

# On the Fixation Process of a Beneficial Mutation in a Variable Environment

Hildegard Uecker<sup>1</sup> and Joachim Hermisson

Mathematics and Biosciences Group, Faculty of Mathematics and Max F. Perutz Laboratories, University of Vienna, A-1090 Vienna, Austria

**ABSTRACT** A population that adapts to gradual environmental change will typically experience temporal variation in its population size and the selection pressure. On the basis of the mathematical theory of inhomogeneous branching processes, we present a framework to describe the fixation process of a single beneficial allele under these conditions. The approach allows for arbitrary time-dependence of the selection coefficient  $s(t)$  and the population size  $N(t)$ , as may result from an underlying ecological model. We derive compact analytical approximations for the fixation probability and the distribution of passage times for the beneficial allele to reach a given intermediate frequency. We apply the formalism to several biologically relevant scenarios, such as linear or cyclic changes in the selection coefficient, and logistic population growth. Comparison with computer simulations shows that the analytical results are accurate for a large parameter range, as long as selection is not very weak.

**F**OR adaptive evolution to proceed, it is not enough that new beneficial mutations enter a population. To complete an adaptive step, these mutations also need to escape stochastic loss due to genetic drift, get established, and finally rise to fixation. The fixation process of beneficial (or neutral or deleterious) alleles is one of the building blocks of population genetic theory and many of the key results on fixation probabilities and times date back to its early days. Two alternative mathematical frameworks have been developed to derive analytical expressions for these quantities: branching processes (Fisher 1922, 1930; Haldane 1927) and diffusion theory (Kimura 1962; Kimura and Ohta 1969). Today, a large body of literature exists to study fixation under various ecological scenarios and genetic conditions (reviewed in Patwa and Wahl 2008), such as the effects of population structure (Whitlock 2003) and spatial heterogeneity (Whitlock and Gomulkiewicz 2005) and interference due to selection on linked loci (Barton 1995) or due to epistatic interaction (Takahashi and Tajima 2005).

In this article, we consider the fixation process in a variable environment, leading to time-dependent selection coefficients and population sizes. Aspects of this problem have already been studied in previous work: In particular, the impact of various scenarios of demographic change (growth, decline, cycles) on the fixation probability has been treated in a series of articles (Ewens 1967; Chia 1968; Kimura and Ohta 1974; Otto and Whitlock 1997; Pollak 2000; Parsons and Quince 2007a; Orr and Unckless 2008). Studies on time-dependent selection mostly concentrate on stochastic fluctuations of the selection coefficient (Jensen 1973; Karlin and Levikson 1974; Takahata and Ishii 1975; Huillet 2011). Since the distribution of the selection coefficients is constant across generations, these models are still time-homogeneous in a probabilistic sense. In contrast, surprisingly little is known when the changes of the selection coefficient  $s = s(t)$  follow an explicit trend. Ohta and Kojima (1968) derive an expression for the fixation probability of a mutation with time-dependent selective advantage in the context of the evolution of chromosomal inversions. Apart from that, only particular functions have been considered: Kimura and Ohta (1969) discuss the case where selection decreases exponentially in time, and Pollak (1966) derives expressions for the fixation probability under two alternating selection pressures. No previous work seems to exist where both population size and selection strength are variable, although this is a generic case under realistic ecological conditions. Also, there does not seem to be

Copyright © 2011 by the Genetics Society of America  
doi: 10.1534/genetics.110.124297

Manuscript received February 17, 2011; accepted for publication May 17, 2011

Supporting information is available online at <http://www.genetics.org/content/suppl/2011/06/06/genetics.110.124297.DC1>.

<sup>1</sup>Corresponding author: Mathematics and Biosciences Group, Faculty of Mathematics, Nordbergstrasse 15, A-1090 Vienna, Austria. E-mail: hildegard.uecker@univie.ac.at

an investigation of fixation or passage times in variable environments.

In the following, we present a formalism to describe the fixation process under a wide range of scenarios of environmental variation. We use branching processes in continuous time to derive analytical approximations for the fixation probability and the passage time needed for the mutant allele to reach some intermediate frequency  $x_c$ . After the introduction of our model, we describe how a general approximation for the fixation probability can be obtained from known mathematical results on inhomogeneous branching processes. Afterward, we discuss applications of this result to several biologically relevant scenarios. In the second part of the article, we introduce and apply a method to calculate the distribution of the passage time needed for a beneficial mutation to reach an intermediate frequency. The method works by combining the stochastic fluctuations from the branching approximation with the deterministic growth of the full model. This technique has been used before (for constant selection and population size) by Desai and Fisher (2007) in a model of clonal evolution. All analytical results are complemented by computer simulations, which are briefly described in a separate section. We close with a short discussion. In the *Appendix*, we discuss a generalized version of the model to include allele-frequency-dependent population demographies. We illustrate how the formalism can be used by applying it to an illustrative ecological scenario: the fixation probability of a “rescue mutation” in a population that is threatened by extinction (*cf.* Orr and Unckless 2008). Additional material is devoted to [Supporting Information, File S1, Figure S1, Figure S2, and Figure S3](#). We discuss in some detail the scope and limits of the approach and the accuracy of the approximation. In section S2 of [File S1](#), we present an alternative treatment to derive the fixation probability in a variable environment from a diffusion approach.

## The Model

We consider a large population of haploid individuals with time-dependent population size  $N_t$ . The population dynamics are modeled as a time-inhomogeneous birth–death process with birth and death rates  $b(t, N_t)$  and  $d(t, N_t)$ :

$$\begin{aligned} N_t \rightarrow N_t + 1 &: b(t, N_t)N_t, \\ N_t \rightarrow N_t - 1 &: d(t, N_t)N_t. \end{aligned} \quad (1)$$

The impact of the changes in the external environment on the population size is reflected in the explicit time-dependence of the rates on  $t$ . The dependence on  $N_t$  accounts for density-dependence [e.g., logistic:  $b(t, N_t) = b_1(t) - b_2(t)N_t$ ]. We call  $r(t, N_t) = b(t, N_t) - d(t, N_t)$  the growth parameter. Obviously, the expected change of  $N_t$  over a small time interval  $dt$  reads

$$E[\Delta N | N_t] = r(t, N_t)N_t dt. \quad (2)$$

Consider now two alleles, a beneficial mutant allele  $A$  and the ancestral (resident) allele  $a$ , that segregate in the population at a single locus. Recurrent mutations in both directions are ignored. In general, birth and death rates might be different for residents and mutants. These rates can depend on time and on the (absolute) frequencies of both allelic types, allowing for general frequency-dependent selection. As a consequence, also the population dynamics depend on the allelic composition and cannot be described by Equation 1 anymore. We discuss this model in the *Appendix*. For the main part of the article, however, we assume that the rates are the same for mutants and residents and that all model parameters are independent of allele frequencies. This means in particular that selection is soft; *i.e.*, changes in the allelic composition due to selection or drift do not interfere with the population dynamics. Population growth and decline of the polymorphic population are then correctly described by Equation 1.

In this setting, selection is modeled as competitive replacement between individuals, which does not change the population size, and is implemented as follows: At per capita rate  $\xi(t, N_t) + s(t, N_t)$ , a mutant additionally reproduces and succeeds in replacing a randomly chosen individual from the population by its offspring. Residents do the same at rate  $\xi(t, N_t)$ . Again, the selective advantage  $s(t, N_t)$  of the mutant may thus depend on the external environment (modeled by the dependence of  $s(t, N_t)$  on  $t$ ) and the population size (modeled by the dependence on  $N_t$ ). Changes in the number of mutants then occur at rates

$$\begin{aligned} n_t \rightarrow n_t + 1 &: (\xi(t, N_t) + s(t, N_t)) \frac{n_t(N_t - n_t)}{N_t} + b(t, N_t)n_t, \\ n_t \rightarrow n_t - 1 &: \xi(t, N_t) \frac{n_t(N_t - n_t)}{N_t} + d(t, N_t)n_t. \end{aligned} \quad (3)$$

The model corresponds to a continuous-time Moran model, but with a population size that may change in time. Putting  $b(t, N_t) = d(t, N_t) = 0$ ,  $\xi(t, N_t) = 1$ , and  $s(t, N_t) = s = \text{const.}$  reproduces the standard Moran model (Moran 1958a,b; Novozhilov *et al.* 2006). The free parameter  $\xi(t, N_t)$  has been introduced to our model to allow for easy interpolation to other models (see below) and additionally to make the analysis of density-dependent competition possible.

To further clarify the relation to other models, we calculate how the frequency of mutants  $x_t := n_t/N_t$  changes over time. Let  $\Delta x$  be its change in an infinitesimal time interval  $dt$ . The expectation and the variance of  $\Delta x$  are calculated to be

$$E[\Delta x | x_t, N_t] = s(t, N_t)x_t(1 - x_t)dt, \quad (4a)$$

$$\text{Var}[\Delta x | x_t, N_t] \approx E[(\Delta x)^2 | x_t, N_t] \approx \frac{x_t(1 - x_t)}{N_{e,t}}dt \quad (4b)$$

with the time-dependent variance effective population size

$$N_{e,t} = \frac{N_t}{2\xi(t, N_t) + b(t, N_t) + d(t, N_t) + s(t, N_t)}. \quad (5)$$

In the last step we approximated  $N_t + 1 \approx N_t$  and  $N_t - 1 \approx N_t$  (see section S3 of File S1 for the derivation of Equations 4a and 4b).

We see that the strength of drift, measured as  $N_{e,t}^{-1}$ , is proportional to the total rate of events in the model. The choice  $2\xi(t, N_t) + b(t, N_t) + d(t, N_t) = 2$  coincides with the strength of drift in the standard Moran model, while  $2\xi(t, N_t) + b(t, N_t) + d(t, N_t) = 1$  is consistent with the scaling in the Wright–Fisher model. In contrast to many diffusion or coalescent approaches, we do not rescale time with the effective population size (which would be impractical since  $N_{e,t}$  itself depends on  $t$ ). Generation time in the continuous-time Moran model is defined as the inverse of the total death rate of an individual,  $(\xi(\tau, N_\tau) + d(\tau, N_\tau))^{-1}$ , and may again depend on time in our model.

## Fixation Probability

### Analytical theory

Following pioneering work by Haldane (1927) and Fisher (1930), there has been a long tradition in population genetics to calculate fixation probabilities by branching process methods (reviewed in Haccou *et al.* 2005 and Patwa and Wahl 2008). For the general time-dependent case, the relevant results have long been known in the mathematical literature (e.g., Kendall 1948; Allen 2011). However, only specific cases (usually in the context of changing population sizes) have been discussed in the population genetics context (Ewens 1967; Otto and Whitlock 1997; Pollak 2000; Wahl and Gerrish 2001). We therefore give a brief outline of the general theory below and show how it applies to the biological problem at hand. Previous results are recovered as special cases.

The branching process approximation is based on the following reasoning: Initially, the fate of a new beneficial mutation arising in a population will be strongly determined by genetic drift. In most cases it will actually get lost again. Once the mutation has survived this early phase, it is, however, almost sure to get fixed given its selective advantage is large enough. To calculate the fixation probability it is therefore often sufficient to consider the stage at which the mutant population size  $n_t$  is still small relative to the total population size  $N_t$ . In this early phase, the mutant individuals suffer nearly independent fates, as do the individuals in a Galton–Watson branching process (this assumption is precisely met in an infinite population). The extinction probability of the latter can therefore be used as an approximation for the probability that the mutation gets lost. Because a mutation in a finite population is in the long term either fixed or lost, the fixation probability is the complementary probability. (We exclude the unbiological case of

a population that increases without bounds, where possibly neither wild types nor mutants become extinct.)

Ignoring terms proportional to  $x_t = n_t/N_t$  in the birth–death model (Equation 3), which corresponds to the limit  $N_t \rightarrow \infty$ , leads to transition rates that are proportional to  $n_t$ . Following standard practice in ecological modeling, we further ignore stochastic fluctuations in the population dynamics. This is done by replacing the stochastic variable  $N_t$  by its deterministic approximation denoted as  $N(t)$ , with dynamics

$$\frac{dN(t)}{dt} = r(t, N(t))N(t). \quad (6)$$

Inserting the deterministic solution  $N(t)$  into the rates for birth, death, and selection reduces the dependence of these rates on  $t$  and  $N_t$  to a dependence on  $t$  only [ $s(t, N_t) \rightarrow s(t, N(t)) =: s(t)$ , etc.]. We arrive at a branching process with time-dependent per capita birth and death rates:

$$\begin{aligned} \text{birth rate} &: \lambda(t) = b(t) + \xi(t) + s(t), \\ \text{death rate} &: \mu(t) = d(t) + \xi(t). \end{aligned} \quad (7)$$

As explained above for the birth–death process, the total rate of events determines the strength of genetic drift, while  $\lambda(t) - \mu(t) = s(t) + r(t)$  corresponds to the absolute expected rate of increase of a small mutant population (its Malthusian fitness parameter).

To derive the extinction probability of this process, we follow Allen (2011, p. 278ff). Let  $p_i(n_0, t)$  be the probability that there are  $i$  individuals at time  $t$  when the process started with  $n_0$  individuals at time  $t = 0$ . Using the Kolmogorov forward equation,

$$\begin{aligned} \frac{dp_i(n_0, t)}{dt} &= \lambda(t)(i-1)p_{i-1}(n_0, t) + \mu(t)(i+1)p_{i+1}(n_0, t) \\ &\quad - (\lambda(t) + \mu(t))ip_i(n_0, t), \end{aligned} \quad (8)$$

a differential equation for the probability generating function  $P_{n_0}(z, t) = \sum_{i=0}^{\infty} p_i(n_0, t)z^i$  of the branching process can be derived (see Allen 2011, p. 279):

$$\begin{aligned} \frac{\partial P_{n_0}(z, t)}{\partial t} &= [\lambda(t)(z^2 - z) + \mu(t)(1 - z)] \frac{\partial P_{n_0}(z, t)}{\partial z}, \\ P_{n_0}(z, 0) &= z^{n_0}. \end{aligned} \quad (9)$$

The solution is known from the mathematical literature (Kendall 1948; Allen 2011, p. 280) and given as

$$P_{n_0}(z, t) = \left( 1 + \frac{1}{e^{-\rho(t)}/(z-1) - \int_0^t \lambda(\tau) e^{-\rho(\tau)} d\tau} \right)^{n_0} \quad (10)$$

with

$$\rho(t) = \int_0^t [\lambda(\tau) - \mu(\tau)] d\tau = \int_0^t [s(\tau) + r(\tau)] d\tau. \quad (11)$$

To keep notation short, we introduce the abbreviations

$$A(t) := e^{-\rho(t)}, \quad (12a)$$

$$B(t) := \int_0^t \lambda(\tau) e^{-\rho(\tau)} d\tau. \quad (12b)$$

The extinction probability  $p_0(n_0, t)$  is immediately obtained from the generating function via

$$\begin{aligned} p_0(n_0, t) = P_{n_0}(0, t) &= \left(1 - \frac{1}{(A+B)(t)}\right)^{n_0} \\ &= \left[\frac{\int_0^t \mu(\tau) e^{-\rho(\tau)} d\tau}{1 + \int_0^t \mu(\tau) e^{-\rho(\tau)} d\tau}\right]^{n_0}. \end{aligned} \quad (13)$$

The probability that the mutation will eventually fix in the population is therefore given by

$$p_{\text{fix}}(n_0) = 1 - \lim_{t \rightarrow \infty} p_0(n_0, t) = 1 - \left(1 - \frac{1}{A+B}\right)^{n_0}, \quad (14)$$

where we introduced

$$\begin{aligned} A+B &= \lim_{t \rightarrow \infty} (A+B)(t) = 1 + \int_0^\infty \mu(t) e^{-\rho(t)} dt \\ &= \int_0^\infty \lambda(t) e^{-\rho(t)} dt. \end{aligned} \quad (15)$$

The last equality is valid for  $\lim_{t \rightarrow \infty} \exp(-\rho(t)) = 0$ . This condition is met in all examples considered below. For a single mutant ( $n_0 = 1$ ), it thus holds:

$$p_{\text{fix}} = \frac{1}{A+B} = \frac{2}{\left[1 + \int_0^\infty (\lambda + \mu)(t) \exp\left(-\int_0^t (s+r)(\tau) d\tau\right) dt\right]} \quad (16a)$$

$$= \frac{2}{\left[1 + \int_0^\infty (N(0)/N_e(t)) \exp\left(-\int_0^t s(\tau) d\tau\right) dt\right]}, \quad (16b)$$

where we have used  $\int_0^t r(\tau) d\tau = \int_0^t (\dot{N}(\tau)/N(\tau)) d\tau = \ln(N(t)/N(0))$  for the last equality and  $N_e(t)$  is defined analogously to Equation 5. A similar expression was also derived by Ohta and Kojima (1968). Restricted to a constant population size and a Poisson offspring distribution, their result is, however, less general.

