Fixation probability and fixation time under strong recurrent mutation

Martin Pontz[1], Remus Stana[1], and Yoav Ram[1*]

[1]School of Zoology, Faculty of Life Sciences, Tel Aviv University, Tel Aviv, Israel
[*]Corresponding author: yoav@yoavram.com
[1]ORCID: YR 0000-0002-9653-4458, MP 0000-0002-9699-2061

July 11, 2023

Abstract

When the mutation rate is high and/or the population size is large, recurrent mutation can lead to multiple, independently generated copies of the same beneficial allele spreading through the population. However, classical analyses of fixation probability and time assume that the mutation rate is low and therefore, that fixation and extinction of a beneficial allele occur faster than the appearance of additional copies. We developed a diffusion equation approximation for the fixation probability and time that accounts for recurrent mutation, incomplete fixation, and fixation from standing genetic variation. Our results show that when the number of new beneficial alleles per generation in the population is greater than one, fixation is guaranteed, and fixation time is significantly lower than expected by the standard approximation. Moreover, we show that fixation time is significantly shorter if the initial allele frequency is greater than 0, or if fixation is defined for an allele frequency lower than 1.

Keywords. mutation, fixation probability, fixation time, diffusion equation approximation, wright-fisher model, evolutionary theory

Competing Interests. The authors declare no competing interests.

Acknowledgements. We thank Dmitri Petrov and Daniel Weissman for discussions and comments. This work was supported in part by the Israel Science Foundation (ISF, YR 552/19), the US–Israel Binational Science Foundation (BSF, YR 2021276), Minerva Center for Live Emulation of Evolution in the Lab (YR), Minerva Stiftung short-term research grant (MP).

Data availability. Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

Introduction

Analyses of the fixation probability and time of beneficial mutations have mostly focused on scenarios in which the appearance of the beneficial allele is rare enough so that it becomes extinct or fixed before the next beneficial allele appears (Kimura, 1962, Kimura and Ohta, 1969). However, adaption sometimes occur in large populations and/or with high mutation rates, leading to high mutation supply (mutation rate $\times$ population size), such that multiple new mutants are produced in every generation, rather than many generations passing between the occurrence of each mutant. Some examples are the rapid appearance of many, selectively similar, beneficial alleles (Karasov et al., 2010, Barroso-Batista et al., 2014, Levy et al., 2015, Nguyen Ba et al., 2019); the quick appearance and fixation of large-effect, rapidly forming, copy number variants (Egan et al., 2007, Yona et al., 2012, Payen et al., 2014, Avecilla et al., 2022); and evidence for mutation rate variation across the genome (Egan et al., 2007, Harpak et al., 2016). Indeed, it has been suggested that "recurrent, parallel mutation modes can profoundly shape the paths taken by evolution and undermine common models of evolutionary genetics" (Press et al., 2019).

Here, we derive an approximation for the fixation probability and the fixation time under the effect of mutation, selection, and genetic drift. Comparing our approximations to simulation results, we find they perform very well. We find that when the mutation supply is high enough (roughly one mutant individual per generation), fixation is certain and rapid. Importantly, it is significantly faster than implied by approximations that neglect mutation, as such approximations perform poorly, underestimating the fixation probability and overestimating the fixation time. We also compare our approximations to those previously derived by Hermisson and Pfaffelhuber (2008) in their analysis of genetic hitchhiking. We find that our approximation is more flexible: it allows one to assume any initial frequency for the beneficial allele—which is likely to appear in more than one individual if the mutation supply is large—and it also allows to approximate the time for the beneficial allele to reach any upper frequency, rather than 1, which is useful for applying the approximation to results of evolutionary experiments or simulations, in which researchers may not want to wait until the beneficial allele is completely fixed in the population.

In the following we present a simple Wright-Fisher model with a single, bi-allelic locus under the effects of selection, recurrent mutation, and drift; we derive diffusion equation approximations for the fixation probability and time in this model; compare these approximations to simulation results and previous approximations; and discuss our results.

Model and Results

Wright-Fisher model

Here, we consider a Wright-Fisher model of a haploid population with non-overlapping generations and a constant population size N. We focus on a single bi-allelic locus with alleles A and a. The frequencies of alleles A and a are x and $1-x$, respectively. The fitness of allele A relative to a is $1+s$, such that $s > 0$ is the selection coefficient of A, and the effect of selection is given by

$$x^s = \frac{x(1+s)}{x(1+s)+(1-x)}. \tag{1}$$

We assume that mutations from a to A occur with rate μ and neglect back-mutations from A to a (the case with back mutations was analyzed by Ewens (2004)). Thus, the effect of mutation is given by

$$x^m = x^s + (1-x^s)\mu. \tag{2}$$

Finally, the effect of random genetic drift is given by

$$x' = n/N, \quad n \sim Bin(N, x^m), \tag{3}$$

such that x' is the frequency of allele A in the next generation.

Diffusion equation approximation

Let $a(x) = \mathbf{E}[x' - x]$ and $b(x) = \mathrm{Var}[x' - x]$ be the expectation and variance of the change in frequency of allele A. We define the scaled selection parameter as $\alpha = Ns$ and the scaled mutation parameter as $\theta = 2N\mu$, which is also called the *mutation supply*. We therefore have (see Ewens, 2004, eq. 5.6)

$$a(x) = \alpha x(1-x) + \theta(1-x)/2, \tag{4a}$$

$$b(x) = x(1-x). \tag{4b}$$

Next, we follow Ewens (2004, ch. 4 and 5) to determine the fixation probability and the fixation time of allele A. In the following results, $\Gamma(x)$ and $\Gamma(x, y, z)$ are the *gamma function* and the *generalized incomplete gamma function*, given by the integral equations $\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}\,\mathrm{d}t$ and $\Gamma(x, y, z) = \int_y^z t^{x-1}e^{-t}\,\mathrm{d}t$. When $x = 1$, the generalized incomplete gamma function can be simplified, $\Gamma(1, y, z) = e^{-y} - e^{-z}$.

