A Unified Treatment of the Probability of Fixation when Population Size and the Strength of Selection Change Over Time

D. Waxman[1]
Centre for Computational Systems Biology, Fudan University, Shanghai 200433, People's Republic of China, and Faculty of Science and Engineering, University of Brighton, Brighton BN2 4GJ, United Kingdom

ABSTRACT The fixation probability is determined when population size and selection change over time and differs from Kimura's result, with long-term implications for a population. It is found that changes in population size are not equivalent to the corresponding changes in selection and can result in less drift than anticipated.

A new mutation in a *finite* population is subject to genetic drift and its ultimate fate is random: it may be either extinction (loss) or complete establishment (fixation). In a randomly mating population, the *typical* outcome for a new mutation is its loss, with fixation occurring only with small probability. Under static conditions (constant population size and constant strength of selection) a new beneficial mutation with small selective advantage, s, in a randomly mating population with discrete generations, has only a small probability of fixation: $\sim 2s$ when reproduction is treated as a branching process and the number of offspring has a Poisson distribution (Haldane 1927). A deleterious mutation has a yet smaller probability of fixation that is not calculable under a branching process. However, despite the relative rarity of fixation among the fates of all mutations, attention is largely focused on this outcome because the fixation of beneficial mutations plays a central role in the long-term adaptation of populations, and the fixation of deleterious mutations, in the absence of recombination, plays an important role in the long-term survival of populations (Muller 1964; Felsenstein 1974). Our understanding of the rate of such phenomena depends sensitively on the probability of fixation, and deviations from its static value, due to time-dependent conditions, are of particular significance. Indeed, there are a variety of reasons, both abiotic and biotic, why population size and the strength of selection do not generally remain constant over time.

Temporal changes, such as systematic trends in the composition or temperature of the atmosphere or oceans over time, although abiotic in nature, often have major implications for biological systems and may force biotic change. For example, atmospheric temperature changes may affect various biological processes within an organism, but also affect the vegetation on which an organism feeds, thereby affecting both selection and carrying capacity. Thus the general situation is complex, with selection fluctuating for multiple reasons; indeed, "… natural selection is very complicated, it is unlikely that the selection coefficient stays constant" (Ohta 1972, p. 307). Additionally, changes in, for example, resource/habitat availability or the density of parasites or predators will generally change the strength of selection as well as the size of a population. Thus generally we should expect variation in population size and the strength of selection.

The Soay sheep provide an illustration of the interplay of the various factors that affect population size and the strength of selection and the interrelation of these two quantities. The Soay sheep are an intensively studied wild mammalian population and their survival is density dependent and closely tied in with the availability of vegetation, whose quality and abundance are highly variable (Clutton-Brock and Pemberton 2003). Parasite population dynamics have been shown to regulate vertebrate populations and, in the Soay sheep, over-winter survival has been

---

doi: 10.1534/genetics.111.129288
Manuscript received December 03, 2010; accepted for publication April 20, 2011
Supporting information is available online at http://www.genetics.org/content/suppl/2011/04/28/genetics.111.129288.DC1.
[1] Address for correspondence: School of Mathematical Sciences, Fudan University, Shanghai 200433, People's Republic of China. E-mail: dwaxman.biology@gmail.com

identified with a response of the host's immune system to parasitic load (Coltman *et al.* 1999). Milner *et al.* (2003) concluded that the 20-year study of Soay sheep clearly demonstrated selection intensity fluctuating in a temporal fashion, and during years of high population density mortality was highest and hence selection most intense.

Theoretical studies of the way that population-size change affects fixation have a long history that includes an approximate diffusion analysis by Kimura and Ohta (1974) for logistic population growth and a much more recent investigation by Otto and Whitlock (1997), who considered various scenarios of population-size change, using a generalization of the branching process of Haldane (1927). Among many other things, the intuition established from the work of Otto and Whitlock (1997) is that when a beneficial mutation segregates in a population of increasing size, it has an increased probability of fixation. Subsequent work extended the calculations to include population-size changes of a stochastic nature, either within a branching-process framework (Engen *et al.* 2009) or for the Moran model (Parsons and Quince 2007; Parsons *et al.* 2010). The treatment of how changes in the strength of selection in finite populations affect fixation includes work of Kimura and Ohta (1972), while correlated fluctuations in the strength of selection, when population size is finite, have been considered by Takahata *et al.* (1975). Lambert (2006) combined drift and branching processes (and hence incorporates stochastic number fluctuations) and has obtained results in the regime of weak selection, which is defined by $4N_e|s| \ll 1$, where N_e denotes the *effective population size*.

The previous work has shown how the probability of fixation is affected by temporal variation in either population size or the strength of selection. In the present work we aim to compare and quantify the different ways that temporal changes in population size and the strength of selection affect fixation. We consider cases where selection may not be weak (*i.e.*, $4N_e|s|$ may *not* be small compared with 1). To this end, we present a unified treatment of such temporal variations on the probability of fixation that covers beneficial, neutral/nearly neutral, and deleterious mutations. To cover this range of selective regimes we work within the framework of the diffusion approximation (Kimura 1955a), where the relative frequency of a gene is treated as a random variable that takes continuous values. This approximation derives its name from the diffusion equation that governs the distribution of the relative gene frequency, and a diffusion analysis has been used to derive many fundamental results in population genetics (Crow and Kimura 1970).

Let us begin with the standard case where a single locus determines fitness in a randomly mating diploid sexual population under static conditions (*i.e.*, constant population size and constant strength of selection). (Asexual populations can also be studied via diffusion analysis; for a recent example, see Waxman and Loewe 2010.) The locus has two alleles, denoted A and a, and selection is on viability and is semidominant, with AA, Aa, and aa genotype individuals having relative fitnesses of $1 + 2s$, $1 + s$, and 1, respectively. Generations are discrete and the processes taking place in one generation are given by the life cycle

$$\begin{array}{ll} \text{adults} & \\ (\text{generation } t) & \\ \downarrow & \text{random mating, following by the death of all adults} \\ \text{zygotes} & \\ \downarrow & \text{viability selection} \\ \text{juveniles} & \\ \downarrow & \text{thinning (number regulation)} \\ \text{adults} & \\ (\text{generation } t{+}1). & \end{array}$$

We assume that each adult contributes to a very large number zygotes, so that viability selection may be treated as being deterministic in character. Juveniles (the individuals that survive viability selection) undergo a nonselective process of ecological thinning that leads to an adult population of N individuals in each generation.

The proportion of all genes at the locus in adults that are allele A is written $X(t)$; this is the relative frequency (henceforth termed frequency) of allele A. Because of the process of thinning in the life cycle, the frequency, $X(t)$, generally varies randomly from generation to generation and may have different values in different replicates of a population. Statistics of $X(t)$ can be described by a Wright–Fisher model (Fisher 1930; Wright 1931). However, to make theoretical progress, we consider an analysis based on the diffusion approximation, using methods of Kimura (1955), McKane and Waxman (2007), and Waxman (2011).

Readers not concerned with the detailed technical aspects of this work may omit the derivations/proofs contained in the main text and the material in supporting information, File S1.

A diffusion analysis is based on the diffusion equation

$$\begin{aligned} -\frac{\partial}{\partial t} f(x,\ t) = &-\frac{1}{4N_e}\frac{\partial^2}{\partial x^2}[x(1-x)f(x,\ t)] \\ &+\frac{\partial}{\partial x}[sx(1-x)f(x,t)] \end{aligned} \qquad (1)$$

for the probability density $f(x,\ t)$ of the frequency of the A allele at time t and frequency x.

