

Population Growth Enhances the Mean Fixation Time of Neutral Mutations and the Persistence of Neutral Variation

D. Waxman¹

Centre for Computational Systems Biology, School of Mathematical Sciences, Fudan University, Shanghai 200433, People's Republic of China

ABSTRACT A fundamental result of population genetics states that a new mutation, at an unlinked neutral locus in a randomly mating diploid population, has a mean time of fixation of $\sim 4N_e$ generations, where N_e is the effective population size. This result is based on an assumption of fixed population size, which does not universally hold in natural populations. Here, we analyze such neutral fixations in populations of changing size within the framework of the diffusion approximation. General expressions are derived for the mean and variance of the fixation time in changing populations. Some explicit results are given for two cases: (i) the effective population size undergoes a sudden change, representing a sudden population expansion or a sudden bottleneck; (ii) the effective population changes linearly for a limited period of time and then remains constant. Additionally, a lower bound for the mean time of fixation is obtained for an effective population size that increases with time, and this is applied to exponentially growing populations. The results obtained in this work show, among other things, that for populations that increase in size, the mean time of fixation can be enhanced, sometimes substantially so, over $4N_{e,0}$ generations, where $N_{e,0}$ is the effective population size at the time the mutation arises. Such an enhancement is associated with (i) an increased probability of neutral polymorphism in a population and (ii) an enhanced persistence of high-frequency neutral variation, which is the variation most likely to be observed.

HOW are properties of mutations at neutral loci, such as the mean time of fixation, modified when population size is not constant? The answer to this question directly influences the persistence of neutral variation in a population and the probability of neutral polymorphism. To put the question in context, we note that the genomes of many organisms contain loci and sites that can be described as neutral, with most substitutions seeming to have a neutral effect on the fitness of the host organism (see, e.g., Kimura 1983; Gillespie 1994, Chap. 6; Nei *et al.* 2010).

When mutation produces a new allele, its long-term outcome (neglecting additional mutations) is either loss or fixation. Henceforth, we restrict attention to mutations at unlinked neutral loci in randomly mating diploid populations. Typically, a new allele at such a locus does not achieve high relative frequencies and becomes lost in a short time

(Kimura and Ohta 1969; Crow and Kimura 1970). A rarer outcome is fixation, which typically takes a relatively long time to occur. With N_e the (variance) effective population size, alleles that start at low relative frequencies take a mean time to fix of $\sim 4N_e$ generations (Kimura and Ohta 1969), under the implicit assumption that the effective population size is constant. An allele that fixes generally achieves intermediate and high relative frequencies on the way to fixation and hence contributes to the neutral variation of the population for a time of the order of the *mean fixation time*. Related to this is neutral polymorphism, which arises from recurrent neutral mutation and random genetic drift. The likelihood of observing neutral polymorphism in a population is governed by the size of the product of mean fixation time and *neutral genomic mutation rate* (Kimura 1983). Thus there are direct ways the mean fixation time affects a population.

The above ideas form a standard part of our understanding of neutral mutations and connect with a larger body of work describing the effects of random genetic drift at the gene and molecular levels (Crow and Kimura 1970; Kimura 1983). Here we present an analysis of random genetic drift in situations involving neutral fixation in populations of

Copyright © 2012 by the Genetics Society of America
doi: 10.1534/genetics.112.139220

Manuscript received February 2, 2012; accepted for publication March 7, 2012

¹Corresponding author: School of Mathematical Sciences, Fudan University, 220 Handan Rd., Shanghai 200433, People's Republic of China. E-mail: davidwaxman@fudan.edu.cn

changing size. Such an analysis is required to establish how fundamental results, such as the “ $4N_e$ ” result for the mean time to fixation, become modified when population size is not constant and whether such results have an impact on neutral effects in a population.

Background

Fixation is a process that takes a random time to occur. For an unlinked neutral locus, the distribution of fixation times is approximately known when population size is constant (Kimura and Ohta 1969). It is recognized, however, that the assumption of fixed population size does not universally hold. Many natural populations have a size that significantly changes over time. In a study of the effects of changing population size on the fixation of nonneutral mutations, Otto and Whitlock (1997) listed the following as possible reasons for changing population numbers in the wild: (i) varying physical conditions, (ii) resource availability, (iii) habit availability, (iv) predator density, and (v) human disturbance. Human beings constitute an example of a population whose size is changing at a varying rate of growth.

To illustrate some of the issues associated with changing population size, let us consider examples involving fixation of a neutral allele.

Examples

Assume first that a single mutant allele arises in a population of fixed size N and that the allele ultimately fixes. From the first appearance of the allele to its fixation, which we assume takes T generations, all $2N$ resident alleles of the population are lost (strictly, the number lost is $2N - 1$, given one initial mutant allele) and are replaced by mutant alleles. The disappearance of resident alleles thus occurs at an average rate of $2N/T$ alleles per generation. Using the estimate $T = 4N_e$ generations leads to an average rate of loss of resident alleles of $N/(2N_e)$ alleles per generation. This simple picture becomes complicated if the mutant allele arises and goes to fixation in a population of changing size. To see this, consider a population whose size always remains close to its carrying capacity, which, for reasons external to the population, changes with time. We assume the carrying capacity exhibits the simplest nontrivial behavior, namely a linear change for a limited period of time. During this limited period, the population size is given by $N(t) = N_0 + \gamma t$, where γ is the constant rate of change of population size. We consider the implications of population sizes that either decrease ($\gamma < 0$) or increase ($\gamma > 0$).

Generally, the effective population size at any time t , written $N_e(t)$, will change over time. For the purposes of this example, we assume the simplest relation between the actual and effective population sizes and take $N_e(t) = \text{constant} \times N(t)$. Such an assumption is consistent with results from static populations. For example, in a population with a biased sex ratio, the quantity N/N_e depends solely on the sex ratio (see, e.g., Gillespie 2004, p. 49).

Let us now consider a mutant allele that arises at time $t = 0$ and that ultimately fixes in a population of linearly changing size. Fixation of the mutant allele occurs when all resident alleles are lost. We thus focus on the overall rate of loss of resident alleles, which arises from a sum of two terms: (i) a contribution caused by population size changes and (ii) a contribution arising from loss of the alleles due to random genetic drift.

In the early stages of the fixation process, residents comprise the vast majority of individuals in the population, and a rate of change of population size of γ (which may be positive or negative) closely corresponds to resident alleles changing at the rate of 2γ alleles per generation and hence of resident alleles being lost at the rate of -2γ alleles per generation. Thus we take -2γ as the contribution of population size change to the overall rate of loss of resident alleles; it may be positive or negative, depending on the sign of γ .

In a population of constant size, the rate of loss of resident alleles, due to random genetic drift, was estimated above to be $N/(2N_e)$ alleles per generation. In a changing population, the corresponding rate of loss of resident alleles will generally differ from this result, but we use $N/(2N_e)$ as an estimate of this rate.

Under population decrease ($\gamma < 0$) the contributions to the overall rate of loss of resident alleles, from random genetic drift and population size change, are $N/(2N_e)$ and $-2\gamma \equiv 2|\gamma|$, respectively, and are both positive. As a consequence the instantaneous rate of loss of resident alleles, when the population size is N , will typically be faster than that of a population of fixed size N .

Under population increase ($\gamma > 0$), the overall rate of loss of resident alleles has a negative contribution from population size change (estimated above to be -2γ) and a positive contribution from random genetic drift [estimated above to be $N/(2N_e)$]. Thus when the population size is N , the total rate of loss of resident alleles is $\sim N/(2N_e) - 2\gamma$, i.e., slower than that of a population of fixed size N . A particularly interesting regime is where the rate of loss of resident alleles due to random genetic drift can be overcome by the effect of population size increase, which increases the number of resident alleles. This will occur when γ is sufficiently large compared with N/N_e that the number of resident alleles does not decrease but typically changes at a net positive rate, i.e., increases. We should then expect that fixation could take much longer to occur. Indeed if the population growth persisted at such a rate indefinitely, the mean time to fixation could, plausibly, become infinite.

From the above examples it is reasonable to infer that the mean time to fixation is generally sensitive to a changing population size.

Methods

To address the issue of neutral fixation in a changing population, we separate the problem into two components:

- i. The first component is a statistic, associated with fixation, which has the properties that its distribution is independent of the pattern of population-size changes over time. The statistic introduced has a distribution that depends on a single parameter (the initial frequency of the mutant allele).
- ii. The second component is a function that encapsulates the way that population size changes over time. This function is independent of the fixation time.

In this way we cleanly separate a *common* feature of fixation (the distribution of the statistic introduced here) from the pattern of size changes of the *particular* population under consideration.

Model

Consider an unlinked neutral locus within the genome of a randomly mating diploid population with discrete generations. We assume the population-size changes over time, such that in generation t there are $N(t)$ adults. The processes that lead from the adults of one generation to the adults of the next generation are taken to be (i) reproduction, shortly followed by the death of all adults; (ii) viability selection; and (iii) nonselective ecological thinning (which is required, on the assumption that more individuals survive selection than can be supported by the environment). The thinning process reduces the number of individuals in the population to the $N(t + 1)$ adults of the next generation. It is step iii where random genetic drift can be identified as occurring. Alternative assumptions about step iii may alter the way random genetic drift affects the population.

Associated with random genetic drift is the *variance effective population size* (Ewens 2004), which we assume also changes with time. The variance effective population size in generation t is denoted by $N_e(t)$. This is determined from an averaging over processes that occur in a *single* generation. Throughout this work we use $N_e(t)$ to refer only to such a *local* quantity, that is, one associated with generation t . In particular, we do not make use of averages of the effective population size (such as the harmonic mean) that summarize properties of the effective population size over multiple generations and hence reflect population-size information that is *nonlocal* in time. For generality, we leave the relation between the actual population size, $N(t)$, and the effective population size, $N_e(t)$, arbitrary. It is very natural, however, to regard the effective population size as a proxy for the actual population size and hence assume that changes in $N(t)$ are accompanied by closely related changes in $N_e(t)$.

