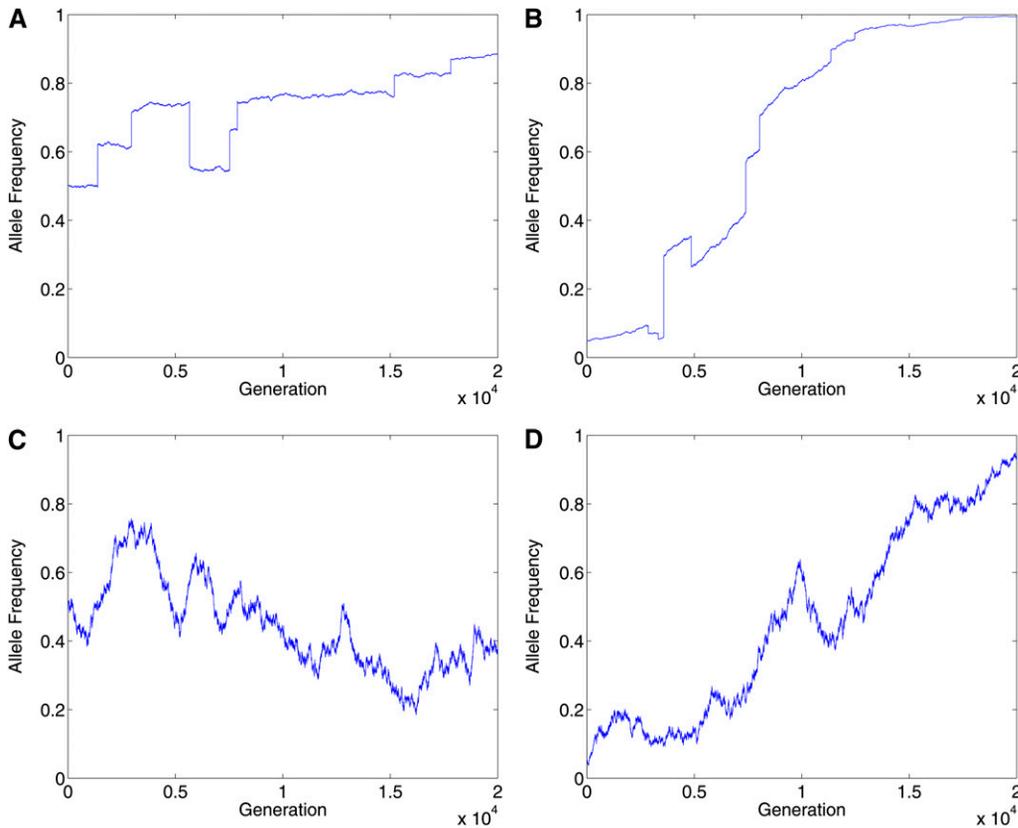


Figure 2 Simulated sample paths for the Eldon-Wakeley Process (panels A and B), and the variance-equivalent Moran processes (panels C and D), with and without selection. Parameters are $N = 2000$, $\gamma = 1$, $\lambda = 0.25$. (A and C) $\beta = 0$; (B and D) $\beta = 0.5$. Note the large difference in fixation probability between the Eldon-Wakeley and Moran models with selection: fixation probability = 1 in panel B, and fixation probability = 0.55 in panel D.

The splitting into three regimes that we observe in the forward-time process is mirrored by an analogous trichotomy in the associated coalescent theory of these processes, as described in Eldon and Wakeley (2006). For $\gamma > 2$, the retrospective genealogy is the classical Kingman coalescent, whereas $\gamma < 2$ is described by a Λ -coalescent with Dirac measure at λ . The critical case $\gamma = 2$ again mixes the two processes. This correspondence between the forward-time and backward-time limits is of course not accidental, but rather reflects the duality theory of Möhle (2001).

An intuitive explanation of forward-time dynamics

There is an intuitive explanation for the trichotomy of regimes discussed above. Recall that a process moving only by Moran steps (*i.e.*, when the offspring distribution P_U is concentrated at 2) requires at least $O(N^2)$ time steps before the allele frequency changes a fixed fraction of the population size N . More generally, in the discrete Eldon-Wakeley model (1), jump events occur every $O(N^\gamma)$ steps, and Moran steps occur in between the jump events. As a result, when $\gamma < 2$, the jumps occur too rapidly for the intervening Moran steps to change the allele frequency substantially—so that, in the continuum limit, allele frequencies are constant in between the jumps. In the case $\gamma > 2$, by contrast, strictly more than $O(N^2)$ Moran steps occur between jump events, and the process is likely to have fixed at 0 or 1 during the intervening time as a result of the Moran steps. Thus the process with $\gamma > 2$ is completely governed, in the limit, by the Moran process. Finally, at $\gamma = 2$ a balance is achieved:

jumps occur every $O(N^2)$ steps, but this time period is of precisely the correct order for the intervening Moran steps to move allele frequencies some fraction of the population size. Thus allele frequencies follow Moran-type drift in between jumps, when $\gamma = 2$.

The preceding discussion applies in the absence of selection (*i.e.*, $\beta = 0$). Adding selection does not change the frequency or the size of jumps. However, when selection is present and $\gamma < 2$, then the periods of constant allele frequency between jumps are replaced with deterministic logistic growth. Similarly, when selection is present and $\gamma = 2$, the periods of Kimuran drift between jumps are replaced by Kimuran drift with directional selection. Sample paths illustrating these behaviors are shown in Figure 2, along with comparisons to the Moran process.

Selection Amplification in the Eldon-Wakeley Process

The forward-time continuum limit (8) allows us to study the Eldon-Wakeley model in the presence of selection, which was not possible from a continuum limit of the (neutral) genealogies alone (Eldon and Wakeley 2006). In this section we focus on the case when large family sizes occur sufficiently often that the Eldon-Wakeley model diverges from the standard Moran model; *i.e.*, $\gamma < 2$. We show that such populations exhibit a remarkable behavior: an advantageous allele has markedly higher probability of fixation relative to the classical Moran process with the same selection pressure and the same variance effective population size.