Analysis of fixation probability

The probability that a beneficial mutant allele fixates, i.e., takes over the population, is an important quantity in population genetics and evolutionary theory (reviewed in Patwa and Wahl, 2008, McCandlish and Stoltzfus, 2014). Early approximations of the fixation probability used branching processes (Haldane, 1927, Eshel, 1981). Modern and more accurate approximations usually use a diffusion process to calculate the probability that the allele frequency x reaches 1 before it reaches 0. However, to the best of our knowledge, the literature does not have an approximation for the fixation probability under directional mutation (i.e., mutation from a to A) with a high mutation rate. The following result provides such an approximation, with proof in Appendix A.

Result 1. *The fixation probability of a beneficial allele initially at a frequency $p > 0$, $\pi(p)$, is the probability that the allele reaches frequency $x = 1$ before it reaches frequency $x = 0$. So, $\pi(p)$ is approximated by one of the following:*

1. *Low mutation supply: If $0 \le \theta < 1$, then*

$$\pi(p) = \frac{\Gamma(1-\theta, 0, 2p\alpha)}{\Gamma(1-\theta, 0, 2\alpha)}. \tag{5}$$

2. *High mutation supply: If $\theta \ge 1$, then $\pi(p) = 1$.*

Remark 1. Note that without mutation, $\theta = 0$, the expression simplifies to the classical approximation (Kimura, 1962),

$$\pi(p) = \frac{1 - e^{-2p\alpha}}{1 - e^{-2\alpha}}. \tag{6}$$

Remark 2. The case distinction in Result 1 follows from the fact that the boundary $x = 0$ is inaccessible if $1 \le \theta$ (see Appendix A.3).

Analysis of fixation time

Next we focus on the fixation time (conditioned on fixation), which has also been previously estimated using a diffusion equation approximation, e.g., by Kimura and Ohta (1969) and Ewens (2004, eq. 4.21). The following result gives an approximation for our model with unidirectional mutation, with proof in Appendix A.

Result 2. *The conditional fixation time of a beneficial allele initially at a frequency $p > 0$, $\bar{t}(p)$, is the waiting time for the allele to reach frequency $x = 1$ conditioned on reaching $x = 1$ before reaching $x = 0$,*

$$\bar{t}(p) = \int_0^1 t(x, p)\, \mathrm{d}x, \tag{7}$$

using one of the following Green's functions, $t(x, p)$:

1. *Low mutation supply: If $0 \le \theta < 1$, then Green's function is*

$$t(x,p) = \frac{2^\theta e^{2\alpha x}\,(\alpha x)^{\theta-1}}{(1-x)\Gamma(1-\theta,0,2\alpha)} \times \begin{cases} t_0(x,p), & \text{if } 0 \le x \le p < 1 \\ t_1(x,p), & \text{if } 0 < p \le x \le 1 \end{cases}, \tag{8}$$

where

$$t_0(x,p) = \frac{\Gamma(1-\theta,2p\alpha,2\alpha)\;\Gamma(1-\theta,0,2x\alpha)^2}{\Gamma(1-\theta,0,2\alpha p)}, \tag{9a}$$

$$t_1(x,p) = \Gamma(1-\theta,2x\alpha,2\alpha)\;\Gamma(1-\theta,0,2x\alpha). \tag{9b}$$

2. *High mutation supply: If $\theta \ge 1$, then Green's function is*

$$t(x,p) = \frac{2^\theta e^{2\alpha x}(\alpha x)^{\theta-1}}{1-x} \times \begin{cases} \Gamma(1-\theta,2p\alpha,2\alpha), & \text{if } 0 \le x \le p < 1, \\ \Gamma(1-\theta,2x\alpha,2\alpha), & \text{if } 0 < p \le x \le 1. \end{cases} \tag{10}$$

Remark 3. Note that due to the change of parameters from s to α and from μ to θ, to get the fixation time in generations we have to multiply the result by the population size, N.

Remark 4. In general, it is not possible to get analytical expressions for $\bar{t}(p)$ (eq. (7)). However, one can numerically integrate these expressions, which serves many practical applications. However, this requires the integration of gamma functions, and should be performed with care. For example, numerical integration with *Wolfram Mathematica* was successful, but using Python with `scipy` failed for some parameter values.

Remark 5. Without mutation, $\theta = 0$, the fixation time $\bar{t}(p)$ has been approximated previously by Kimura and Ohta (1969) (see also Ewens, 2004, Durrett, 2008). We present this result and a closed-form solution that uses exponential integral functions in Appendix B.

Approximation for strong selection

Hermisson and Pfaffelhuber (2008, abbreviated *HP* henceforth) have approximated the conditional fixation time in our model with recurrent mutation, assuming strong selection (large α).

In the following, $H_x = \sum_{n=1}^{\infty} \frac{x}{n(n+x)}$, which is an analytic continuation of the harmonic numbers for real-values x:

$$H_x = \sum_{k=1}^{x} \frac{1}{k} = 1 + \frac{1}{2} + \ldots + \frac{1}{x} + \left(\frac{1}{1+x} + \ldots\right) - \left(\frac{1}{1+x} + \ldots\right) = \sum_{n=1}^{\infty} \frac{1}{n} - \frac{1}{n+x} = \sum_{n=1}^{\infty} \frac{x}{n(n+x)}.$$

We now repeat their result using our notation.