The result depends on two independent parameters, which are compositions of three biologically relevant factors: the strength of selection given by  $s(t)$ , the combined rate of birth and death events  $(\lambda + \mu)(t)$ , and the changes in the total population size [modeled by  $r(t)$  or  $N(t)$ ]. In the above equation we formulated the result via two different combinations of these three variables. In the first version (Equation 16a) it is expressed in terms of the absolute rate

of increase of mutants in the population  $(s + r)(t)$  and the total rate  $(\lambda + \mu)(t)$  at which events happen, which defines the time scale of the problem and also quantifies the influence of drift. In the second formulation (Equation 16b) of the result we combined the birth and death rates and the changing population size with the time-dependent variance effective population size. The second decisive parameter is the selection coefficient of the mutation. Depending on the question to be answered, one or the other version is more favorable. In the first version, the correspondence between a mutation with time-dependent selective advantage and a mutation in a population of changing size can be easily seen: a mutation with time-dependent selective advantage  $s(t)$  in a population of (on average) constant size  $b(t) = d(t) = \text{const.}$  has the same chance to reach fixation as a mutation with constant selective advantage  $s_0$  in a population with time-dependent death rate  $d(t)$  and growth parameter  $r(t) = s(t) - s_0$ . [Note, however, that  $s_0$  must be  $>0$ , such that fixation is (almost) certain once the mutation has survived genetic drift.] The second version is closer to the traditional view in population genetics. It is advantageous if the variance effective population size is directly given and allows, in particular, also for the treatment of discontinuous changes in the population size.

We see that the fixation probability is independent of many details of the individual-level dynamics of the original process (Equation 3), which depends on four rates  $[b(t, N_t), d(t, N_t), s(t, N_t), \text{ and } \xi(t, N_t)]$ . Further aspects of the stochastic model are ignored by our deterministic approximation for the population dynamics. In particular, the analytical results become independent of the particular form of density regulation. As an example, consider three scenarios: (1) a population with inherently constant size  $N_0$  with  $b(t, N_t) = d(t, N_t) = 0$  and  $\xi(t, N_t) = c$ ; (2) a population with density regulation according to  $b(t, N_t) = c + \rho(1 - N_t/N_0)$  and  $d(t, N_t) = c$ , with initial size  $N_0$ , and  $\xi(t, N_t) = 0$ ; and (3) a population without density regulation with  $b(t, N_t) = d(t, N_t) = c$ , initial size  $N_0$ , and  $\xi(t, N_t) = 0$ . In all three cases, the deterministic dynamics of the population size are the same ( $N(t) = N_0$ ) and the analytical predictions for the fixation probability coincide for arbitrary  $s(t, N_t)$ . Simulation results of all three scenarios indeed showed no significant difference, justifying the approximation (see section S1 of [File S1](#)). This observation agrees with findings by Parsons and Quince (2007a) that demographic stochasticity does not significantly influence the fixation probability of advantageous alleles [Parsons and Quince (2007a) discuss this issue for a population that starts in the vicinity of the dynamic equilibrium]. For concreteness, we use the notion of “constant population size” in the following to refer to the case of a strictly constant population size ( $b(t, N_t) = d(t, N_t) = 0$ ). We further set  $\xi(t) = \xi = \text{const.}$  in all applications. If not stated otherwise, we use  $\xi = 1$  (corresponding to the Moran model scaling) for the results shown in the figures.

A necessary condition for the branching approximation to yield meaningful results is that the fate of the mutation is decided as long as the mutant frequency  $x_t = n_t/N_t$  is still small. As  $x_t$  increases, mutants are no longer independent in the original birth–death process (Equation 3), and both processes differ significantly from each other. In particular, the birth–death process has a second absorbing boundary at  $x_t = 1$ . As a consequence, neutral or even deleterious mutations, too, can become fixed by genetic drift. In the corresponding branching model, an upper absorbing boundary does not exist. Consequently, neutral or deleterious alleles must go extinct in the long term. For an allele with a general time-dependent selection coefficient, the branching process approximation is thereby valid only if the allele is “sufficiently beneficial on average.” We can formalize this condition as follows: Since divergence of the (first) integral in Equation 15 leads to the wrong prediction of a zero fixation probability, we need to require that the integral converges. If we assume that  $\mu(t)$  is bounded below by a constant  $C_\mu > 0$ , a necessary condition for the convergence of the integral is divergence of the integral in Equation 11. More precisely, if we assume that  $\mu(t)$  has bounds  $C_\mu > 0$ ,  $C^\mu$  such that  $C_\mu < \mu(t) < C^\mu$ , it must hold that  $\lim_{t \rightarrow \infty} \rho(t) > C \ln[t]$  for some constant  $C > 1$ . A prominent example where this condition is not fulfilled is a mutation with an exponentially decreasing selection coefficient in a population of constant size [Kimura and Ohta (1969), cf. also Pollak (1966) and Ohta and Kojima (1968) for mutations that lose their advantage over time]. The condition is, however, not sufficient to obtain good results: It is also necessary that the contribution of times beyond the initial phase be negligible. In particular, changes in the environment at times larger than the fixation time must not have an impact on the results. As shown in section S1 of File S1, the approximation will usually be excellent if the estimate of  $p_{\text{fix}}$  from Equation 16 fulfills  $p_{\text{fix}} N \gtrsim 10$ . By a similar reasoning it is clear that the requirement that the mutation be beneficial all the time is at this point unnecessarily strict. It is sufficient that the extinction probability is in both processes—the original birth–death process and the branching approximation—negligible once the number of mutants has reached a certain size.

## Applications

**Constant selection and constant population size:** Setting  $\lambda(t) = \xi + s$  and  $\mu(t) = \xi$  yields for the fixation probability

$$p_{\text{fix}} = \frac{s}{\xi + s} \approx \frac{s}{\xi} \approx 2s \frac{N_e}{N}. \quad (17)$$

The well-known result found by Haldane (1927) is therefore reproduced for  $N_e = N$ . In the special case of a constant environment, it is possible to calculate the exact fixation probability from the transition matrix defined via Equation 3 (see Ewens 2004, p. 90). One obtains

$$p_{\text{fix}}^{(\text{exact})} = \frac{(s/(\xi + s))}{1 - (\xi/(\xi + s))^N} \approx \frac{(s/(\xi + s))}{1 - \exp(-(s/(\xi + s))N)}, \quad (18)$$

which shows that the branching approximation is very accurate for large values of  $sN/\xi$ . Furthermore, it is possible to calculate the fixation probability of a deleterious mutation with selective disadvantage  $-s$ . To do so, we switch the roles of the transition rates defined in Equation 3. Every path leading to fixation  $n = N$  has now a chance to be realized, which is by a factor of  $(\xi + s)^{-N}$  lower than the corresponding path for a beneficial mutation. It immediately follows that

$$p_{\text{fix}}(-s) = \frac{p_{\text{fix}}(s)}{(\xi + s)^N} \approx \frac{p_{\text{fix}}(s)}{\exp(sN)}. \quad (19)$$

**Constant selection and changing population size:** Otto and Whitlock (1997) analyzed the fixation probability for several important scenarios of demographic change. Our general result (Equation 16b) contains all those scenarios, among arbitrary others. In addition, it provides some insight into the influence of the individual-level dynamics on the fixation probability. As an example, we consider a population that follows logistic growth (or decline) until it has reached its new carrying capacity  $K$ . There are different ways to describe these global dynamics at the individual level. We discuss two possibilities, which arise naturally in a biological context: The first one assumes that a decreasing availability of resources per individual leads to a lower birth rate while the death rate stays constant (e.g., fertility is reduced). The second one assumes that the same circumstances lead to a higher death rate while the birth rate stays constant. The selection coefficient  $s$  of the beneficial mutant is constant in both cases. In the first scenario, the birth and death rates are given by

$$b(t, N_t) = b + r \left(1 - \frac{N_t}{K}\right), \quad (20)$$

$$d(t, N_t) = b.$$

The total population size thus changes according to

$$N(t) = \frac{KN_0}{(K - N_0)e^{-rt} + N_0}, \quad (21)$$

and using Equation 16b yields

$$p_{\text{fix}}^{(1)} = \frac{s(r + s)}{s(s + r) + (b + \xi)(s + r\gamma_0)} \quad (22)$$

with  $\gamma_0 = N_0/K$ .

In the second scenario, birth and death rates are then given by

$$b(t, N_t) = b + r, \quad (23)$$

$$d(t, N_t) = b + r \frac{N_t}{K}.$$

While the deterministic dynamics at the population level are the same for both scenarios (given by Equation 21), this is not true for the fixation probability. From Equation 16b, we obtain for the fixation probability in the second scenario

$$p_{\text{fix}}^{(2)} = \frac{s(r+s)}{s(r+s) + (b+\xi)(s+r\gamma_0) + \gamma_0 r(r+s)} \leq p_{\text{fix}}^{(1)}, \quad (24)$$

i.e., a reduced probability relative to the first scenario. This is explained by the fact that genetic drift is stronger in the second scenario. For small values of  $s$  and  $r$  both fixation probabilities are approximately the same,  $p_{\text{fix}}^{(1)} \approx p_{\text{fix}}^{(2)} \approx s(r+s)/[(b+\xi)(s+r\gamma_0)]$ , and reproduce the result found by Otto and Whitlock (1997) if we choose  $b+\xi = 0.5$  such that the variance of the increase in the mutant number  $\text{Var}[\Delta n | n_t]$  is the same as in their model.

Note that for a sudden jump in population size, the result of our theory coincides with the result derived by Otto and Whitlock (1997) (where selection is effective during the change in population size), while Wahl and Gerrish (2001) consider a slightly different situation (where selection is switched off during the bottleneck; cf. Patwa and Wahl 2008).

**Linearly increasing selection:** When environmental conditions develop continuously in a given direction, the selective advantage of a mutation may gradually increase during the fixation process. Let us assume that the total population size stays constant and that the selection coefficient of the mutation linearly increases in time, thus  $s(t) = s_0 + s_1 t$ . We obtain the following expression for  $p_{\text{fix}}$  from Equation 16a,

$$p_{\text{fix}} = \left[ 1 + \xi \sqrt{\frac{\pi}{2s_1}} e^{s_0^2/2s_1} \text{erfc}\left(\frac{s_0}{\sqrt{2s_1}}\right) \right]^{-1}, \quad (25)$$

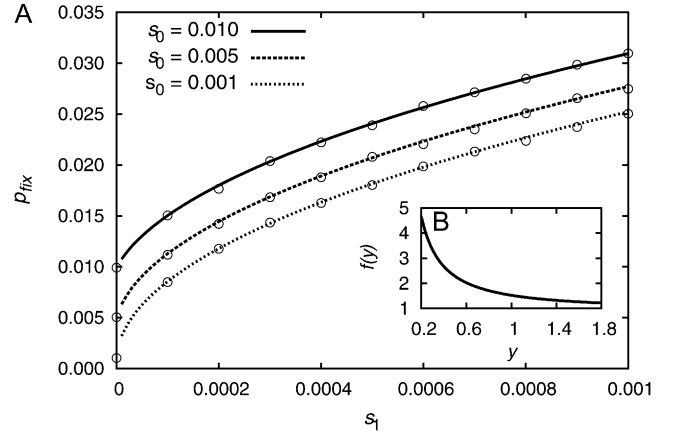
where  $\text{erfc}(x) = (2/\pi) \int_x^\infty e^{-t^2} dt$  is the complementary error function. For the special case  $s_0 = 0$ , the result simplifies,

$$p_{\text{fix}}(s_0 = 0) = \left[ 1 + \xi \sqrt{\frac{\pi}{2s_1}} \right]^{-1} \approx \frac{\sqrt{s_1}}{\xi} \sqrt{\frac{2}{\pi}}. \quad (26)$$

In Figure 1 the analytical results are compared to simulation results, showing very good agreement. The fixation probability increases significantly with  $s_1$ . Generally,

$$\frac{p_{\text{fix}}(s(t) = s_0 + s_1 t)}{p_{\text{fix}}(s(t) = s_0)} = \frac{\xi + s_0}{s_0 + \xi \sqrt{(\pi s_0^2/2s_1)} e^{s_0^2/2s_1} \text{erfc}(s_0/\sqrt{2s_1})} \approx \left[ \sqrt{\frac{\pi}{2}} y e^{y^2/2} \text{erfc}\left(\frac{y}{\sqrt{2}}\right) \right]^{-1} =: f(y) \quad (27)$$

with  $y := s_0/\sqrt{s_1}$ . The function  $f(y)$  is shown in the inset of Figure 1. For  $y = 1$  it is evaluated to be  $f(1) \approx 1.52$ . This



**Figure 1** (A) Fixation probability for linearly increasing selection strength  $s(t) = s_0 + s_1 t$  in dependence of  $s_1$  for various values of  $s_0$ . Simulations were performed for a population of 100,000 individuals, and each simulation point is the average over  $10^6$  runs. (B) Function  $f(y)$  (see Equation 27). The fixation probability increases significantly with increasing  $s_1$ .

means that if  $y \approx 1$ , i.e.,  $s_0 \approx \sqrt{s_1}$ , we observe an increase in the fixation probability of  $\sim 50\%$  in comparison to a mutation with constant selective advantage. E.g., for initially moderately strong selection  $s_0 = 0.01$  the fixation probability is still increased by  $\approx 50\%$  if selection increases as slightly as  $s(t) = 0.01 + 0.0001t$ .