In the following, we restrict ourselves to a discussion of the case of positive selection; the techniques presented can be easily adapted to investigate the effects of deleterious mutations. As with the standard Moran model, the Eldon–Wakeley process X_t will eventually absorb at either 0 or 1, and the fixation probability, as a function of initial allele frequency x , can be expressed as $\pi(x) = \mathbb{P}(X_\infty = 1 \mid X_0 = x)$. From the general theory of Markov processes, this quantity is the unique stationary solution of the backward equation (8),

$$\beta x(1-x)\pi'(x) + G_D\pi(x) = 0, \quad (10)$$

with boundary conditions $\pi(0) = 0$, $\pi(1) = 1$. Equation 10 is simply the analog of the diffusion equation Kimura used to compute the fixation probability of the Moran model (Kimura 1955). However, since G_D is now a complicated operator, (10) does not possess a simple closed-form solution. Nonetheless, without explicitly solving (10), we can still use it to study properties of the fixation probability π . To begin, we can derive a lower bound on the fixation probability $\pi(x)$ for a mutant initially at frequency x (see *Appendix*), which takes the form, for positive selective values $\beta \geq 0$,

$$\pi(x) \geq x^{p_r}, \quad (11)$$

where $p = p_r$ is the positive root of the equation

$$S(p) = \beta p - 1 + (1-\lambda)^p = 0. \quad (12)$$

The bound (11) provides significant information about the fixation probability of an adaptive allele. For fixed λ , the root p_r is a decreasing function of selection pressure β , and so our bound on the fixation probability increases with the strength of selection, as expected. When $\beta > \lambda$, then $p_r < 1$ and π has infinite derivative at $x = 0$, whereas the analogous fixation curve in the standard Moran model has finite derivative at zero, for all fitness values. In other words, for sufficiently strong selection ($\beta > \lambda$) the fixation probability of a rare mutant is much higher under the Eldon–Wakeley model than under the Moran model with the same selection pressure and the same variance effective population size. Thus, the Eldon–Wakeley process amplifies the impact of selection, relative to the standard Moran model, even though the two processes have the same effective population size.

The extent of selection amplification can be quite dramatic. In fact, the fixation of the advantageous allele can even be ensured, in some regimes. For the Eldon–Wakeley model with $\gamma < 2$ there exists a *finite* value of the selection pressure β , which we call β^* , for which p_r decreases all the way to zero, and hence $\pi(x) = 1$ for all $x > 0$. This value occurs at precisely

$$\beta^* = -\log(1-\lambda) \quad (13)$$

(see *Appendix*). In other words, alleles introduced at any positive fraction of the population size will *assuredly* fix

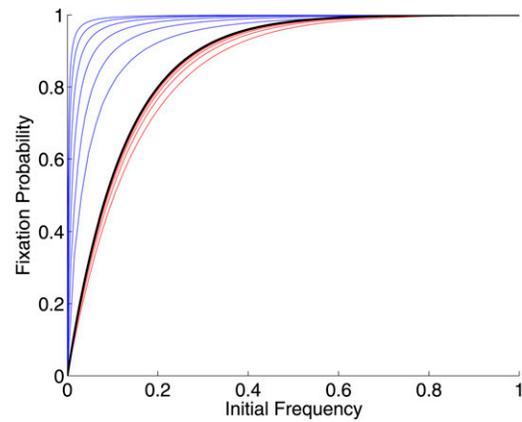


Figure 3 Fixation probability of an advantageous mutant as a function of its initial frequency, for various population sizes. Blue curves, bottom to top: Eldon–Wakeley model ($\gamma = 1$, $\lambda = 0.5$) with $N = 64, 128, 256, 512, 1024, 2048$. Red curves: variance-equivalent Moran models at the same population sizes. Black curve: Kimura’s classical expression for the fixation probability, derived by the diffusion approximation to the Moran model. Scaled selection pressure is fixed at $\beta = 0.5$.

when the selection pressure exceeds the finite threshold, β^* , despite the presence of genetic drift. Figure 2B shows a typical sample path illustrating this behavior: there, $\beta = 0.5$ exceeds the threshold $\beta^* = -\log(0.75)$. Moreover, we demonstrate in the *Appendix* that the value β^* is indeed critical: if $\beta < \beta^*$, then fixation is not ensured for some starting frequency x . Thus β^* gives the precise value of selection pressure under which assured fixation occurs for all positive initial frequencies x , in the Eldon–Wakeley process with $\gamma < 2$. By contrast, assured fixation never occurs in the classical Moran model or in the Eldon–Wakeley model with $\gamma > 2$.

Assured fixation of an advantageous mutant is a property of the continuum limit; in the discrete model, assured fixation is achieved only as the population size increases to infinity. Figure 3 displays an example of this phenomenon, in which $\pi(x)$ is plotted as a function of initial frequency x with fixed selection pressure $\beta > \beta^*$, for both the Eldon–Wakeley model and a variance-equivalent Moran model (in the sense that the reference Moran model is chosen to have the exact same variance effective population size as the corresponding Eldon–Wakeley model, for each population size N). As the population size gets large, the Eldon–Wakeley fixation curves approach $\pi(x) \equiv 1$, whereas the equivalent Moran fixation probability approaches Kimura’s standard equation: $\pi(x) = (1 - \exp(\beta x))/(1 - \exp(\beta))$.

Fixation of new mutants

The analysis above focused on the fate of beneficial mutants initiated at a fixed fraction of the population size. We are naturally also interested in the fate of a new mutant, initially at frequency $1/N$. While the continuum theory cannot be used formally to obtain results on the fixation probability of new mutants, the extremely large derivative of $\pi(x)$ near $x = 0$ suggests that selection amplification also occurs when

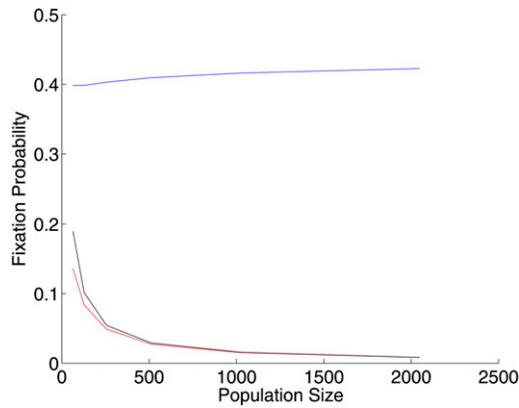


Figure 4 Fixation probabilities of a new mutant as a function of population size. Blue: Eldon–Wakeley model ($\gamma = 1.1$, $\lambda = 0.5$). Red: variance-equivalent Moran model. Black: Kimura’s classical expression for the fixation probability, derived by the diffusion approximation to the Moran model. Selection pressure is fixed at $s = 1$.

the Eldon–Wakeley process is initiated at a single mutant. We have investigated this behavior numerically.