Result 3. *Suppose the beneficial allele is initially absent ($p = 0$). Then the fixation time $\bar{t}(0)$ can be approximated by one of the following.*

1. *Low mutation supply: If $0 < \theta < 1$, then*

$$\bar{t}(0) = \frac{2}{\alpha}\left(\log(2\alpha) + \gamma_e\right) + O\left(\frac{\log(\alpha)}{\alpha^2}\right). \tag{11}$$

2. *High mutation supply: If $\theta \geq 1$, then*

$$\bar{t}(0) = \frac{2}{\alpha}\left(\log(2\alpha) + \gamma_e\right) - \frac{H_{\theta-1}}{\alpha} + O\left(\frac{\log(\alpha)}{\alpha^2}\right) + \frac{1}{\theta}O\left(\alpha e^{-\alpha}\right). \tag{12}$$

Intrigued by the Figure 3b, we assumed that the reduction in fixation time is independent of μ if we assume fixation to be at a value $z < 1$. By informed guessing and testing various parameter combinations against eq. (7) with the appropriate Green's function, we conclude that the difference is roughly

$$t^- = \frac{\log(2(1-z)\alpha) + (2-z)\gamma_e}{\alpha}. \tag{13}$$

Thus, to quickly calculate the time until the population reaches a fixation cut-off z, one can use the formula

$$\bar{t}(0) - t^- = \frac{\log\left(\frac{2\alpha}{1-z}\right) + z\gamma_e + \mathbb{1}_{\theta\in(1,\infty)}H_{\theta-1}}{\alpha}, \tag{14}$$

where $\mathbb{1}_{x\in A}$ denotes the indicator function that is one if x is an element of the set A and zero otherwise. The insets of Figure 3 show the very small relative error of eq. (14) compared to eq. (7).

Unfortunately, we did not find a and adjustment for an initial number of mutants $n_0 > 1$.

Numerical results: Fixation probability

We quantitatively compare our approximation (eq. (5), $\theta \geq 0$), the classical approximation (eq. (6), $\theta = 0$), and simulation results (eqs. (1) to (3)), in Figure 1. For small population sizes, the fixation probability goes to 1, because the initial number of copies of the beneficial allele is not far from the population size. For intermediate population sizes, roughly, $100 < N < 100,00$ the fixation probability is roughly $2s \cdot pN$ (Haldane, 1927, eq. 1.0). Importantly, for larger population sizes in which $\theta \geq 1$, our approximation (blue lines) increases to 1 as the mutational supply ($\theta = 2N\mu$) increases, whereas Kimura's approximation (orange lines) remains constant, underestimating the fixation probability (compare with markers).

Numerical results: Fixation time

Again, we provide a quantitative comparison of the above approximations and simulation results (eqs. (1) to (3)) in Figure 2. Several observations are worth mentioning. First, the classical approximation (eq. (B1)) is very good when the population size is small (i.e., mutation load is low, $\theta \leq 1$, left of the vertical line). Second, with a large population size (i.e., high mutation load, $\theta \geq 1$, right of the line), the classical approximation overestimates the fixation time. Third, both our approximation (Result 2) and the *HP* approximation (Result 3) perform well, even when the classical approximation fails. Fourth, although the *HP* approximation assumes strong selection, it performs well even for very weak selection ($s = 3 \cdot 10^{-6}$), although not as well as our approximation (compare blue and orange lines in Figure 2b).

However, the *HP* approximation overestimates the fixation time when two implicit assumptions are not met. First, it approximates the time for the beneficial allele to reach 100% frequency. However,

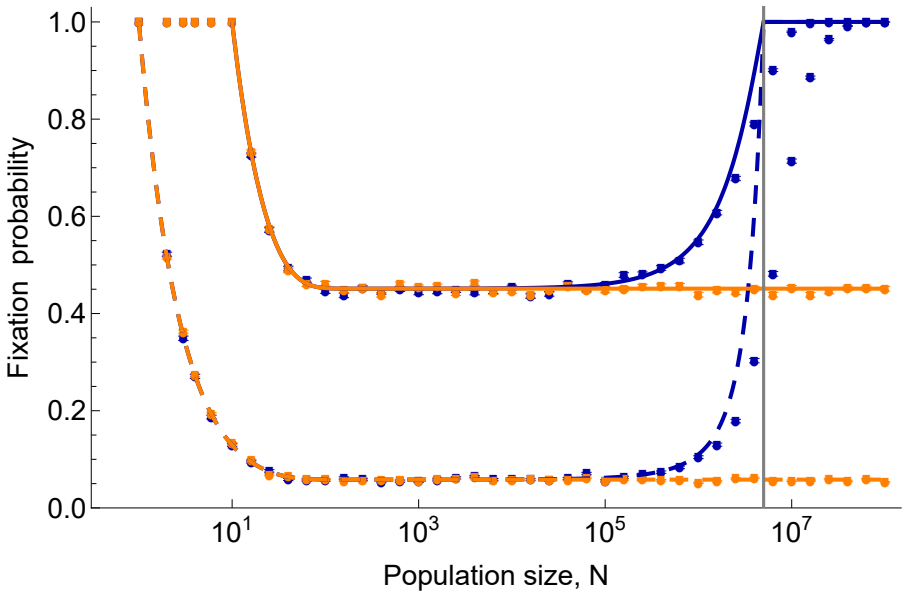

Figure 1: Fixation probability for increasing population size. Markers for results of simulations of the Wright-Fisher model (eqs. (1) to (3); error bars denote 50% CI.), lines for diffusion equation approximations: blue and orange for model with mutation ($\mu = 1 \cdot 10^{-7}$; Result 1) and without mutation ($\mu = 0$; eq. (6)), respectively. Lines and markers at the bottom are for fixation from a single copy of the beneficial allele ($p = 1/N$); lines and markers at the top are for fixation from 10 copies of the beneficial allele ($p = 10/N$). Gray vertical line indicates the population size $N = \mu/2$ at which $\theta = 1$. Here, the selection coefficient is $s = 0.03$.

waiting for complete fixation of the beneficial allele may take too long in many applications, including both simulations and experiments. In such cases, our approximation can still be used by simply changing the upper bound of the integral in eq. (7) from 1 to the required frequency z, such that we compute $\bar{t}_z(p) = \int_0^z t(x, p)\,\mathrm{d}x$. Comparing this and the *HP* approximation to results of simulations in which we wait for the beneficial allele to reach 95% frequency, rather than 100%, we find that our approximation performs well whereas the *HP* approximation overestimates so much that it is actually better to use the classical approximation (eq. (B1)), again by changing the upper bound of the second integral from 1 to z (Figure 3).