Equation 1 can be solved to determine the probability of fixation of the A allele at long times, where the only possible outcomes for the locus are the A allele fixing or being lost. In terms of the frequency $X(t)$ we have

$$\lim_{t\to\infty} X(t) = \begin{cases} 1, & \text{if the } A \text{ allele fixes} \\ 0, & \text{if the } A \text{ allele is lost.} \end{cases}$$

The probability of occurrence of these different outcomes depends on the frequency, p, of the A allele at the initial time $t = 0$. The fixation probability can be written as $P_{fix}(p) = \lim_{t\to\infty} E_p[X(t)]$, where $E_p[\ldots]$ denotes an average over replicate populations when the A allele frequency has the value p at time $t = 0$. Thus $E_p[\ldots]$ is a shorthand for the conditional expectation $E[\ldots \mid X(0) = p]$.

The diffusion approximation of the fixation probability follows from Equation 1 and was found by Kimura (1955b) to be

$$P_{fix}(p) = \lim_{t\to\infty} E_p[X(t)] = \frac{1-e^{-Sp}}{1-e^{-S}}, \tag{2}$$

where $S = 4N_e s$. [The result of Equation 2 was derived assuming discrete generations. A closely related but different result is obtained if it is assumed from the outset that generations are overlapping (Moran 1958).] The numerical errors in the approximation of Equation 2 are remarkably small, even for haploid populations of size 12 (equivalent to diploid populations of size 6) as demonstrated by Ewens (1963).

The fixation probability of a single copy of an A allele in a population of census size N is obtained by setting $P = 1/(2N)$ in Equation 2. When the population size is such that $S = 4N_e s$ is large ($S >> 1$) but $Sp \equiv 2N_e s/N$ is small ($Sp << 1$), we arrive at the approximation $P_{\text{fix}}(p) \simeq 2N_e s/N$, which, when $N_e = N$, coincides with the leading term, $2s$, of a branching process (Haldane 1927). Thus for a constant strength of selection and a constant population size, branching processes are valid for beneficial mutations in populations of suitably large size ($S >> 1$). Diffusion results have a broader range of applicability that includes beneficial, deleterious and neutral/nearly neutral mutations in populations of essentially arbitrary size; in such an approach, there are essentially no restrictions on the *sign* and *size* of S).

When the effective population size and the selection coefficient depend on the time t, the parameter S acquires time dependence and becomes $S(t) = 4N_e(t)s(t)$. We initially proceed under the assumption that all changes in the composite quantity $S(t)$ are deterministic in character and have the property that they cease after a finite time, which we denote T. That is to say for times $\geq T$ we take $S(t)$ to have a constant value, and schematically

$$S(t) = 4N_e(t)s(t) = \begin{cases} \text{arbitrary}, & t < T \\ S_\infty, & t \geq T. \end{cases} \tag{3}$$

Such an assumption on $S(t)$ is not greatly restrictive. It could, for example, describe the situation where the strength of selection remains constant, while population size alone changes for a finite time before achieving a constant value. While logistic growth of a population does not precisely cease after a finite time, the population size can be well approximated as achieving its carrying capacity after a finite time and hence closely fits within the framework of Equation 3. Beyond such a case, Equation 3 could also describe environmental change of finite duration, which affects both population size and selection.

Generalization of $P_{\text{fix}}(p)$

The generalization of Equation 2 to the case of time-dependent population sizes, $N_e(t)$, and time-dependent selection coefficients, $s(t)$, can be obtained from some basic considerations. For deterministic changes in $N_e(t)$ and $s(t)$ that lead to the composite quantity $S(t) = 4N_e(t)s(t)$ having the form in Equation 3, the generalization of Equation 2 is

$$P_{fix}(p) = \frac{1-E_p\left[e^{-S_\infty X(T)}\right]}{1-e^{-S_\infty}}. \tag{4}$$

(the derivation of Equation 4 is given in the following paragraph and an alternative derivation is given in Part 1 of File S1). The only statistic required in Equation 4 involves $X(T)$, the random value of the A allele frequency at time T. The statistic in question, $E_p[e^{-S_\infty X(T)}]$, represents the average of $e^{-S_\infty X(T)}$ over all replicate populations where the initial frequency is p at time $t = 0$ [*i.e.*, having $X(0) = p$].

Derivation

To derive Equation 4 we note that as in the static case, the fixation probability can be written as $P_{fix}(p) = \lim_{t\to\infty} E_p[X(t)]$. Conditioning on the value of $X(t)$ at time T [*i.e.*, *after* changes in $N_e(t)$ and $s(t)$ have taken place], we have $P_{fix}(p) = \lim_{t\to\infty} E_p[X(t)] = \lim_{t\to\infty} E_p[E[X(t)|X(T)]] = E_p[\lim_{t\to\infty} E[X(t)| X(T)]]$. The quantity $\lim_{t\to\infty} E[X(t)|X(T)]$ appearing in the last expression can be written as $\lim_{t\to\infty} E_{X(T)}[X(t)]$ and follows from Equation 2 with the substitution $p \to X(T)$. We obtain $\lim_{t\to\infty} E_{X(T)}[X(t)] = (1-e^{-S_\infty X(T)})/(1-e^{-S_\infty})$ and hence the generalization of Equation 2 is $P_{fix}(p) = E_p[(1-e^{-S_\infty X(T)})/(1-e^{-S_\infty})]$, which is equivalent to Equation 4.

We note that the solution for $P_{\text{fix}}(p)$ in Equation 4, which directly follows from a diffusion analysis, does not generally agree with the forms assumed by Kimura and Ohta (1972, 1974) for situations of changing selection strength or changing population size. Additionally, Equation 4 will not be compatible with a branching-process approach when S_∞ is not large ($S_\infty \lesssim 1$) or indeed when S_∞ is zero or negative; there is limited validity to a branching-process treatment.

Before considering the numerical estimation of the fixation probability from Equation 4, we investigate some limiting cases and properties of Equation 4. Some of the limiting cases allow us to verify that Equation 4 is in accordance with well-known/well-understood results.

Limiting cases

1. When the time-interval T, over which all change occurs, tends to zero there is a negligibly short period of time dependence in the parameters. It can be shown (see Part 2 of File S1) that as a consequence, the frequency $X(T)$ is unaffected by these changes and tends to its initial value: $X(T) \to p$. Equation 4 then collapses to Kimura's result, Equation 2, for a population of constant size and constant selection coefficient.
2. When the problem is static, in the sense that neither population size nor the strength of selection changes over time [*i.e.*, $N_e(t) = N_e$ and $s(t) = s$], it should follow that Equation 4 coincides with Kimura's result, Equation 2,

which covers the static case. In this case, T can be taken to have any nonnegative value and it can be shown (see Part 3 of File S1) that under static conditions, the expectation appearing in Equation 4, namely $E_p[e^{-S_\infty X(T)}]$, is independent of T and hence equals its value at $T = 0$, which is $e^{-S_\infty p}$. As a consequence, Equation 4 collapses to Kimura's result. The property of $E_p[e^{-S_\infty X(T)}]$ of being independent of T is a well-known Martingale property of the diffusion approximation of the Wright–Fisher model; it appears to have been first identified as a property of the diffusion approximation by Ewens (1964).

3. When S_∞ becomes vanishingly small ($S_\infty \to 0$), Equation 4 reduces to $P_{\text{fix}}(p) = E_p[X(T)]$. Thus when there is selective neutrality *after* time T, the probability of fixation equals the mean allele frequency at time T.
4. When S_∞ becomes large and positive ($S_\infty \to +\infty$), the result in Equation 4 depends only on the probability that $X(T) \neq 0$. Equation 4 leads to $P_{\text{fix}}(p) = 1 - P_{\text{loss}}(T;\, p)$, where $P_{\text{loss}}(T;\, p)$ is the probability of loss of the A allele by time T given it had a frequency of p at time $t = 0$ (the proof is given in the following paragraph). Thus in an environment that has strongly positive selection after time T, the probability of fixation is determined by the probability that the A allele has not been lost *by* time T. Any A alleles that are present in a population after time T will subsequently fix.