Figure 4 shows the fixation probability of a new mutant with fixed selective advantage s , as a function of population size, for both the Eldon–Wakeley model and the variance-equivalent Moran model. Note that since the offspring variance $\sigma_N^2 \rightarrow \infty$, the fixation probability of the Moran sequence goes to zero. The Eldon–Wakeley process, on the other hand, exhibits an asymptotically positive fixation probability, and it displays much higher fixation probability than the Moran model for all population sizes. As Figure 4 demonstrates, a novel advantageous mutant that would almost certainly be lost in a large population, under the Moran model, has a substantial probability of fixation under the Eldon–Wakeley model—despite the fact that the two models have the exact same effective population size.

Figure 5 shows the fixation probability of new mutants as a function of their selective advantage s , again for an Eldon–Wakeley model and for a variance-equivalent Moran process. For all values of s , the Eldon–Wakeley model amplifies the efficiency of selection relative to the Moran model and especially so for larger values of s . By contrast, the Moran fixation probability is relatively small and approximately linear with s in the regime explored.

Figure 6 shows the fixation probability of a new mutant as a function of the model parameters λ and γ , with the population size and selection coefficient fixed. In Figure 6A, when $\gamma < 2$, the fixation probability is a decreasing function of the jump size λ . Here, offspring variance increases with λ , which decreases the fixation probability. A similar situation occurs in Figure 6B, where the jump frequency γ is varied, and all other parameters are fixed. For $\gamma > 2$, the process is similar to the classic Moran model and it has fixation probability $\approx s$. As the rate of bottleneck events increases (smaller γ), the offspring variance decreases and so the fixation probability declines.

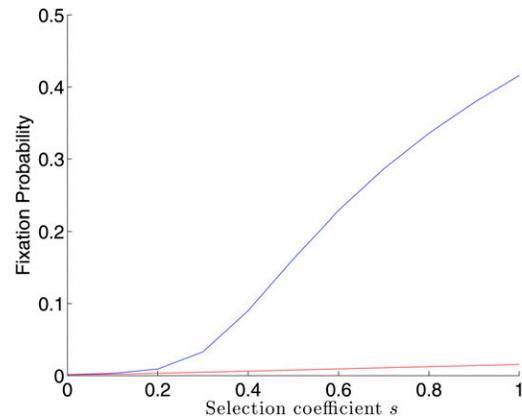


Figure 5 Fixation probability of a new mutant as a function of the selection coefficient. Blue: Eldon–Wakeley model ($\gamma = 1.1$, $\lambda = 0.5$). Red: variance-equivalent Moran model. Population size $N = 1000$.

A fitness estimate for Pacific oysters

Eldon and Wakeley (2006) applied their model of large family sizes to sampled genetic data from Pacific oysters (Hedgecock 1994). They estimated the model parameters $\theta = 2\mu N^{\gamma-1}$ and λ as $\hat{\theta} = 0.0308$ and $\hat{\lambda} = 0.0775$, suggesting that occasionally the offspring of a single oyster accounts for $\sim 8\%$ of the oyster population in a given generation, which is presumably possible for pelagic organisms. Given these model estimates, we can investigate the properties of selection amplification in such an oyster population.

Equation 13 can be used to estimate a critical selection coefficient that ensures a high probability that an advantageous mutant will fix, assuming $\gamma < 2$. Substituting $\beta = sN^{\gamma-1}$, we find that the critical selective value is

$$s^* = \frac{-\log(1 - \lambda)}{N^{\gamma-1}}. \quad (14)$$

Given the estimates of Eldon and Wakeley (2006) for Pacific oysters, and assuming a typical per-locus mutation rate of $\mu = 10^{-5}$, we find that $N^{\gamma-1} = 1.54 \times 10^3$, and so $s^* = 5.24 \times 10^{-5}$. This critical value of the selection coefficient is quite small, in the sense that such a mutant would have very little chance of fixing under the standard Moran model. However, under an Eldon–Wakeley model for $\gamma < 2$, the same mutant has a reasonable probability of fixation, provided the population size is large enough. Thus, the offspring mechanism of oysters not only changes the expected lengths of gene genealogies (Eldon and Wakeley 2006), but also radically alters the behavior of advantageous mutants, compared to the standard Moran model.

Discussion

Eldon and Wakeley observed that a skew in an organism’s offspring distribution can drastically change the predictions for neutral genetic diversity in a population (Eldon and Wakeley 2006). Their analysis was based on the genealogical properties of a sample from the population. Here we

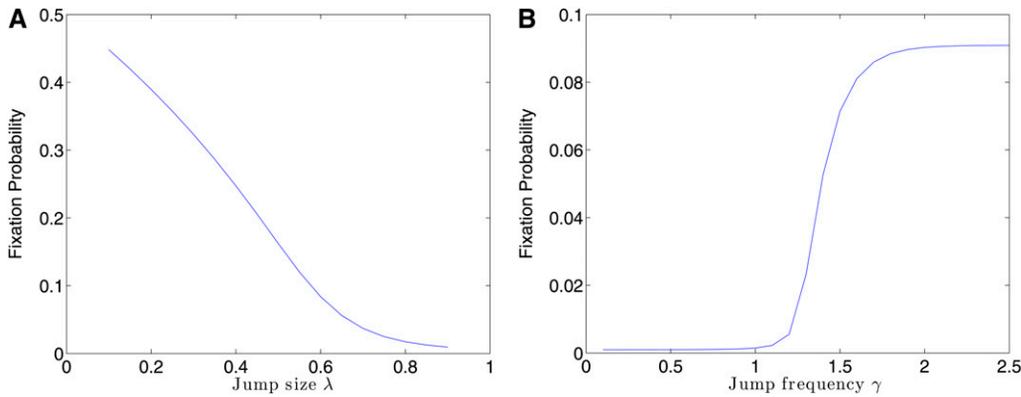


Figure 6 Fixation probability of a new mutant as a function of model parameters. (A) $\gamma = 1$, $N = 1000$, $s = 1$. (B) $\lambda = 0.5$, $N = 1000$, $s = 0.1$.