In addition, the *HP* approximation also assumes that the beneficial allele is initially absent, that is, $p = 0$. However, this may be unrealistic with a high mutation supply, which is expected to generate high standing variation. Consider a population evolving under stress, such as heat or drug treatment, until the fixation of a beneficial allele relieve the stress. Assuming this allele was neutral or deleterious before the stress, its frequency would have been roughly μ or μ/s, respectively. Under strong mutation, both values could be much larger than zero, meaning that fixation proceeds from an initial frequency $x \gg 0$ to $x = 1$. Our approximation (Result 2) explicitly takes the positive initial frequency of the beneficial allele into consideration as a parameter p, and indeed, it performs well when the initial frequency is greater than zero (Figure 4). However, the fixation time decreases as the initial frequency increases, and thus the *HP* approximation overestimates the fixation time in these cases, whereas the classical approximation performs well, as long as the mutation supply is low ($\theta \leq 1$), as it can also take the initial frequency as parameter (eq. (B1)).

Discussion

Recurrent mutation is defined here as the repeated production of a mutant genotype such that multiple independently produced mutants are co-segregating in the population. This can occur when the

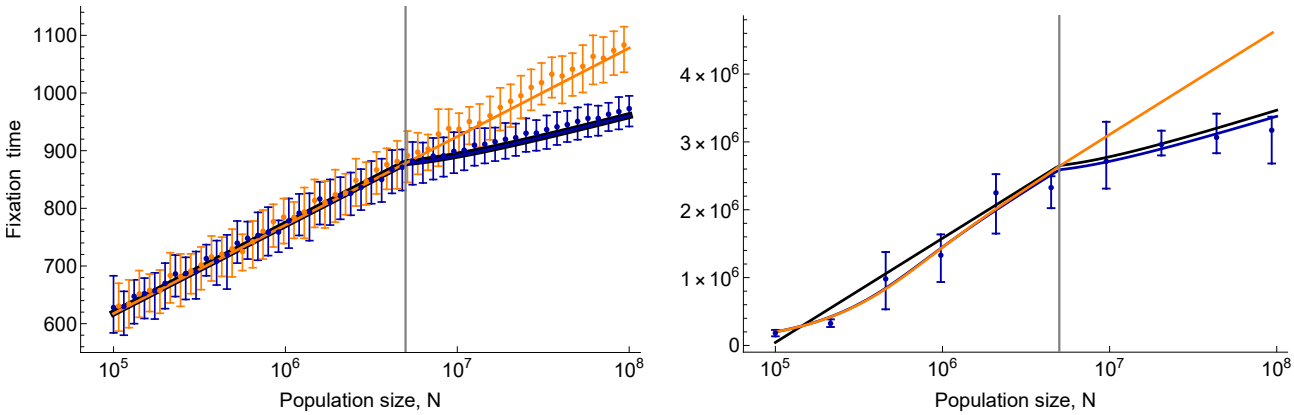

Figure 2: Fixation time for increasing population size. Time for the beneficial allele to go from one copy, $x = 1/N$, to full fixation, $x = 1$, conditioned on fixation (i.e., $x \neq 0$). Markers for results of simulations of the Wright-Fisher model (eqs. (1) to (3); error bars denote 50% CI), lines for diffusion equation approximations: blue and orange for model with ($\mu = 1 \cdot 10^{-7}$; Result 2) and without mutation ($\mu = 0$; eq. (B1)), respectively. Black line (sometimes overlapping the blue line) is the *HP* approximation (eqs. (11) and (12)). In both panels the fixation time with and without mutation diverges at $\theta = 1$, with the former being lower. **(a)** Strong selection, $s = 0.03$. **(b)** Weak selection, $s = 0.000003$. *HP* approximation fits well, only overestimating by a bit for majority of population sizes.

mutation supply is high due to large population sizes and/or high mutation rates. Note that here the term "mutation" includes not only single-nucleotide substitutions, but any genetic modification, such as focal or segmental duplications, aneuploidy, transposition, or short tandem repeats, all of which may have high mutation rates compared to point mutations (Press et al., 2019). Recurrent mutation has been somewhat neglected in the evolutionary theory literature due to the common assumption that mutations are rare, although previous studies have investigated its effect on the site-frequency spectrum (Harpak et al., 2016, Jenkins and Song, 2011, Wakeley et al., 2023), soft vs. hard selective sweeps (Pennings and Hermisson, 2006), genetic hitchhiking (Hermisson and Pfaffelhuber, 2008), and estimation of recombination rates (McVean et al., 2002).

Here we have developed a diffusion equation approximation for the fixation probability and time of a beneficial mutant allele under the effects of selection, drift, and importantly, recurrent mutation ($\theta = 2N\mu > 1$). We find that if recurrent mutations occur, then our approximation (Results 1 and 2) for the fixation probability and time is superior to the classic approximations (Kimura, 1962, Kimura and Ohta, 1969), which can significantly overestimate both quantities. Our approximation is flexible, as it considers both the evolutionary parameters (selection coefficient s, mutation rate μ, population size N), and the boundary conditions: the initial and final mutant frequencies (p and z, respectively). Nevertheless, if these boundary conditions are chosen such that the mutant increases in frequency from absence ($p = 0$) to full fixation ($z = 1$), then the previously developed *HP* approximations (eqs. (11) and (12)) are sufficient. Otherwise, our approximations are more accurate and flexible.