**Periodically changing selection:** Cyclic environmental changes, such as seasonal changes or cyclic climate fluctuations (like the El Niño phenomenon), are frequent in nature. In the following, we consider strictly periodic changes, although our theory does not rely on this condition. Let us assume that again the total population size stays unaffected, but that the selective advantage of a mutation changes periodically with time; e.g.,

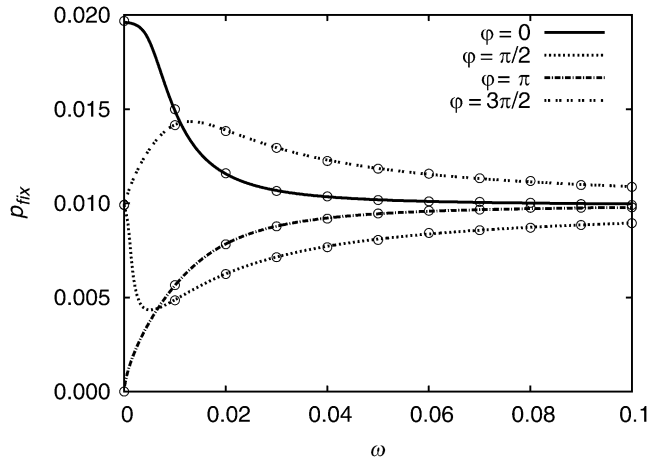
$$s(t) = s_0 + (s_{\text{max}} - s_0) \cos(\omega t + \varphi). \quad (28)$$

Depending on the parameter values, it is thus possible that the mutation is disadvantageous at certain periods of time. We obtain for the fixation probability:

$$p_{\text{fix}} = \left[ 1 + \xi \int_0^\infty \exp \left[ -s_0 t - (s_{\text{max}} - s_0) \times \left( t + \frac{1}{\omega} \sin(\omega t + \varphi) - \frac{1}{\omega} \sin(\varphi) \right) \right] dt \right]^{-1}. \quad (29)$$

The integral can be evaluated only numerically. Comparison to simulated data (see Figure S4) shows that the theory provides an accurate prediction of the fixation probability also for scenarios in which the mutation temporarily becomes disadvantageous. For  $s_0 \leq 0$ , however, it predicts a fixation probability of zero and therefore underestimates the true value.

To slightly reduce the parameter space, we concentrate now on the special case  $s_{\text{max}} = 2s_0$ , i.e.,



**Figure 2** Comparison between the analytical theory and simulation results for the fixation probability in the case of periodically changing selection  $s(t) = s_0(1 + \cos(\omega t + \phi))$ ,  $s_0 = 0.01$ . Simulations were performed for a population of 100,000 individuals, and every simulation point is the average over  $10^7$  runs.

$$s(t) = s_0(1 + \cos(\omega t + \phi)). \quad (30)$$

Figure 2 shows how the fixation probability changes with  $\omega$  for various values of  $\phi$ . For small and intermediate values of  $\omega$ , the value of  $\phi$  has a strong impact on the result. If  $\omega$  increases, the fixation probability converges to the value for a mutation with selective advantage  $s(t) = s_0$  for all values of  $\phi$ . This behavior is further illustrated in Figure 3, which shows the fixation probability and the initial selection strength in dependence of  $\phi$  for various values of  $\omega$ . In the limit  $\omega \rightarrow 0$ , the fixation probability equals  $s(0)/(1 + s(0)) \approx s(0)$ . It therefore follows approximately the curve of  $s(0)$  for small values of  $\omega$ . For large values it converges to  $s_0$  for all values of  $\phi$ . In an intermediate regime  $s_0/10 \leq \omega \leq 10s_0$ , more complex behavior is found: Here, not only the initial value  $s(0)$  but also the following time development of the selection strength [i.e., not only  $s(0) = s_0(1 + \cos(\phi))$ , but  $\phi$  itself] becomes important. The extrema of  $p_{\text{fix}}(\phi)$  are attained for smaller values of  $\phi$  than the extrema of  $s(0)$ . A straightforward calculation shows furthermore that

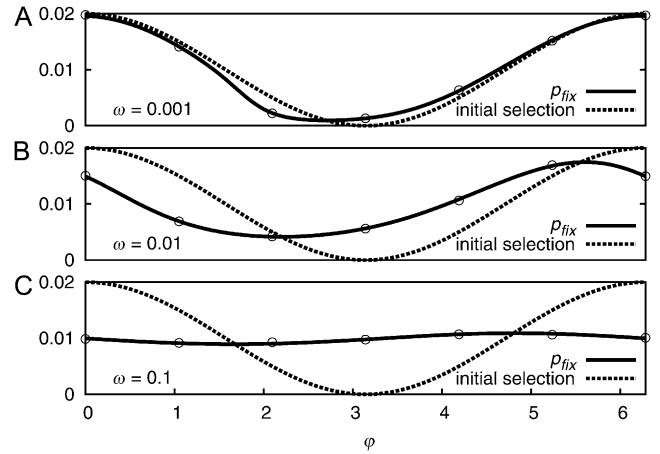
$$p_{\text{fix}}(\phi_e) = \frac{s_0(1 + \cos(\phi_e))}{1 + s_0(1 + \cos(\phi_e))}, \quad (31)$$

where  $\phi_e$  is the value where the extremum is attained. We thus see that the fixation probability at the extrema  $\phi_e$  is the same for cyclic selection and constant selection with  $s = s(0) = s_0(1 + \cos(\phi_e))$ .

### Time to reach an intermediate frequency $x_c$

#### Analytical theory

In models of the adaptive process, it is often necessary to know the time that it takes for a successful mutation to become established and to reach a certain threshold frequency (e.g., Desai and Fisher 2007; Kopp and Hermisson



**Figure 3** (A–C) Fixation probability and initial selection strength  $s(0) = s_0(1 + \cos(\phi))$  in dependence of  $\phi$  for various values of  $\omega$  when the selection coefficient changes according to  $s(t) = s_0(1 + \cos(\omega t + \phi))$ ,  $s_0 = 0.01$ . Again, one observes the expected limiting behavior for small and large values of  $\omega$ . For intermediate values of  $\omega$  a nontrivial behavior is found. Simulations were performed for a population of 100,000 individuals, and each simulation point is the average over  $10^6$  runs.

2009a). It is well known that the deterministic model of the allele frequency increase yields a poor approximation even for the expected value of this time. The reason is that a mutation that has survived the initial phase will on average have grown faster during this phase than predicted by the deterministic model. When the frequency is finally large enough that stochasticity can be neglected and the path be modeled deterministically, the frequency will thus be larger than if it had *always* grown following the deterministic path. In that phase, it is well described by the deterministic path that is not started from a single individual, but from an (on average) larger “effective initial population size.” The method presented here builds on work by Cohn and Jagers (1994) and Desai and Fisher (2007). It consists of subsuming all the stochasticity of the path under this effective initial population size as a single random variable and then modeling the path deterministically. The procedure consists of two steps: In a first step the distribution of the effective initial population size is estimated via a branching process. In a second step the deterministic approximation of the full birth–death model is used to describe the allele frequency path starting from the effective initial population size.

Let us again consider the phase in which the mutation is rare and in which the dynamics can be described by the time-inhomogeneous branching process (Equation 7). The key to the method is that it is possible to separate stochastic fluctuations and deterministic growth for this process. Here, deterministic growth coincides with the time-development of the expected number of mutants:

$$E[n_t] = n_0 \exp(\rho(t)). \quad (32)$$

In the following, we restrict to the case  $n_0 = 1$ . Define now a new random variable:

$$v_t := \frac{n_t}{E[n_t]} = n_t \exp(-\rho(t)). \quad (33)$$

$v_t$  describes all the stochasticity that has accumulated in the branching process until time  $t$ . Crucially, a theorem by Cohn and Jagers (1994) guarantees that for  $t \rightarrow \infty$ ,  $v_t$  converges (almost surely) to a positive random variable  $v$  that summarizes the entire stochasticity of the process. Since  $n_t = v \exp(\rho(t))$  in the limit  $t \rightarrow \infty$ , we can interpret  $v$  as the random initial population size of an ensemble of deterministically growing paths that approximate the original process  $n_t$  for large  $t$ .

For the fixation process, in particular, we are interested in the distribution of  $v$  conditioned on non-extinction. We proceed as follows: From the probability generating function  $P(z, t) \equiv P_1(z, t)$  (see Equation 10), we obtain for the probability  $p_n(t) \equiv p_n(1, t)$  to have  $n$  individuals at time  $t$  (proof by induction, see section S3 of File S1):

$$p_n(t) = \frac{1}{n!} \left. \frac{d^n P(z, t)}{dz^n} \right|_{z=0} = \frac{B^{n-1}(t) A(t)}{(A+B)^{n+1}(t)}, \quad n \geq 1. \quad (34)$$

Conditioning on non-extinction [which requires the biologically meaningful condition  $p_0(t) \neq 1$ ] leads to

$$\text{Prob}[n, t | \text{not extinct}] = \frac{p_n(t)}{1 - p_0(t)} = \frac{A(t)}{B(t)} \left( \frac{B(t)}{(A+B)(t)} \right). \quad (35)$$

By induction, we find

$$E[n_t^k | \text{not extinct}] = \frac{1}{1 - p_0(t)} E[n_t^k] \quad \text{and} \quad E[n_t^k] \sim e^{k\rho(t)}, \quad (36)$$

i.e., for large times the  $k$ th moment grows in the conditioned as well as in the unconditioned process as  $e^{k\rho(t)}$ . In particular, we obtain for the expected value

$$E[n_t | \text{not extinct}] = \frac{1}{1 - p_0(t)} e^{\rho(t)}. \quad (37)$$

[Since  $1/(1 - p_0(t)) \geq 1$ , a comparison with Equation 32 confirms that on average the conditioned path grows faster than the unconditioned path.] The distribution function of  $v_t | \text{not extinct}$  is immediately obtained from the distribution function of  $n_t | \text{not extinct}$ :

$$P(n_t \leq n_0 | \text{not extinct}) = 1 - \left( \frac{B(t)}{(A+B)(t)} \right)^{n_0} \quad (38)$$

$$\Rightarrow P(v_t \leq v_0 | \text{not extinct})$$

$$= 1 - \exp \left[ v_0 \exp(\rho(t)) \ln \left( \frac{B(t)}{(A+B)(t)} \right) \right]. \quad (39)$$

As  $\lim_{t \rightarrow \infty} \exp(\rho(t)) \ln(B(t)/(A+B)(t)) = -1/(A+B) = -p_{\text{fix}}$ , we obtain in the limit  $t \rightarrow \infty$  the stationary distribution

$$P(v \leq v_0 | \text{not extinct}) = 1 - \exp(-p_{\text{fix}} v_0). \quad (40)$$

This effective initial mutant population size may now be used as the starting value for the deterministic solution of the full model. If the birth and death rates are the same for mutants and residents, the latter can be obtained from Equation 4a and is given by a generalized logistic,

$$x(t) = \frac{\exp(\beta(t))}{N_0/v - 1 + \exp(\beta(t))}, \quad (41)$$

where  $x(t)$  is the mutant frequency and

$$\beta(t) = \int_0^t s(\tau) d\tau. \quad (42)$$

We want to calculate the time needed to reach an intermediate frequency  $x_c$ . Here, intermediate frequency means a frequency at which the dynamics are well described by Equation 41, i.e., not too close to either 0 or 1. Using Equation 41, we can express the random variable  $T_{x_c}$  defined via  $x(T_{x_c}) = x_c$  in terms of  $v$ :

$$T_{x_c} = \begin{cases} \beta^{-1} \left( \ln \left( \frac{x_c(N_0 - v)}{(1 - x_c)v} \right) \right), & v < x_c N_0 \\ 0, & v \geq x_c N_0. \end{cases} \quad (43)$$

For the definition of  $T_{x_c}$  to be unique for all  $x_c$ , the condition  $s(t) > 0$  (up to single points) is necessary. It is possible that the process grows so quickly that the effective initial mutant population size is already larger than  $x_c N$  [cf. the exponential tail in the distribution function for the effective initial population size  $v$  (Equation 40)]. In this case, we have set  $T_{x_c} = 0$  so that the distribution of  $T_{x_c}$  will have a mass on  $T_{x_c} = 0$ . However, if  $x_c$  is not too small, this is very unlikely and the mass will be small. As the deterministic path (Equation 41) ignores stochastic fluctuations, there is only a single passage time in the deterministic approximation in contrast to the stochastic path in which the frequency  $x_c$  can be hit several times.  $T_{x_c}$  as defined in Equation 43 is best interpreted as the average over all times at which the path crosses the frequency  $x_c$  from lower to higher values.

By a parameter transformation ( $v \rightarrow T_{x_c}$  using Equation 43), the distribution of  $T_{x_c}$  is found from Equation 40:

$$P(T_{x_c} \leq T) = \exp \left[ - \frac{p_{\text{fix}} N_0}{((1 - x_c)/x_c) \exp(\beta(T)) + 1} \right]. \quad (44)$$

Equation 44 constitutes the main result of this section.

$\beta(t)$  is a function of  $s(t)$ , and  $p_{\text{fix}}$  can be expressed in terms of  $s(t)$  and  $N_e(t)$  (Equation 16b). Like the fixation probability, the distribution of  $T_{x_c}$  therefore depends on

two parameters that summarize the time-dependence of the model. For the fixation probability, we have seen in Equation 16a that there is a second way to summarize the dynamics in terms of the absolute rate of increase  $(s + r)(t)$  and the total rate of birth and death events. This is not possible for the distribution of  $T_{x_c}$ , linked to the fact that variable selection and demographic change are not equivalent in this case. While the selection function  $s(t)$  has a strong influence on the deterministic part of the frequency path  $x(t)$ , changes in the total population size (with equal effect on mutants and residents) have no influence on the latter.