The question we address concerns features of the fixation process when the actual and effective population sizes have *known historical variation* over time. Thus we view $N(t)$ and $N_e(t)$ as being completely determined and not having any stochasticity. Accordingly, it would be inappropriate to carry out any averaging over these functions. Addressing questions of a different nature may require population-size averaging (see, e.g., Parsons and Quince 2007; Engen *et al.* 2009; Parsons *et al.* 2010; Uecker and Hermisson 2011; Waxman 2011c).

Let the locus in question have two alleles, denoted A and a , with a the wild-type allele and A a mutant allele that is the focus of interest.

In the adults of generation t , the proportion of all genes at the locus that are allele A , written $X(t)$, is the relative frequency of the A allele (henceforth termed the *frequency*). The frequency $X(t)$ generally exhibits random variation over time, due to the uncertainties of individuals contributing their genes to the following generation. The statistical properties of $X(t)$ can be obtained by considering many different replicates of a population with identical population-size behaviors. Initially, all replicates start with the same frequency of the A allele, but after the initial time different replicates may, due to stochasticity of the dynamics, have different values of the frequency. An extension of the Wright–Fisher model (Fisher 1930; Wright 1931), which takes into account changes in population size over time, can be used to calculate statistical results for $X(t)$, but this is mathematically complicated and normally leads only to numerical results. To make theoretical progress we perform an analysis on the basis of the diffusion approximation, where time t and frequency $X(t)$ are both treated as continuous quantities. The frequency $X(t)$ becomes a continuous random variable and is described by a probability density whose value at frequency x and time t is $f(x, t)$. The diffusion approximation derives its name from the *diffusion* equation governing $f(x, t)$. For an unlinked neutral locus in a randomly mating diploid population this reads $-\partial f(x, t)/\partial t = -[4N_e(t)]^{-1} \partial^2[x(1-x)f(x, t)]/\partial x^2$ (Crow and Kimura 1970), where, as stated above, $N_e(t)$ is the effective population size that is appropriate for generation t .

Let the A allele have a frequency of y at the time of its initial appearance, which we take throughout this work to be $t = 0$. Thus if n copies of the A allele are present in a population at the initial time, then the initial frequency is given by $y = n/(2N(0))$. At large times the A allele will have either fixed, with probability y , or been lost, with probability $1 - y$. These probabilities hold irrespective of the way that the effective population size changes after appearance of the mutation (see *Appendix A*).

Statistic

More complex information about the evolutionary process is contained in the description of the random *time* it takes for fixation or loss of the A allele to occur. If fixation ultimately occurs, the random time to fixation, denoted T_{fix} , has a probability distribution that depends on the behavior of the effective population size over time. We have found it advantageous to work with a statistic that is associated with the fixation time, but whose distribution is *independent* of the behavior of $N_e(t)$, and hence has an *invariant form*. The statistic, denoted R , is defined by

$$R = \int_0^{T_{\text{fix}}} \frac{dv}{4N_e(v)} \quad (1)$$

and explicitly involves T_{fix} . In *Appendix B* we establish that under the diffusion approximation, the statistic R has the

property that its distribution depends upon just one parameter, namely the frequency of the A allele at its time of appearance ($t = 0$). Let us briefly note the implications of this property for two populations where the A allele starts with the same initial frequency and ultimately fixes, but the two populations have different effective sizes over time. Despite differences in $N_e(t)$, all statistical properties of the two populations that do not explicitly depend on $N_e(t)$ and involve only R (for example, the mean or variance of R) will be identical for both populations. The independence of the distribution of R on the behavior of $N_e(t)$ contrasts sharply with statistical properties of fixation times, which will generally be different for the two populations. Before we illustrate the different statistical properties of the fixation time in populations with different $N_e(t)$, it is necessary to fully characterize R .

Generally, all statistical properties of R are summarized by its cumulative distribution function, $F_y(r)$, which depends on the initial frequency (y) as a parameter. The cumulative distribution function is defined as the probability

$$F_y(r) = \text{Prob}(R \leq r | X(0) = y) \quad (2)$$

with $0 \leq r < \infty$. In *Appendix B* we establish that the diffusion approximation yields

$$F_y(r) \approx 1 - 2 \sum_{n=0}^{\infty} \frac{(1-y)(2n+3)}{(n+1)(n+2)} (-1)^n C_n^{(3/2)} (1-2y)e^{-(n+1)(n+2)r}, \quad (3)$$

where $C_n^{(3/2)}(y)$ is a particular polynomial of y [a Gegenbauer polynomial of order $3/2$ and degree n (Abramowitz and Stegun 1964)]. In the practically important case of low initial frequencies ($y \rightarrow 0$) we write the cumulative distribution as $F_0(r)$ and this has the simple approximation

$$F_0(r) \approx \begin{cases} \left(\frac{\pi}{r}\right)^{3/2} e^{r/4 - \pi^2/4r}, & 0 \leq r \leq 2 \\ 1 - 3e^{-2r}, & r > 2 \end{cases} \quad (4)$$

(see *Appendix B*). The expressions in Equations 3 and 4 follow from existing results in the literature (Kimura 1955; Waxman 2011b).

For an initial frequency of y , the mean and variance of R are written as $E_y[R]$ and $\text{Var}_y(R)$. These have simple forms under the diffusion approximation (see *Appendix C* for details):

$$E_y[R] = \int_0^{\infty} [1 - F_y(r)] dr \approx -\frac{(1-y)\ln(1-y)}{y} \quad (5)$$

and

$$\text{Var}_y(R) \approx \frac{2(1-y)\ln(1-y)}{y} - \left(\frac{(1-y)\ln(1-y)}{y}\right)^2 + \frac{\pi^2}{3} - 2 \text{dilog}(1-y), \quad (6)$$

where $\text{dilog}(x) = \int_1^x (\ln(u)/(1-u)) du$ denotes the dilogarithm function (Abramowitz and Stegun 1964).

The results of Equations 5 and 6 reduce, in the limit of small y ($y \rightarrow 0$), to $E_0[R] = 1$ and $\text{Var}_0(R) \approx (\pi^2/3 - 3)$.

In *Appendix D* we use simulations and the results of Equations 5 and 6 to illustrate the prediction of the diffusion approximation, that statistical properties of R depend only on the initial frequency of the A allele and not on the behavior of the effective population size.

Results

Statistical properties of the time to fixation

Equation 1 relates the statistic R to another random quantity—the fixation time, T_{fix} . Equation 1 can always be solved (by either analytical or numerical methods) to yield T_{fix} as a function of R and we write this relation as

$$T_{\text{fix}} = h(R). \quad (7)$$

Expectations involving the fixation time follow from knowledge of $h(r)$ and $F_y(r)$. For example, with a an arbitrary positive number, the mean of $(T_{\text{fix}})^a$, namely $E_y[(T_{\text{fix}})^a]$, can be written as $E_y[(T_{\text{fix}})^a] \approx \int_0^{\infty} [h(r)]^a [\partial F_y(r)/\partial r] dr$. This result can be used to derive expressions for statistical properties of T_{fix} such as its mean and variance (see *Appendix E* for details). In particular, the mean and the variance of the time to fixation are given by

$$E_y[T_{\text{fix}}] \approx \int_0^{\infty} h(r) \frac{\partial F_y(r)}{\partial r} dr = \int_0^{\infty} 4N_e(h(r)) [1 - F_y(r)] dr \quad (8)$$

$$\text{Var}_y(T_{\text{fix}}) = 2 \int_0^{\infty} h(r) \frac{dh(r)}{dr} [1 - F_y(r)] dr - \left(\int_0^{\infty} \frac{dh(r)}{dr} [1 - F_y(r)] dr \right)^2. \quad (9)$$

Equation 8 is the generalization of the constant population result to effective population sizes that are time dependent.

When the effective population size is constant, Equation 8 leads to $E_y[T_{\text{fix}}] \approx -4N_e \times (1-y)\ln(1-y)/y$ and for a small initial frequency ($y \rightarrow 0$) this result reduces to the standard result $E_0[T_{\text{fix}}] \approx 4N_e$ (Kimura and Ohta 1969). Furthermore, for a constant population size and $y \rightarrow 0$, the variance reduces to the known result $\text{Var}_0(T_{\text{fix}}) \approx (\pi^2/3 - 3) \times (4N_e)^2$ (Kimura 1970).

The form of the function $h(r)$, which appears in Equations 7–9 and relates T_{fix} to R , depends specifically on the *history* of $N_e(t)$, from the time the mutant allele appears (at $t = 0$) onward. We next provide details of two examples, following from known forms of $N_e(t)$, to illustrate the behavior of the mean and variance of T_{fix} . We also present two different inequalities for $E_y[T_{\text{fix}}]$ that do not require knowledge of the full history of $N_e(t)$. One inequality applies when $N_e(t)$ is known to lie within a finite range of values for all times,

and the other inequality provides a lower bound on $E_y[T_{\text{fix}}]$ when the population size increases over time.

Examples

Example 1: For the first example, consider the situation where $N_e(t)$ has a discontinuous jump in size at time τ (see curve b in Figure D1, Appendix D):

$$N_e(t) = \begin{cases} N_{e,0}, & t \leq \tau, \\ N_{e,1}, & t > \tau. \end{cases} \quad (10)$$

In Appendix F it is shown that Equations 1, 7, and 8 lead to a mean fixation time of

$$E_y[T_{\text{fix}}] \approx 4N_{e,0} \int_0^\rho [1 - F_y(r)] dr + 4N_{e,1} \int_\rho^\infty [1 - F_y(r)] dr, \quad (11)$$

where ρ is a measure of the time of population-size jump:

$$\rho = \frac{\tau}{4N_{e,0}}. \quad (12)$$

Note that because $N_{e,0}$ and $N_{e,1}$ are arbitrary, Equation 11 can apply to a population experiencing a *sudden expansion* ($N_{e,1} > N_{e,0}$) or a *sudden bottleneck* ($N_{e,1} < N_{e,0}$).