Proof. The result $P_{\text{fix}}(p) = 1 - P_{\text{loss}}(T;\, p)$ follows since $\lim_{S_\infty \to \infty}(1-e^{-S_\infty X(T)})/(1-e^{-S_\infty})$ corresponds to the indicator function $1_{\{X(T)>0\}}$, which has the value of unity when $X(T) > 0$ and vanishes when $X(T)=0$. This indicator function coincides with $1 - 1_{\{X(T)=0\}}$ and thus Equation 4 becomes $P_{\text{fix}}(p) = 1 - E_p[1_{\{X(T)=0\}}] = 1 - P_{\text{loss}}(T;\, p)$, where $P_{\text{loss}}(T;\, p)$ is the probability that $X(T)=0$, given $X(0)=p$.

5. When S_∞ becomes large and negative ($S_\infty \to -\infty$), the result in Equation 4 depends only on the probability that $X(T) = 1$. Equation 4 leads to $P_{\text{fix}}(p) = P_{\text{fix}}(T;\, p)$, namely the probability of fixation of the A allele by time T, given it had a frequency of p at time $t = 0$ (the proof is given in the following paragraph). Thus in an environment that has strongly negative selection after time T, the probability of fixation is determined by fixations that occur up to time T; after time T any A alleles present in a population will be lost.

Proof. The result $P_{\text{fix}}(p) = P_{\text{fix}}(T;\, p)$ follows since $\lim_{S_\infty \to -\infty}(1-e^{-S_\infty X(T)})/(1-e^{-S_\infty}) = \lim_{S_\infty \to -\infty} e^{-|S_\infty|(1-X(T))}$, which corresponds to the indicator function $1_{\{X(T)=1\}}$. Equation 4 then becomes $P_{\text{fix}}(p) = E_p[1_{\{X(T)=1\}}] = P_{\text{fix}}(T;\, p)$, where $P_{\text{fix}}(T;\, p)$ is the probability that $X(T) = 1$, given $X(0) = p$.

The above results may give the impression that only the time dependence of the parameter $S(t) = 4N_e(t)s(t)$ is of significance for the fixation probability, $P_{\text{fix}}(p)$, of Equation 4, and that $P_{\text{fix}}(p)$ is not sensitive to the way that $N_e(t)$ and $s(t)$ *separately* change. This is not generally true. To illustrate this, let us consider two cases where either $N_e(t)$ or $s(t)$ linearly increases by a factor of f in time T and then remains constant after this:

Case a:

$$N_e(t) = \begin{cases} \left[1 + \frac{(f-1)t}{T}\right] N_0, & t < T \\ fN_0, & t \geq T, \end{cases} \qquad s(t) = \begin{cases} s_0, & t < T \\ s_0, & t \geq T. \end{cases}$$

Case b:

$$N_e(t) = \begin{cases} N_0, & t < T \\ N_0, & t \geq T, \end{cases} \qquad s(t) = \begin{cases} \left[1 + \frac{(f-1)t}{T}\right] s_0, & t < T \\ fs_0, & t \geq T. \end{cases}$$

Both case a and case b lead to

$$S(t) = \begin{cases} 4\left[1 + \frac{(f-1)t}{T}\right] N_0 s_0, & t < T \\ 4fN_0 s_0 & t \geq T, \end{cases}$$

yet they lead to *different* values for the fixation probability. As a concrete example of the fixation probabilities that can arise for case a and case b, consider the parameter values $N_0 = 10, f = 10, T = 10, s_0 = 0.04$, and $P = 1/(2N_0) = 0.05$. From simulations we find that case a leads to a fixation probability of $P_{\text{fix}}(p) \simeq 0.363$ while case b leads to $P_{\text{fix}}(p) \simeq 0.275$. [All simulations carried out in this work are made within the framework of a Wright–Fisher model (Fisher 1930; Wright 1931). In such a framework, selection is treated as a deterministic process, and only the random sampling of individuals without regard to type, *i.e.*, the process of random genetic drift, is treated stochastically.] In this example, $S(t)$ increases by a factor of 10 in each case but when it is just $N_e(t)$ that varies over time, the probability of fixation is ~30% larger than when just $s(t)$ varies.

The results of cases a and b illustrate the more general phenomenon that, as far as fixation is concerned, the way that $S(t) = 4N_e(t)s(t)$ varies over time is not the full story: the probability of fixation generally depends on the source of the variation of $S(t)$.

The phenomenon that a varying N_e or a varying s has nonequivalent effects on the probability of fixation follows ultimately from the way that a changing N_e modifies the *timescale* of random genetic drift. Since we shall compare populations with the same initial size, $N_e(0)$, we take $N_e(0)$ as a fixed parameter and find it convenient to take the timescale associated with genetic drift to be

$$\tau = \int_0^t \frac{N_e(0)}{N_e(u)} du. \tag{5}$$

We call τ the *drift time*. We can express t as a function of the drift time, τ, and write $t = t(\tau)$. The form of $t(\tau)$ follows from solving Equation 5 for t. Let us now explain the advantage of using the drift time, τ, instead of the actual time, t.

Assuming a time-dependent population size and a time-dependent strength of selection, we transform the diffusion equation, Equation 1, so that the drift time τ (Equation 5) is used *instead* of the actual time t. It can be shown (see Part 4 of File S1) that under such a transformation, the quantity playing the role of the strength of selection in the resulting diffusion equation is $R(\tau) = 4N_e(t(\tau))s(t(\tau))$ and we call $R(\tau)$ the *overall strength of selection*. The quantity $N_e(t)$ does not appear elsewhere in the transformed diffusion equation. In a static problem, the probability of fixation is determined by just the initial frequency, p, and the overall strength of selection (see Equation 2). In a more general case, where $N_e(t)$ and $s(t)$ exhibit deterministic change over time, the probability of fixation is determined by p, and the entire history of $R(\tau)$, from $\tau = 0$ onward.

We note that when population size is *static* ($N_e(t) = N_e(0)$), the drift time and the actual time coincide: $t(\tau) = \tau$, independent of any variation of s. Thus for a static population size but a changing strength of selection, the overall strength of selection is

$$R_0(\tau) = 4N_e(0)s(\tau), \qquad \text{static } N_e. \tag{6}$$

By contrast, when N_e varies with time, the drift time does not generally coincide with the actual time: $t(\tau) \neq \tau$. In such a case, the overall strength of selection, when s has the fixed value $s(0)$, is

$$R_1(\tau) = 4N_e(t(\tau))s(0), \quad \text{varying } N_e. \tag{7}$$

There is a key difference between Equations 6 and 7: the time that s depends on in Equation 6 is simply the time we have adopted for the diffusion equation, namely τ. By contrast, in Equation 7, the time that N_e depends upon is $t(\tau)$. It can be shown (see Part 4 of File S1) that irrespective of whether $N_e(t)$ increases with time or whether it decreases with time, the quantity $N_e(t(\tau))$ generally satisfies

$$N_e(t(\tau)) > N_e(\tau) \quad \text{for all } \tau > 0. \tag{8}$$

This inequality stems directly from the modification of the timescale induced by an N_e that exhibits either increase or decrease. [For an $N_e(t)$ that changes *monotonically*, the inequality in Equation 8 becomes replaced with $N_e(t(\tau)) \geq N_e(\tau)$.] It means the population size $N_e(t(\tau))$ in the transformed diffusion equation is *larger* than the value of the population size that we might believe is relevant at time τ, namely $N_e(\tau)$.

All other things being equal, a population whose size is larger than anticipated exhibits less drift than anticipated. This is the reason case a above, for a positively selected allele, leads to a larger probability of fixation than case b. Similarly, the inequality in Equation 8 indicates that a negatively selected allele, when present in a population that increases, or one that decreases, will have a reduced probability of fixation (due to less drift) compared with the case where population size is static and all variation occurs in the strength of selection.