have analyzed the forward-time properties of the same model introduced by Eldon and Wakeley. To do so, we developed a continuum approximation, which generalized the diffusion approximation of Kimura. Our analysis of the Eldon–Wakeley process in forward time has uncovered three distinct regimes of behavior, analogous to the distinct coalescent limits discovered by Eldon and Wakeley.

The virtue of a forward-time analysis is that it allows us to study the effects of natural selection, in a population with reproductive skew. We have found that the skewed offspring distribution greatly amplifies the effects of selection. In particular, when the frequency of large family sizes is sufficiently large ($\gamma < 2$), the fixation probability of an advantageous allele is far greater in the Eldon–Wakeley model than it is in a corresponding Moran or Wright–Fisher model with the same selection coefficient and variance-effective population size. Moreover, in this regime the fixation of a beneficial allele can even be ensured in the (continuum limit of the) Eldon–Wakeley model—a phenomenon that never occurs in the standard population-genetic model or its diffusion limit.

Our results on the Eldon–Wakeley model can be seen as a special case of general studies initiated in Der *et al.* (2011), where it was shown that, under minimal assumptions on the structure of drift, the Moran model possesses the *strongest* type of genetic drift and minimizes the fixation probability of advantageous mutants. The intuition developed there, in terms of the strength of higher-order moments of the offspring distribution and their tendency to lengthen the holding time of states, also holds for the Eldon–Wakeley model. As Equation 3 shows, the Eldon–Wakeley process can be viewed a mixture between two processes: (1) a Moran component, which is effectively a low-variance (high N_e) process, and (2) a rarer jump component, which is a high-variance (low N_e) process. These two components are mixed in such a way that, as an aggregate, they possess the same variance as the standard Moran process. Nonetheless, selection amplification occurs because new mutants can rise to an appreciable frequency through the low-variance Moran component, which becomes logistic growth in the limit of large populations. The high-variance bottleneck events disrupt this trajectory and reduce the fixation probability, but

not sufficiently often enough for a commensurate decrease in the fixation probability of the advantageous mutant.

Accuracy of the continuum limit

Our analysis has relied in parts on a continuum approximation of the discrete process. For a given model, these approximations are guaranteed to be accurate for large population size, N . The accuracy of this representation, however, is not uniform over the space of model parameters λ and γ .

Some idea of the accuracy of the limit can be obtained by recalling the offspring variance (Equation 4) of the process:

$$\sigma_N^2 = 2N^{-1} + \lambda^2 N^{-\gamma+1}, \quad N \rightarrow \infty. \quad (15)$$

When the term $2N^{-1}$, corresponding to the Moran component, is small relative to the jump term $\lambda^2 N^{-\gamma+1}$, we can expect the $\gamma < 2$ continuum limit to be accurate, and vice versa for $\gamma > 2$. In other words, Kimura’s diffusion is a good approximation as γ becomes large, and the λ -process is a good approximation as γ becomes small. Near $\gamma = 2$, however, very large values of the population size are required for the continuum limit on either side to be an accurate representation of the discrete process. Large values of population size are also required when $\gamma < 2$ and the jump-size λ is small or, conversely, when $\gamma > 2$ and the jump-size λ is large.

Regardless of the accuracy of the continuum approximation, the Eldon–Wakeley process exhibits marked selection amplification and a qualitatively different behavior from the Moran model, even at modest population sizes. These results are demonstrated in Figures 2–5, which were produced by simulations of the discrete Eldon–Wakeley Markov chains, as opposed to using their continuum approximation.

Alternative mechanisms of selection

We introduced fitness differentials into the Eldon–Wakeley model, using a form of viability selection. This involved biasing the allele frequencies logistically before applying the action of genetic drift and has the interpretation of introducing a preparatory stage in which not all individuals survive to the moment of reproduction. This form of selection has

the virtue that it changes the continuum limit simply by adding the usual advective term $\beta x(1-x)(d/dx)$ to the limiting generator.

Another way to incorporate fitness differences is fertility selection, where the sampling of gametes is biased itself. In this model, individuals of type 1 would have $(1 + s)$ times greater likelihood of being selected to reproduce than type 2 individuals, per capita. It can be shown that in this version of selection, the continuum limit still involves adding only the selective term $\beta x(1-x)(d/dx)$, in the regime $1 < \gamma < 2$. For $\gamma \leq 1$, however, fertility selection alters the neutral drift component of the continuum limit as well. It can be demonstrated, however, that all of our results concerning amplification still hold: indeed the selection amplification is as strong or even *stronger* under fertility selection, compared to viability selection.

The absorption time and coming down from infinity

There is an important connection between the absorption time of the forward-time process and the concept of “coming down from infinity” (Pitman 1999; Schweinsberg 2000) in the associated coalescent. A coalescent initialized with a partition possessing an infinite number of blocks comes down from infinity when there are only finitely many blocks in the coalescent partition for all positive times. Let τ be the time to absorption for the continuum Eldon–Wakeley model, for $\gamma < 2$, started at some interior state. When $\lambda < 1$, the process cannot jump from an interior state to the boundary, but rather only to another interior state. Since the time between jumps is independent and identically exponentially distributed, and since the process is driven by the strictly interior logistic curve between jumps, it follows that the time to absorption is almost surely infinite. Note that this is in contrast to the Moran model, whose diffusion limit has a finite absorption time. This pathology occurs only in the continuum model, however; indeed it can be shown that the discrete Eldon–Wakeley model has finite absorption time with expectation at most $O(N^\gamma \log N^\gamma)$ steps (Der *et al.* 2011).