The diffusion equation is often used in evolutionary theory to describe evolutionary dynamics, that is, to describe the change in allele frequency over time in a population due to the effects of natural selection, mutation, and random genetic drift. Diffusion equation approximations have been derived for the probability and waiting time for an allele to fix in a population since the classical studies by Kimura (Kimura, 1962, Kimura and Ohta, 1969) under various models (Ewens, 2004, Durrett, 2008). However, the literature is missing an explicit analysis of the diffusion equation approximation in the case of a beneficial allele under the effect of strong mutation, in which the mutant allele continues to be generated even after the first copy has appeared; existing analyses assume selection without mutation (Kimura, 1962, Kimura and Ohta, 1969), uni-directional mutation without selection, and bi-directional

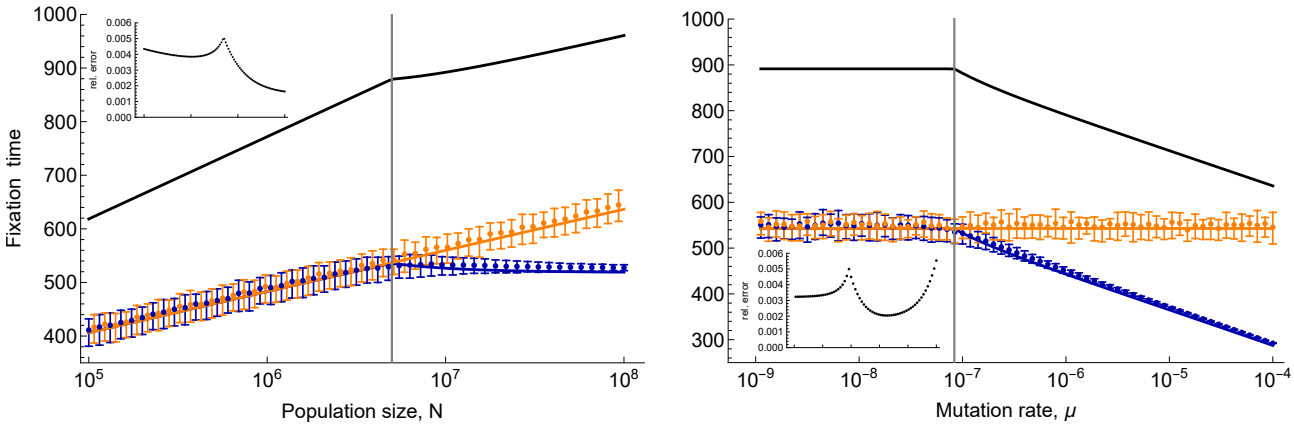

Figure 3: Time to reach 95% for increasing population size and mutation rate. Time for the beneficial allele to go from one copy, $x = 1/N$, to 95%, or $x = 0.95$, rather than full fixation. Markers for results of simulations of the Wright-Fisher model (eqs. (1) to (3); error bars denote 50% CI), lines for diffusion equation approximations. Blue and orange for model with ($\mu = 1 \cdot 10^{-7}$; Result 2) and without mutation ($\mu = 0$; eq. (B1)), respectively, integrated numerically from 0 to 0.95 rather than to 1. Black line (sometime overlapping the blue line) is the *HP* approximation (eqs. (11) and (12)), which assumes full fixation (100%) rather than high frequency (95%), and therefore overestimates the fixation time. Here, the selection coefficient is $s = 0.03$; **(a)** Mutation rate is $\mu = 1 \cdot 10^{-7}$; **(b)** Population size is $N = 6 \cdot 10^6$. The insets show the relative error of the adjusted HP approximation eq. (14) compared to the blue line.

weak mutation and selection (Ewens, 2004). Thus, our results close a gap in the theoretical literature on the fixation of beneficial alleles.

When the frequency of the mutant is low, a high mutation rate may be high enough to be a major factor in its evolution (i.e. its increase in frequency). When the mutant reaches a high frequency, its further evolution will be driven by natural selection. However, in some scenarios, the high mutation rate implies that other mutants are also likely to appear, leading to rapid adaptation, clonal interference, and soft sweeps. In these cases, the focal mutant is unlikely to reach fixation before the appearance of additional mutations, either on other backgrounds, leading to clonal interference, or on the background of the focal mutant, leading to further adaptation. Thus, the assumptions underlying our approximations may not be met. Still, these approximations are superior to the classic approximations, and can be used to estimate if the focal mutant has enough time to fix or reach some target establishment frequency before other beneficial mutants appear.