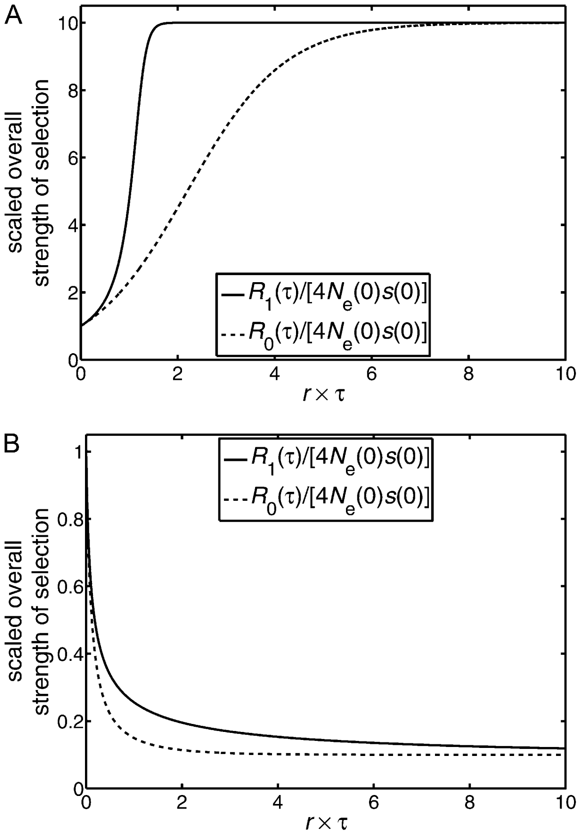

Figure 1 Scaled overall strength of selection. We present plots of the *scaled* "overall strength of selection," namely $R(\tau) = 4N_e(t(\tau))s(t(\tau))$ divided by its initial value $4N_e(0)s(0)$. Different scenarios of change are illustrated, where the selection coefficient $s(t)$ is positive for all t. We used the two functions $a(t)$ and $b(t)$ to produce this figure. The function $a(t)$ corresponds to logistic *growth* by a factor of 10; *i.e.*, $a(t) = (1/10 + 9/10e^{-rt})^{-1}$, where r is a positive constant, and hence $a(0) = 1$ while $a(\infty) = 10$. The function $b(t)$ corresponds to logistic *decay* by a factor of 10; *i.e.*, $b(t) = (10 - 9e^{-rt})^{-1}$ and hence $b(0) = 1$ while $b(\infty) = 1/10$. (A) We compare (i) $R_0(\tau)$ for a fixed effective population size and a logistically increasing selection coefficient ($N_e(t) = N_e(0)$, $s(t) = s(0)a(t)$), with (ii) $R_1(\tau)$ for a logistically increasing effective population size and a fixed selection coefficient ($N_e(t) = N_e(0)a(t)$, $s(t) = s(0)$). (B) We compare (iii) $R_0(\tau)$ for a fixed effective population size and a logistically decreasing selection coefficient ($N_e(t) = N_e(0)$, $s(t) = s(0)b(t)$), with (iv) $R_1(\tau)$ for a logistically decreasing effective population size and a fixed selection coefficient ($N_e(t) = N_e(0)b(t)$, $s(t) = s(0)$). In A, an increasing population size leads to a *larger overall strength of selection* than that of an increasing selection coefficient. In B, a decreasing population size leads again to a *larger overall strength of selection* than that of a decreasing selection coefficient. These properties result from either an increasing population size or a decreasing population size modifying the natural timescale of the diffusion equation in such a way that there is less drift than might be anticipated (see Equation 8).

In Figure 1 we illustrate the forms of the "overall strengths of selection," $R_0(\tau)$ and $R_1(\tau)$, of Equations 6 and 7 to show how different the overall strength of selection can be under "fixed N_e, varying s" and "varying N_e, fixed s." For definiteness, Figure 1 is restricted to alleles with positive selection coefficients.

When $N_e(t)$ exhibits periods of both increase and decrease, no inequality of the form in Equation 8 generally holds.

Table 1 Illustrative numerical results for the probability of fixation when the composite quantity $S(t) = 4N_e(t)s(t)$ changes with time

Data set	Method	T	N_0	N_∞	s_0	s_∞	S_0	S_∞	P_{fix}	Cost/rep
1	Direct simulation	20	50	50	0.005	0.050	1	10	0.0687	12
2	Finite *T* simulation	20	50	50	0.005	0.050	1	10	0.0689	4
3	Direct simulation	20	50	500	0.005	0.005	1	10	0.0817	115
4	Finite *T* simulation	20	50	500	0.005	0.005	1	10	0.0818	9
5	Direct simulation	200	50	50	0.005	0.050	1	10	0.0320	11
6	Finite *T* simulation	200	50	50	0.005	0.050	1	10	0.0321	10
7	Direct simulation	200	50	500	0.005	0.005	1	10	0.0451	51
8	Finite *T* simulation	200	50	500	0.005	0.005	1	10	0.0447	18
9	Direct simulation	20	500	500	0.005	0.050	10	100	0.0672	18
10	Finite *T* simulation	20	500	500	0.005	0.050	10	100	0.0671	4
11	Direct simulation	20	500	5000	0.005	0.005	10	100	0.0814	192
12	Finite *T* simulation	20	500	5000	0.005	0.005	10	100	0.0812	8
13	Direct simulation	200	500	500	0.005	0.050	10	100	0.0298	14
14	Finite *T* simulation	200	500	500	0.005	0.050	10	100	0.0302	11
15	Direct simulation	200	500	5000	0.005	0.005	10	100	0.0428	90
16	Finite *T* simulation	200	500	5000	0.005	0.005	10	100	0.0425	17

We take $S(t)$ to start at the positive value S_0 at time $t = 0$ and then to linearly increase to the value $S_\infty = 10S_0$ by time T and then remain constant at the value S_∞ for all times larger than T. Two different methods are used to estimate the probability of fixation when initially there is only a single copy of an *A* allele: (i) direct simulation, where we "follow" each replicate population until either fixation or loss occurs and (ii) simulations based on Equation 4, where we follow each replicate population for, at most, T generations. The column labeled "Cost/rep" gives the mean number of generations a replicate population was followed in a simulation. In the simulations, 5×10^5 replicate populations were used, and we adopted a Wright–Fisher model where the only place in the life cycle where randomness occurs is in the thinning of the number of individuals to $N_e = N$ adults. The initial values of N_e and s are N_0 and s_0, while the final values are N_∞ and s_∞; data sets 1, 2, 5, 6, 9, 10, 13, and 14 correspond to N_e fixed and s changing with time. It is evident from the table that there are differences in the fixation probability, depending on whether N_e changed with time, at fixed s, or s changed with time at fixed N_e.

Estimation of the probability of fixation

Consider now how we would estimate the probability of fixation, for a case where potentially complicated deterministic changes of $s(t)$ and $N_e(t)$ take place up to a finite time, T. A *direct* approach would simply be to *simulate* the behavior of many replicates of a population. Each replicate population would need to be "followed" for a sufficient number of generations until either fixation or loss occurred. The fraction of such populations that fix is an estimate of the fixation probability. By contrast, using Equation 4 it is necessary to follow replicate populations for, at most, only T generations. On average it will require less than T generations, since fixation or loss will occur in some replicate populations prior to time T, and no further change will occur in such populations. An estimate of the fixation probability follows from such simulations by using the average of $e^{-S_\infty X(T)}$ in Equation 4. This procedure may be significantly shorter than the *direct* approach, depending on the parameter values in the problem (Table 1 illustrates differences in the time required). Alternatively, the quantity $E_p[e^{-S_\infty X(T)}]$ that appears in Equation 4 can be estimated from numerical solution of the backward diffusion equation (see Part 5 of File S1). This thus provides an alternative route to estimation of the fixation probability.