To calculate moments of the distribution, we perform the parameter transformation

$$T_{x_c} \rightarrow Y := \frac{p_{\text{fix}} N_0}{((1 - x_c)/x_c) \exp(\beta(T_{x_c})) + 1}. \quad (45)$$

Solving for  $T_{x_c}$  yields

$$T_{x_c} = \beta^{-1} \left( \ln \left[ \frac{x_c}{1 - x_c} \left( \frac{p_{\text{fix}} N_0 - Y}{Y} \right) \right] \right); \quad (46)$$

and we obtain for the moments

$$\langle T_{x_c}^n \rangle = \int_0^{p_{\text{fix}} N_0 x_c} \left[ \beta^{-1} \left( \ln \left[ \frac{x_c}{1 - x_c} \left( \frac{p_{\text{fix}} N_0 - Y}{Y} \right) \right] \right) \right]^n \exp(-Y) dY. \quad (47)$$

For constant selection and therefore  $\beta(t) = st$  we obtain for the mean of  $T_{x_c}$ :

$$\begin{aligned} \langle T_{x_c} \rangle = & \frac{1}{s} \left[ \ln[p_{\text{fix}} N_0] + \gamma + \ln \left[ \frac{x_c}{1 - x_c} \right] - \frac{1}{p_{\text{fix}} N_0} - \left( \frac{1}{p_{\text{fix}} N_0} \right)^2 \right. \\ & \left. + \mathcal{O} \left( \left( \frac{1}{p_{\text{fix}} N_0} \right)^3 \right) \right]. \end{aligned} \quad (48)$$

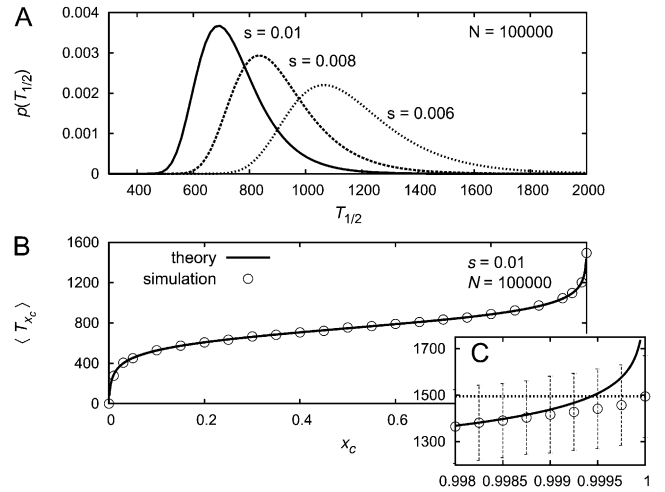
### Applications

**Constant selection and constant population size:** For the distribution of the passage time  $T_{x_c}$ , we obtain with  $p_{\text{fix}} = s/(\xi + s) \approx 2s(N_e/N)$  (see Equation 17)

$$\begin{aligned} P(T_{x_c} \leq T) = & \exp \left[ -\frac{s}{\xi + s} \frac{N}{((1 - x_c)/x_c) \exp(sT) + 1} \right] \\ \approx & \exp \left[ -\frac{2sN_e}{((1 - x_c)/x_c) \exp(sT) + 1} \right]. \end{aligned} \quad (49)$$

Plots of the probability density for  $x_c = 0.5$  and various values of  $s$  are depicted in Figure 4A. With increasing selection strength the distribution gets narrower and is shifted to the left.

In the particular situation of constant selection and population size, additional results can be obtained. First,



**Figure 4** (A) Probability density of the time to reach frequency  $x_c = 0.5$  in the case of constant selection and constant population size. With decreasing selection strength the density function becomes broader and is shifted to the right. (B) Mean time to reach a frequency  $x_c$  in dependence of  $x_c$ . Theoretical results are compared to simulation results. Each simulation point is the average over 1000 runs. (C) Amplification of the region of high frequencies. Simulations show the mean average passage time. One sees that the theoretical predictions become less good for  $x_c > 1 - 1/\alpha$ . The dotted line is the result for the whole fixation time obtained by the convolution, i.e., two times the mean time needed to reach the frequency  $x_c = 0.5$ . The bars around the simulation points denote the standard deviation of  $T_{x_c}$  as obtained from the simulations. In contrast, the standard error of the simulations is much smaller and vanishes in the symbols.

we note that  $T_{x_c}$  can also be interpreted as the age of a derived allele that is currently found at frequency  $x_c$  in the population. This is a consequence of the time-homogeneity of the model. Second, the distribution of the time to reach fixation at  $x = 1$  starting from a frequency  $(1 - x_c)$  equals the distribution of  $T_{x_c}$  (see section S3 of File S1 for an explanation). In particular, the times needed from 0 to 0.5 and from 0.5 to 1 follow the same distribution. Since both random variables are independent, we obtain the probability density  $\tilde{p}(t_{\text{fix}})$  of the whole fixation time from the density  $p(T_{1/2})$  as

$$\tilde{p}(t_{\text{fix}}) = \int_0^{t_{\text{fix}}} p(T_{1/2}) p(t_{\text{fix}} - T_{1/2}) dT_{1/2}. \quad (50)$$

An alternative expression for the distribution of  $t_{\text{fix}}$  has been derived before by Wang and Rannala (2004), using diffusion techniques. We note that our approximation (Equation 50) is simpler than the series expansion in terms of Gegenbauer polynomials, using the eigenvalues of the oblate spheroidal angular function and the intermediate coefficients of the spheroidal harmonics obtained by these authors. As discussed in section S1 of File S1, it is nevertheless highly accurate if selection is not very weak or the population size small. The cumulants of the fixation time are just two times the cumulants of the time to reach frequency 0.5. For the mean fixation time, in particular, we obtain

$$\langle t_{\text{fix}} \rangle = \frac{2}{s} \left( \ln \left( \frac{sN}{\xi + s} \right) + \gamma - \frac{\xi + s}{sN} - \frac{(\xi + s)^2}{(sN)^2} + \mathcal{O} \left( \frac{(\xi + s)^3}{(sN)^3} \right) \right), \quad (51)$$

where  $\gamma \approx 0.577$  is the Euler constant.

For the higher cumulants, approximations of leading order in  $sN$  can be obtained by the approximation  $\exp(sT) + 1 \approx \exp(sT)$  in the distribution function (Equation 49) and by extending the integral (Equation 47) to infinity. We state the results for the variance, the skewness, and the kurtosis,

$$\kappa_2 = \langle t_{\text{fix}}^2 \rangle - \langle t_{\text{fix}} \rangle^2 \approx \frac{1}{3} \frac{\pi^2}{s^2}, \quad (52a)$$

$$\kappa_3 \approx \frac{4\zeta(3)}{s^3} > 0, \quad (52b)$$

$$\kappa_4 \approx \frac{2}{15} \frac{\pi^4}{s^4} > 0, \quad (52c)$$

where  $\zeta(z)$  is the Riemann Zeta function and  $\zeta(3) \approx 1.202$  is the Apéry constant. It can be shown that all cumulants of order  $\geq 2$  are to leading order of the form

$$\kappa_j = c_j \left( \frac{1}{s} \right)^j \quad (53)$$

with some numerical constant  $c_j$ .

Since  $s > 0$  is required for the branching approximation, deviations from the exact values must be expected for weak selection. To estimate the precision of our approximation in this case, we compare the result for the mean fixation time to the result obtained from the diffusion approximation. For this purpose we scale time with  $N(t_{\text{fix}} \rightarrow \tau_{\text{fix}} = t_{\text{fix}}/N)$  and take the diffusion limit (i.e.,  $s \rightarrow 0$ ,  $N \rightarrow \infty$  with  $\alpha := sN/\xi$  remaining constant) of Equation 51. We obtain

$$\langle \tau_{\text{fix}} \rangle = \frac{2}{\alpha} \left( \ln(\alpha) + \gamma - \frac{1}{\alpha} - \frac{1}{\alpha^2} + \mathcal{O} \left( \frac{1}{\alpha^3} \right) \right). \quad (54)$$

This is in perfect agreement with the approximation for the mean fixation time found in Hermisson and Pennings (2005), using the diffusion approximation. As shown in section S2 of [File S1](#), the deviation of our branching process approximation from the exact diffusion result is exponentially small in  $\alpha/2$ .

Figure 4, B and C, shows the mean time to reach a frequency  $x_c$  in dependence of  $x_c$ . Significant deviations between the theoretical predictions of Equation 49 and simulation results are found only for values of  $x_c$  that are extremely close to 1 (or 0). Generally, it holds that the approximation becomes less good for  $x_c \gtrsim 1 - 1/\alpha$  (or  $x_c \lesssim 1/\alpha$ ), as in this regime the deterministic frequency path (Equation 41), which we used in the derivation, is not a good approximation to the real frequency path of the mutation (see Ewens 2004, p. 167f).

It is interesting to note that the distribution of  $T_{x_c}$  is the same for beneficial mutations with selective advantage  $s$  and deleterious mutations with selective disadvantage  $-s$  (cf. Ewens 2004): For every point of a particular path leading from  $n = 1$  to  $n = n_c$ , the sum of the birth and the death rate is the same for both scenarios. This implies that the distribution of the run time of this particular path is the same. The chance to be realized differs for each such path by a constant factor  $(\xi + s)^{-n_c}$ . If we therefore condition on  $n = n_c$  being reached, the distribution of the time to reach  $n_c$  is the same for beneficial and deleterious mutations.

**Constant selection and changing population size:** Using the expression for  $p_{\text{fix}}$  calculated for a logistically growing population, we immediately get the distribution of the passage time  $T_{x_c}$ . [Figure S5](#) shows a comparison to simulation results for the second scenario. With Equation 48 we obtain the result for the mean of  $T_{x_c}$ :

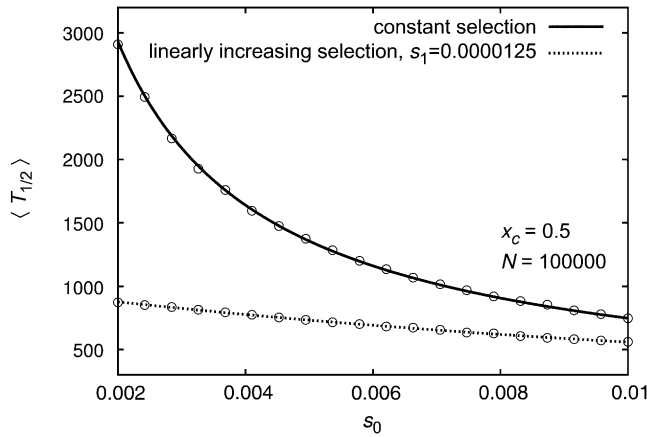
$$\langle T_{x_c} \rangle \approx \frac{1}{s} \left[ \ln \left[ N_0 \frac{s(r+s)}{s(r+s) + (b+\xi)(s+r\gamma_0) + \gamma_0 r(r+s)} \right] + \gamma + \ln \left[ \frac{x_c}{1-x_c} \right] \right]. \quad (55)$$

**Linearly increasing selection:** Putting the expression obtained for  $p_{\text{fix}}$  (Equation 25) into Equation 44, one obtains the distribution of the time  $T_{x_c}$  at which a fraction  $x_c$  of the population consists of mutant individuals. The distribution can be used to calculate the mean of  $T_{x_c}$  numerically. A comparison to simulation results for the distribution and the mean of  $T_{x_c}$  in dependence of  $s_1$  can be found in [Figure S6](#). From [Figure 5](#), it can be seen that for small values of  $s_0$  even a very slight increase of the selection strength over time leads to a drastically smaller mean value of  $T_{x_c}$  in comparison to an environment with constant selection  $s = s_0$ .

**Periodically changing selection:** [Figure 6](#) shows the mean time to reach frequency  $x_c = 0.5$  for a periodic selection coefficient  $s(t) = s_0(1 + \cos(\omega t + \phi))$ . As for the fixation probability, the initial selection strength  $s(0)$  is decisive if the environment changes very slowly;  $\phi$  itself gains importance with increasing (but still small)  $\omega$ . If the environment changes fast,  $\langle T_{x_c} \rangle$  becomes independent of  $\phi$  and converges to its value for constant selection  $s(t) = s_0$ . Convergence is much faster than for the fixation probability.

## Simulations

To test our analytical results, we performed individual based computer simulations for which we use a Gillespie algorithm (Gillespie, 1977). Events happen at rate  $(2\xi(t, N_t) + s(t, N_t))n_t(N_t - n_t)/N_t + b(t, N_t)N_t + d(t, N_t)N_t$ , where  $\xi(t, N_t)$ ,  $s(t, N_t)$ ,  $b(t, N_t)$  and  $d(t, N_t)$  are assumed to be constant between events. Once the time of an event is fixed, which kind of event takes place is determined using the respective probabilities. Subsequently, the values of  $N_t$



**Figure 5** Mean time to reach frequency  $x_c = 0.5$  in dependence of  $s_0$ , where a scenario with constant selection and a scenario with linearly increasing selection ( $s_1 = 1/80,000$ ) are compared. The mean fixation time is significantly smaller in the case of linearly increasing selection than in the case of constant selection. Each simulation point is the average over 1000 runs.

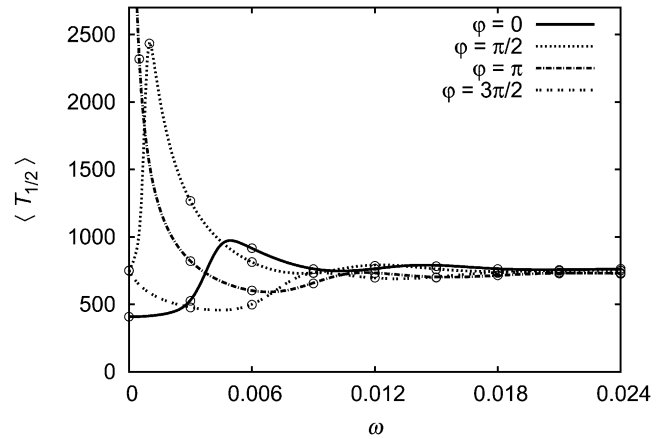
and  $n_t$  are updated and the rates are set to their new values (additional update steps between events did not change the results).

For most of the simulation runs, we used the first passage time of a given frequency  $x_c$  to determine  $T_{x_c}$ , i.e., we neglected fluctuations around  $x_c$ , which for large values of  $\alpha$  and not too small or large values of  $x_c$  has no significant effect. For the data shown in panel C of Figure 4 and in Figure S2 and Figure S3, where we pushed the boundaries of the theory, we took the average over all passage times, more precisely the average over all times at which the path crossed the frequency  $x_c$  from lower to higher values (cf. section on the analytical theory).

For all simulation results, the number of runs was chosen sufficiently large that the standard error bars vanish in the symbols. All programs were written in C, making use of the Gnu Scientific Library (Galassi *et al.*, 2009). We used Mathematica (Wolfram Research, Champaign, USA) for all numerical evaluation of integrals.

## Discussion

Adaptation is the evolutionary response of a population to an environmental challenge: variation in the environmental conditions leads to altered selection pressures and changing population sizes. In nature, these changes occur on all time scales, from rapid shifts within a single generation to long-term geological trends. It therefore seems natural that models on the genetics of adaptation, too, should account for the ecological dynamics that drive the process. In contrast to this expectation, however, the vast majority of published studies with a genetic focus assume a constant environment (reviewed, e.g., in Orr 2010). They rely on the idea that a fast—almost instantaneous—change in the environment is followed by a period of environmental stasis. If



**Figure 6** Expected value of  $T_{1/2}$ . Selection is periodic according to  $s(t) = s_0(1 + \cos(\omega t + \phi))$  with  $s_0 = 0.01$ . The total population size is  $N = 100,000$ . Each simulation point is the average over 2000 runs.

this period is long compared to the total time it takes for one or several beneficial mutations to appear and to rise to fixation, ecology and evolution are effectively decoupled. In many cases of ecological interest, however, this separation of time scales is not appropriate. In this case, the parameters of the evolutionary model (such as selection coefficients or population sizes) are turned into time-dependent variables, which must be determined from an underlying ecological model.