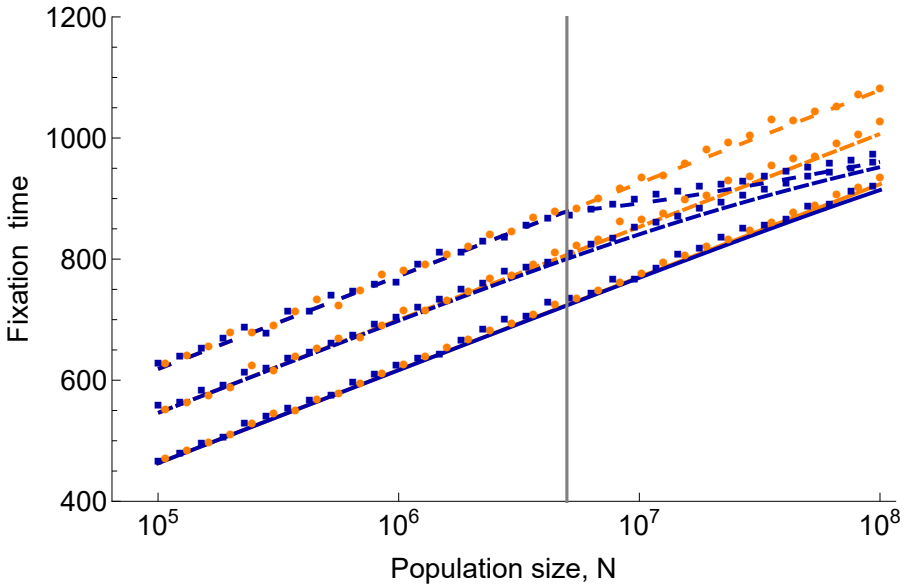

Figure 4: Fixation time with various initial frequencies. From top to bottom are shown fixation times for initial frequencies of the beneficial allele: top, $p = 1/N$; middle, $p = 100/N$; bottom, $p = 1,000/N$. Markers for results of simulations of the Wright-Fisher model (eqs. (1) to (3); error bars denote 50% CI), lines for diffusion equation approximations. Blue and orange for model with ($\mu = 10^{-7}$; Result 2) and without mutation ($\mu = 0$; eq. (B1)), respectively. Importantly, the fixation time decreases as the initial frequency increases. The approximations fit the simulations well, and the difference between models with and without mutation shrink as the initial frequency increases. Error bars are small and similar to those in Figure 2, and not shown to avoid over-plotting. Here, the selection coefficient is $s = 0.03$.

Appendices

Appendix A Proofs of Result 1 and Result 2

A.1 Boundary types

There are two important properties that characterize the boundary: (i) a boundary is *accessible* if there is a positive probability that it can be reached in finite time from a given interior point, and (ii) is *absorbing* if the process remains forever at the boundary after reaching the boundary.

Here, we determine the boundary type of $x = 0$ and $x = 1$.This is nicely demonstrated by Durrett (2008), which not only differs in notation from Ewens (2004), but also provides helpful examples. We focus on a specific example (Durrett, 2008, example 7.33, pg. 295) that studies a Wright-Fisher diffusion with mutations in both directions, but without selection. It is explained that adding selection does not change the (in)finiteness of the integrals. If $0 \leq \theta < 1$, then $x = 0$ is accessible but not absorbing, whereas it is not accessible for $\theta \geq 1$. In our model there is no mutation in the other direction, from A to a, so the boundary $x = 1$ is absorbing.

Therefore, for $0 \leq \theta < 1$ there is a positive probability $1 - \pi(p) > 0$ that population initially with frequency p of allele A reaches $x = 0$ before it reaches $x = 1$. However, mutation will eventually introduce the extinct A allele and thus this state, $x = 0$, is not absorbing. Therefore, to continue with the derivation of the fixation times, we have to consider the two cases: $0 \leq \theta < 1$ and $1 \leq \theta$.

A.2 Boundary x = 0 is accessible, $0 \leq \theta < 1$

An important quantity that we will use in the following analyses is $\psi(x)$. It is given in eq. 4.16 in (Ewens, 2004). Substituting $a(x)$ and $b(x)$ from eq. (4) we have

$$\psi(x) = \exp\left(-2\int_x \frac{a(y)}{b(y)}\,\mathrm{d}y\right) = e^{-2\alpha x}x^{-\theta}, \tag{A1}$$

Note that ψ is different from Ψ, which denotes the digamma function.

If $\theta < 1$, we use eqs. 4.22 and 4.23 of Ewens (2004) to calculate Green's function,

$$t(x,p) = \frac{2}{b(x)\psi(x)} \times \begin{cases} \big(1-\pi(p)\big)\int_0^x \psi(y)\,\mathrm{d}y, & \text{if } 0 \leq x \leq p < 1, \\ \pi(p)\int_x^1 \psi(y)\,\mathrm{d}y, & \text{if } 0 < p \leq x \leq 1, \end{cases} \tag{A2}$$

where $\pi(p)$ is the fixation probability. Substituting $\pi(p)$ from eq. (5) into eq. (A2) gives the following Green's function,

$$t^*(x,p) = \frac{2^\theta e^{2x\alpha}(\alpha x)^{\theta-1}}{(1-x)\,\Gamma(1-\theta,0,2\alpha)} \times \begin{cases} t_0(x,p), & \text{if } 0 \leq x \leq p < 1, \\ t_1(x,p), & \text{if } 0 < p \leq x \leq 1, \end{cases} \tag{A3}$$

where

$$t_0(x,p) = \Gamma(1-\theta,2p\alpha,2\alpha)\;\Gamma(1-\theta,0,2x\alpha) \tag{A4a}$$

$$t_1(x,p) = \Gamma(1-\theta,2x\alpha,2\alpha)\;\Gamma(1-\theta,0,2p\alpha). \tag{A4b}$$

All derivations are done with the help of the computer algebra software *Wolfram Mathematica*.

So far the fixation time was calculated for a population that randomly reaches one of the absorbing states, either $x = 0$ or $x = 1$. In this study we want to focus on the time until the population hits and is absorbed in $x = 1$, that is, we are interested in the conditional fixation time, conditioned on reaching $x = 1$ before $x = 0$.

In equation 4.48 in Ewens (2004, pg. 146) we have the formula to compute the sojourn times conditioned that the process hits $x = 1$ first, which gives the following Green's function,

$$t(x,p) = t^*(x,p)\,\frac{\pi(x)}{\pi(p)} = t^*(x,p)\,\frac{\Gamma(1-\theta,0,2x\alpha)}{\Gamma(1-\theta,0,2p\alpha)}, \tag{A5}$$

by substituting eq. (5) for $\pi(p)$. Substituting eq. (A3) into eq. (A5) gives eq. (8).