The different approaches to the calculation of the fixation probability are illustrated with an example in Table 1.

Stochastic fluctuations

So far we have presented results of the probability of fixation for cases where the time-dependent changes in the population size and the strength of selection are deterministic in character. Let us now point out a generalization of these results that includes stochastic fluctuations in population size and the strength of selection.

We note that Lambert (2006) obtained results in the regime of weak selection, when there is fluctuating population size and random genetic drift, by combining branching and Wright–Fisher processes, while Parsons *et al.* (2010) considered the effects of fluctuating population size, in a quasi-neutral (*i.e.*, weak selection) situation, where different alleles have the same ratio of intrinsic birth to death rates. The work of Parsons and Quince (2007) covers the nonneutral regime and includes density dependence and fluctuations in population size that arise from uncorrelated stochastic births and deaths. By contrast, Karlin and Levikson (1974) considered the case where stochastic changes in $N_e(t)$ and $s(t)$ have temporal correlations that relate the values of these quantities in adjacent generations. These authors found that various statistics of $N_e(t)$ and $s(t)$ make contributions to the drift and diffusion coefficients of the diffusion equation. Here, we make the alternative assumption that the stochastic fluctuations of $N_e(t)$ and $s(t)$ have temporal autocorrelations that decay slowly, over *very many* generations. It is then possible to account for these fluctuations using the approach of Takahata *et al.* (1975); see also Huerta-Sanchez *et al.* (2008) where two different models of autocorrelation are incorporated into stochastic fluctuations of selection.

To generalize Equation 4, we first assume there are both deterministic changes and stochastic fluctuations in $N_e(t)$ and $s(t)$ for times $t \leq T$, but only stochastic fluctuations for times $t > T$. Then the appropriate generalization of Equation 4 for this case is $P_{fix}(p) = E_p[\bar{\Psi}_1(X(T))]$, where $\bar{\Psi}_1(x)$ is an eigenfunction of an *averaged* backward diffusion operator: see Part 6 of File S1 for further details.

To summarize: in this work we have presented results, based on the diffusion approximation, which generalize Kimura's result for the probability of fixation to cases where population size and the strength of selection are time dependent. We have provided results when the changes are deterministic and also shown their generalization when there are stochastic fluctuations with temporal autocorrelations that decay over many generations. This work has implications for the long-term adaptation of populations, demonstrating that while temporal variation in population size and the strength of selection both affect the probability of fixation, the changes are not equivalent and that generally, a population size that either increases or decreases will lead to less drift and hence less fixation of deleterious mutations and greater fixation of beneficial mutations than would otherwise be anticipated. There are many possible scenarios where population size and the strength of selection change, and the result of Equation 4 allows an efficient way to explore the implications of these for fixation.

Acknowledgments

I thank Warren Ewens, Jianfeng Feng, Guillaume Martin, Yannis Michalakis, and Andy Overall for stimulating discussions and the latter for very helpful suggestions on the manuscript. I also thank the two anonymous reviewers for comments and suggestions that substantially improved this work.

Note: See Uecker and Hermisson (pp. 915–930) in this issue, for a related work.

Literature Cited

Coltman, D. W., J. G. Pilkington, J. A. Smith, and J. M. Pemberton, 1999 Parasite-mediated selection against inbred Soay sheep in a free-living, island population. Evolution 53: 1259–1267.

Clutton-Brock, T. H., and J. M. Pemberton, 2003 Soay Sheep: Dynamics and Selection in an Island Population. Cambridge University Press, Cambridge, UK.

Crow, J. F., and M. Kimura, 1970 An Introduction to Population Genetics Theory. Harper & Row, New York.

Engen, S., R. Lande, and B. Saether, 2009 Fixation probability of beneficial mutations in a fluctuating population. Genet. Res. 91: 73–82.

Ewens, W. J., 1963 Numerical results and diffusion approximations in a genetic process. Biometrika 50: 241–249.

Ewens, W. J., 1964 The pseudo-transient distribution and its uses in genetics. J. Appl. Probab. 1: 141–156.

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

Fisher, R. A., 1930 The Genetical Theory of Natural Selection. Clarendon Press, Oxford.

Haldane, J. B. S., 1927 A mathematical theory of natural and artificial selection. V. Selection and mutation. Proc. Camb. Philos. Soc. 23: 838–844.

Huerta-Sanchez, M., R. Durrett, and C. D.. Bustamante, 2008 Population genetics of polymorphism and divergence under fluctuating selection. Genetics 178: 325–337.

Karlinson, S., and B. Levikson, 1974 Temporal fluctuations in selection intensities: Case of small population size. Theor. Pop. Biol. 6: 383–412.

Kimura, M., 1955a Stochastic processes and distribution of gene frequencies under natural selection. Cold Spring Harbor Symp. Quant. Biol. **20:** 33–53.

Kimura, M., 1955b Solution of a process of random genetic drift with a continuous model. Proc. Natl. Acad. Sci. USA 41: 141–150.

Kimura, M., and T. Ohta, 1972 Probability of fixation of a mutant gene in a finite population when selective advantage decreases with time. Genetics 65: 525–534.

Kimura, M., and T. Ohta, 1974 Probability of gene fixation in an expanding finite population. Proc. Natl. Acad. Sci. USA 71: 3377–3379.

Lambert, A., 2006 Probability of fixation under weak selection: a branching process unifying approach. Theor. Popul. Biol. 69: 419–441.

Milner, J. M., S. D. Albon, L. E. B. Kruk, and J. M. Pemberton, 2003 Selection on phenotype, pp. 190–216 in *Soay Sheep: Dynamics and Selection in an Island Population*, edited by T. H. Clutton-Brock and J. M. Pemberton. Cambridge University Press, Cambridge, UK.

Moran, P. A. P., 1958 Random processes in genetics. Math. Proc. Camb. Philos. Soc. 54: 60–71.

McKane, A. J., and D. Waxman, 2007 Singular solutions of the diffusion equation of population genetics. J. Theor. Biol. 247: 849–858.

Muller, H. J., 1964 The relation of recombination to mutational advance. Mutat. Res. 1: 2–9.

Ohta, T., 1972 Population size and the rate of evolution. J. Molec. Evol 1: 305–314.

Otto, S. P., and M. Whitlock, 1997 The probability of fixation in populations of changing size. Genetics 146: 723–733.

Parsons, T. L., and C. Quince, 2007 Fixation in haploid populations exhibiting density dependence I: the non-neutral case. Theor. Popul. Biol. 72: 121–135.

Parsons, T. L., C. Quince, and J. B. Plotkin, 2010 Some consequences of demographic stochasticity in population genetics. Genetics 185: 1345–1354.

Takahata, N., K. Ishii, and H. Matsuda, 1975 Effect of temporal fluctuation of selection coefficient on gene frequency in a population. Proc. Natl. Acad. Sci. USA 72: 4541–4545.

Waxman, D., 2011 Comparison and content of the Wright–Fisher model of random genetic drift, the diffusion approximation, and an intermediate model. J. Theor. Biol. 269: 79–87.

Waxman, D., and L. Loewe, 2010 A stochastic model for a single click of Muller's ratchet. J. Theor. Biol. 264: 1120–1132.

Wright, S., 1931 Evolution in Mendelian populations. Genetics 16: 97–159.