Likewise, it can be demonstrated that the Eldon–Wakeley model for $\gamma < 2$ corresponds to a generalized coalescent that “does not come down from infinity,” so that the time to coalescence for the entire population tends to be longer than that of a Moran process (Birkner and Blath 2009). This lengthening of fixation and absorption times does not appear to be merely accidental, and we speculate that for a general population process, there is a close relationship between the characteristics of (1) selection amplification, (2) forward-time absorption, and (3) the property of coming down from infinity.

Model generalizations

Our analysis has focused on the case where the offspring distribution P_U is a mixture of two components: a Moran component and a bottleneck component where the reproducing individual contributes a fixed fraction λ of the population size to the succeeding generation (Eldon and

Wakeley 2006). More generally, the offspring distribution P_U can be conceived as a mixture of bottleneck components over a range of jump sizes λ , where λ is drawn from some probability measure Λ on $[0, 1]$. Under this scenario, the continuum limit for the neutral Eldon–Wakeley processes takes the full form of a Λ -process:

$$G_{DU}(x) = \sigma^2 \int_0^1 \frac{xu(x + (1-x)\lambda) - u(x) + (1-x)u(x - \lambda x)}{\lambda^2} d\Lambda(\lambda). \quad (16)$$

Each of the continuum limits described above is a special case of this general formula: when $\gamma > 2$, Λ is a Dirac function concentrated at 0 and $\sigma^2 = 2$, giving the standard limit for the Moran model; when $\gamma < 2$, Λ is a Dirac function concentrated at the point λ with $\sigma^2 = \lambda^2$; and finally when $\gamma = 2$, Λ is mixture of atoms at 0 and at λ .

In general, the full Λ -process also exhibits selection amplification relative to the Kimura diffusion model. The strength of this amplification can be characterized by the number β^* we have defined above—the selective value required for assured fixation of a mutation initiated at positive frequency. One can demonstrate that the following formula holds for β^* ,

$$\beta^* = -\sigma^2 \int_0^1 \frac{\log(1-\lambda)}{\lambda^2} d\Lambda(\lambda), \quad (17)$$

which reduces to our special case (Equation 13) when Λ is a Dirac measure at λ . Note that when Λ has any mass at zero, then $\beta^* = \infty$ —that is, any process for which the continuum limit possesses a nontrivial Moran component cannot exhibit assured fixation for finite values of the selective pressure β .

Acknowledgments

The authors thank two anonymous reviewers for their feedback, which improved the manuscript. J.B.P. acknowledges funding from the Burroughs Wellcome Fund, the David and Lucile Packard Foundation, the James S. McDonnell Foundation, the Alfred P. Sloan Foundation, and grant D12AP00025 from the U.S. Department of the Interior and Defense Advanced Research Projects Agency.

Literature Cited

- Birkner, M., and J. Blath, 2009 Measure-valued diffusions, general coalescents and population genetic inference. *Trends Stoch. Anal.* 351: 329–363.
- Bustamante, C. D., J. Wakeley, S. Sawyer, and D. L. Hartl, 2001 Directional selection and the site-frequency spectrum. *Genetics* 159: 1779–1788.
- Cannings, C., 1974 The latent roots of certain Markov chains arising in genetics: a new approach, I. Haploid Models. *Adv. Appl. Prob.* 6(2): 260–290.
- Der, R., 2010 A theory of generalised population processes. Ph.D. Thesis, ProQuest, Philadelphia.

- Der, R., C. L. Epstein, and J. B. P. Plotkin, 2011 Generalized population processes and the nature of genetic drift. *Theor. Popul. Biol.* 80: 80–99.
- Desai, M. M., and J. B. Plotkin, 2008 The polymorphism frequency spectrum of finitely many sites under selection. *Genetics* 180: 2175–2191.
- Donnelly, P., and T. Kurtz, 1999 Particle representations for measure-valued population models. *Ann. Probab.* 27: 166–205.
- Eldon, B., and J. Wakeley, 2006 Coalescent processes when the distribution of offspring number among individuals is highly skewed. *Genetics* 172: 2621–2633.
- Eldon, B., and J. Wakeley, 2008 Linkage disequilibrium under skewed offspring distribution among individuals in a population. *Genetics* 178: 1517–1532.
- Eldon, B., and J. Wakeley, 2009 Coalescence times and F_{ST} under a skewed offspring distribution among individuals in a population. *Genetics* 181: 615–629.
- Ewens, W. J., 1972 The sampling theory of selectively neutral alleles. *Theor. Popul. Biol.* 3: 87–112.
- Ewens, W. J., 2004 *Mathematical Population Genetics*, Ed. 2. Springer-Verlag, New York.
- Fay, J. C., and C.-I. Wu, 2000 Hitchhiking under positive Darwinian selection. *Genetics* 155: 1405–1413.
- Fisher, R. A., 1958 *The Genetical Theory of Natural Selection*. Dover, New York.
- Fu, Y. X., and W. H. Li, 1992 Statistical tests of neutrality of mutations. *Genetics* 133: 693–709.
- Hedgecock, D., 1994 Does variance in reproductive success limit effective population size of marine organisms? pp. 122–134 in *Genetics and Evolution of Aquatic Organisms*, edited by A. R. Beaumont. Chapman and Hall, London.
- Hudson, R. R., M. Kreitman, and M. Aguadé, 1987 A test of neutral molecular evolution based on nucleotide data. *Genetics* 116: 153–159.
- Karlin, S., and J. McGregor, 1964 Direct product branching processes and related Markov chains. *Proc. Natl. Acad. Sci. USA* 51: 598–602.
- Kimura, M., 1955 Solution of a process of random genetic drift with a continuous model. *Proc. Natl. Acad. Sci. USA* 41: 144–150.
- Kingman, J. F. C., 1982 The coalescent. *Stoch. Proc. Appl.* 13: 235–248.
- Lessard, S., 2007 An exact sampling formula for the Wright-Fisher model and a solution to a conjecture about the finite-island model. *Genetics* 177: 1249–1254.
- McDonald, J., and M. Kreitman, 1991 Adaptive protein evolution at the Adh locus in *Drosophila*. *Nature* 351: 652–654.
- Möhle, M., 2001 Forward and backward diffusion approximations for haploid exchangeable population models. *Stoch. Proc. Appl.* 95: 133–149.
- Moran, P. A. P., 1958 Random processes in genetics. *Proc. Camb. Philos. Soc.* 54: 60–71.
- Pitman, J., 1999 Coalescents with multiple collisions. *Ann. Probab.* 27: 1870–1902.
- Sagitov, S., 1999 The general coalescent with asynchronous mergers of ancestral lines. *J. Appl. Probab.* 36: 1116–1125.
- Sargsyan, O., and J. Wakeley, 2008 A coalescent process with simultaneous multiple merger for approximating the gene genealogies of many marine organisms. *Theor. Popul. Biol.* 74: 104–114.
- Sawyer, S. A., and D. L. Hartl, 1992 Population genetics of polymorphism and divergence. *Genetics* 132: 1161–1176.
- Schweinsberg, J., 2000 A necessary and sufficient condition for the Λ -coalescent to come down from infinity. *Electron. Comm. Probab.* 5: 1–11.
- Schweinsberg, J., 2003 Coalescent processes obtained from supercritical Galton-Watson processes. *Stoch. Proc. Appl.* 106: 107–139.
- Tajima, F., 1989 Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123: 585–595.
- Wright, S., 1931 Evolution in Mendelian populations. *Genetics* 16: 97–159.
- Yang, Z., and J. P. Bielawski, 2000 Statistical methods for detecting molecular adaptation. *Trends Ecol. Evol.* 15: 496–503.