In Appendix F we also derive an expression for the variance of T_{fix} , which has a more complex form than that of Equation 11. To gain intuition about Equation 11, let us consider some informative extreme limits.

If $N_{e,0} \rightarrow \infty$ with $N_{e,1}$ held constant, Equation 11 reduces to $E_y[T_{\text{fix}}] = \tau + 4N_{e,1}E_y[R]$ and corresponds to fixation occurring only after time τ and, then, only from a frequency of y . This has the explanation that an effectively infinite effective population size up to time τ prevents any change of the allele frequency from its initial value; hence fixation occurs only after the population size has become finite. Alternatively, if $N_{e,1} \rightarrow \infty$ with $N_{e,0}$ held constant, then the mean time to fixation diverges. This arises because there is a finite probability that a mutant allele does not fix before time τ ; the allele then finds itself in an effectively infinite population and will not fix in any finite time. In the methods presented in this work, this result arises because $4N_{e,1}$, in Equation 11, is multiplied by the finite probability that R will be larger than ρ .

In the limit of low initial frequencies ($y \rightarrow 0$), the mean and variance of T_{fix} become $E_0[T_{\text{fix}}]$ and $\text{Var}_0(T_{\text{fix}})$, respectively, and we can obtain approximations for these quantities. For small, intermediate, and relatively large values of ρ we find that

$$E_0[T_{\text{fix}}] \approx \begin{cases} \rho \times 4N_{e,0} + (1 - \rho) \times 4N_{e,1}, & \rho \ll 1, \\ 0.50 \times 4N_{e,0} + 0.50 \times 4N_{e,1}, & \rho \approx 0.51, \\ 0.80 \times 4N_{e,0} + 0.20 \times 4N_{e,1}, & \rho = 1, \\ \left(1 - \frac{3e^{-2\rho}}{2}\right) \times 4N_{e,0} + \frac{3e^{-2\rho}}{2} \times 4N_{e,1}, & \rho \geq 2 \end{cases} \quad (13)$$

(see Appendix F for details.) An inspection of Equation 13 indicates that when the scaled time of population-size jump, ρ , is small ($\rho \ll 1$, corresponding to $\tau \ll N_{e,0}$), the mean fixation time approximately equals $4N_{e,1}$ and hence is largely determined by the population size achieved at *later* times. This is naturally attributed to most replicate populations not having fixed the A allele by the time of the population-size jump. The intermediate value of ρ used in Equation 13, namely $\rho \approx 0.51$, was chosen so that $4N_{e,0}$ and $4N_{e,1}$ make *equal* contributions to $E_0[T_{\text{fix}}]$. The result obtained is very similar to the result obtained by using this value of ρ in the small ρ result in the preceding line. For relatively large values of ρ ($\rho \geq 2$, corresponding to $\tau \geq 8N_{e,0}$) the mean fixation time approximately equals $4N_{e,0}$ and hence is largely determined by the population size achieved at *early* times; this is naturally attributed to most replicate populations having fixed the A allele by the time of the population-size jump.

The variance of T_{fix} can also be approximated for low initial frequencies ($y \rightarrow 0$). We find (see Appendix F for details)

$$\text{Var}_0(T_{\text{fix}}) \approx \begin{cases} \left(\frac{\pi^2}{3} - 3\right) \times (4N_{e,1})^2, & \rho \ll 1, \\ 0.27 \times (4N_{e,1})^2 + 0.01 \times (4N_{e,0})(4N_{e,1}), & \rho \approx 0.51, \\ 0.05 \times (4N_{e,0})^2 + 0.08 \times (4N_{e,0})(4N_{e,1}) + 0.16 \times (4N_{e,1})^2, & \rho = 1, \\ \left(\frac{\pi^2}{3} - 3 + \frac{3}{2}(1 - 2\rho)e^{-2\rho}\right) \times (4N_{e,0})^2 + \frac{3}{2}e^{-2\rho} \times (4N_{e,1})^2, & \rho \geq 2. \end{cases} \quad (14)$$

The variance obtained for $\rho \approx 0.51$ is close to the result obtained using this value in the small ρ result (since $\pi^2/3 - 3 \approx 0.29$ is close to 0.27). It is somewhat surprising that the variance, when $\rho = 1$, has the largest weighting given to the *later* population size ($N_{e,1}$), given that the mean of T_{fix} for this value of ρ (see Equation 13) has the largest weighting given to the contribution of the *earlier* population size ($N_{e,0}$).

Example 2: For the second example, consider the situation outlined at the start of this article, where the effective population size was assumed to change linearly, for a finite period of time, and then maintain a constant value. We present only the details of an analysis for an *increasing* population. A formal solution is justified in this case since the issues involved are potentially complex (see Background).

We write

$$N_e(t) = \begin{cases} N_{e,0} + \gamma_e t, & t \leq \tau, \\ N_{e,0} + \gamma_e \tau, & t > \tau, \end{cases} \quad (15)$$

where the rate of change, γ_e , is taken to be positive. We find (see Appendix G for details)

$$E_y[T_{\text{fix}}] \approx 4N_{e,0} \times \left(\int_0^{\tau_0} e^{4\gamma_e r} [1 - F_y(r)] dr + e^{4\gamma_e \tau_0} \int_0^\infty [1 - F_y(r)] dr \right), \quad (16)$$

where $r_0 = (4\gamma_e)^{-1} \log(1 + \gamma_e \tau / N_{e,0})$. Appendix G contains an expression for the variance of T_{fix} , which again has a more complex form than that of the mean.

Restricting considerations to low initial frequencies, where the mean and variance of T_{fix} become $E_0[T_{\text{fix}}]$ and $\text{Var}_0(T_{\text{fix}})$, respectively, the results for these can, using Equation 4, be numerically evaluated. The dependence of $E_0[T_{\text{fix}}]$ on the duration of linear change, τ , is affected by the rate of change of the effective population size, γ_e , as illustrated in Figure 1A. The dependence of the variance on τ is also affected by γ_e , and to facilitate direct comparison with $E_0[T_{\text{fix}}]$, the square root of the variance is plotted as a function of τ in Figure 1B.

Approximate results can also be obtained for the dependence of $E_0[T_{\text{fix}}]$ and $\text{Var}_0(T_{\text{fix}})$ on τ . When γ_e has a fixed value and τ is sufficiently large that it satisfies both $\gamma_e \tau / (4N_{e,0}) > e^{8\gamma_e} - 1$ and $\gamma_e \tau / (4N_{e,0}) \gg 1$, we find the following approximate τ dependences:

$$E_0[T_{\text{fix}}] \approx k_1 + k_2 \left(\frac{\gamma_e \tau}{N_{e,0}} \right)^p \quad (17)$$

$$\text{Var}_0(T_{\text{fix}}) \approx k_3 + k_4 \left(\frac{\gamma_e \tau}{N_{e,0}} \right)^p + k_5 \left(\frac{\gamma_e \tau}{N_{e,0}} \right)^{2p} + k_6 \left(\frac{\gamma_e \tau}{N_{e,0}} \right)^{1+p}, \quad (18)$$

where k_1, k_2, \dots, k_6 are independent of τ and p is the parameter

$$p = \frac{\gamma_e - 1/2}{\gamma_e} \quad (19)$$

(see Appendix G for details). The three powers of τ appearing in Equations 17 and 18, namely p , $2p$, and $1 + p$, are all negative for $\gamma_e < 1/4$, but for $1/4 < \gamma_e < 1/2$, p and $2p$ are negative while $1 + p$ is positive, and for $\gamma_e > 1/2$ all three powers are positive. This leads to $E_0[T_{\text{fix}}]$ and $\text{Var}_0(T_{\text{fix}})$ having dependences on τ (when τ is large), which is sensitive to the value of γ_e . When $\gamma_e < 1/4$, both $E_0[T_{\text{fix}}]$ and $\text{Var}_0(T_{\text{fix}})$ vary as τ to a negative power, indicating a weak dependence on τ . However, when $1/4 < \gamma_e < 1/2$, it follows that $E_0[T_{\text{fix}}]$ varies as τ to a negative power while $\text{Var}_0[T_{\text{fix}}]$ has a term varying as τ to a positive power. Thus the variance has a potentially strong dependence on τ in this regime. Finally, when $\gamma_e > 1/2$, both $E_0[T_{\text{fix}}]$ and $\text{Var}_0(T_{\text{fix}})$ vary as positive powers of τ , indicating that both have potentially a strong dependence on τ .