Communicating editor: L. M. Wahl

GENETICS

Supporting Information
http://www.genetics.org/content/suppl/2011/04/28/genetics.111.129288.DC1

A Unified Treatment of the Probability of Fixation when Population Size and the Strength of Selection Change Over Time

D. Waxman

DOI: 10.1534/genetics.111.129288

FILE S1

SUPPORTING INFORMATION

The Supporting Information presented here is separated into background material followed by six separate parts. In Part 1 an alternative derivation is given for the probability of fixation when the effective population size and the strength of selection may change up to time T, but do not change beyond time T. In Part 2 a limiting case of the probability of fixation is determined, while in part 3 it is shown how Equation (4) of the main text, for the fixation probability, simplifies when conditions are static. In Part 4 a useful transformation of the diffusion equation is made and a property of the population size in this equation is exposed. In Part 5 details are given of a way to numerically evaluate the quantity $E_p[e^{-S_\infty X(T)}]$ (which appears in Equation (4) of the main text). Lastly, in Part 6, we give details of the generalisation of Equation (4) of the main text to include stochastic fluctuations in population size and the strength of selection.

Background

Consider a single locus that determines fitness in a population of randomly mating diploid individuals of (variance) effective size $N_e(t)$ at time t. The locus has two alleles, denoted A and a, and is subject to semidominant selection where, at time t, the relative fitnesses of AA, Aa and aa genotype individuals are $1+2s(t)$, $1+s(t)$ and 1, respectively. Until Part 6 of the Supporting Information, we will assume that $N_e(t)$ and $s(t)$ change deterministically.

We write the relative frequency (henceforth referred to as the *frequency*) of allele A at time t as $X(t)$. Given an initial A allele frequency of y at time u, the probability density of the A allele frequency at later time t, at frequency x, is written as $K(x,t|y,u)$. Under a diffusion approximation $K(x,t|y,u)$ obeys the forward diffusion equation

$$-\frac{\partial}{\partial t}K(x,t|y,u) = -\frac{1}{4N_e(t)}\frac{\partial^2}{\partial x^2}\left[x(1-x)K(x,t|y,u)\right] + s(t)\frac{\partial}{\partial x}\left[x(1-x)K(x,t|y,u)\right] \tag{S1}$$

(Kimura 1955) and is subject to the initial condition $K(x,u|y,u)=\delta(x-y)$ where $\delta(x)$ denotes a Dirac delta function of argument x.

The result of the main text for the fixation probability, Equation (4), is given here for completeness:

$$P_{\text{fix}}(p) = \frac{1-E_p[e^{-S_\infty X(T)}]}{1-e^{-S_\infty}} \tag{S2}$$

where p is the frequency at time $t=0$, S_∞ is the value of $4N_e(t)s(t)$ when $t\geq T$ and $E_p[...]$ denotes the conditional expectation $E[...|X(0)=p]$.

In what follows, we shall repeatedly make use of the relation

$$E_p\left[e^{-S_\infty X(T)}\right] \equiv E\left[e^{-S_\infty X(T)}|X(0)=p\right] = \int_0^1 e^{-S_\infty x}K(x,T|p,0)dx \tag{S3}$$

where the last equality follows since $K(x,T|p,0)$ is the probability density, at frequency x, of $X(T)$, conditional on $X(0)=p$.

PART 1

To give more insight into Equation (4) of the main text (reproduced in Equation (S2) above), and for use in Part 6 of the Supporting Information, we give another derivation of Equation (S2).

We use a property of $K(x,t|y,u)$ that follows from Equation (S1) being first order in time derivatives and linear, namely

$$K(x,t|y,u) = \int_0^1 K\ \text{[illegible]}\ )K(z,r|y,u)dz \tag{S4}$$

where $u \leq r \leq t$. Equation (S4) is often known as the Chapman-Kolmogorov equation.

To determine the fixation probability, given an A allele frequency of p at time $t = 0$, we determine the behaviour of $K(x,t|p,0)$ at large values of t. We work under the assumption that $s(t)$ and $N_e(t)$ do not change when the time is larger than a specific time, T, hence

$$4N_e(t)s(t) = S_\infty \text{ (a constant)} \qquad \text{for } t \geq T. \tag{S5}$$

We take Equation (S5) into account when making specific choices of r, y and u in Equation (S4):

$$K(x,t|p,0) = \int_0^1 K(x,t|z,T)K(z,T|p,0)dz. \tag{S6}$$

The factor $K(z,T|p,0)$ in Equation (S6) is potentially complicated since it is determined by the time-dependent changes of $N_e(t)$ and $s(t)$ that occur from time 0 to time T. By contrast, $K(x,t|z,T)$ applies for the range of times where the strength of selection and the population size have achieved constant values.

In the work of McKane and Waxman (2007) and Waxman (2011), 'zero current' boundary conditions were imposed on the solution of the forward diffusion equation[1]. Zero current boundary conditions, in contrast to the approach adopted by Kimura (1955), ensure that probability is conserved and lead to an interpretation of Equation (S1) that is consistent with underlying the Wright-Fisher model for all x including $x = 0$ and $x = 1$, *i.e.*, including fixation and loss (see Waxman 2011). We apply these boundary conditions in the present context.

For times $t > T$, where the strength of selection and the population size have achieved the constant values s_∞ and N_∞, we can write

$$K(x,t|z,T) = \sum_{n=0}^{\infty} \Phi_n(x)\Psi_n(z)e^{-\lambda_n(t-T)} \tag{S7}$$

where $\Phi_n(x)$ ($\Psi_n(z)$) is an eigenfunction of the forward (backward) diffusion operator associated with eigenvalue λ_n:

$$-\frac{1}{4N_\infty}\frac{d^2}{dx^2}\left[x(1-x)\Phi_n(x)\right] + s_\infty\frac{d}{dx}\left[x(1-x)\Phi_n(x)\right] = \lambda_n\Phi_n(x) \tag{S8}$$

$$-\frac{y(1-y)}{4N_\infty}\frac{d^2}{dy^2}\Psi_n(y) - s_\infty y(1-y)\frac{d}{dy}\Psi_n(y) = \lambda_n\Psi_n(y). \tag{S9}$$

For $K(x,t|z,T)$ to be a solution of the appropriate diffusion equation, the eigenfunctions appearing in Eqs. (S7), (S8) and (S9) are orthogonal and normalised (Waxman 2011) in the sense

$$\int_0^1 \Phi_n(x)\Psi_m(x)dx = \delta_{m,n} \tag{S10}$$

where $\delta_{m,n}$ is a Kronecker delta, which equals 1 when $m = n$ and vanishes otherwise.

A key feature of the analysis of McKane and Waxman (2007) and Waxman (2011) is that under 'zero current' boundary conditions, there are *two* eigenfunctions of the forward and backward eigenvalue equations

---

[1] Zero current boundary conditions correspond to the probability current density $-\frac{1}{4N_e(t)}\frac{\partial}{\partial x}\left[x(1-x)K(x,t|y,u)\right] + s(t)x(1-x)K(x,t|y,u)$ vanishing at $x = 0$ and $x = 1$. Since $K(x,t|y,u)$ is a probability density, it is an integrable function of x with the consequence that $x(1-x)K(x,t|y,u)$ vanishes at $x = 0$ and $x = 1$. Accordingly, zero current boundary conditions can be taken as requiring that $\frac{\partial}{\partial x}\left[x(1-x)K(x,t|y,u)\right]$ vanishes at $x = 0$ and $x = 1$. See McKane and Waxman (2007) and Waxman (2011) for further details.

with zero eigenvalue. These are given the labels $n = 0$ and $n = 1$ and as $t \to \infty$ only terms with these labels persist in Equation (S7):

$$\lim_{t\to\infty} K(x,t|z,T) = \Phi_0(x)\Psi_0(z) + \Phi_1(x)\Psi_1(z). \tag{S11}$$

We take[2] $\Phi_0(x) = \delta(x)$ and $\Phi_1(x) = \delta(1-x)$: these are probability densities associated with the A allele frequency having the precise values 0 and 1, respectively, and correspond to loss and fixation of the A allele. Using Equation (S11) in Equation (S6), we arrive at $\lim_{t\to\infty} K(x,t|p,0) = \Phi_0(x)\int_0^1 \Psi_0(z)K(z,T|p,0)dz + \Phi_1(x)\int_0^1 \Psi_1(z)K(z,T|p,0)dz$ and the coefficient of $\Phi_1(x)$ in this expression is the probability of fixation. We thus find

$$P_{\text{fix}}(p) = \int_0^1 \Psi_1(z)K(z,T|p,0)dz = E_p\left[\Psi_1(X(T))\right]. \tag{S12}$$