In a series of recent publications by several authors, the impact of the ecological dynamics on the genetics of adaptation has been studied for the so-called moving optimum model (Bello and Waxman 2006; Collins *et al.* 2007; Kopp and Hermisson 2007, 2009a,b). The consensus, at least for this model, is that the adaptive process is strongly affected by the dynamics of the selective environment. In this article, we present a more detailed treatment of the most basic aspect of the genetics of adaptation: the fixation process of a single beneficial mutation in a variable environment. All relevant parameters of the process, i.e., selection coefficient, population size, and the variance in offspring number (or, equivalently, the birth and death rates of wild types and mutants) may depend explicitly on time. An example of how this can result from an explicit ecological model is given in the *Appendix*, where we discuss the fixation probability of a “rescue mutation” in a population that is otherwise doomed to extinction. Again, the results are strongly influenced by the ecological dynamics, leading to qualitative differences relative to previous studies that assume constant selection (Orr and Unckless 2008).

Even if the external environment is constant, individual alleles in a population can experience variable selection pressures if multiple selected alleles segregate in the population and if these alleles interfere due to either epistasis or linkage (cf. Hartfield and Otto 2011). Examples include the evolution of compensatory mutations or classical problems of clonal interference, where the beneficial mutation rate is

high relative to the recombination rate. Another potential application is adaptive gene flow across a genetic barrier, where adaptations need to be purged from linked deleterious alleles by recombination.

Using branching process techniques, we obtain analytical approximations for the fixation probability and the distribution of the time for the mutant to reach a given frequency  $x_c$  in the population.

### Fixation probability

The derivation of fixation probabilities of rare beneficial mutations from a supercritical branching process is a standard approach in population genetics. In particular, results have previously been obtained for the most important modes of demographic changes (e.g., Otto and Whitlock 1997; Pollak 2000). A look into the mathematical literature reveals, however, that a general formalism with arbitrary time-dependent birth and death rates has been available since the work of Kendall (1948). In contrast to most of the previous studies in population genetics, which use a branching process in discrete time (following Haldane 1927), this approach is based on a continuous-time framework, which simplifies some technical aspects. Our adaptation of this formalism to the genetic context leads to a compact, yet general formula for the fixation probability  $p_{\text{fix}}$  (Equations 16a and 16b) that covers previous results as special cases. In section S2 of [File S1](#), we show how an analogous result can also be obtained using the diffusion approach.

It turns out that the ecological dynamics affect  $p_{\text{fix}}$  through two independent variables. In two alternative formulations of the result, these can be either the time-dependent birth and death rates of rare mutants (Equation 16a) or the selection coefficient  $s(t)$  and the variance effective population size  $N_e(t)$  (Equation 16b). Our applications to various scenarios show that relative to the case with constant selection  $s_0$ , consistent changes in the selection coefficient  $\Delta s$  per generation of the order of  $\Delta s \geq s_0^2$  have a strong effect on  $p_{\text{fix}}$ . This is to be expected since, for constant selection, the fate of a new mutation (fixation or loss) is decided once the mutant allele reaches a frequency on the order of  $(Ns_0)^{-1}$ . According to Equation 48, this will take, on average, in the order of  $s_0^{-1}$  generations. The observation is of practical importance since it shows that predictions about fixation probabilities cannot be based on short-term fitness assays. Unmeasurably small fitness changes across generations may have a large effect. The limitations of the branching approach are the usual ones: The fate of the mutation must be decided while the mutant frequency is small and the independence assumption of the branching process applies. Deviations from the simulation results are found in the case of mutations that are almost neutral on average, with fixation probabilities  $p_{\text{fix}} \lesssim 10/N$  (cf. [File S1](#)).

### Time $T_{x_c}$ to reach frequency $x_c$

The usual method in population genetics to derive fixation times or (more generally) first-passage times for mutant alleles to reach certain frequencies is diffusion theory. Since this proves difficult for the time-inhomogeneous case, we again turn to a branching process approach. We can use the fact that (almost) all stochasticity of a mutant trajectory  $x(t)$  is due to its early phase while the mutant number is still small. We can therefore (approximately) describe this stochasticity in the branching framework and combine it with the deterministic growth of the full process. For constant selection and population sizes, this idea has previously been used by Desai and Fisher (2007) in the context of a model for clonal evolution.

For the general case, the crucial step of the method has again been anticipated in the mathematical literature (Cohn and Jagers 1994). There, it has been shown that a clean separation exists for the random variable  $n_t$  of an inhomogeneous supercritical branching process into a growth function that describes the asymptotic growth and a time-constant random variable  $\nu$  that describes the stochastic fluctuations in the large-time limit. We can interpret  $\nu$  as the effective initial size of the mutant population. It turns out that the distribution function of  $\nu$ ,  $P(\nu \leq \nu_0 | \text{not extinct}) = 1 - \exp(-p_{\text{fix}}\nu_0)$  (Equation 40), is pleasingly simple even in a variable environment. In particular, the impact of the ecological dynamics is conveniently summarized in the fixation probability  $p_{\text{fix}}$ . When this initial size is combined with the deterministic growth of the full model, an approximation for the distribution function for  $T_{x_c}$  is obtained by a simple transformation of the probability density (Equation 44).

Computer simulations show that the results from the branching approach are usually highly accurate as long as selection is not very weak. This is also confirmed by a comparison with previous results for the expected value of the fixation time  $T_1$  for constant  $s$  and  $N$  (Kimura and Ohta 1969; Ewens 2004) derived from the diffusion approximation: The order of the error term for the average fixation time is  $o(\exp(-Ns/2))$  (see [File S1](#)). Previous results also exist for the entire probability density of  $T_1$  in a constant environment (Wang and Rannala 2004). Here, the branching approximation leads to a much simpler analytical result.

As a major caveat of the method, we note that for variable  $s$  or  $N$ , we do not obtain a closed expression for the time to reach full fixation  $x_c = 1$ . More generally, the approximation is accurate only as long as  $1/(sN) \lesssim x_c \lesssim 1 - 1/(sN)$ . For  $x_c$  close to 0, estimates can easily be found directly from the conditioned branching process, if needed. For mutant frequencies near 1, the fixation process enters another stochastic phase when the wild-type alleles become rare. Although this phase can be described, in principle, as a subcritical branching process, its details depend on the random time when this stochastic phase is entered, which complicates the treatment. In natural systems, an allele

frequency of  $x_c = 1$  is not an absorbing state and may never be reached in the face of back mutation, migration, or ongoing adaptation. For many practical applications, other thresholds (such as 5%, 50%, or 95%) are therefore more relevant and have previously been used. For example, Desai and Fisher (2007) use a threshold of  $x_c = 1/(Ns)$  to characterize their establishment time  $\tau_{\text{est}}$ . Kopp and Hermisson (2007, 2009a) use  $x_c = 0.5$  as the critical value where the mutant comes to dominate the population in an analysis of the order and step sizes of adaptive substitutions.

Although we have formulated our results for haploids, they also apply to diploids as long as the mutant is not fully recessive (i.e.,  $h_s$  must be sufficiently large on average). Only the selection coefficients (or birth and death rates) of heterozygotes enter into the stochastic part of the model described by the branching process. In contrast, the deterministic frequency path used in the derivation of  $T_{x_c}$  depends on the fitness values of both heterozygous and homozygous mutants. In the purely recessive case, the dynamics of rare mutants can no longer be described by a supercritical branching process. Diffusion methods (for constant selection) show that also the scaling of the fixation time with the selection parameters is altered in this case (Ewing *et al.* 2011).

To summarize, our results show that inhomogeneous branching processes provide a powerful framework to describe the fixation process under a wide range of ecological scenarios and genetic conditions. We therefore hope that the methods and results can provide a step to combine ecological and genetic modeling traditions to study the genetics of adaptation under realistic ecological conditions.

*Note:* After acceptance of this article we became aware of a recent preprint by Waxman on the fixation process under variable selection and demography [see D. Waxman (pp. 907–913) in this issue]. Both articles are complementary: While our approach focuses on analytical approximations for the fixation process of a definitely beneficial mutant, Waxman (2011) presents numerical methods to derive fixation probabilities for alleles with arbitrary selection coefficients.

## Acknowledgments

We thank Nick Barton, Ulf Dieckmann, Peter Pfaffelhuber, and Matthew Hartfield for helpful discussions and Lindi Wahl and two anonymous referees for many useful comments. This work was made possible with financial support from the Vienna Science and Technology Fund and from the Deutsche Forschungsgemeinschaft, Research Unit 1078, *Natural selection in structured populations*.

## Literature Cited

Allen, L. J. S., 2011 *An Introduction to Stochastic Processes with Applications to Biology*, Ed. 2. Pearson Education, Upper Saddle River, NJ.

- Barton, N. H., 1995 Linkage and the limits to natural selection. *Genetics* 140: 821–841.
- Bello, Y., and D. Waxman, 2006 Near-periodic substitutions and the genetic variance induced by environmental change. *J. Theor. Biol.* 239: 152–160.
- Chia, A. B., 1968 Random mating in a population of cyclic size. *J. Appl. Probab.* 5: 21–30.
- Cohn, H., and P. Jagers, 1994 General branching processes in varying environment. *Ann. Appl. Probab.* 4(1): 184–193.
- Collins, S. J., J. de Meaux, and C. Acquisti, 2007 Adaptive walks toward a moving optimum. *Genetics* 176: 1089–1099.
- Desai, M. M., and D. S. Fisher, 2007 Beneficial mutation-selection balance and the effect of linkage on positive selection. *Genetics* 176: 1759–1798.
- Ewens, W. J., 1967 The probability of survival of a new mutant in a fluctuating environment. *Heredity* 22: 438–443.
- Ewens, W. J., 2004 *Mathematical Population Genetics*, Ed. 2. Springer-Verlag, New York.
- Ewing, G., J. Hermisson, P. Pfaffelhuber, and J. Rudolf, 2011 Selective sweeps for recessive alleles and for other modes of dominance. *J. Math. Biol.* (in press).
- Fisher, R. A., 1922 On the dominance ratio. *Proc. R. Soc. Edinb.* 42: 321–341.
- Fisher, R. A., 1930 The distributions of gene ratios for rare mutations. *Proc. R. Soc. Edinb.* 50: 204–219.
- Galassi, M., J. Davies, J. Theiler, B. Gough, G. Jungman, *et al.*, 2009 *GNU Scientific Library Reference Manual*, Ed. 3. Network Theory Ltd., Bristol, UK.
- Gillespie, D. T., 1977 Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* 81(25): 2340–2361.
- Haccou, P., P. Jagers, and V. A. Vatutin, 2005 *Branching Processes. Variation, Growth and Extinction of Populations*. Cambridge University Press, Cambridge, UK.
- Haldane, J. B. S., 1927 A mathematical theory of natural and artificial selection. V. Selection and mutation. *Proc. Camb. Philos. Soc.* 23: 838–844.
- Hartfield, M., and S. P. Otto, 2011 Recombination and hitchhiking of deleterious alleles. *Evolution* (in press).
- Hermisson, J., and P. S. Pennings, 2005 Soft sweeps: molecular population genetics of adaptation from standing genetic variation. *Genetics* 169: 2335–2352.
- Huillet, T., 2011 On the Karlin-Kimura approaches to the Wright-Fisher diffusion with fluctuating selection. *J. Stat. Mech.* 2: P02016.
- Jensen, L., 1973 Random selective advantages of genes and their probabilities of fixation. *Genet. Res.* 21(3): 215–219.
- Karlin, S., and B. Levikson, 1974 Temporal fluctuations in selection intensities: case of small population size. *Theor. Popul. Biol.* 6: 383–412.
- Kendall, D. G., 1948 On the generalized “birth-and-death” process. *Ann. Math. Stat.* 19(1): 1–15.
- Kimura, M., 1962 On the probability of fixation of mutant genes in a population. *Genetics* 47: 713–719.
- 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.
- Kimura, M., and T. Ohta, 1974 Probability of gene fixation in an expanding finite population. *Proc. Natl. Acad. Sci. USA* 71: 3377–3379.
- Kopp, M., and J. Hermisson, 2007 Adaptation of a quantitative trait to a moving optimum. *Genetics* 176: 715–719.
- Kopp, M., and J. Hermisson, 2009a The genetic basis of phenotypic adaptation I: fixation of beneficial mutations in the moving optimum model. *Genetics* 182: 233–249.
- Kopp, M., and J. Hermisson, 2009b The genetic basis of phenotypic adaptation II: fixation of beneficial mutations in the moving optimum model. *Genetics* 183: 1453–1476.

Moran, P. A. P., 1958a The effect of selection in a haploid genetic population. *Math. Proc. Camb. Philos. Soc.* 54: 463–467.

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

Novozhilov, A. S., G. P. Karev, and E. V. Koonin, 2006 Biological applications of the theory of birth-and-death processes. *Brief. Bioinform.* 7(1): 70–85.

Ohta, T., and K.-I. Kojima, 1968 Survival probabilities of new inversions in large populations. *Biometrics* 24(3): 501–516.

Orr, H. A., 2010 The population genetics of beneficial mutations. *Philos. Trans. R. Soc. B* 365: 1195–1201.

Orr, H. A., and R. L. Unckless, 2008 Population extinction and the genetics of adaptation. *Am. Nat.* 172(2): 160–169.

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

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

Parsons, T. L., and C. Quince, 2007b Fixation in haploid populations exhibiting density dependence II: the quasi-neutral case. *Theor. Popul. Biol.* 72: 468–479.

Patwa, Z., and L. M. Wahl, 2008 The fixation probability of beneficial mutations. *J. R. Soc. Interface* 5: 1279–1289.

Pollak, E., 1966 Some effects of fluctuating offspring distributions on the survival of genes. *Biometrika* 53(3/4): 391–396.

Pollak, E., 2000 Fixation probabilities when the population size undergoes cyclic fluctuations. *Theor. Popul. Biol.* 57: 51–58.

Takahasi, K. R., and F. Tajima, 2005 Evolution of coadaptation in a two-locus epistatic system. *Evolution* 59(11): 2324–2332.

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

Wahl, L. M., and P. J. Gerrish, 2001 The probability that beneficial mutations are lost in populations with periodic bottlenecks. *Evolution* 55: 2606–2610.

Wang, Y., and B. Rannala, 2004 A novel solution for the time-dependent probability of gene fixation or loss under natural selection. *Genetics* 168: 1081–1084.

Waxman, D., 2011 A unified treatment of the probability of fixation when population size and the strength of selection change over time. *Genetics* 188: 907–913.

Whitlock, M. C., 2003 Fixation probability and time in subdivided populations. *Genetics* 164: 767–779.

Whitlock, M. C., and R. Gomulkiewicz, 2005 Probability of fixation in a heterogeneous environment. *Genetics* 171: 1407–1417.