Bounds on the mean

The above results are based on explicit forms for $N_e(t)$. To make use of Equation 8, for the mean time to fixation, requires full knowledge of $N_e(t)$, so that the function $h(r)$ can be determined. We can circumvent the requirement of full knowledge of $N_e(t)$ at the cost of obtaining bounds, not precise values. The bounds obtained are based on the diffusion approximation and hence are not rigorous inequalities for $E_y[T_{\text{fix}}]$:

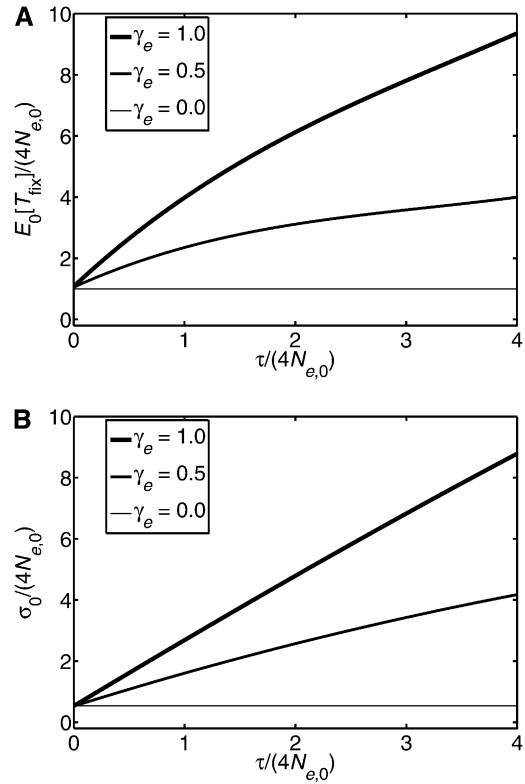


Figure 1 The case where the effective population size has a period of linear growth of duration τ before remaining constant (Equation 15). (A) Illustration of the behavior of the mean time to fixation, as a function of τ , for mutant alleles with low initial frequency. It follows from Equation 16 that if both $\gamma_e \tau / N_{e,0}$ and $\tau / N_{e,0}$ are small ($\ll 1$), then $E_0[T_{\text{fix}}] \approx 4N_e(\tau) \approx 4(N_{e,0} + \gamma_e \tau)$; i.e., $E_0[T_{\text{fix}}]$ increases linearly with τ . For larger τ it is apparent from A that $E_0[T_{\text{fix}}]$ increases with τ , but more slowly than linearly. (B) Illustration of the behavior of the square root of the variance, $\sigma_0 = \sqrt{\text{Var}_0(T_{\text{fix}})}$, as a function of τ , for mutant alleles with low initial frequency.

1. For the first bound, note that Equation 8 can be written in the alternative form $E_y[T_{\text{fix}}] \approx 4\bar{N}_e \times E_y[R]$, where $E_y[R]$ is given in Equation 5, and \bar{N}_e represents a particular sort of average of $N_e(t)$, namely $\bar{N}_e = \int_0^\infty N_e(h(r)) [1 - F_y(r)] dr / \int_0^\infty [1 - F_y(r)] dr$. If $N_e(t)$ is known to always lie within the range $N_{e,\min}$ to $N_{e,\max}$ it immediately follows that \bar{N}_e , by virtue of being an average, must lie in the range $N_{e,\min} < \bar{N}_e < N_{e,\max}$. It follows that the mean time of fixation satisfies

$$4N_{e,\min} \times E_y[R] < E_y[T_{\text{fix}}] < 4N_{e,\max} \times E_y[R].$$

2. A different bound can be obtained for an *increasing* population. A mutation that appears at time $t = 0$ has a mean time of fixation that satisfies

$$E_y[T_{\text{fix}}] > \left(\frac{1}{L} \int_0^L 4N_e(v) dv \right) \times E_y[R], \quad (20)$$

where $L = 4N_e(0) \times E_y[R]$ (see Appendix H for details). Since $E_y[R] \leq 1$, the maximum value of L is $4N_e(0)$. In

the case of small initial frequencies we have $E_y[R] \approx E_0[R] = 1$ and hence $L \approx 4N_e(0)$.

The inequality in Equation 20 involves the average $L^{-1} \int_0^L 4N_e(v)dv$, which requires knowledge of $N_e(t)$ only over times $t = 0$ to $t = L$. If, for example, the effective population size exhibits exponential growth over this time interval, with $N_e(t) = N_e(0)e^{at}$, then Equation 20 leads to $E_y[T_{\text{fix}}]/4N_e(0) > (e^{aL} - 1) E_y[R]/aL$. If $a \times L$ is appreciable, then neutral mutations should have a mean time to fixation that is large relative to $4N_e(0)$.

Discussion

In this work we have considered statistical properties of the time it takes mutations at unlinked *neutral* loci to fix in populations of changing effective size. Such alleles contribute to the neutral variation of a population during the random time they are segregating. Even in populations of constant size, neutral alleles constitute a special case since they take longer (on average) to fix than beneficial or equally deleterious alleles under genic selection (Maruyama 1977, pp. 80 and 81). In the present work we show that the mean and variance of the time to fixation of neutral alleles are sensitive to the behavior of the effective population size. In particular, the mean time to fixation can be significantly increased beyond the constant population result when the effective size increases over time. In a population that increases over time, alleles that start with a low initial frequency have a mean time to fixation that is predicted to be $>4N_e(0)$, where $N_e(0)$ is the effective population size *at the time the mutation arises*. We use two of the results obtained in this work to illustrate this:

- i. Equation 17 indicates how long-term *linear growth* of the effective population size affects the mean time of fixation. If linear population growth continues for a very long time ($\tau \rightarrow \infty$), then Equation 17 shows that $\gamma_e = 1/2$ is a critical value of the growth rate, which separates *small* and potentially *large* values of $E_0[T_{\text{fix}}]$. When actual and effective population sizes coincide, the critical value of the growth rate (*i.e.*, $\gamma_e = 1/2$) corresponds to an increase by one new individual every *two* generations. This is a modest rate of growth. The *variability* of the fixation time also depends on the value of γ_e . When population growth continues for a very long time ($\tau \rightarrow \infty$), the variance, $\text{Var}_0[T_{\text{fix}}]$, exhibits different behavior from the mean. The variance changes character at $\gamma_e = 1/4$ from *small* to potentially *large* values [due to the term $(\gamma_e \tau / N_{e,0})^{1+p}$ in Equation 18]. This last feature, which corresponds to an increase of the population by one new individual every *four* generations, when actual and effective population sizes coincide, also illustrates that the mean and variance of T_{fix} can depend on parameters in quite different ways.
- ii. Equation 20 shows that a large mean time of fixation results if the population grows an appreciable amount in the first $4N_e(0)$ generations. For example, with exponential growth at rate a we obtained $E_y[T_{\text{fix}}]/4N_e(0) > (e^{aL} - 1)E_y[R]/aL$,

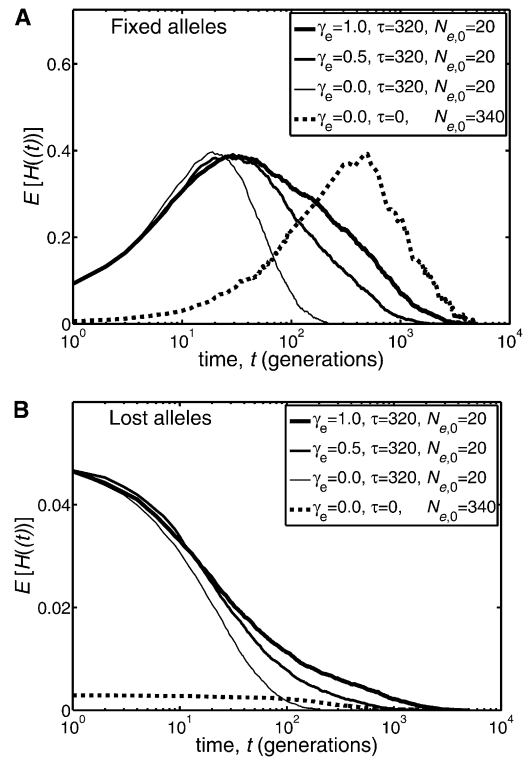


Figure 2 Illustration of the behavior of the mean heterozygosity $E[H(t)]$ [where $H(t) = 2X(t)(1 - X(t))$] as a function of the time, t , when the effective population size has a period of linear growth, as in Equation 15. The results presented were derived from simulations of 2×10^4 replicate populations, which were run until either fixation or loss occurred. Binomial sampling was used, and hence actual and effective population sizes coincided. A single copy of a mutant allele was assumed present at time $t = 0$. The heterozygosities plotted are conditional on the mutant alleles either fixing (A) or being lost (B). The solid lines correspond to an initial effective population size of $N_{e,0} = 20$ and a duration of linear growth of $\tau = 16 \times N_{e,0} = 320$ generations. For comparison, the dashed line corresponds to a population of *constant* effective size, whose value is the *largest* size achieved by a population with $\gamma_e = 1$ and $N_{e,0} = 20$; namely, $N_{e,0} = 20 + 320 = 340$.

where $L = 4N_e(0)E_y[R]$. Thus if the growth rate has the value $a = 0.05$ per generation, then an initial effective size as small as $N_e(0) = 20$ leads, for small initial frequencies, to $E_0[T_{\text{fix}}] > 10^3$ generations, which is $>4N_e(0)$ by more than a factor of 10.

The time a mutation takes to fix in a population is closely related to the persistence of variation in that population. A measure of *variation* in a population (which in the present work arises from the segregation of neutral alleles) is given by the mean heterozygosity, $E[H(t)]$, where $H(t) = 2X(t)(1 - X(t))$. To illustrate the relation between fixation and the persistence of variation consider a population with a period of linear growth as in Equation 15, with γ_e positive. Simulations of the heterozygosity are plotted as a function of time in Figure 2, for the case where a single *A* allele is initially present in the population.

For a fixed duration of linear change, τ , different values of the linear growth rate, γ_e , were used in Figure 2. The

mean heterozygosity is very different for mutant alleles that fix compared with those that are lost. Mutant alleles that fix lead to a relatively high level of mean heterozygosity (~ 0.4), which has a persistence that increases with γ_e . Mutant alleles that are lost lead to a mean heterozygosity that is relatively small and that persists for substantially less time than that of alleles that fix. For comparison, Figure 2 also contains a curve showing the heterozygosity of a population of constant effective size, whose value is the largest size achieved by the population with $\gamma_e = 1$ and $N_{e,0} = 20$, namely $N_{e,0} = 340$ (dashed curve). The comparison is partially confounded because the initial frequency used in this constant population corresponds to a single copy of a mutant allele and hence has the value $1/(2 \times 340)$, while the initial frequency in all of the populations that changed size (solid curves) was $1/(2 \times 20)$. Figure 2A suggests that the maximum value that the mean heterozygosity achieves, for a mutant allele that ultimately fixes, is insensitive to the population size; however, the time the heterozygosity persists, and the actual pattern realized, is sensitive to the behavior of the effective population size. The evidence from Figure 2A is that an increasing population enhances the persistence of the neutral variation associated with the most observable mutant alleles, that is to say those that achieve intermediate and high frequencies and proceed to fix. No such phenomenon is strongly apparent in the variation associated with alleles that typically remain at low frequencies and are ultimately lost (Figure 2B).

