Haldane's Probability of Mutant Survival is Not the Probability of Allele Establishment

Ivan Krukov* **and A.P. Jason de Koning*,†**

*University of Calgary, Cumming School of Medicine, Department of Biochemistry and Molecular Biology Graduate Program (Bioinformatics stream), Calgary, Alberta, Canada, †University of Calgary, Cumming School of Medicine, Department of Medical Genetics, and Alberta Children's Hospital Research Institute, Calgary, Alberta, Canada

ABSTRACT Haldane notably showed in 1927 that the probability of fixation for an advantageous allele is approximately $2s$, for selective advantage s. This widely known result is variously interpreted as either the fixation probability or the establishment probability, where the latter is considered the likelihood that an allele will survive long enough to have effectively escaped loss by drift. While Haldane was concerned with escape from loss by drift in the same paper, in this short note we point out that: 1) Haldane's 'probability of survival' is analogous to the probability of fixation in a Wright-Fisher model (as also shown by others); and 2) This result is unrelated to Haldane's consideration of how common an allele must be to 'probably spread through the species'. We speculate that Haldane's survival probability may have become misunderstood over time due to a conflation of terminology about surviving drift and 'ultimately surviving' (*i.e.*, fixing). Indeed, we find that the probability of establishment remarkably appears to have been overlooked all these years, perhaps as a consequence of this misunderstanding. Using straightforward diffusion and Markov chain methods, we show that under Haldane's assumptions, where establishment is defined by eventual fixation being more likely that extinction, the establishment probability is actually $4s$ when the fixation probability is $2s$. Generalizing consideration to deleterious, neutral, and adaptive alleles in finite populations, if establishment is defined by the odds ratio between eventual fixation and extinction, k, the general establishment probability is $(1+k)/k$ times the fixation probability. It is therefore $4s$ when $k=1$, or $3s$ when $k=2$ for beneficial alleles in large populations. As k is made large, establishment becomes indistinguishable from fixation, and ceases to be a useful concept. As a result, we recommend establishment be generally defined as when the odds of ultimate fixation are greater than for extinction ($k=1$, following Haldane), or when fixation is twice as likely as extinction ($k=2$).

KEYWORDS establishment, fixation, diffusion approximation, Wright-Fisher model, computational population genetics

Introduction

Established alleles are those that have persisted long enough in a population so that they are unlikely to be lost by chance, and can thus be said to have escaped loss by drift. The concept of establishment traces back at least to Fisher (1922, p. 419). Fisher argued that the fate of a new positively selected mutant is determined more by random drift when it is rare than if it persists long enough to spread through much of the population. Later, Haldane (1927) sought to "consider the course of events in a population where the new factor is present in such numbers as to be in no danger of extinction by mere bad luck". In that work, Haldane used a branching process formulation to derive his famous probability of eventual fixation for a beneficial mutant with selection coefficient s ($P_{\text{Fix}} \approx 2s$), which was termed the "probability of survival". Perhaps because of context and the use of this term, which comes from the theory of branching processes, $2s$ is now commonly interpreted as both the fixation probability (*e.g.*, Kimura 1962; Otto and Whitlock 1997) *and* as the establishment probability (*i.e.*, the probability that an allele

doi: 10.1534/genetics.XXX.XXXXXX
Manuscript compiled: Sunday 14th July, 2019
*,†Corresponding authors: University of Calgary, Cumming School of Medicine, Department of Biochemistry and Molecular Biology. Email: (IK) ikryukov@ucalgary.ca, (APJdK) jdekonin@ucalgary.ca

will reach a sufficiently high frequency to have effectively escaped loss by drift) (*e.g.*, Gerrish and Lenski 1998; Peck 1994; Messer and Petrov 2013). However, while fixation implies prior establishment, establishment should not necessarily imply fixation with complete certainty, or else it is at best a redundant concept. The probability of establishment should thus nearly always be greater than the probability of fixation, and the difference between the two should vary according to how exactly we define establishment.

A reading of Haldane's paper (Haldane 1927) supports this view. In particular, after deriving the probability of fixation, Haldane went on to argue that a dominant allele achieving a population count of $\log(2)/2s$ will "probably" spread through the species. Although it was not stated explicitly, this is an approximation to the number of initial mutants that make the probability of fixation equal to $1/2$, so that exceeding this threshold implies better than even odds that the allele will go on to become fixed (derivation below). Here we clarify that, unlike in the branching process formulation (Ewens 2004, p. 29), fixation and survival to establishment can be easily differentiated in a Wright-Fisher model. A Wright-Fisher perspective will thus clarify that Haldane's $2s$ is unrelated to his establishment count, since the probability of fixation given the establishment count of $\log(2)/2s$ is $1/2$ (not $2s$). Under Haldane's assumptions, we will see that the establishment probability of a beneficial allele is rather approximated by $4s$, and is therefore off by a factor of 2 compared to the probability of fixation. This result applies to both haploid and diploid populations with zygote/heterozygote fitness of $1+s$. When heterozygote fitness is defined as $(1+hs)$ for $h = 1/2$, the probabilities of fixation and establishment are rather s and $2s$ respectively.