Communicating editor: L. M. Wahl

## Appendix

### Fixation in General Ecological Models

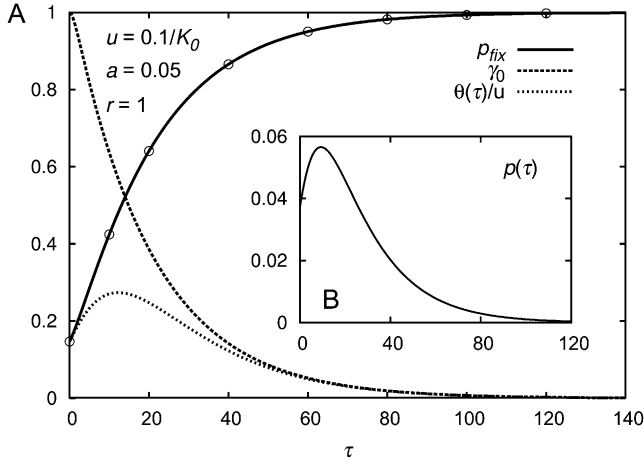
In this *Appendix*, we describe how the branching process approach can be applied to a general ecological scenario where the genetic composition of the population and the population dynamics are mutually dependent. Assume that at time  $t$  a population consists of  $w_t$  wild-type “residents” and  $n_t$  mutants. Both types reproduce and die at different per capita rates, which may depend on the number of residents and mutants,  $w_t$  and  $n_t$ , and on external factors that are independent of the population, but also may change over time. A general framework, which takes the full dynamics with all kinds of interactions between types into account, is given by the following scheme of transition rates:

$$\begin{aligned}
 n_t \rightarrow n_t + 1 : & \quad n_t \left[ b_1^{(m)}(t) + b_2^{(m)}(t)n_t + b_3^{(m)}(t)w_t + \dots \right], \\
 n_t \rightarrow n_t - 1 : & \quad n_t \left[ d_1^{(m)}(t) + d_2^{(m)}(t)n_t + d_3^{(m)}(t)w_t + \dots \right], \\
 w_t \rightarrow w_t + 1 : & \quad w_t \left[ b_1^{(w)}(t) + b_2^{(w)}(t)w_t + b_3^{(w)}(t)n_t + \dots \right], \\
 w_t \rightarrow w_t - 1 : & \quad w_t \left[ d_1^{(w)}(t) + d_2^{(w)}(t)w_t + d_3^{(w)}(t)n_t + \dots \right].
 \end{aligned}
 \tag{A1}$$

Higher-order interaction terms with transition rates proportional to any polynomial in  $n_t$  and  $w_t$  can be added as needed. The model corresponds to a general two-dimensional birth–death process with time-dependent coefficients. For such a model, the corresponding deterministic evolution of  $n_t$  and  $w_t$  is generically given by a system of two coupled

differential equations. In most cases, it is impossible to solve this system analytically. Since the deterministic frequency path of the mutant is needed for the theory developed to calculate  $T_{x_c}$ , derivations for  $T_{x_c}$  must resort to numerical solutions to apply the methods outlined in the main text. In contrast, an analytical approximation for the fixation probability (and thus also the distribution of the effective initial population size  $\nu$ , cf. Equation 40) is often still possible.

In the branching limit of rare mutant individuals, all interactions of either residents or mutants with (other) mutant individuals can be neglected. Mathematically, this corresponds to the approximation  $n_t \approx 0$  in the birth and death rates of the residents and negligence of all terms of order  $\mathcal{O}(n_t) > 1$  in the transition rates of the mutants. As a consequence, the evolution of the residents is decoupled from the mutants. In the deterministic limit, the time development of the resident population size  $w(t) \approx N(t)$  is then given by an ordinary differential equation that can frequently be solved explicitly. We can then describe the dynamics of rare mutants by a one-dimensional branching process. Inserting the deterministic solution for  $w(t)$  into all mutant transition rates of linear order in the mutant number  $n$ , the theory developed in this article can be applied by choosing  $\lambda(t)$  and  $\mu(t)$  appropriately. Note, however, that the first expression for  $p_{\text{fix}}$  (Equation 16a) must be used, since the step leading from Equation 16a to Equation 16b is not possible if the growth rates of mutants and wild types are not the same. Previous results on fixation probabilities in two-dimensional birth–death chains (with constant coefficients) have been obtained by Parsons and Quince (2007a). Their approach is based on singular perturbation



**Figure A1** (A) Fixation probability  $p_{fix}(\tau)$  of a single rescue mutation, the ratio  $\gamma_0(\tau) = N(\tau)/K_0$ , and the product of the two, which characterizes the rate of successful mutants, in dependence of the time  $\tau$  since the onset of the selection pressure. (B) Density function  $p(\tau)$  that determines the probability that the successful rescue mutation arises in the time interval  $[\tau, \tau + d\tau]$  conditioned on population survival. While the fixation probability per mutant increases with  $\tau$ , the latter peaks at an intermediate value for not too large mutation probabilities  $u$ . Simulations were performed for a population of 100,000 individuals, and every simulation point is the average over  $10^6$  runs.

methods and has the advantage that it can be applied to neutral and deleterious mutation, too (Parsons and Quince 2007a,b), which is beyond the reach of the branching model. However, it cannot easily be extended to the time-inhomogeneous case that is our focus here.

### Rescue mutation

Consider, as an example, the following scenario: A population is subject to a new and severe selection pressure (e.g., new insecticide, drug, parasite, competitor, ...). As a consequence, the population size decreases, and selection will drive the population to extinction unless a rescue mutation can be established that confers immunity to its carriers. A similar scenario has been discussed by Orr and Unckless (2008).

A plausible ecological model for this situation could be as follows: The original resident population evolves under logistic population dynamics with growth parameter  $r$  and time-dependent carrying capacity  $K(t)$ . Due to the environmental challenge, the carrying capacity decreases exponentially like  $K(t) = K_0 \exp(-at)$ . At the individual level, the model can be described in the scheme of Equation A1,

$$b_1^{(w)}(t) = r, \quad d_2^{(w)}(t) = d_3^{(w)}(t) = r \frac{1}{K(t)}, \quad (A2)$$

with all other coefficients  $b_i^{(w)}(t)$  and  $d_i^{(w)}(t)$  equal to zero. As long as the number of beneficial mutants is small, the wild-type population size changes according to

$$w(t) = \frac{K_0(a+r)}{a \exp(-rt) + r \exp(at)}. \quad (A3)$$

Assume now that a beneficial mutation—if it survives stochastic loss—restores population size to its original carrying capacity  $K_0 = K(0)$ ; i.e.,

$$b_1^{(m)}(t) = r, \quad d_2^{(m)}(t) = d_3^{(m)}(t) = r \frac{1}{K_0} \quad (A4)$$

(and again all other coefficients are equal to zero). The fixation probability of a single mutation arising  $\tau$  time units after the decline in the carrying capacity started is then given by

$$p_{fix}(\tau) = 2 \left[ 1 + \int_0^\infty r \left( 1 + \frac{w(\tau+t)}{K_0} \right) \exp \left[ - \int_0^t r \left( 1 - \frac{w(\tau+x)}{K_0} \right) dx \right] dt \right]^{-1}. \quad (A5)$$

In Figure A1 the fixation probability of a single mutation and the ratio  $\gamma_0(\tau) = w(\tau)/K_0 = N(\tau)/K_0$  are shown in dependence of  $\tau$ . Note that the fixation probability increases with  $\tau$ : While the mutant does not have a higher intrinsic growth rate than the resident (in a variant of the model, it could even be lower), it thrives due to relaxed competition as the resident population declines. However, for a given mutation probability per time unit  $u$  from resident to mutant, the new total rate of beneficial mutations is proportional to the resident population size and thus decreases with  $\tau$ . The rate at which new successful mutations arise is given by  $\theta_{fix}(\tau) = uw(\tau)p_{fix}(\tau)$ , which reaches a maximum at some intermediate time  $\tau^*$  (see Figure A1). The probability that a successful mutation arises before time  $T$  is

$$P(t \leq T) = 1 - \exp \left[ - \int_0^T \theta_{fix}(\tau) d\tau \right]. \quad (A6)$$

The probability that a successful mutation arises at all, i.e., that the population is rescued from extinction, is therefore given by

$$P_{rescue} = 1 - \exp \left[ - \int_0^\infty \theta_{fix}(\tau) d\tau \right]. \quad (A7)$$

This probability is equal to 1 if and only if the integral  $\int_0^\infty \theta_{fix}(\tau) d\tau$  diverges, i.e., if the rate  $\theta_{fix}(\tau)$  declines sufficiently slowly. Otherwise the probability that the population is saved from extinction by a rescue mutation is  $<1$ , as was also found by Orr and Unckless (2008). For the parameter values used in Figure A1 the survival probability of the population is calculated to be  $P_{rescue} \approx 0.39$ . Given the population is rescued, the probability that the rescuing mutation arises in the infinitesimal time interval  $[\tau, \tau + d\tau]$  is given by  $p(\tau)d\tau$  with

$$p(\tau) = \frac{\theta_{fix}(\tau) \exp \left[ - \int_0^\tau \theta_{fix}(\hat{\tau}) d\hat{\tau} \right]}{P_{rescue}}. \quad (A8)$$

For not too large mutation probabilities ( $u \lesssim 0.5/K_0$  for the chosen parameter values), this function has a maximum at an intermediate value of  $\tau = \tau^{**}$ , as also shown in Figure A1. This result is in contrast to Orr and Unckless (2008), who find a monotonic decrease in the (conditioned) probability that a successful mutation arises in a given generation independently of the mutation rate. This is due to the assumption in Orr and Unckless (2008) that the selection co-

efficient of mutants (and thus the fixation probability  $p_{\text{fix}}$ ) of a single mutant is constant over time. In our model, selection pressure on mutants is not simply included by assumption, but results from the underlying ecological dynamics. As such, it explicitly depends on time. The scenario by Orr and Unckless (2008) is approximately met only for large  $\tau > \tau^{**}$  when competition from the resident population gets very small.

# GENETICS

**Supporting Information**

<http://www.genetics.org/content/suppl/2011/06/06/genetics.110.124297.DC1>

## **On the Fixation Process of a Beneficial Mutation in a Variable Environment**

**Hildegard Uecker and Joachim Hermisson**

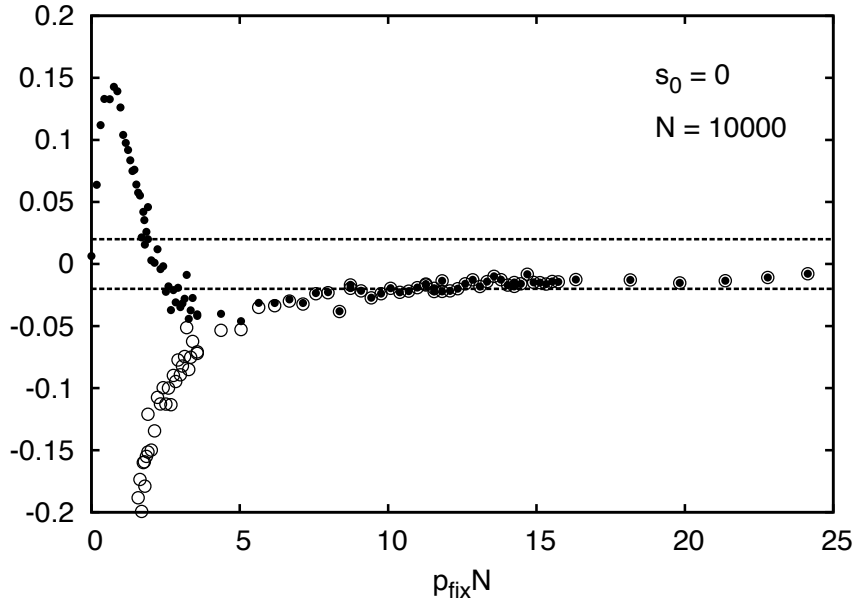


Figure S1: Relative deviation between analytical and simulation results for an allele with selective advantage  $s(t) = s_1 t$ . The empty dots are obtained from branching process approximation, the small filled dots from the corrected version Eq. (S1). One sees that the corrected version provides a much better approximation for small values of  $p_{fix}N$ ; for large values of  $p_{fix}N$ , the results coincide. Simulations were performed for a population of  $N = 10\,000$  individuals, and each simulation point is the average over  $5 \cdot 10^7$  runs.

## S1 Accuracy of the approximation

**Weak selection.** As pointed out in the main text, deviations from the exact solution are expected for weak selection.

For an allele with constant selective advantage and  $N = const.$ , a comparison to the exact solution is immediately possible (see main text).

As a further example, we consider the fixation probability of an allele with  $s(t) = s_1 t$  and examine the relative error between theoretical prediction and simulation results in dependence of  $p_{fix}N$  (see Figure S1). A relative error of less than 2% is found for  $p_{fix}N < 10$ . The results for small values of  $p_{fix}N$  can be improved if we correct the branching approximation by a factor  $1/(1 - \exp(-p_{fix}N))$ :

$$p_{fix}^* = \frac{p_{fix}}{1 - \exp(-p_{fix}N)}. \quad (S1)$$

This heuristic approximation is inspired by Eq. (18). The relative deviation from the simulation results is added to Figure S1. It is seen that the approximation is considerably improved.

In Figure S2, we compare the distribution function of  $T_{1/2}$  to simulation results for various values of  $s$  and  $N$  where  $s$  and  $N$  are constant. While agreement is excellent for high values of  $\alpha$ , deviations increase for decreasing values of  $\alpha$ . A maximum absolute deviation of  $\geq 0.05$  between theory and simulations is found for  $\alpha = Ns \lesssim 10$ .

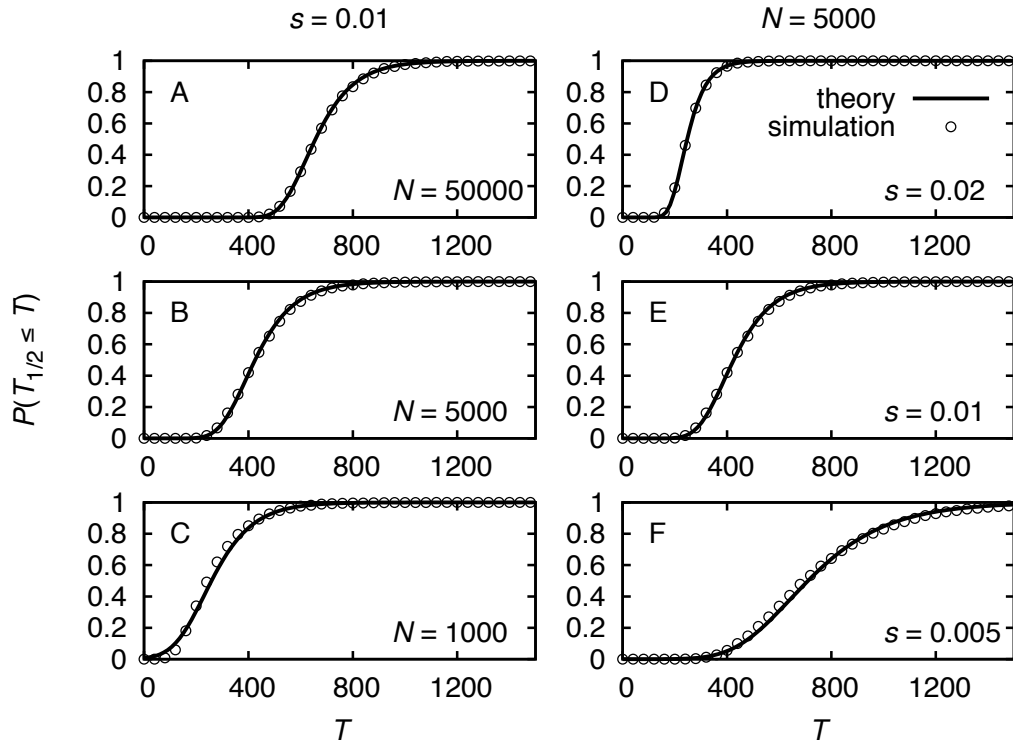


Figure S2: Distribution of the time to reach frequency  $x_c = 0.5$  for various values of  $s$  and  $N$ , where  $s$  and  $N$  are constant. Analytical and simulation results are compared. Simulation results are averaged over 5000 runs. A maximum absolute deviation of  $\geq 0.05$  between theory and simulations is found for  $\alpha = Ns \lesssim 10$  (see panel C).