Communicating editor: Y. S. Song

Appendix

In the sections that follow we give an outline of the arguments involved, without specifying all the mathematical details.

Description of the Jump Matrix

The exact specification of the forward-time jump matrix \mathbf{J} appearing in (3) is as follows. Let Y be a random variable with distribution corresponding to the i th row of \mathbf{J} . Conditioned on the event E_1 that a type 1 individual is sampled for extreme reproduction,

$$Y|E_1 = \lambda N + H_1, \quad (\text{A1})$$

where H_1 is a random variable with the hypergeometric distribution of sample size $N - 1$, initial configuration $i - 1$, and $N - \lambda N$ draws. Conditioned on the event E_2 that a type 2 individual is sampled for extreme reproduction,

$$Y|E_2 = H_2, \quad (\text{A2})$$

where H_2 is a hypergeometric random variable with sample size $N - 1$, initial configuration i , and $N - \lambda N$ draws. The full variable Y is then a mixture of $Y | E_1$ and $Y | E_2$, with respective weights i/N and $1 - i/N$.

Derivation of the Limiting Generator

Let $\mathbf{P}_N = \mathbf{S}_N \mathbf{Q}_N$ be the transition matrix of the Eldon–Wakeley process with selection, where we have made explicit the dependence on population size N . We choose to study a continuum limit whose state space corresponds to allele frequencies lying in $[0, 1]$; *i.e.*, we scale the discrete state space by the population size N . With this choice, the continuum representation for the sequence \mathbf{P}_N under time-scaling c_N involves finding an operator G acting on smooth functions u for which

$$\lim_{N \rightarrow \infty} c_N^{-1} (\mathbf{P}_N - \mathbf{I}) u_N = Gu, \quad (\text{A3})$$

where u_N are samples from the smooth test function u on the unit interval, so that

$$u_N = (u(0), u(1/N), u(2/N), \dots, u(1)).$$

More precisely, it is necessary to prove that for any smooth function u on $[0, 1]$, one has

$$\lim_{N \rightarrow \infty} \sup_{i \in \mathcal{I}} |c_N^{-1} ((\mathbf{P}_N - \mathbf{I}) u_N)(i) - Gu(i/N)| = 0, \quad (\text{A4})$$

where \mathcal{I} is the index set $\{0, \dots, N\}$. From here on we use the expression (A3) as a short-hand for (A4).

We derive the conditions on c_N that result in a nontrivial continuum limit. Let us first consider the process in the absence of selection, so that $\mathbf{P}_N = \mathbf{Q}_N$. The neutral drift component \mathbf{Q}_N is a convex combination of the Moran process \mathbf{M}_N and a jump process \mathbf{J}_N : $\mathbf{Q}_N = (1 - N^{-\gamma}) \mathbf{M}_N + N^{-\gamma} \mathbf{J}_N$. Thus

$$\lim_N c_N^{-1} (\mathbf{P}_N - \mathbf{I}) u_N = \lim_N c_N^{-1} ((1 - N^{-\gamma}) (\mathbf{M}_N - \mathbf{I}) u_N + N^{-\gamma} (\mathbf{J}_N - \mathbf{I}) u_N). \quad (\text{A5})$$

Case 1: $\gamma > 2$: From the classical theory of the Moran model, we know that if the state space is scaled by N , then $\lim_N c_N^{-1} (\mathbf{M}_N - \mathbf{I}) u_N$ is noninfinite only if $c_N^{-1} = O(N^2)$. Considering the above equation, if c_N^{-1} grows strictly slower than N^2 , then the first term goes to zero, as must the second term, since $\|(\mathbf{J}_N - \mathbf{I}) u_N\| \leq 2\|u\| < \infty$, and the second term will be $o(1)$ as $\gamma > 2$. Thus, the only scaling that leads to a nontrivial limit must be $c_N^{-1} = N^2$. In this case, the second term still vanishes, and so

$$\lim_{N \rightarrow \infty} c_N^{-1} (\mathbf{P}_N - \mathbf{I}) u_N = \lim_{N \rightarrow \infty} N^2 (\mathbf{M}_N - \mathbf{I}) u_N = x(1-x)u''(x), \quad (\text{A6})$$

where the second equality follows by the classical diffusion approximation for the Moran model.