To establish the connection between the persistence of variation from *mutant alleles that proceed to fix* and the *mean fixation time*, consider the mean heterozygosity, given that fixation ultimately occurs. We write this as $E[H(t) | \text{Fix}]$. The only mutant alleles contributing to $E[H(t) | \text{Fix}]$ are those that fix *after* time t (i.e., that have $T_{\text{fix}} > t$). Accordingly, we have $E[H(t) | \text{Fix}] = E[H(t) | \text{Fix}, T_{\text{fix}} > t] \times \text{Prob}(T_{\text{fix}} > t)$, and from the simulations for Figure 2A, we find that $\text{Prob}(T_{\text{fix}} > t)$ is the factor in this expression that is *largely responsible* for the observed decrease of $E[H(t) | \text{Fix}]$ with t . That is, for the range of t where $E[H(t) | \text{Fix}]$ decreases, we find that $E[H(t) | \text{Fix}] \approx \text{constant} \times \text{Prob}(T_{\text{fix}} > t)$. It follows that the persistence of $E[H(t) | \text{Fix}]$ is largely determined by the behavior of $\text{Prob}(T_{\text{fix}} > t)$. Additionally, $E[T_{\text{fix}}]$ is fully determined from knowledge of $\text{Prob}(T_{\text{fix}} > t)$. Thus a changing population size can be viewed as most fundamentally affecting $\text{Prob}(T_{\text{fix}} > t)$, which then affects other quantities such as $E[H(t) | \text{Fix}]$ and $E[T_{\text{fix}}]$. From the simulations, an increasing population size is found to shift the decay of $\text{Prob}(T_{\text{fix}} > t)$ to larger values of t , and this explains why an increased mean time of fixation is accompanied by an enhanced persistence of variation.

This work has concerned itself with neutral mutations. The degree to which alleles affect the fitness of their carriers is determined by the environment. Selective neutrality may be a transient property, because under environmental changes, originally neutral loci may become selected. The associated variation then becomes converted to nonneutral

variation that becomes a target for selection, as has been discussed in the literature for, e.g., alleles associated with 6PG allozymes (Dykehuizen and Hartl 1980). Such nonneutral variation will generally be significant for the continued survival of a population. It could allow rapid adaptation to a changed environment or, perhaps, result in a very high load that leads to extinction. Either way, the potential to evolve and the direction of evolution are closely associated with the presence of neutral alleles in a population and hence on the time they take to achieve fixation.

Acknowledgments

I thank Sally Otto, Andy Overall, Joel Peck, and Tianqi Zhu for stimulating discussions and helpful suggestions for the manuscript. I also thank two anonymous reviewers for their comments and suggestions, which have improved this work.

Literature Cited

- Abramowitz, M., and I. Stegun, 1964 *Handbook of Mathematical Functions*. Dover, New York.
- Crow, J. F., and M. Kimura, 1970 *An Introduction to Population Genetics Theory*. Harper & Row, New York.
- Dykehuizen, D. E., and D. L. Hartl, 1980 Evolution of selective neutrality: further considerations. Selective neutrality of 6PGD allozymes in *E. coli* and the effects of genetic background. *Genetics* 96: 801–817.
- 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., 2004 *Mathematical Population Genetics*, Ed. 2. Springer, New York.
- Fisher, R. A., 1930 *The Genetical Theory of Natural Selection*. Clarendon Press, Oxford.
- Gillespie, J. H., 1994 *The Causes of Molecular Evolution*. Oxford University Press, New York.
- Gillespie, J. H., 2004 *Population Genetics: A Concise Guide*, Ed. 2. Johns Hopkins University Press, Baltimore.
- Gradsteyn, I. S., and I. M. Ryzhik, 1980 *Table of Integrals, Series and Products*. Academic Press, London.
- Kimura, M., 1955 Solution of a process of random genetic drift with a continuous model. *Proc. Natl. Acad. Sci. USA* 41: 141–150.
- Kimura, M., 1970 The length of time required for a selectively neutral mutant to reach fixation through random frequency drift in a finite population. *Genet. Res.* 15: 131–133.
- Kimura, M., 1983 *The Neutral Theory of Molecular Evolution*. Cambridge University Press, Cambridge, UK.
- Kimura, M., and T. Ohta, 1969 The average number of generations until fixation of a mutant gene in a finite population. *Genetics* 61: 763–771.
- Maruyama, T., 1977 *Stochastic Problems in Population Genetics* (Lecture Notes in Biomathematics, Vol. 17). Springer, Berlin.
- McKane, A. J., and D. Waxman, 2007 Singular solutions of the diffusion equation of population genetics. *J. Theor. Biol.* 247: 849–858.
- Nei, M., Y. Suzuki, and M. Nozawa, 2010 The neutral theory of molecular evolution in the genomic era. *Annu. Rev. Genomics Hum. Genet.* 11: 265–289.

- 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.
- Uecker, H., and J. Hermisson, 2011 On the fixation process of a beneficial mutation in a variable environment. *Genetics* 188: 915–930.
- Waxman, D., 2011a 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., 2011b A compact result for the time-dependent probability of fixation at a neutral locus. *J. Theor. Biol.* 274: 131–135.
- Waxman, D., 2011c A unified treatment of the probability of fixation when population size and the strength of selection change over time. *Genetics* 188: 907–913.
- Wright, S., 1931 Evolution in Mendelian populations. *Genetics* 16: 97–159.

Communicating editor: N. A. Rosenberg

Appendix A: Preliminaries and Probability of Fixation

To obtain analytical results for the Wright–Fisher model of an unlinked neutral locus, as described in the main text, we use the diffusion approximation. Under this approximation time and allele frequencies are approximated as continuous quantities. The probability density of the frequency of the *A* allele at time *t* and frequency *x*, written $f(x, t)$, obeys the diffusion equation

$$-\frac{\partial}{\partial t} f(x, t) = -\frac{1}{4N_e(t)} \frac{\partial^2}{\partial x^2} [x(1-x)f(x, t)] \quad (\text{A1})$$

(Crow and Kimura 1970), where $N_e(t)$ is the variance effective population size at time *t*. We solve Equation A1 subject to the condition that at time $t = 0$ all replicate populations start with the *A* allele at a frequency of *y*. This corresponds to the initial probability density

$$f(x, 0) = \delta(x - y), \quad (\text{A2})$$

where $\delta(x)$ denotes a Dirac delta function of argument *x*.

To make the solution $f(x, t)$ of Equation A1 mathematically well defined, it is necessary to impose conditions at the boundaries $x = 0$ and $x = 1$. We require that probability is conserved for all times ($\int_0^1 f(x, t) dx = 1$) and this leads us to impose the condition that the probability current, defined by $j(x, t) = -(1/4N_e(t))(\partial/\partial x)[x(1-x)f(x, t)]$, vanishes at the boundaries (McKane and Waxman 2007):

$$j(0, t) = 0, j(1, t) = 0. \quad (\text{A3})$$

A consequence of Equation A3 is that fixation and loss are included in the solution of Equation A1 in the form of Dirac delta functions located at $x = 0$ and $x = 1$ (McKane and Waxman 2007; Waxman 2011a). The *total integrated probability density*, which includes contributions of the delta functions at the boundaries, then has the value of unity for all times. This property of the solution of the diffusion equation is in accordance with the probability distribution of the Wright–Fisher model, which has a *sum over all frequency states* of unity.

Proof

We now prove that the probability of the ultimate fixation or loss of a neutral allele is independent of the way that the effective population size, $N_e(t)$, changes over time, subject to a mild restriction on $N_e(t)$.

Proceeding, we write $\bar{x}^n(t)$ for the expected value $E_y[X^n(t)] \stackrel{\text{def}}{=} E[X^n(t)|X(0) = y]$ and have $E_y[X^n(t)] = \int_0^1 x^n f(x, t) dx$. Multiplying Equation A1 by x or x^2 , integrating over all x , and using Equation A3 yields

$$\frac{d\bar{x}(t)}{dt} = 0 \quad (\text{A4})$$

and

$$\frac{d\bar{x}^2(t)}{dt} = \frac{\bar{x}(t) - \bar{x}^2(t)}{2N_e(t)}. \quad (\text{A5})$$

The initial frequency, $X(0)$, has the precise value of y since $f(x, t)$ satisfies Equation A2. Defining

$$Q(t) = \int_0^t \frac{1}{2N_e(v)} dv, \quad (\text{A6})$$

the solutions of Equations A4 and A5 are $\bar{x}(t) = y$ and $\bar{x}^2(t) = y - y(1-y)e^{-Q(t)}$. If

$$\lim_{t \rightarrow \infty} Q(t) = \infty, \quad (\text{A7})$$

then $\lim_{t \rightarrow \infty} \bar{x}^2(t) = y$. That is, at large t , $\bar{x}^2(t)$ converges to $\bar{x}(t) = y$. The only way this dispersionless convergence can come about is if $\lim_{t \rightarrow \infty} f(x, t) = (1-y)\delta(x) + y\delta(1-x)$. The coefficient of $\delta(1-x)$ in this expression is the long time fixation probability (McKane and Waxman 2007) and equals the initial frequency, y , of the A allele. Similarly, the long time probability of loss is $1-y$. These long time results for the probabilities of fixation and loss are independent of $N_e(t)$ and hence apply for static or time-dependent N_e , providing only that the condition of Equation A7 holds. This condition is fairly mild: for example, any $N_e(t)$ that is bounded from above [$N_e(t) < \infty$ for all t] will lead to Equation A7 holding. Populations that grow too rapidly at large t , such as exponentially growing populations, will violate Equation A7, but this should not be problematic since such indefinite growth is not biologically feasible.