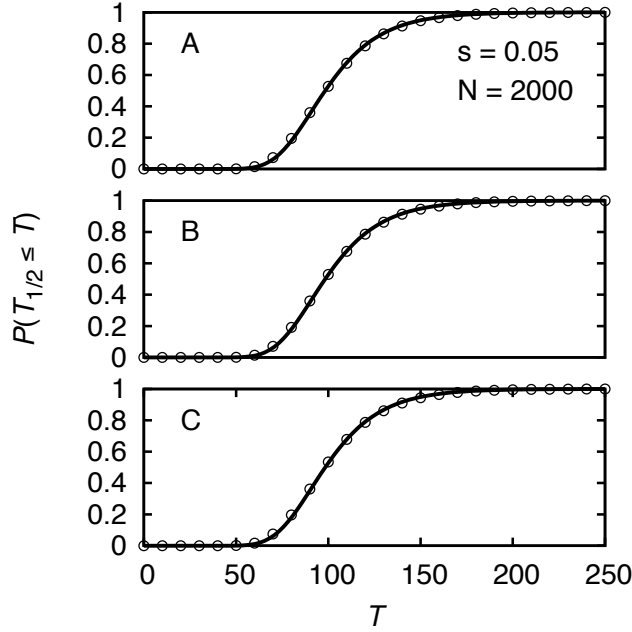


Figure S3: Distribution function of  $T_{1/2}$  for three different modes of density regulation. The selection strength is  $s = 0.05$  and the initial population size  $N_0 = 2000$ . A: Inherently constant population size, i.e.  $b(t, N_t) = d(t, N_t) = 0$ , with  $\xi(t, N) = 1$ . B: Lotka-Volterra dynamics, i.e.  $b(t, N_t) = b + \rho(1 - N_t/K)$  and  $d(t, N_t) = b$ , with  $b = 1.0$ ,  $r = 1.0$  and  $K = N_0 = 2000$ . C: Freely fluctuating population size, i.e.  $b(t, N_t) = d(t, N_t) = b$ , with  $b = 1.0$ . Simulations are averaged over 50 000 runs. The maximum absolute deviation is of the same order of magnitude for all three scenarios.

**Mode of density-regulation.** To test whether the mode of density-regulation influences the outcome, some exemplary simulations for three different scenarios of density regulation were performed ( $b(t, N_t) = d(t, N_t) = \text{const.}$ , Lotka-Volterra dynamics, i.e.  $b(t, N_t) = b + \rho(1 - N_t/K)$  and  $d(t, N_t) = b$ , and  $b(t, N_t) = d(t, N_t) = 0$  with accordingly chosen  $\xi$ ). Figure S3 shows a comparison of the simulation results to the theoretical curve. We see that demographic stochasticity has only a slight effect (note that we used a relatively low (initial) population size of  $N_0 = 2000$  for the examples to make demographic stochasticity strong).

## S2 Diffusion approximation

Pioneered by Kimura (1957), the diffusion approximation has been established as the second main approach for the analytical analysis of the fixation process. In this appendix, we show how an approximation for the fixation probability with variable selection and population size can be derived within the diffusion framework. We also use diffusion results for the mean fixation time for constant  $s$  and  $N$  to estimate the error of the branching process approach in this case.

As in the main text, we assume that the population inhabits a variable environment. All key model parameters may explicitly depend on time. In particular, let  $s = s(t)$  be the selection coefficient of the  $A$  allele and  $N = N(t)$  the

total population size. Finally, the variance in offspring number is  $\sigma^2 = \sigma^2(t)$ . Following standard practice, we can combine  $N$  and  $\sigma^2$  in the variance effective population size,  $N_e(t) = N(t)/\sigma^2(t)$ .

We can formulate the model as a diffusion process as follows. Let  $g_t(x \rightarrow x + \epsilon)$  be the transition probability that the allele frequency changes from  $x$  at time  $t$  to  $x + \epsilon$  at time  $t + dt$ . The diffusion is characterized by the mean and the variance of the transition probability in the limit of small  $dt$  and  $\epsilon$ :

$$E_{x,t}[\epsilon]dt = \int_{-\infty}^{\infty} \epsilon g_t(x \rightarrow x + \epsilon) d\epsilon, \quad (S2a)$$

$$\text{Var}_{x,t}[\epsilon]dt = \int_{-\infty}^{\infty} \epsilon^2 g_t(x \rightarrow x + \epsilon) d\epsilon + \mathcal{O}[dt^2]. \quad (S2b)$$

To leading order in  $N_e^{-1}$ ,

$$E_{x,t}[\epsilon] = s(t)x(1-x), \quad (S3a)$$

$$\text{Var}_{x,t}[\epsilon] = \frac{\sigma^2(t)}{N(t)}x(1-x) = \frac{x(1-x)}{N_e(t)}. \quad (S3b)$$

Define  $f(x, \tau|p, t)$  as the probability density for an allele frequency of  $x$  at time  $t + \tau$ , given that the frequency at time  $t$  was  $p$ . Then

$$\begin{aligned} f(x, \tau + dt|p, t) &= \int_{-\infty}^{\infty} g_t(p \rightarrow p + \epsilon) f(x, \tau|p + \epsilon, t + dt) d\epsilon \\ &\approx f(x, \tau|p, t + dt) \\ &\quad + \left( \frac{\partial f(x, \tau|p, t + dt)}{\partial p} E_{p,t}[\epsilon] + \frac{1}{2} \frac{\partial^2 f(x, \tau|p, t + dt)}{\partial p^2} \text{Var}_{p,t}[\epsilon] \right) dt. \end{aligned} \quad (S4)$$

In the limit  $dt \rightarrow 0$ , we obtain the Kolmogorov backward equation for the conditional density  $f(x, \tau|p, t)$  of a time-inhomogeneous diffusion:

$$\frac{\partial f}{\partial \tau} - \frac{\partial f}{\partial t} = s(t)p(1-p) \frac{\partial f}{\partial p} + \frac{p(1-p)}{2N_e(t)} \frac{\partial^2 f}{\partial p^2}. \quad (S5)$$

If  $s(t)$  and  $N_e(t)$  are both constant,  $\partial f/\partial t = 0$  and (S5) reduces to the backward equation of the classical homogeneous case. With time dependence, the state of the population can no longer be described by the allele frequency  $x$  alone;  $f$  depends on both,  $x$  and  $t$ . Note that we measure time on a generations scale, not in  $2N$  generations, as it is often done in diffusion theory.

Denote now as  $P(p, t)$  the fixation probability of an allele with frequency  $p$  in the population at time  $t$ . Since fixation is the probability of eventual absorption of the diffusion at the boundary  $x = 1$ , we can write  $P(p, t)$  in terms of the transition probability as

$$P(p, t) = \lim_{\substack{\tau \rightarrow \infty \\ y \rightarrow 1}} \int_y^1 f(x, \tau|p, t) dx. \quad (S6)$$

Integrating (S5) over any frequency interval  $[x_1, x_2]$ , we see that the Kolmogorov backward equation also holds for the corresponding probabilities. For the probability of eventual fixation, in particular, the  $\tau$  dependence vanishes and we obtain (Kimura and Ohta, 1974)

$$-\dot{P}(p, t) = s(t)p(1-p)P'(p, t) + \frac{p(1-p)}{2N_e(t)}P''(p, t), \quad (S7)$$

where  $\dot{P}$  and  $P'$  denote derivatives with respect to  $t$  and  $p$ , respectively. In contrast to the time-homogeneous case, there is no exact solution to the time-inhomogeneous equation. However, it is possible to derive an approximate solution for small  $p$  by setting  $1 - p \approx 1$  in (S7). This approximation was first used by Kimura and Ohta (1974) to derive the fixation probability for a logistically growing population. As we will see, a full solution of the general time-dependent model is possible under the same assumption. We use the following Ansatz,

$$P(p, t) = 1 - \exp(-Q(t)p). \quad (\text{S8})$$

Substituting this relation into the approximated PDE (ignoring  $1 - p$  terms in Eq. (S7)) leads to the following ODE for  $Q(t)$ ,

$$\frac{dQ(t)}{dt} = -s(t)Q(t) + \frac{Q^2(t)}{2N_e(t)}. \quad (\text{S9})$$

A general solution to this differential equation can be found,

$$Q(t) = \frac{2}{\int_t^\infty \frac{1}{N_e(\tau)} \exp\left(-\int_t^\tau s(y)dy\right) d\tau}, \quad (\text{S10})$$

and we obtain an approximation for the fixation probability from Eq. (S8). If the fixation process starts from a single copy of the derived allele that enters the population at time  $t = 0$ , in particular, we find

$$p_{fix,diff} := P(1/N(0), 0) = 1 - \exp(-Q(0)/N(0)) \approx Q(0)/N(0). \quad (\text{S11})$$

This expression may be directly compared to the result for the fixation probability  $p_{fix}$  (16b) from the branching process derivation. We find that

$$p_{fix} = \frac{p_{fix,diff}}{1 + p_{fix,diff}/2}. \quad (\text{S12})$$

Both expressions are thus equal to leading order in  $p_{fix,diff}$  (i.e. to leading order in the selection strength).

The approximate solution from the diffusion implies the absence of competition among mutant alleles (which appears in (S7) through the  $p^2$  terms). This is equivalent to the independence assumption of the branching process. Note also that the approximation is constructed in a way that it fulfills the boundary condition  $P(0, t) = 0$ , but not the condition  $P(1, t) = 1$ . It is therefore only valid if the initial frequency of the allele,  $p$ , is sufficiently small.

For constant selection, the diffusion equation (with correct boundary conditions at  $p = 0$  and  $p = 1$ ) can be solved exactly. In particular, the derivation of the mean fixation time is possible (Kimura and Ohta, 1969; Ewens, 2004). This allows for a comparison of our approximate solution with the exact diffusion result in this case. From Eq. (47), we obtain the following expression for the mean fixation time in the diffusion limit ( $s \rightarrow 0$ ,  $\alpha = sN/\xi = \text{const.}$ ):

$$\langle \tau_{fix} \rangle^{(1)} = \frac{2}{\alpha} \left( \int_0^{\alpha/2} \ln\left(\frac{\alpha - Y}{Y}\right) \exp(-Y) dY \right). \quad (\text{S13})$$

Within the diffusion framework, the mean fixation time is given by (see Ewens, 2004, p.140ff and p.167ff),

$$\begin{aligned} \langle \tau_{fix} \rangle^{(2)} &= \frac{2}{\alpha[\exp(\alpha) - 1]} \int_0^1 \frac{[\exp(\alpha x) - 1][\exp(\alpha(1-x)) - 1]}{x} dx \\ &= \frac{2}{\alpha[\exp(\alpha) - 1]} \int_0^\alpha \frac{\exp(\alpha) - \exp(Y) - \exp(\alpha - Y) + 1}{Y} dY \\ &\approx \frac{2}{\alpha} \left( \int_0^\alpha \frac{1 - \exp(-Y)}{Y} dY + \exp(-\alpha) \int_0^\alpha \frac{1 - \exp(Y)}{Y} dY \right). \end{aligned} \quad (\text{S14})$$

In the last step, we approximated  $(\exp(\alpha) - 1)^{-1} \approx \exp(-\alpha)$ . This approximation is of order  $o(\exp(\alpha))$ . By integration by parts,

$$\int_0^\alpha \ln\left(\frac{\alpha - Y}{Y}\right) \exp(-Y) dY = \int_0^\alpha \frac{1 - \exp(-Y)}{Y} dY + \exp(-\alpha) \int_0^\alpha \frac{1 - \exp(Y)}{Y} dY. \quad (\text{S15})$$

We thus obtain for the error term  $\Delta(\alpha)$  of the branching result:

$$\Delta(\alpha) = \langle \tau_{fix} \rangle^{(2)} - \langle \tau_{fix} \rangle^{(1)} = \frac{2}{\alpha} \int_{\alpha/2}^\alpha \ln\left(\frac{\alpha - Y}{Y}\right) \exp(-Y) dY + o(\exp(-\alpha)). \quad (\text{S16})$$

Since

$$\begin{aligned} \exp\left(\frac{\alpha}{2}\right) \Delta(\alpha) &= \frac{2}{\alpha} \int_{\alpha/2}^\alpha \ln\left(\frac{\alpha - Y}{Y}\right) \exp\left[-\left(Y - \frac{\alpha}{2}\right)\right] dY \\ &= \int_0^1 \ln\left(\frac{1 - Y}{1 + Y}\right) \exp\left(-\frac{\alpha}{2} Y\right) dY \xrightarrow{\alpha \rightarrow \infty} 0, \end{aligned} \quad (\text{S17})$$

we conclude that  $\Delta(\alpha)$  is of order  $o(\exp(-\alpha/2))$ .