The eigenfunction $\Psi_1(y)$ of the backward eigenvalue equation has zero eigenvalue; it obeys Equation (S9) with $n = 1$ and $\lambda_1 = 0$. The eigenfunction is subject to $\Psi_1(0) = 0$ and $\Psi_1(1) = 1$ which result from Equation (S10) with $m = 1$ and $n = 0$ or $n = 1$. It follows that $\Psi_1(y) = \dfrac{1-e^{-S_\infty y}}{1-e^{-S_\infty}}$. We can thus write Equation (S12) as $P_{\text{fix}}(p) = E_p\left[\dfrac{1-e^{-S_\infty X(T)}}{1-e^{-S_\infty}}\right] = \dfrac{1-E_p\left[e^{-S_\infty X(T)}\right]}{1-e^{-S_\infty}}$.

PART 2

In this part of the Supporting Information we obtain the limiting case of Equation (S2) when $T \to 0$.

Given that $K(x,t|p,0)$ obeys the forward diffusion equation, Equation (S1), a direct calculation, assuming $[N_e(t)]^{-1}$ and $|s(t)|$ remain bounded for $0 \le t \le T$, yields $\left|E_p\left[e^{-S_\infty X(T)}\right] - e^{-S_\infty p}\right| = O(T)$ and hence as $T \to 0$, Equation (S2) collapses to $P_{\text{fix}}(p) = \dfrac{1-e^{-S_\infty p}}{1-e^{-S_\infty}}$ which is Equation (2) of the main text.

PART 3

When population size and the strength of selection are independent of time ($N_e(t) = N_e$ and $s(t) = s$) the expectation $E_p\left[e^{-S_\infty X(T)}\right]$ appearing in Equation (S2) takes a simple form which allows Equation (S2) to be significantly simplified.

To establish the form of $E_p\left[e^{-S_\infty X(T)}\right]$ we use Equation (S3) and obtain an equation for the quantity $\int_0^1 e^{-S_\infty x}K(x,T|p,0)dx$ by multiplying Equation (S1) (with t replaced by T), by $e^{-S_\infty x}$ and integrating over all x. When N_e and s are independent of time we obtain

$$\begin{aligned} -\frac{\partial}{\partial T}\int_0^1 e^{-S_\infty x}K(x,T|p,0)dx &= -\frac{1}{4N_e}\int_0^1 e^{-S_\infty x}\frac{\partial^2}{\partial x^2}\left[x(1-x)K(x,T|p,0)\right]dx \\ &\quad + s\int_0^1 e^{-S_\infty x}\frac{\partial}{\partial x}\left[x(1-x)K(x,T|p,0)\right]dx. \end{aligned} \tag{S13}$$

Integrating the first term on the right hand side by parts twice, the second term once by parts, and using $4N_e s = S_\infty$, leads to

$$-\frac{\partial}{\partial T}\int_0^1 e^{-S_\infty x}K(x,T|p,0)dx = -\left[\frac{e^{-S_\infty x}}{4N_e}\frac{\partial}{\partial x}\left[x(1-x)K(x,T|p,0)\right]\right]_{x=0}^{x=1}. \tag{S14}$$

---

[2] There are two eigenfunctions associated with zero eigenvalue. Different linear combinations of these eigenfunctions also have zero eigenvalue. In the present work we have made the particular choice of these eigenfunctions made by Waxman (2011); an alternative, but equivalent choice of these eigenfunctions has been made by McKane and Waxman (2007).

The right hand side vanishes under the 'zero current' boundary conditions imposed by McKane and Waxman (2007) and Waxman (2011). We then have $\frac{\partial}{\partial T}\int_0^1 e^{-S_\infty x}K(x,T|p,0)dx \equiv \frac{\partial}{\partial T}E_p\left[e^{-S_\infty X(T)}\right] = 0$ corresponding to $E_p\left[e^{-S_\infty X(T)}\right]$ being independent of T. We can take $T = 0$ and arrive at

$$E_p\left[e^{-S_\infty X(T)}\right] = e^{-S_\infty p}. \tag{S15}$$

Using this last result in Equation (S2) shows that when parameters are time-independent, the fixation probability reduces to $P_{\rm fix}(p) = \frac{1-e^{-S_\infty p}}{1-e^{-S_\infty}}$, *i.e.*, Equation (2) of the main text.

The property of Equation (S15), that under static conditions $E_p\left[e^{-S_\infty X(T)}\right]$ is independent of T (and hence equals its value when $T = 0$) is a Martingale property of the diffusion approximation that appears to have been first identified by Ewens (1964). It can also be derived by noting that $e^{-S_\infty y}$ is a linear superposition of $\Psi_0(y)$ and $\Psi_1(y)$, the two eigenfunctions of the backward equation with zero eigenvalue.

PART 4

Transformation of the diffusion equation

In this part of the Supporting Information, we transform the diffusion equation, Equation (S1), by replacing the time t by the 'drift time' given in Equation (5) of the main text:

$$\tau = \int_0^t \frac{N_e(0)}{N_e(u)}du \tag{S16}$$

and determine a key property of $N_e(t)$.

We shall compare populations with the same value of $N_e(0)$ and hence consider $N_e(0)$ a fixed parameter.

We can, in principle, solve Equation (S16) for t and obtain it as a function of τ, which we write as $t(\tau)$. Defining

$$\widetilde{K}(x,\tau|y,\sigma) = K(x,t(\tau)|y,t(\sigma)) \tag{S17}$$

we find that Equation (S1) becomes

$$\begin{aligned} -4N_e(0)\frac{\partial}{\partial\tau}\widetilde{K}(x,\tau|y,\sigma) &= -\frac{\partial^2}{\partial x^2}\left[x(1-x)\widetilde{K}(x,\tau|y,\sigma)\right] \\ &\quad +R(\tau)\frac{\partial}{\partial x}\left[x(1-x)\widetilde{K}(x,\tau|y,\sigma)\right] \end{aligned} \tag{S18}$$

where

$$R(\tau) = 4N_e(t(\tau))s(t(\tau)). \tag{S19}$$

By virtue of its position in Equation (S18) the quantity $R(\tau)$ encapsulates selection and population size in a single term that we call the 'overall strength of selection'. Note that the time arguments of N_e and s in Equation (S19) are $t(\tau)$ and hence are determined by the relationship between t and τ of Equation (S16), *i.e.*, are determined by the way that N_e varies over time.

Equation (S18) is the diffusion equation that is obtained from the original diffusion equation, Equation (S1), when it is transformed to depend on the drift time, τ.

Property of $N_e(t(\tau))$

Let us now demonstrate a property of the quantity $N_e(t(\tau))$ appearing in Equation (S19). We consider the two cases, where $N_e(t)$ either increases with t or where it decreases with t.

1) Increasing $N_e(t)$ (*i.e.*, $dN_e(t)/dt > 0$). This immediately yields (i) $N_e(t_1) > N_e(t_2)$ for $t_1 > t_2$. From Equation (S16) we obtain (ii) $t(\tau) > \tau$. It follows from (i) and (ii) that $N_e(t(\tau)) > N_e(\tau)$.

2) Decreasing $N_e(t)$ (*i.e.*, $dN_e(t)/dt < 0$). This immediately yields (i) $N_e(t_1) < N_e(t_2)$ for $t_1 > t_2$. From Equation (S16) we obtain (ii) $\tau > t(\tau)$. It follows from (i) and (ii) that $N_e(t(\tau)) > N_e(\tau)$.