Appendix B: Establishing Properties of the R Statistic

In this *Appendix* we establish that under the diffusion approximation, the statistic

$$R = \int_0^{T_{\text{fix}}} \frac{dv}{4N_e(v)} \quad (\text{B1})$$

has a probability distribution that depends upon only a single parameter, namely the frequency of the A allele at time $t = 0$. In particular, the probability density of R is independent of the form of $N_e(t)$.

To establish the claimed property of R , let $A(t | y, u)$ denote the probability that $X(t) = 1$, given that $X(u) = y$. The probability $A(t | y, u)$ appears within part of the solution of the forward diffusion equation and therefore also obeys the backward diffusion equation

$$\frac{\partial}{\partial u} A(t | y, u) = -\frac{y(1-y)}{4N_e(u)} \frac{\partial^2}{\partial y^2} A(t | y, u) \quad (\text{B2})$$

and is subject to the initial and boundary conditions $A(t | y, t) = 0$ for $0 < y < 1$, $A(t | 0, u) = 0$ for $u \leq t$, and $A(t | 1, u) = 1$ for $u \leq t$.

By changing the time variable in the backward diffusion equation, Equation B2, from u to the quantity $\int_u^t (dv/4N_e(v))$, it may be verified that the initial and boundary conditions depend only on y and $\int_u^t (dv/4N_e(v))$ and do not depend on t . Accordingly, $A(t | y, u)$ depends only on y and $\int_u^t (dv/4N_e(v))$ and we can therefore write

$$A(t | y, u) = G_y \left(\int_u^t \frac{dv}{4N_e(v)} \right) \quad (\text{B3})$$

for some function $G_y(r)$. Properties of $G_y(r)$ follow from properties of $A(t | y, u)$. From Equations B2 and B3 it follows that $G_y(r)$ obeys

$$-\frac{\partial}{\partial r} G_y(r) = -y(1-y) \frac{\partial^2}{\partial y^2} G_y(r). \quad (\text{B4})$$

Furthermore, $G_y(r)$ is subject to the initial condition

$$G_y(0) = 0 \quad \text{for } 0 < y < 1 \quad (\text{B5})$$

and the boundary conditions

$$\begin{aligned} G_0(r) &= 0 \quad \text{for } r \geq 0 \\ G_1(r) &= 1 \quad \text{for } r \geq 0. \end{aligned} \tag{B6}$$

Additionally, $G_y(r)$ also satisfies

$$G_y(\infty) = y \tag{B7}$$

since $G_y(\infty) \equiv A(\infty | y, u) = y$ is the probability of ultimate fixation of a neutral allele (see *Appendix A*).

We can determine $G_y(r)$ by solving Equation B4 subject to Equations B5 and B6. However, this is not necessary since we can infer the form of $G_y(r)$ from an existing result in the literature. We note that the probability $A(t | y, u)$, when $u = 0$ and for a population of fixed size, has been found by Kimura (1955) and depends only on t and N_e in the combination $t/(4N_e)$. Because of Equation B3, we can take Kimura's solution and replace $t/(4N_e)$ by r to obtain the form for $G_y(r)$. It is often convenient to work with the quantity $G_y(r)/y$. Accordingly, we define

$$F_y(r) = \frac{G_y(r)}{y} \tag{B8}$$

and then find from Kimura's form for $A(t | y, u)$ that

$$F_y(r) = 1 - \sum_{n=0}^{\infty} \frac{2(1-y)(2n+3)}{\lambda_n} (-1)^n C_n^{(3/2)}(y) (1-2y) \exp(-\lambda_n r), \tag{B9}$$

where $\lambda_n = (n+1)(n+2)$ and $C_n^{(3/2)}(y)$ denotes a Gegenbauer polynomial in y of order $3/2$ and degree n (Abramowitz and Stegun 1964). In the limit of small y we write $F_y(r)$ as $F_0(r)$. In this regime a simple approximate form for $A(t | y, u)$ is known (Waxman 2011b) and hence we have the simple approximation for $F_0(r)$:

$$F_0(r) \approx \begin{cases} \left(\frac{\pi}{r}\right)^{3/2} e^{r/4 - \pi^2/(4r)}, & 0 \leq r \leq 2, \\ 1 - 3e^{-2r}, & r > 2. \end{cases} \tag{B10}$$

Since the process of fixation among replicate populations is irreversible (once a replicate population has fixed, it remains fixed), the quantity $A(t | y, u)$ is the probability that fixation has actually occurred within the time interval u to t . That is, $A(t | y, u)$ is the probability that fixation has occurred by time t . It follows that for populations where fixation ultimately occurs, the random time at which fixation occurs, written T_{fix} , obeys

$$\text{Prob}(T_{\text{fix}} \leq t | X(u) = y) = \frac{A(t | y, u)}{A(\infty | y, u)}. \tag{B11}$$

Strictly, the probability on the left-hand side should indicate that it is conditional on fixation ultimately occurring. Given, however, that T_{fix} is not defined unless fixation will ultimately occur, we adopt the simpler notation where this conditioning is left implicit.

Equation B11 can be written for any $u \leq t$ as $\text{Prob}(T_{\text{fix}} \leq t | X(u) = y) = G_y(\int_u^t (dv/4N_e(v)))/y = F_y(\int_u^t (dv/4N_e(v)))$ and completely equivalently, we have

$$\text{Prob}\left(\int_u^{T_{\text{fix}}} \frac{dv}{4N_e(v)} \leq \int_u^t \frac{dv}{4N_e(v)} \middle| X(u) = y\right) = F_y\left(\int_u^t \frac{dv}{4N_e(v)}\right). \tag{B12}$$

Restricting ourselves to the special case of $u = 0$, so T_{fix} is the amount of time that elapses before the mutant allele fixes (if $u \neq 0$, the amount of time that elapses before fixation is $T_{\text{fix}} - u$), we can write Equation B12 as

$$\text{Prob}(R \leq r | X(0) = y) = F_y(r), \tag{B13}$$

where R is the random variable given in Equation B1.

We thus have established that the function of r given by $F_y(r)$ is, by virtue of Equation B13, the *cumulative distribution function* of the random variable R , and this depends on only one parameter, namely y , the initial frequency of the A allele. Differentiating $F_y(r)$ with respect to r yields the probability density of r , and this also depends on the single parameter y .

Appendix C: Deriving Expressions for the Mean and Variance of R

In this *Appendix* we determine expressions for the mean and variance of the statistic R defined in Equation 1 of the main text.

Moments of R are expectations of the form $E[R^n | X(0) = y]$ with $n = 1, 2, \dots$, which we write in the shorter form $E_y[R^n]$. The probability density of R is the derivative, with respect to r , of the cumulative distribution function of R . Here, following *Appendix B*, we write the cumulative distribution function of R as $G(r, y)/y$. Accordingly,

$$\begin{aligned} E_y[R^n] &= \int_0^\infty r^n \frac{\partial}{\partial r} \left(\frac{G_y(r)}{y} \right) = \int_0^\infty r^n \frac{\partial}{\partial r} \left(\frac{G_y(r)}{y} - 1 \right) dr \\ &= \frac{n}{y} \int_0^\infty r^{n-1} [y - G_y(r)] dr. \end{aligned} \quad (C1)$$

We write

$$E_y[R^n] = \frac{n}{y} J_{n-1}(y), \quad (C2)$$

where

$$J_n(y) \stackrel{\text{def}}{=} \int_0^\infty r^n [y - G_y(r)] dr. \quad (C3)$$

Because of Equation B6, we take

$$J_n(0) = 0, J_n(1) = 0. \quad (C4)$$

From Equation B4 we can establish equations that the $J_n(y)$ obey. We write Equation B4 as $-y(1-y)(\partial^2/\partial y^2)[y - G_y(r)] = -(\partial/\partial r)[y - G_y(r)]$. Multiplying this equation by r^n and integrating from $r = 0$ to $r = \infty$ yields $-y(1-y)(d^2/dy^2)J_n(y) = -\int_0^\infty r^n (\partial/\partial r)[y - G_y(r)] dr$. An integration by parts yields

$$-y(1-y) \frac{d^2}{dy^2} J_n(y) = [-r^n [y - G_y(r)]]_{r=0}^{r=\infty} + nJ_{n-1}(y). \quad (C5)$$

When $n = 0$, there is a contribution from the boundary term on the right-hand side at $r = 0$, while for $n > 0$ there is no such contribution. It follows that

$$-y(1-y) \frac{d^2}{dy^2} J_0(y) = y \quad (C6)$$

and

$$-y(1-y) \frac{d^2}{dy^2} J_n(y) = nJ_{n-1}(y), \quad n = 1, 2, 3, \dots \quad (C7)$$

The solution to Equation C6, subject to Equation C4, is $J_0(y) = -(1-y)\ln(1-y)$. The corresponding solution to Equation C7 for $n = 1$ is $J_1(y) = (1-y)\ln(1-y) + y[\pi^2/6 - \text{dilog}(1-y)]$, where $\text{dilog}(x)$ denotes the dilogarithm function: $\text{dilog}(x) = \int_1^\infty (\ln(t)/(1-t)) dt$ (Abramowitz and Stegun 1964). We have not found a particularly simple expression for $J_2(y)$; however, a small y expansion can be determined if required.