Case 2: $\gamma < 2$: Let us analyze the jump term. Fix a number $0 < x < 1$ and consider the row $i = [xN]$. From the earlier description of the jump matrix, we can decompose

$$\mathbf{J}_N u_N(i) = \frac{i}{N} \mathbb{E}u\left(\frac{Y_1}{N}\right) + \left(1 - \frac{i}{N}\right) \mathbb{E}u\left(\frac{Y_2}{N}\right), \quad (\text{A7})$$

where $Y_1 = Y | E_1$ and $Y_2 = Y | E_2$ are the hypergeometric distributions described earlier in the *Appendix*. Computing the mean and the variance of the hypergeometric distributions, one finds that $\text{var}(Y_1/N) = O(N^{-1})$, $\text{var}(Y_2/N) = O(N^{-1})$, and hence go to zero; thus

$$\lim_{N \rightarrow \infty} \mathbb{E}u\left(\frac{Y_1}{N}\right) = \lim_{N \rightarrow \infty} u\left(\mathbb{E}\frac{Y_1}{N}\right) = u(x + (1-x)\lambda) \quad (\text{A8})$$

$$\lim_{N \rightarrow \infty} \mathbb{E}u\left(\frac{Y_2}{N}\right) = \lim_{N \rightarrow \infty} u\left(\mathbb{E}\frac{Y_2}{N}\right) = u(x - \lambda x). \quad (\text{A9})$$

The above, in conjunction with (A7) proves that

$$\lim_{N \rightarrow \infty} (\mathbf{J}_N - \mathbf{I})u_N = xu(x + (1-x)\lambda) - u(x) + (1-x)u(x - \lambda x). \quad (\text{A10})$$

Now let us consider the appropriate scaling regime. From (A5) and the Moran theory, we still have the constraint $c_N^{-1} = O(N^2)$. But because $\gamma < 2$, and because of the result (A10), any scaling c_N^{-1} strictly larger than N^γ must make the second term of (A5) unbounded. On the other hand, using a Taylor expansion, one finds $\|(\mathbf{M}_N - \mathbf{I})u_N\| \leq C\|u''\|/N^2$, and so if we take c_N^{-1} to be a sequence strictly smaller than N^γ , then all terms must converge to zero. Thus, the only scaling leading to a nontrivial continuum limit is $c_N^{-1} = N^\gamma$. This choice leads to

$$\lim_{N \rightarrow \infty} c_N^{-1}(\mathbf{P}_N - \mathbf{I}) = \lim_{N \rightarrow \infty} (\mathbf{J}_N - \mathbf{I})u_N = xu(x + (1-x)\lambda) - u(x) + (1-x)u(x - \lambda x). \quad (\text{A11})$$

Case 3: $\gamma = 2$: The mixture continuum limit follows by combining cases 1 and 2. Here the only scaling leading to a nontrivial limit is $c_N^{-1} = N^2$, and we have

$$\lim_{N \rightarrow \infty} c_N^{-1}(\mathbf{P}_N - \mathbf{I})u_N = x(1-x)u''(x) + xu(x + (1-x)\lambda) - u(x) + (1-x)u(x - \lambda x). \quad (\text{A12})$$

To summarize the time scalings of the preceding cases, c_N must have the asymptotic behavior

$$c_N^{-1} = \begin{cases} N^2, & \gamma > 2 \\ N^\gamma, & 0 < \gamma \leq 2. \end{cases} \quad (\text{A13})$$

Adding selection

First we address the component of the process due to selection alone. We seek an operator G_S acting on smooth functions such that

$$\lim_{N \rightarrow \infty} c_N^{-1}(\mathbf{S}_N - \mathbf{I})u_N = G_S u. \quad (\text{A14})$$

Fix a frequency $0 < x < 1$. As discussed in the main text, \mathbf{S}_N is the process that biases allele frequencies according to fitness differences. For a population process for which each transition corresponds to a single reproductive event, it can be described as moving from state $i = [xN]$ to state j with probability 1, where j is given by

$$j = \frac{(1 + s/N)([xN]/N)}{(1 + s/N)([xN]/N) + (1 - [xN]/N)}. \quad (\text{A15})$$

We have

$$c_N^{-1}((\mathbf{S}_N - \mathbf{I})u_N)(i) = c_N^{-1}\left(u\left(\frac{(1 + s/N)([xN]/N)}{(1 + s/N)([xN]/N) + (1 - [xN]/N)}\right) - u\left(\frac{[xN]}{N}\right)\right). \quad (\text{A16})$$

Since for all $\gamma > 0$, c_N^{-1} tends to infinity, this implies, in conjunction with the above equation, that we must have $s/N \rightarrow 0$. Using a Taylor expansion in u ,

$$c_N^{-1}((\mathbf{S}_N - \mathbf{I})u_N)(i) = c_N^{-1} \frac{(s/N)([xN]/N)(1 - [xN]/N)}{(1 + s/N)[xN]/N + (1 - [xN]/N)} u' \left(\frac{[xN]}{N} \right) + c_N^{-1} O \left(\left(\frac{s}{N} \right)^2 \right). \quad (\text{A17})$$

The limit above is noninfinite only if $c_N^{-1}sN^{-1} = O(1)$, which can now in light of the preceding analysis on c_N be expressed as the selection scaling limits:

$$s = O(N^{-1}), \quad \gamma > 2 \quad (\text{A18})$$

$$s = O(N^{-\gamma+1}), \quad \gamma \leq 2. \quad (\text{A19})$$

With this selection scaling, and writing $sN^{-1}c_N^{-1} \rightarrow \beta$ for the limit, we have

$$\lim_N c_N^{-1}((\mathbf{S}_N - \mathbf{I})u_N)u_N = \beta x(1-x)u'(x) = G_S u(x). \quad (\text{A20})$$

Finally, we can address the product $\mathbf{P}_N = \mathbf{S}_N \mathbf{Q}_N$. From the algebraic identity

$$c_N^{-1}(\mathbf{S}_N \mathbf{Q}_N - \mathbf{I}) = c_N^{-1}(\mathbf{S}_N - \mathbf{I}) + c_N^{-1} \mathbf{S}_N (\mathbf{Q}_N - \mathbf{I}) \quad (\text{A21})$$

and the fact that $\mathbf{S}_N \rightarrow \mathbf{I}$, one sees that the total generator is just a sum of the parts due to selection and pure drift,

$$\lim_{N \rightarrow \infty} c_N^{-1}(\mathbf{S}_N \mathbf{Q}_N - \mathbf{I})u_N = \lim_N c_N^{-1}(\mathbf{S}_N - \mathbf{I})u_N + \lim_N c_N^{-1}(\mathbf{Q}_N - \mathbf{I})u_N \quad (\text{A22})$$

$$= G_S u + G_D u, \quad (\text{A23})$$

where G_S is given by (A20), and G_D is given by the operators in (A6), (A10), or their mixture (A12), depending on γ .