Thus in both cases we have $N_e(t(\tau)) > N_e(\tau)$ and this generally holds for an $N_e(t)$ that exhibits only increase or only decrease.

PART 5

The quantity $E_p[e^{-S_\infty X(T)}]$ that appears in Equation (S2) can be determined by numerically solving a diffusion equation. To establish this we first note that $K(x,t|y,u)$ not only obeys Equation (S1) but also the backward equation

$$\frac{\partial}{\partial u}K(x,t|y,u) = -\frac{y(1-y)}{4N_e(u)}\frac{\partial^2}{\partial y^2}K(x,t|y,u) - s(u)y(1-y)\frac{\partial}{\partial y}K(x,t|y,u) \tag{S20}$$

subject to

$$\begin{aligned} K(x,t|y,t) &= \delta(x-y) \\ K(x,t|0,u) &= \delta(x) \\ K(x,t|1,u) &= \delta(1-x) \end{aligned} \tag{S21}$$

The second and third conditions in Equation (S21) follow from the first condition since $K(x,t|0,u)$ and $K(x,t|1,u)$ are independent of u, by virtue of Equation (S20), and hence may be evaluated at $u = t$.

If we multiply $K(x,T|y,u)$ by $e^{-S_\infty x}$ and integrate from $x = 0$ to $x = 1$ we obtain the result $G(y,u) \overset{\text{def}}{\equiv} \int_0^1 e^{-S_\infty x}K(x,T|y,u)dx = E[e^{-S_\infty X(T)}|X(u) = y]$. The quantity $G(y,u)$ is often known as the Laplace-Stieltjes transform of $X(t)$, and in the present case it is evaluated at the solution of Equation (S1). The equation for $G(y,u)$ follows from Equation (S20) by multiplying $e^{-S_\infty x}$ and integrating from $x = 0$ to $x = 1$. It reads

$$\frac{\partial}{\partial u}G(y,u) = -\frac{y(1-y)}{4N_e(u)}\frac{\partial^2}{\partial y^2}G(y,u) - s(u)y(1-y)\frac{\partial}{\partial y}G(y,u) \tag{S22}$$

and is subject to (i) $G(y,T) = e^{-S_\infty y}$, (ii) $G(0,u) = 1$ and (iii) $G(1,u) = e^{-S_\infty}$ which follow from Equation (S21).

Equation (S22), subject to the initial condition (i), and the boundary conditions (ii) and (iii) is a well defined mathematical problem for $G(y,u)$ that can be solved by a standard numerical technique such as the Crank Nicholson method (see e.g., Press *et al.* 2007). Thus we 'integrate backwards' from $u = T$ to $u = 0$ and obtain $G(y,0) = E[e^{-S_\infty X(T)}|X(0) = y] \equiv E_y[e^{-S_\infty X(T)}]$ which, when used in Equation (S2), yields a numerical estimate of the fixation probability.

PART 6

The result in Equation (S2) for the fixation probability has a wider applicability than just for Equation (S1). In a more general case, we assume we assume there are both deterministic changes and stochastic fluctuations in $N_e(t)$ and $s(t)$ for times $t \leq T$, but only stochastic fluctuations for times $t > T$. Then the diffusion operator may be 'time averaged' and results in a time-independent form for $t > T$ that differs from Equation (S1). Time averaging over correlated fluctuations in the strength of selection was first carried out by Takahata *et al.* (1975) and we adopt the same approach here, implicitly assuming that the correlations persist over many generations.

To demonstrate this in the simplest way possible, we transform the diffusion equation using the drift time $\tau = \int_0^t \frac{1}{4N_e(u)}du$ (which differs from the drift time defined in Equation (S16) by an overall constant factor). This drift time eliminates the factor $4N_e(0)$ on the left hand side of Equation (S18) but otherwise leaves it unchanged in form. As already stated, the quantity $R(\tau)$, by virtue of its position in the transformed diffusion equation plays the role of an effective strength of selection. We can directly follow the approach of Takahata *et al*. (1975), making similar assumptions:

(i) we assume that $R(\tau)$ has bounded fluctuations around a mean value of $\bar{R}$

(ii) with overbars denoting ensemble averages over fluctuations, we assume the correlations of $R(\tau)$ obey $\int_0^\tau \overline{\left(R(\sigma+\sigma_1)-\bar{R}\right)\left(R(\sigma)-\bar{R}\right)}d\sigma = V$ for all σ_1 larger than some small correlation time.

The average of $\widetilde{K}(x,\tau|y,\sigma)$ over fluctuations is written $\bar{K}(x,\tau|y,\sigma)$ and this obeys (cf. Takahata *et al*. 1975)

$$\begin{aligned} -\frac{\partial}{\partial\tau}\bar{K}(x,\tau|y,\sigma) &= -\frac{\partial^2}{\partial x^2}\left\{\left[x(1-x)+Vx^2(1-x)^2\right]\bar{K}(x,\tau|y,\sigma)\right\} \\ &\quad +\frac{\partial}{\partial x}\left\{\left[\bar{R}x(1-x)+Vx(1-x)(1-2x)\right]\bar{K}(x,\tau|y,\sigma)\right\}. \end{aligned} \tag{S23}$$

The eigenvalue equation of the backward diffusion operator associated with Equation (S23) is

$$-y(1-y)\left[1+Vy(1-y)\right]\frac{d^2}{dy^2}\overline{\Psi}(y)-y(1-y)\left[\bar{R}+V(1-2y)\right]\frac{d}{dy}\overline{\Psi}(y)=\lambda\overline{\Psi}(y). \tag{S24}$$

To employ the result in Equation (S12) within Part 1 of the Supporting Information, we require the eigenfunction $\overline{\Psi}_1(y)$ with eigenvalue 0 that obeys $\overline{\Psi}_1(0)=0$ and $\overline{\Psi}_1(1)=1$. This is $\overline{\Psi}_1(y)=\dfrac{1-\left(\frac{1-y/\alpha_+}{1-y/\alpha_-}\right)^\beta}{1-\left(\frac{1-1/\alpha_+}{1-1/\alpha_-}\right)^\beta}$ where $\alpha_\pm=\dfrac{1\pm\sqrt{1+4/V}}{2}$ and $\beta=\dfrac{\bar{R}}{V\sqrt{1+4/V}}$. The form of $\overline{\Psi}_1(y)$ is equivalent to Equation (27) of TAKAHATA *et al*. (1975) under the substitutions $\bar{s}\rightarrow\bar{R}$ and $N\rightarrow 1/4$.

The generalisation of Equation (S2), that includes fluctuations in $N_e(t)$ and $s(t)$ is then given by $P_{\text{fix}}(p)=E_p\left[\,\overline{\Psi}_1(X(T))\right]$.

LITERATURE CITED IN THE SUPPORTING INFORMATION

Ewens, W. J., 1964 The pseudo-transient distribution and its uses in genetics. J. Appl. Probab. 1: 141–156.

Kimura, M., 1955 Stochastic processes and distribution of gene frequencies under natural selection. Cold Spring Harbour Symp. Quant. Biol. 20: 33–53.

McKane, A. J. and Waxman, D., 2007 Singular solutions of the diffusion equation of population genetics. J. Theor. Biol. 247: 849–858.

Press, W. H., Teukolsky, S. A, and Vetterling, W. T., 2007 *Numerical Recipes 3rd Edition: The Art of Scientific Computing.* Cambridge University Press, Cambridge.

Takahata, N., Ishii, K. and Matsuda, H., 1975 Effect of temporal fluctuation of selection coefficient on gene frequency in a population. Proc. Natl. Acad. Sci. USA 72: 4541–4545.

Waxman, D., 2011 Comparison and content of the Wright–Fisher model of random genetic drift, the diffusion approximation, and an intermediate model. J. Theor. Biol. 269: 79–87.