We can now determine $E_y[R]$, which equals $J_0(y)/y$:

$$E_y[R] = \int_0^\infty [1 - F_y(r)] dr = -\frac{(1-y)\ln(1-y)}{y}. \quad (C8)$$

Also, we have $E_y[R^2] = 2J_1(y)/y$, i.e.,

$$E_y[R^2] = 2 \frac{(1-y)\ln(1-y)}{y} + \frac{\pi^2}{3} - 2\text{dilog}(1-y), \quad (\text{C9})$$

where $\text{dilog}(x) = \int_1^x (\ln(u)/(1-u))du$ denotes the dilogarithm function (Abramowitz and Stegun 1964). The result for the variance in the main text follows directly by combining the results of Equations C8 and C9:

$$\begin{aligned} \text{Var}_y(R) &= E_y[R^2] - (E_y[R])^2 \\ &\simeq \frac{2(1-y)\ln(1-y)}{y} - \left(\frac{(1-y)\ln(1-y)}{y} \right)^2 + \frac{\pi^2}{3} - 2\text{dilog}(1-y). \end{aligned} \quad (\text{C10})$$

Appendix D: Illustrating the Independence of the Distribution of R on Population Behavior

In this *Appendix* we present a set of results for the statistic R of Equation 1 to illustrate the independence of its distribution on the way population size changes over time, as predicted by the diffusion approximation. To obtain data we carried out individually based simulations of a neutral population that evolves in discrete time. These simulations used binomial sampling of individuals in discrete time, which results in the effective and actual population sizes coinciding: $N_e(t) = N(t)$. In Figure D1, five quite different forms of $N(t)$ have been adopted, ranging from logistic growth to logistic decay.

For each case of $N(t)$ shown in Figure D1, the fate of the mutants was followed in 10^6 replicate populations. As expected, a fraction very close to y of these replicate populations underwent fixation. For a given population behavior, the times of fixation were recorded in the different replicate populations where fixation occurred. If fixation occurred in a replicate population at time t_{fix} , then the statistic corresponding to R of Equation 1 was taken to be the discrete analogue of the integral, namely $\sum_{t=0}^{t_{\text{fix}}-1} (1/4N_e(t))$. With a single copy of a mutant allele initially present, the mean times to fixation for the population behaviors a–e in Figure D1 differed among themselves by considerable amounts, $>200\%$ (see Table D1). The mean values of the R statistic for the different population behaviors were all within 2% of the value predicted by Equation 5. We also used Equation 6 to predict the variance of R for Figure 1. The variances of R determined from cases a–e in Figure D1 were all within 7% of the predicted value—see Table D1.

To test the quality of the diffusion results for the mean time for fixation, we calculated the percentage error in this quantity, relative to the simulation results, for two cases. For the standard case of a constant population (curve d in Figure D1) we used the diffusion result $E[T_{\text{fix}}] = 4N_e E_y[R]$. For a discontinuous jump in the population size (curve b in Figure D1) we used the diffusion result Equation E5, employing Equation 4 of the main text, thereby approximating the initial frequency by zero. For both cases, the errors in the calculated mean times of fixation were found to be $<1\%$.

Appendix E: Deriving Expressions for the Mean and Variance of the Time to Fixation

In this *Appendix* we establish expressions for the mean and the variance of the time to fixation.

Equation 7 of the main text reads $T_{\text{fix}} = h(R)$ and we first determine $E_y[(T_{\text{fix}})^a]$ with a an arbitrary positive number. With $F_y(r)$ the cumulative probability distribution of R at value r (see Equation 2), we have

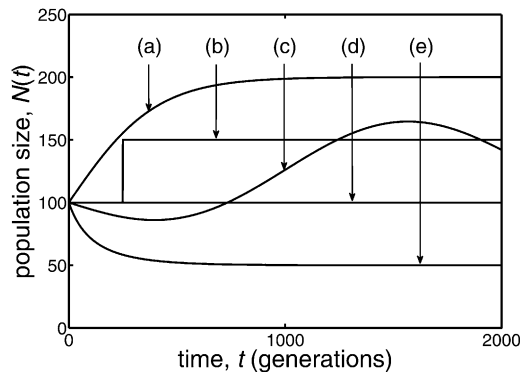


Figure D1 Five quite different examples of population size behavior, $N(t)$, are illustrated. In all cases the initial population size was 100. Case a corresponds to logistic growth, with an intrinsic growth rate of $1/200$ and the population size eventually doubling to 200. In case b the population discontinuously jumps to 150 individuals at generation 250, while case c has a variable population size that was derived from the sum of a constant and several different sine functions. In case d the population size remains constant, while in case e the population undergoes logistic decay, with an intrinsic decay rate of $1/200$ and the population size eventually halving to 50.

Table D1 Results from individually based simulations for the fixation of a neutral mutant allele are given for the population behaviors illustrated in Figure D1

	$E_y[R]$ (% error)	$\text{Var}_y[R]$ (% error)	$E_y[T]$
a: Logistic growth	1.001 (-0.338)	0.294 (-1.268)	625.6
b: Discontinuous jump	0.995 (0.248)	0.291 (-0.294)	479.2
c: Variable	0.978 (2.027)	0.272 (6.658)	361.1
d: Constant	0.993 (0.453)	0.280 (3.348)	397.2
e: Logistic decay	0.990 (0.770)	0.292 (-0.593)	268.8

The initial population size was 100 in all cases, and other parameters are given in the Figure D1 legend. The initial frequency of the mutant allele was $y = 1/200$ and 10^6 replicate populations were simulated for each of the five behaviors illustrated in Figure D1. The expectations and variances given are estimates from the simulations. For the mean and variance of R , the percentage errors, given in parentheses, are the errors of the diffusion results (given in Equations C8 and C10), relative to the simulation results.

$$\begin{aligned}
 E_y[(T_{\text{fix}})^a] &\approx E_y[h^a(R)] = \int_0^\infty h^a(r) \frac{\partial F_y(r)}{\partial r} dr = - \int_0^\infty h^a(r) \frac{\partial}{\partial r} [1 - F_y(r)] dr \\
 &= a \int_0^\infty h^{a-1}(r) \frac{dh(r)}{dr} [1 - F_y(r)] dr.
 \end{aligned} \tag{E1}$$

Taking $a = 1$ in Equation E1 leads to the following expression for the mean time to fixation:

$$E_y[T_{\text{fix}}] \approx \int_0^\infty h(r) \frac{\partial F_y(r)}{\partial r} dr = \int_0^\infty \frac{dh(r)}{dr} [1 - F_y(r)] dr. \tag{E2}$$

There is an alternative way of writing the second form of this expression. We note that the function $h(r)$ is formally defined by

$$r = \int_0^{h(r)} \frac{dv}{4N_e(v)}. \tag{E3}$$

Differentiating Equation E3 with respect to r leads to

$$\frac{dh(r)}{dr} = 4N_e(h(r)) \tag{E4}$$

and using this result in Equation E2 yields an alternative expression for the mean:

$$E_y[T_{\text{fix}}] \approx \int_0^\infty 4N_e(h(r)) [1 - F_y(r)] dr. \tag{E5}$$

To determine an expression for the variance we use $\text{Var}_y(T_{\text{fix}}) = E_y[(T_{\text{fix}})^2] - (E_y[T_{\text{fix}}])^2$ and set $a = 1$ and $a = 2$ in Equation E1 to obtain Equation 9 of the main text.

Appendix F: Derivation of the Results for a Population-Size Jump Given in Equations 13 and 14 of the Main Text

In this *Appendix* we determine exact and approximate results for the mean time of fixation and its variance when the effective population size discontinuously jumps at time τ from $N_{e,0}$ to $N_{e,1}$, as given in Equation 10 of the main text.

To begin, we note that the function $h(r)$, which relates T_{fix} to R , is formally defined by Equation E3. It may be verified that with $N_e(t)$ given by Equation 10, $h(r)$ is given by

$$h(r) = \begin{cases} 4N_{e,0}r, & r \leq \rho, \\ 4N_{e,0}\rho + 4N_{e,1}(r - \rho), & r > \rho, \end{cases} \tag{F1}$$

where $\rho = \tau/(4N_{e,0})$ is a scaled measure of the time of the population-size jump.

The mean of T_{fix} can be obtained from Equations E2 and F1 and takes the form

$$E_y [T_{\text{fix}}] = A_0 \times 4N_{e,0} + A_1 \times 4N_{e,1}, \quad (\text{F2})$$

where

$$A_0 = \int_0^\rho [1 - F_y(r)] dr, \quad A_1 = \int_\rho^\infty [1 - F_y(r)] dr \equiv 1 - A_0. \quad (\text{F3})$$

The variance of T_{fix} follows from Equations 9 and F1 and takes the form

$$\begin{aligned} \text{Var}_y(T_{\text{fix}}) = & (B_{00} - A_0^2) \times (4N_{e,0})^2 + (B_{01} - 2A_0A_1) \times (4N_{e,0})(4N_{e,1}) \\ & + (B_{11} - A_1^2) \times (4N_{e,1})^2, \end{aligned} \quad (\text{F4})$$

where

$$\begin{aligned} B_{00} = & 2 \int_0^\rho r [1 - F_y(r)] dr, \quad B_{01} = 2\rho \int_\rho^\infty [1 - F_y(r)] dr \equiv 2\rho A_1, \\ B_{11} = & 2 \int_\rho^\infty (r - \rho) [1 - F_y(r)] dr. \end{aligned} \quad (\text{F5})$$

We derive results for the case of vanishingly small initial frequencies ($y \rightarrow 0$) and make extensive use of the following $y \rightarrow 0$ forms of Equations C8 and C9:

$$\int_0^\infty [1 - F_0(r)] dr = 1, \quad \int_0^\infty r [1 - F_0(r)] dr \equiv \frac{1}{2} E_0 [R^2] = \frac{1}{2} \left(\frac{\pi^2}{3} - 2 \right). \quad (\text{F6})$$

Derivation of results for $y \rightarrow 0$ and $\rho \ll 1$

We used Equation 4 of the main text and obtained the following approximate results, which all have corrections of $O(e^{-\pi^2/(4\rho)})$: $A_0 \approx \rho$, $A_1 \approx 1 - \rho$, $B_{00} \approx \rho^2$, $B_{01} \approx 2\rho(1 - \rho)$. For B_{11} we write $B_{11} = 2(\int_0^\infty r [1 - F_0(r)] dr - \int_0^\rho r [1 - F_0(r)] dr) - 2\rho A_1$, which leads to $B_{11} = \pi^2/3 - 2 - 2\rho + \rho^2 + O(e^{-\pi^2/(4\rho)})$. Combining these results leads to $E_0[T_{\text{fix}}] \approx \rho \times 4N_{e,0} + (1 - \rho) \times 4N_{e,1}$ and $\text{Var}_0(T_{\text{fix}}) \approx (\pi^2/3 - 3) \times (4N_{e,1})^2$.