### S3 Additional explanations

**Derivation of Eq. (4a) and (4b).** To calculate the expected frequency change  $E[\Delta x | x_t, N_t]$  in an infinitesimal time interval  $dt$ , we need to consider all events that change the frequency  $x_t = n/N_t$  of mutants in the population. These events are summarized in Table 1. The second and third column give the probability of the event and the change  $\Delta x$  induced by it. Taking all events, their respective probabilities and effects in account, we obtain:

$$\begin{aligned} E[\Delta x | x_t, N_t] &= \left[ (\xi(t, N_t) + s(t, N_t)) \frac{n(N_t - n)}{N_t} \left( \frac{n+1}{N_t} - \frac{n}{N_t} \right) + \xi(t, N_t) \frac{n(N_t - n)}{N_t} \left( \frac{n-1}{N_t} - \frac{n}{N_t} \right) \right. \\ &\quad + b(t, N_t) n \left( \frac{n+1}{N_t+1} - \frac{n}{N_t} \right) + b(t, N_t) (N_t - n) \left( \frac{n}{N_t+1} - \frac{n}{N_t} \right) + d(t, N_t) n \left( \frac{n-1}{N_t-1} - \frac{n}{N_t} \right) \\ &\quad \left. + d(t, N_t) (N_t - n) \left( \frac{n}{N_t-1} - \frac{n}{N_t} \right) \right] dt \\ &= s(t, N_t) x_t (1 - x_t) dt \end{aligned} \quad (\text{S18})$$

Since  $E[\Delta x | x_t, N_t]^2$  is of order  $\mathcal{O}((dt)^2)$ , it holds for the variance of  $\Delta x$ :

$$\text{Var}[\Delta x | x_t, N_t] = E[(\Delta x)^2 | x_t, N_t] - E[\Delta x | x_t, N_t]^2 \approx E[(\Delta x)^2 | x_t, N_t]. \quad (\text{S19})$$

event	probability	$\Delta x$
a mutant replaces a resident	$(\xi(t, N_t) + s(t, N_t)) \frac{n(N_t - n)}{N_t} dt$	$\left( \frac{n+1}{N_t} - \frac{n}{N_t} \right)$
a resident replaces a mutant	$\xi(t, N_t) \frac{n(N_t - n)}{N_t} dt$	$\left( \frac{n-1}{N_t} - \frac{n}{N_t} \right)$
a mutant is born	$b(t, N_t) n dt$	$\left( \frac{n+1}{N_t+1} - \frac{n}{N_t} \right)$
a resident is born	$b(t, N_t) (N_t - n) dt$	$\left( \frac{n}{N_t+1} - \frac{n}{N_t} \right)$
a mutant dies	$d(t, N_t) n dt$	$\left( \frac{n-1}{N_t-1} - \frac{n}{N_t} \right)$
a resident dies	$d(t, N_t) (N_t - n) dt$	$\left( \frac{n}{N_t-1} - \frac{n}{N_t} \right)$

Table 1: Summary of the events that change the frequency  $x$  of mutants in the population.

We use again Table 1 for the calculation:

$$\begin{aligned}
E[(\Delta x)^2 | x_t, N_t] &= \left[ (\xi(t, N_t) + s(t, N_t)) \frac{n(N_t - n)}{N_t} \left( \frac{n+1}{N_t} - \frac{n}{N_t} \right)^2 + \xi(t, N_t) \frac{n(N_t - n)}{N_t} \left( \frac{n-1}{N_t} - \frac{n}{N_t} \right)^2 \right. \\
&\quad + b(t, N_t) n \left( \frac{n+1}{N_t+1} - \frac{n}{N_t} \right)^2 + b(t, N_t) (N_t - n) \left( \frac{n}{N_t+1} - \frac{n}{N_t} \right)^2 \\
&\quad \left. + d(t, N_t) n \left( \frac{n-1}{N_t-1} - \frac{n}{N_t} \right)^2 + d(t, N_t) (N_t - n) \left( \frac{n}{N_t-1} - \frac{n}{N_t} \right)^2 \right] dt \\
&= \left[ \frac{1}{N_t} (2\xi(t, N_t) + s(t, N_t)) x_t (1 - x_t) + \frac{1}{N_t} \frac{n(N_t - n)}{(N_t + 1)^2} b(t, N_t) + \frac{1}{N_t} \frac{n(N_t - n)}{(N_t - 1)^2} d(t, N_t) \right] dt \\
&\approx \left[ \frac{1}{N_t} (2\xi(t, N_t) + s(t, N_t)) x_t (1 - x_t) + \frac{1}{N_t} x_t (1 - x_t) b(t, N_t) + \frac{1}{N_t} x_t (1 - x_t) d(t, N_t) \right] dt \\
&= \frac{2\xi(t, N_t) + b(t, N_t) + d(t, N_t) + s(t, N_t)}{N_t} x_t (1 - x_t) dt = \frac{x_t (1 - x_t)}{N_{e,t}} dt,
\end{aligned} \tag{S20}$$

where we approximated  $N_t + 1 \approx N_t$  and  $N_t - 1 \approx N_t$  in the course of the calculation.

**Derivation of Eq. (34).** In order to proof Eq. (34), we proof the following relation from which Eq. (34) follows immediately by choosing  $z = 0$ :

$$\frac{d^n P(z, t)}{dz^n} = n! \left[ \frac{B^{n-1}(t)}{(A(t) - B(t)(z - 1))^n} + \frac{B^n(t)(z - 1)}{(A(t) - B(t)(z - 1))^{n+1}} \right], \quad n \geq 1 \tag{S21}$$

with the probability generating function (cf. Eq. (10))

$$P(z, t) = 1 + \frac{z - 1}{A(t) - B(t)(z - 1)}. \tag{S22}$$

We carry out a proof by induction. Eq. (S21) builds our induction hypothesis.

- Base case  $n = 1$ :

$$\frac{dP(z, t)}{dz} = \frac{1}{A(t) - B(t)(z - 1)} + \frac{B(t)(z - 1)}{(A(t) - B(t)(z - 1))^2}.$$

- Inductive step  $n \rightarrow n + 1$ :

$$\begin{aligned}
\frac{d^{n+1}P(z, t)}{dz^{n+1}} &= \frac{d}{dz} \frac{d^n P(z, t)}{dz^n} \\
&= \frac{d}{dz} \left[ \frac{B^{n-1}(t)}{(A(t) - B(t)(z-1))^n} + \frac{B^n(t)(z-1)}{(A(t) - B(t)(z-1))^{n+1}} \right] \\
&= n! \left[ n \frac{B^{n+1}(t)B(t)}{(A(t) - B(t)(z-1))^{n+1}} + \frac{B^n(t)}{(A(t) - B(t)(z-1))^{n+1}} + (n+1) \frac{B^n(t)B(t)(z-1)}{(A(t) - B(t)(z-1))^{n+2}} \right] \\
&= (n+1)! \left[ \frac{B^n(t)}{(A(t) - B(t)(z-1))^{n+1}} + \frac{B^{n+1}(t)(z-1)}{(A(t) - B(t)(z-1))^{n+2}} \right].
\end{aligned}$$

It follows:

$$p_n(t) = \frac{1}{n!} \left. \frac{d^n P(z, t)}{dz^n} \right|_{z=0} = \frac{B^{n-1}(t)}{(A+B)^n(t)} - \frac{B^n(t)}{(A+B)^{n+1}(t)} = \frac{B^{n-1}(t)A(t)}{(A+B)^{n+1}(t)}, \quad n \geq 1. \quad (\text{S23})$$

□

**Symmetry of the frequency path in a constant environment.** Let  $\Omega_1$  be the set of paths which lead from  $n_0 = 1$  to  $n = n_c - 1$ , and  $\Omega_2$  be the set of paths which lead from  $n = N - n_c$  to  $n = N - 1$ . Then there exists a bijection between these two sets  $f : \Omega_1 \rightarrow \Omega_2$  which can basically be constructed by mirroring the path at  $x_c = 0.5$ . Concretely: Be  $\omega_1 = (n_1, \dots, n_l) \in \Omega_1$  where  $n_1 = 1$  and  $n_l = n_c - 1$ . Define then  $f(\omega_1) = \omega_2 = (m_1, \dots, m_l)$  by  $m_i = N - n_{l+1-i}$ . Consider now the paths  $\omega'_1 = (n_1, \dots, n_l, n_c)$  and  $\omega'_2 = (m_1, \dots, m_l, N)$ . It now holds: (1) The probability that path  $\omega'_1$  is realized equals the probability that  $\omega'_2$  is realized as both make the same number of steps in both directions. (2) The distribution of the run-time along both paths  $\omega'_1$  and  $\omega'_2$  is identical because the transition rates of the Moran model defined in Eq. (3) are symmetric. From (1) and (2) it follows that the distribution of the time to reach fixation starting from  $(N - n_c + 1)/N \approx (1 - x_c)$  is the same as the distribution of the time to reach frequency  $x_c = n_c/N$  starting from  $1/N \approx 0$ . (This is true whether one considers the first, the last or the mean passage time; the latter is approximated by our theory.)

## References

- Ewens, W. J., 2004 *Mathematical Population Genetics*. Springer, New York et al., second edition.
- Kimura, M., 1957 Some problems of stochastic processes in genetics. *The Annals of Mathematical Statistics* **28**: 882–901.
- 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.
- Kimura, M., and T. Ohta, 1974 Probability of gene fixation in an expanding finite population. *Proceedings of the National Academy of Sciences USA* **71**: 3377 – 3379.

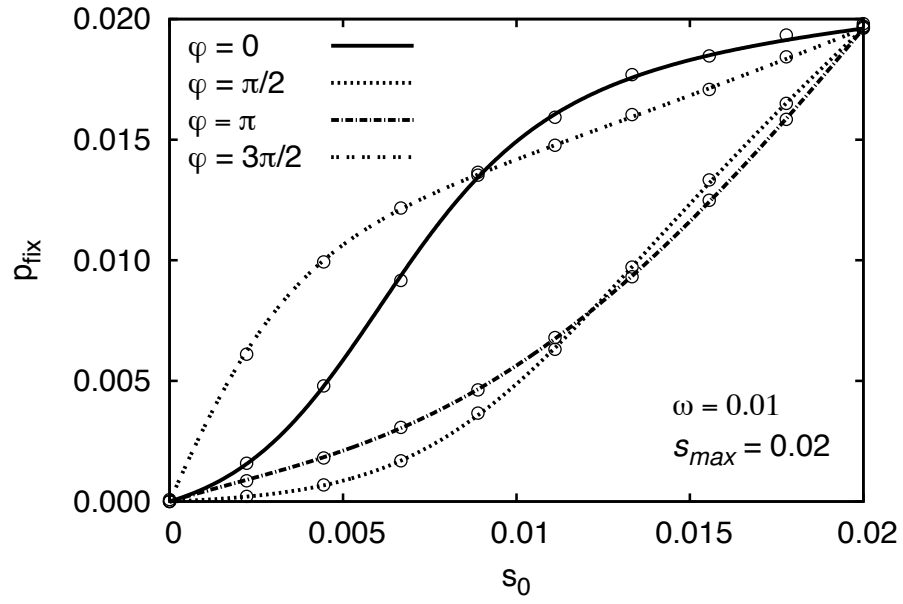


Figure S4: Fixation probability for a mutation with periodically changing selection strength  $s(t) = s_0 + (s_{max} - s_0) \cos(\omega t + \varphi)$  in dependence of  $s_0$  for various values of  $\varphi$ ;  $s_{max} = 0.02$ ,  $\omega = 0.01$ . Small values of  $s_0$  correspond to relatively long and pronounced periods of selective disadvantage. As  $s_0$  approaches  $s_{max}$  the fixation probability tends to  $\approx 2s_{max}$  as expected. Comparison to simulated data shows that the theory provides an accurate prediction of the fixation probability also for scenarios in which the mutation temporarily gets disadvantageous. For  $s_0 \leq 0$ , however, it predicts a fixation probability of zero and therefore underestimates the true value. Simulations were performed for a population of 100 000 individuals, and each simulation point is the average over  $10^6$  runs.

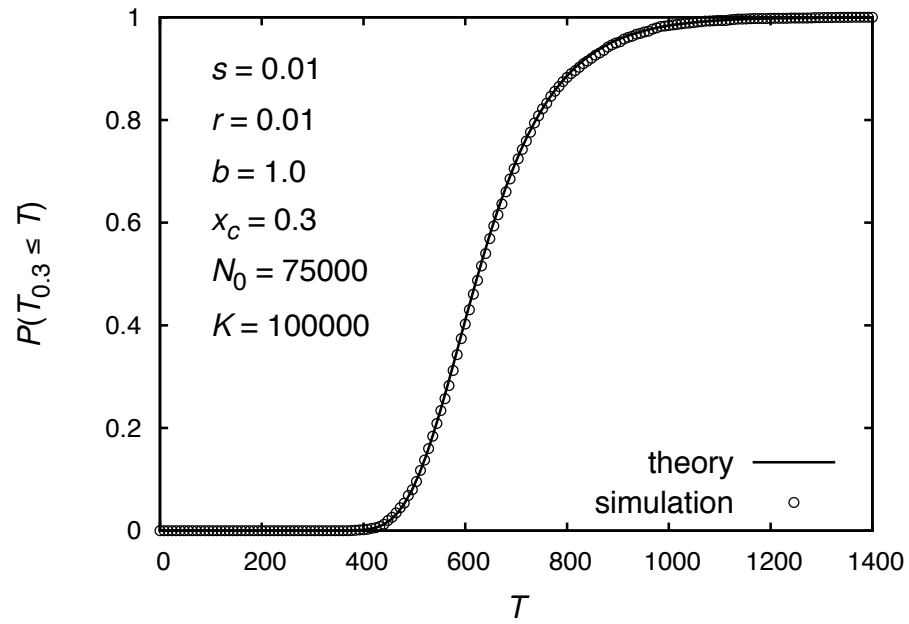


Figure S5: Distribution of the time to reach frequency  $x_c = 0.3$  in a population which grows logistically according to the second scenario with  $\xi = 0$  (see text). Analytical and simulation results are compared. The simulation curve is the average over 5000 runs.

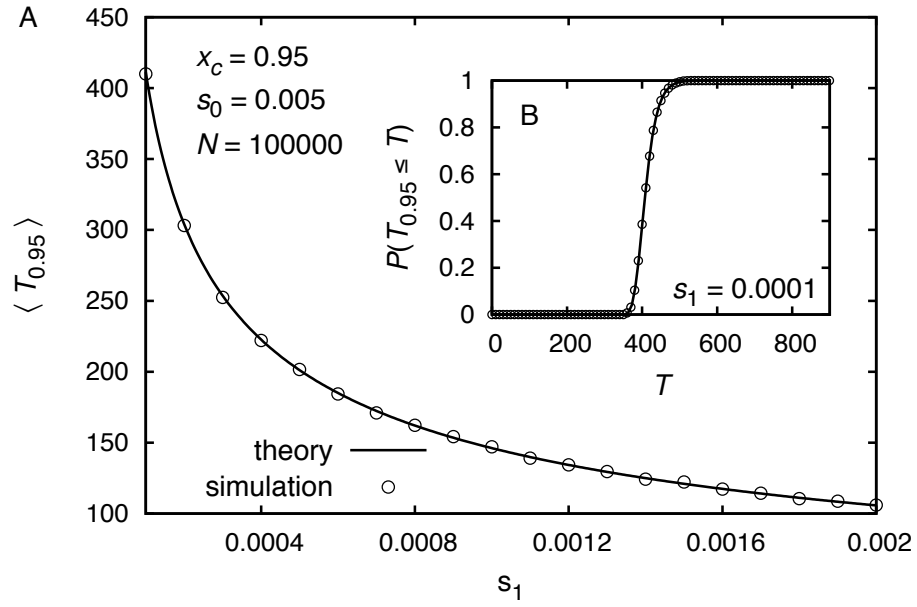


Figure S6: A: Mean time to reach frequency  $x_c = 0.95$  for linearly increasing selection  $s(t) = s_0 + s_1 t$  in dependence of  $s_1$ . Every simulation point is the average over 100 runs. B: Distribution function for a fixed value of  $s_1$ . For the simulation curve 5000 runs were performed.