A.3 Boundary x = 0 is not accessible, $1 \leq \theta$

If $\theta \leq 1$, only $x = 1$ is an accessible boundary state and we have to use eqs. 4.40 and 4.41 from Ewens (2004),

$$t(x,p) = \frac{2}{b(x)\,\psi(x)} \times \begin{cases} \int_p^1 \psi(y)\,\mathrm{d}y, & \text{if } 0 \leq x \leq p < 1, \\ \int_x^1 \psi(y)\,\mathrm{d}y, & \text{if } 0 < p \leq x \leq 1. \end{cases} \tag{A6}$$

Inserting eq. (4) and eq. (A1) into eq. (A6) and solving with *Wolfram Mathematica*, we get eq. (10).

A.4 Green's function $t(x,p)$ at $\theta = 1$

We showed above that there have to be two different derivations for Green's functions, depending on θ. However, at $\theta = 1$, the expressions coincide; the limit from below is a bit more complex, but can be calculated with *Wolfram Mathematica*. We get

$$\lim_{\theta\uparrow 1} t(x,p) = \lim_{\theta\downarrow 1} t(x,p) = \frac{2e^{2x\alpha}}{1-x} \times \begin{cases} \Gamma(0,2p\alpha,2\alpha), & \text{if } 0 \leq x \leq p < 1, \\ \Gamma(0,2x\alpha,2\alpha), & \text{if } 0 < p \leq x \leq 1. \end{cases} \tag{A7}$$

Appendix B Approximation of fixation time without mutation

We set $\theta = 0$ in eqs. (A5), (8) and (9) and use *Wolfram Mathematica* to find a representation of Green's function, $t(x, p)$, with hyperbolic functions, which is shorter than a representation with exponentials (e.g. Kimura and Ohta (1969, eq. 17)). Thus, the expected conditional fixation time without mutation is approximated by

$$\bar{t}(p) = \int_0^p I_1 \, \mathrm{d}x + \int_p^1 I_2 \, \mathrm{d}x, \tag{B1}$$

where

$$I_1 = \frac{2\left(\coth(p\alpha) - \coth(\alpha)\right)\ \sinh^2(x\alpha)}{(1-x)x\alpha},$$

$$I_2 = \frac{2\mathrm{csch}(\alpha)\ \sinh(x\alpha)\ \sinh\left((1-x)\alpha\right)}{(1-x)x\alpha}.$$

Wolfram Mathematica further gives an explicit result for eq. (B1) in terms of hyperbolic integrals, but we prefer to show a non-hyperbolic version,

$$\begin{aligned}
\int_0^p I_1 \, \mathrm{d}x &= \int_0^p \frac{\exp(-2x\alpha)\left(\exp(2p\alpha) - \exp(2\alpha)\right)\left(1 - \exp(2x\alpha))^2\right)}{\left(1 - \exp(2\alpha)\right)\left(1 - \exp(2p\alpha)\right)(1-x)x\alpha} \mathrm{d}x \\
&= \frac{\exp(2p\alpha) - \exp(2\alpha)}{(1 - \exp(2\alpha)(1 - \exp(2p\alpha))\alpha}\left(E_1 + L_1\right), \\
\int_p^z I_2 \, dx &= \int_p^z \frac{\exp(-2x\alpha)\left(\exp(2x\alpha) - \exp(2\alpha)\right)\left(-1 + \exp(2x\alpha))\right)}{\left(1 - \exp(2\alpha)\right)(1-x)x\alpha} \mathrm{d}x \\
&= \frac{E_2^z + L_2^z}{\left(1 - \exp(2\alpha)\right)\alpha}
\end{aligned}$$

where

$$\begin{aligned}
E_1 &= \exp(-2\alpha)\Gamma\big(0, -2\alpha, -2(1-p)\alpha\big) + \\
&\quad \exp(2\alpha)\Gamma\big(0, 2\alpha, 2(1-p)\alpha\big) - \big(Ei(2p\alpha) + Ei(-2p\alpha)\big) \\
L_1 &= 2(\log\left(\frac{2p\alpha}{1-p}\right) + \gamma_e) \\
E_2^z &= \exp(2\alpha)\Big(\Gamma\big(0, 2p\alpha, 2(1-p)\alpha\big) - \Gamma\big(0, 2z\alpha, 2(1-z)\alpha\big)\Big) + \\
&\quad \Gamma\big(0, -2\alpha, -2(1-p)\alpha\big) - \Gamma\big(0, -2z\alpha, -2(1-z)\alpha\big) \\
L_2^z &= \big(1 + \exp(2\alpha)\big)\log\left(\frac{p(1-z)}{z(1-p)}\right) \\
E_2 &= \exp(2\alpha)\Big(\Gamma\big(0, 2p\alpha, 2(1-p)\alpha\big) + Ei(-2\alpha)\Big) + \Gamma\big(0, -2p\alpha, -2(1-p)\alpha\big) + Ei(2\alpha) \\
L_2 &= \big(1 + \exp(2\alpha)\big)\big(\log\left(\frac{p}{(2\alpha(1-p)}\right) - \gamma_e\big)
\end{aligned}$$

Here, E_2^z and L_2^z represent the expressions if the upper bound is $z < 1$. Since just setting $z = 1$ does not work, one needs to take the limit $z \to 1$ in $E_2^z + L_2^z$. This results in $E_2 + L_2$.

References

Motoo Kimura. On the probability of fixation of mutant genes in a population. *Genetics*, 47(6):713–9, 1962.

Motoo Kimura and Tomoko Ohta. The Average Number of Generations until Fixation of a Mutant Gene in a Finite Population. *Genetics*, 61(3):763–71, 1969.

Talia Karasov, Philipp W Messer, and Dmitri A. Petrov. Evidence that adaptation in Drosophila is not limited by mutation at single sites. *PLoS Genettics*, 6(6):e1000924, 2010. doi: 10.1371/journal.pgen.1000924.