Derivation of results for $y \rightarrow 0$ and $\rho \approx 0.51$ or $\rho = 1$

Results for these cases were obtained using Equation 4 of the main text and numerically determining the coefficients in Equations F3 and F5.

Derivation of results for $y \rightarrow 0$ and $\rho \geq 2$

We used Equation 4 of the main text and obtained the following results: $A_0 = \int_0^\infty [1 - F_0(r)] dr - \int_\rho^\infty [1 - F_0(r)] dr \approx 1 - (3/2)e^{-2\rho}$, $A_1 = 1 - A_0 \approx (3/2)e^{-2\rho}$, $B_{00} = 2 \int_0^\infty r [1 - F_0(r)] dr - 2 \int_\rho^\infty r [1 - F_0(r)] dr \approx \pi^2/3 - 2 - (3/2)(1 + 2\rho)e^{-2\rho}$, $B_{01} \approx 3\rho e^{-2\rho}$, and $B_{11} \approx (3/2)e^{-2\rho}$. Combining these results gives $E_0[T_{\text{fix}}] = (1 - (3/2)e^{-2\rho}) \times 4N_{e,0} + (3/2)e^{-2\rho} \times 4N_{e,1}$ and on neglecting terms of $O(e^{-4\rho})$ gives $\text{Var}_0(T_{\text{fix}}) \approx (\pi^2/3 - 3 + (3/2)(1 - 2\rho)e^{-2\rho}) \times (4N_{e,0})^2 + (3/2)e^{-2\rho} \times (4N_{e,1})^2$.

Appendix G: Mean and Variance of the Time to Fixation When N_e Has a Period of Linear Growth

In this *Appendix* we consider a situation outlined in the main text, where population size increases linearly, for a finite period of time, and then remains constant, as given in Equation 15 of the main text. We give results only for the case of vanishingly small initial frequencies ($y \rightarrow 0$).

In terms of the quantity r_0 defined by

$$r_0 = \frac{1}{4\gamma_e} \log \left(1 + \frac{\gamma_e \tau}{N_{e,0}} \right) \quad (\text{G1})$$

the function $h(r)$ defined by Equation E3 can be verified to be given by

$$h(r) = \begin{cases} \frac{N_{e,0}}{\gamma_e} (e^{4\gamma_e r} - 1), & r \leq r_0, \\ \frac{N_{e,0}}{\gamma_e} (e^{4\gamma_e r_0} - 1) + 4N_{e,0}e^{4\gamma_e r_0}(r - r_0), & r > r_0. \end{cases} \quad (\text{G2})$$

To determine the mean time to fixation, we use Equations G2 and E2 to obtain

$$E_0[T_{\text{fix}}] = 4N_{e,0} \times (A_0 + A_1), \quad (\text{G3})$$

where

$$A_0 = \int_0^{r_0} e^{4\gamma_e r} [1 - F_0(r)] dr, \quad A_1 = e^{4\gamma_e r_0} \int_{r_0}^{\infty} [1 - F_0(r)] dr. \quad (\text{G4})$$

The variance follows from Equations G2 and 9 and takes the form

$$\text{Var}_0(T_{\text{fix}}) = (4N_{e,0})^2 \times [B_0 + B_1 - (A_0 + A_1)^2], \quad (\text{G5})$$

where

$$\begin{aligned} B_0 &= 2 \int_0^{r_0} e^{4\gamma_e r} \frac{e^{4\gamma_e r} - 1}{4\gamma_e} [1 - F_0(r)] dr \\ B_1 &= 2e^{4\gamma_e r_0} \int_{r_0}^{\infty} \left[\frac{e^{4\gamma_e r_0} - 1}{4\gamma_e} + e^{4\gamma_e r_0}(r - r_0) \right] [1 - F_0(r)] dr. \end{aligned} \quad (\text{G6})$$

We derive results for *just the* τ dependence of $E_0[T_{\text{fix}}]$ and $\text{Var}_0(T_{\text{fix}})$. The results obtained are restricted to large τ , such that r_0 (given in Equation G1) > 2 .

We make extensive use of Equation 4 of the main text.

With c_1 and c_2 constants with respect to τ we can write

$$A_0 = \int_0^2 e^{4\gamma_e r} [1 - F_0(r)] dr + \int_2^{r_0} e^{4\gamma_e r} [1 - F_0(r)] dr = c_1 + \frac{3e^{(4\gamma_e - 2)r_0}}{4\gamma_e - 2};$$

$$A_1 = \frac{3}{2} e^{(4\gamma_e - 2)r_0};$$

$$B_0 = 2 \int_0^2 e^{4\gamma_e r} \frac{e^{4\gamma_e r} - 1}{4\gamma_e} [1 - F_0(r)] dr + 2 \int_2^{r_0} e^{4\gamma_e r} \frac{e^{4\gamma_e r} - 1}{4\gamma_e} [1 - F_0(r)] dr = c_2 + \frac{3}{2\gamma_e} \left(\frac{e^{(8\gamma_e - 2)r_0}}{8\gamma_e - 2} - \frac{e^{(4\gamma_e - 2)r_0}}{4\gamma_e - 2} \right);$$

$$B_1 = 2e^{4\gamma_e r_0} \int_{r_0}^{\infty} \left[\frac{e^{4\gamma_e r_0} - 1}{4\gamma_e} + e^{4\gamma_e r_0}(r - r_0) \right] 3e^{-2r} dr = \frac{3}{4\gamma_e} e^{(4\gamma_e - 2)r_0} [(1 + 2\gamma_e)e^{4\gamma_e r_0} - 1].$$

These lead to

$$\frac{E_0[T_{\text{fix}}]}{4N_{e,0}} \approx c_1 + \frac{3\gamma_e e^{(4\gamma_e - 2)r_0}}{2\gamma_e - 1} \quad (\text{G7})$$

$$\frac{\text{Var}_0(T_{\text{fix}})}{(4N_{e,0})^2} \approx c_2 + \frac{3}{2\gamma_e} \left(\frac{e^{(8\gamma_e-2)r_0}}{8\gamma_e-2} - \frac{e^{(4\gamma_e-2)r_0}}{4\gamma_e-2} \right) + \frac{3}{4\gamma_e} \left[(1+2\gamma_e)e^{(8\gamma_e-2)r_0} - e^{(4\gamma_e-2)r_0} \right] - \left(c_1 + \frac{3\gamma_e e^{(4\gamma_e-2)r_0}}{2\gamma_e-1} \right)^2. \quad (\text{G8})$$

Assuming that τ is sufficiently large that $\gamma_e \tau \gg N_{e,0}$, Equation G1 allows us to write $e^{4\gamma_e r_0} \approx \gamma_e \tau / N_{e,0}$. Using this result, and the parameter p defined by $p = (2\gamma_e - 1)/2\gamma_e \equiv (\gamma_e - 1/2)/\gamma_e$, we find that the τ dependence in Equations G7 and G8 takes the form given in Equations 17 and 18 of the main text.

Appendix H: Derivation of an Inequality for $E_y[T_{\text{fix}}]$

In this *Appendix* we establish an inequality for $E_y[T_{\text{fix}}]$ that applies when $N_e(t)$ increases with t .

We write the condition that $N_e(t)$ increases with t as

$$dN_e(t)/dt > 0. \quad (\text{H1})$$

We note that because of Equations E4 and H1, the function $h(r)$ is convex ($d^2h(r)/dr^2 > 0$). We can thus employ Jensen's inequality, which for $\phi(r)$ convex reads $E[\phi(R)] > \phi(E[R])$ (Gradsteyn and Ryzhik 1980), and hence obtain $E_y[h(R)] > h(E_y[R])$. Since $E_y[T_{\text{fix}}] = E_y[h(R)]$ (using Equation 7 of the main text), we have

$$E_y[T_{\text{fix}}] > h(E_y[R]). \quad (\text{H2})$$

We write this inequality in a different way that allows additional development. From Equation E4 we obtain $h(r) = \int_0^r 4N_e(h(v))dv$ and hence Equation H2 is equivalent to

$$E_y[T_{\text{fix}}] > \int_0^{E_y[R]} 4N_e(h(v))dv. \quad (\text{H3})$$

Because of Equation H1, we have $N_e(v) > N_e(0)$ for $v > 0$ and using Equation E3 gives $r = \int_0^{h(r)} (dv/4N_e(v)) < \int_0^{h(r)} (dv/4N_e(0))$; *i.e.*, $h(r) > 4N_e(0)r$. Using Equation H1 again then yields $4N_e(h(v)) > 4N_e(4N_e(0)v)$, which can be used in Equation H3 to yield $E_y[T_{\text{fix}}] > \int_0^{E_y[R]} 4N_e(4N_e(0)v)dv$. The change of variable $4N_e(0)v \rightarrow v$ gives $E_y[T_{\text{fix}}] > (1/4N_e(0)) \int_0^{4N_e(0)E_y[R]} 4N_e(v)dv$, which is equivalent Equation 20 of the main text.