Selection Amplification

We consider here the continuum process when $\gamma < 2$, so that the generator for the Eldon–Wakeley process X_t has the form

$$Gu(x) = \beta x(1-x)u'(x) + xu(x + (1-x)\lambda) - u(x) + (1-x)u(x - \lambda x). \quad (\text{A24})$$

We study this operator under positive selection $\beta \geq 0$.

The fixation probability $\pi(x) = \mathbb{P}(X_\infty = 1 \mid X_0 = x)$ as a function of initial frequency x is a harmonic function relative to G and satisfies the equation $G\pi = 0$ with $\pi(0) = 0$, $\pi(1) = 1$. To examine its properties we make use of the following maximum principle.

Theorem 1

Suppose that a function u , continuous on $[0, 1]$, and continuously differentiable on $(0, 1)$, satisfies $Gu(x) \geq 0$ for $x \in (0, 1)$. Then $\max_{x \in [0, 1]} u(x) = \max\{u(0), u(1)\}$.

Proof: This result follows from general results on Markov processes, but we provide a direct verification here. Say $u(c) = \max_{x \in [0, 1]} u(x) = M$, and c is an interior point. We prove $u(1) = M$. For if $u(1) < M$, define the function $z(x) = x^2 - c^2$. A computation reveals that $Gz(x) = x(1-x)(\lambda^2 + 2\beta x)$, which is strictly positive on the open interval. Now look at $w(x) = u(x) + \epsilon z(x)$, for $\epsilon > 0$. We have the strict inequality $Gw(x) > 0$ throughout the open interval, and $w(0) < M$, $w(c) = M$, and $w(1) = u(1) + \epsilon z(1) < M$ whenever ϵ is sufficiently small. It follows now that w must have an interior global maximum. Let x_0 be the location of that interior maximum. But then

$$x_0 w(x_0 + (1-x_0)\lambda) - w(x_0) + (1-x_0)w(x_0 - \lambda x_0) \quad (\text{A25})$$

$$= x_0(w(x_0 + (1-x_0)\lambda) - w(x_0)) + (1-x_0)(w(x_0 - \lambda x_0) - w(x_0)) \quad (\text{A26})$$

$$\leq 0 \quad (\text{A27})$$

which contradicts $Gw(x) > 0$ on the open interval. ■

Now as a corollary we have the following comparison principle.

Corollary 1

Let π be the fixation probability as above, and assume that it is continuous on $[0, 1]$ and continuously differentiable on $(0, 1)$. Suppose that u also has the same regularity, $Gu \geq 0$, and $u(0) = 0$, $u(1) = 1$. Then $u(x) \leq \pi(x)$. Similarly if $Gu \leq 0$ with the same boundary conditions, then $u(x) \geq \pi(x)$.

Proof: Since $G\pi = 0$, $G(u - \pi) \geq 0$, and $u - \pi$ has value zero at the endpoints. Applying the maximum principle above, we find that $u - \pi \leq 0$. If $Gu \leq 0$, look at $G(\pi - u)$ to get the result. ■

We note that while the proof above had to assume π satisfied a regularity condition, it is not in fact necessary for the conclusion to hold.

Corollary 1 is the result we use to establish lower and upper bounds on the fixation probability, by finding functions u super- and subharmonic relative to G , with the boundary conditions $u(0) = 0$ and $u(1) = 1$.

To do this, we decompose the operator G as

$$Gu(x) = G_0u(x) - x(G_0 + I)u(x) + xu(x + (1 - x)\lambda), \quad (\text{A28})$$

where $G_0u(x) = \beta xu'(x) - u(x) + u(x - \lambda x)$. Now observe that x^p is an eigenfunction of G_0 for every $p \geq 0$,

$$G_0x^p = S(p)x^p, \quad (\text{A29})$$

where

$$S(p) = \beta p - 1 + (1 - \lambda)^p. \quad (\text{A30})$$

The function S has generally two roots, one at $p = 0$ and the other a nontrivial root $p_r(\beta)$ that is positive for small β . As β increases, p_r decreases, and there is a critical point β^* , for which the nontrivial root passes through zero. This occurs at precisely $S'(0) = 0$, which can be solved to obtain $\beta^* = -\log(1 - \lambda)$.

For $\beta < \beta^*$, p_r is strictly positive. Now it is easy to show the following.

Theorem 2

If $u(x) = x^{p_r}$, then $Gu(x) \geq 0$, and so u is a lower bound on π .

Proof: Using the decomposition (A28), we find

$$Gu(x) = S(p_r)x^{p_r} - x(S(p_r)x^{p_r} + x^{p_r}) + x(x + (1 - x)\lambda)^{p_r} \quad (\text{A31})$$

$$= -x^{p_r+1} + x(x + (1 - x)\lambda)^{p_r} \geq 0. \quad (\text{A32})$$

That u is now a lower bound on π follows from Corollary 1. ■

Theorem 2 establishes the result (11) and also shows, since $p_r \rightarrow 0$ as $\beta \uparrow \beta^*$, that the selection parameter $\beta = \beta^*$ must produce assured fixation. To show that β^* is critical, i.e., that assured fixation does not occur when the selection pressure is strictly smaller than β^* , one can demonstrate the following.

Theorem 3

Let $\beta < \beta^*$. Then constants $\rho > 0$ and $C > 0$ exist such that $u(x) = Cx^\rho + (1 - C)x$ satisfies $Gu \leq 0$.

We defer the proof of this theorem here. It follows from Corollary 1 that the u constructed in Theorem 3 is an upper bound on π , and hence π must be strictly < 1 in a neighborhood of 0. Thus $\beta = \beta^*$ is indeed the critical selection pressure for assured fixation.