João Barroso-Batista, Ana Sousa, Marta Lourenço, Marie-Louise Bergman, Daniel Sobral, Jocelyne Demengeot, Karina B. Xavier, and Isabel Gordo. The First Steps of Adaptation of *Escherichia coli* to the Gut Are Dominated by Soft Sweeps. *PLoS Genet.*, 10(3):e1004182, 2014. doi: 10.1371/journal.pgen.1004182.

Sasha F. Levy, Jamie R. Blundell, Sandeep Venkataram, Dmitri A. Petrov, Daniel S. Fisher, and Gavin Sherlock. Quantitative evolutionary dynamics using high-resolution lineage tracking. *Nature*, 519 (7542):181–186, 2015. doi: 10.1038/nature14279.

Alex N. Nguyen Ba, Ivana Cvijović, José I. Rojas Echenique, Katherine R. Lawrence, Artur Rego-Costa, Xianan Liu, Sasha F. Levy, and Michael M. Desai. High-resolution lineage tracking reveals travelling wave of adaptation in laboratory yeast. *Nature*, 575(7783):494–499, 2019. doi: 10.1038/s41586-019-1749-3.

Chris M. Egan, Srinath Sridhar, Michael Wigler, and Ira M. Hall. Recurrent DNA copy number variation in the laboratory mouse. *Nature Genetics*, 39(11):1384–1389, 2007. doi: 10.1038/ng.2007.19.

Avihu H. Yona, Yair S Manor, Rebecca H. Herbst, Gal H. Romano, Amir Mitchell, Martin Kupiec, Yitzhak Pilpel, and Orna Dahan. Chromosomal duplication is a transient evolutionary solution to stress. *Proceedings of the National Academy of Sciences of the United States of America*, 109(51): 21010–5, 2012. doi: 10.1073/pnas.1211150109.

Celia Payen, Sara C Di Rienzi, Giang T Ong, Jamie L Pogachar, Joseph C Sanchez, Anna B Sunshine, M K Raghuraman, Bonita J Brewer, and Maitreya J. Dunham. The Dynamics of Diverse Segmental Amplifications in Populations of *Saccharomyces cerevisiae* Adapting to Strong Selection. *G3 Genes|Genomes|Genetics*, 4(3):399–409, 2014. doi: 10.1534/g3.113.009365.

Grace Avecilla, Julie N Chuong, Fangfei Li, Gavin Sherlock, David Gresham, and Yoav Ram. Neural networks enable efficient and accurate simulation-based inference of evolutionary parameters from adaptation dynamics. *PLOS Biology*, 20(5):e3001633, 2022. ISSN 1545-7885. doi: 10.1371/journal.pbio.3001633.

Arbel Harpak, Anand Bhaskar, and Jonathan K. Pritchard. Mutation Rate Variation is a Primary Determinant of the Distribution of Allele Frequencies in Humans. *PLoS Genetics*, 12(12):1–22, 2016. doi: 10.1371/journal.pgen.1006489.

Maximilian Oliver Press, Ashley N. Hall, Elizabeth A. Morton, and Christine Queitsch. Substitutions Are Boring: Some Arguments about Parallel Mutations and High Mutation Rates. *Trends in Genetics*, 35(4):253–264, 2019. doi: 10.1016/j.tig.2019.01.002. URL https://doi.org/10.1016/j.tig.2019.01.002.

Joachim Hermisson and Peter Pfaffelhuber. The pattern of genetic hitchhiking under recurrent mutation. *Electronic Journal of Probability*, 13:2069–2106, 2008. doi: 10.1214/EJP.v13-577.

Warren J. Ewens. *Mathematical Population Genetics*. Springer, New York, 2 edition, 2004.

Zaheerabbas Patwa and Lindi M. Wahl. The fixation probability of beneficial mutations. *Journal of The Royal Society Interface*, 5(28):1279–89, 2008. doi: 10.1098/rsif.2008.0248.

David M. McCandlish and Arlin Stoltzfus. Modeling evolution using the probability of fixation: History and implications. *The Quarterly Review of Biology*, 89:225–252, 9 2014. doi: 10.1086/677571.

J. B.S. Haldane. A Mathematical Theory of Natural and Artificial Selection, Part V: Selection and Mutation. *Mathematical Proceedings of the Cambridge Philosophical Society*, 23(7):838–844, 1927. doi: 10.1017/S0305004100015644.

Ilan Eshel. On the survival probability of a slightly advantageous mutant gene with a general distribution of progeny size-A branching process model. *Journal of Mathematical Biology*, 12(3):355–362, 1981. doi: 10.1007/BF00276922.

Richard Durrett. *Probability Models for DNA Sequence Evolution*. Springer New York, New York, NY, 2008. ISBN 978-0-387-78168-6. doi: 10.1007/978-0-387-78168-6.

Paul A. Jenkins and Yun S. Song. The effect of recurrent mutation on the frequency spectrum of a segregating site and the age of an allele. *Theoretical Population Biology*, 80(2):158–173, 2011. doi: 10.1016/j.tpb.2011.04.001.

John Wakeley, Wai-tong Louis Fan, Evan Koch, and Shamil Sunyaev. Recurrent mutation in the ancestry of a rare variant. *Genetics*, (March):1–47, mar 2023. doi: 10.1093/genetics/iyad049.

Pleuni S. Pennings and Joachim Hermisson. Soft sweeps II–molecular population genetics of adaptation from recurrent mutation or migration. *Molecular Biology & Evolution*, 23(5):1076–84, may 2006. doi: 10.1093/molbev/msj117.

Gil McVean, Philip Awadalla, and Paul Fearnhead. A coalescent-based method for detecting and estimating recombination from gene sequences. *Genetics*, 160(3):1231–1241, 2002. doi: 10.1093/genetics/160.3.1231.