We begin by reviewing Haldane's treatment and its assumptions. We next consider some related arguments by Gillespie (2004), before moving to a diffusion approach and eventually to a direct analysis of the discrete-time Wright-Fisher model (where we are not required to assume weak mutation, weak selection, or large population size). Moving to a full Markov chain treatment is important for validation, and because the cases where establishment are most of interest occur when population mutation rates may be very large and thus could violate assumptions that diffusion approximations usually require (de Koning and de Sanctis 2018). We focus on the general case where mutants may be deleterious, neutral, or advantageous and show that a diffusion approximation to the establishment probability has a pleasing simplicity when defined appropriately. We conclude that both the probability and rate of establishment (Messer and Petrov 2013) are different from what has been previously understood by as much as a factor of 2.

Establishment count in a branching-process model

Haldane used a branching process formulation to consider the ultimate survival of a mutant allele. Assuming individuals leave a Poisson-distributed number of offspring in the next generation, and that the number of mutant offspring has mean $1+s$, the probability that the mutant population will eventually go extinct in the limit of $t \to \infty$ can be determined by considering the probability that a newly arisen mutant in the current generation, t, will eventually go extinct, $P_{\text{Ext}}(X_t = 1)$ (where X_t indicates the number of mutants present in generation t). Noting that $P_{\text{Ext}}(X_t = 1)$ is equivalent to $1 - P_{\text{Fix}}(X_t = 1)$, this can be expressed by writing the probability that an allele will leave i mutant offspring in the next generation times the probability that every one of the i mutants will eventually go extinct, integrated over all i:

$$\begin{aligned} 1 - P_{\text{Fix}}(X_t = 1) &= \sum_{i=0}^{\infty} e^{-(1+s)} \frac{(1+s)^i}{i!} (1 - P_{\text{Fix}}(X_{t+1} = 1))^i \\ &= e^{-(1+s)P_{\text{Fix}}(X_{t+1}=1)} \end{aligned}$$

By noting that the probability of fixation of a newly arisen mutant should be the same in each generation under constant population size and selective effect (Otto and Whitlock 1997), we can set $P_{\text{Fix}}(X_t = 1)$ and $P_{\text{Fix}}(X_{t+1} = 1)$ to $P_{\text{Fix}}(X_0 = 1)$ and solve for s in terms of P_{Fix}. This yields

$$s = -\frac{P_{\text{Fix}}(X_0 = 1) + \log(1 - P_{\text{Fix}}(X_0 = 1))}{P_{\text{Fix}}(X_0 = 1)}.$$

Suppressing the dependence on X_t for convenience, we take the Taylor series expansion of the solution for s around $P_{\text{Fix}} = 0$,

$$s = \frac{P_{\text{Fix}}}{2} + \frac{P_{\text{Fix}}^2}{3} + \frac{P_{\text{Fix}}^3}{4} + O(P_{\text{Fix}}^4)$$

and obtain Haldane's famous result for the probability of fixation of a beneficial allele when using the leading term:

$$P_{\text{Fix}} \approx 2s$$

As noted above, Haldane next derived a minimum allele count, $c^* = \log(2)/2s$, that if exceeded would ensure that the allele will probably spread through the species (*i.e.*, establish). To see how he obtained this result, we begin with an approximation to the probability that c mutants will eventually go extinct based on the above result, $(1-2s)^c$ (Haldane 1927; Otto and Whitlock 1997). This approximation works well when s is small and positive, and indeed it corresponds to Kimura's extinction probability from diffusion theory for a starting frequency of $c/(2N)$ in the limit of infinite population size (and to a first order approximation when s is small; not shown). To find the minimum number of starting copies such that fixation and extinction are equally likely, c^*, we use Haldane's expression for the extinction probability and set this to $(1-2s)^{c^*} = \frac{1}{2}$, so that c^* represents the count at which fixation becomes more likely than extinction when it is exceeded. Solving for c^* yields

$$c^* = -\frac{\log(2)}{\log(1-2s)}$$

Taking the first term in a series expansion of this result around $s = 0$ then gives

$$\begin{aligned} c^* &= \frac{\log(2)}{2s} - \frac{\log(2)}{2} - \frac{\log(2)s}{6} - \frac{\log(2)s^2}{6} + O(s^3) \\ &\approx \frac{\log(2)}{2s} \end{aligned}$$

recovering Haldane's result. We will refer to this quantity as Haldane's establishment count. Throughout this paper, we refer to the minimum population frequency required to become established as the *establishment frequency* (f^*), and the corresponding allele count as the *establishment count* (c^*). Based on the above definitions, a population that achieves exactly Haldane's establishment count will have a fixation probability of $1/2$ (not $2s$, which is the fixation probability assuming we started with a single mutant copy).

Establishment frequency in a Wright-Fisher diffusion ($s > 0, N \to \infty$)

Kimura (1962, 1964) considered the fixation probability in a diffusion approximation to a Wright-Fisher model including selection and drift in a series of celebrated papers. Given an initial mutant allele frequency $X_0 = f_0$, a population size of N diploid reproducing individuals, selection coefficient s and heterozygote fitness $1 + s$, Kimura's probability of fixation is given by:

$$P_{\text{Fix}}(X_0 = f_0) = \frac{1 - e^{-4Nsf_0}}{1 - e^{-4Ns}} \tag{1}$$

For a single starting copy ($f_0 = 1/(2N)$), this yields

$$P_{\text{Fix}}(X_0 = \frac{1}{2N}) = \frac{1 - e^{-2s}}{1 - e^{-4Ns}}$$

As Kimura (1962, eq. 11) noted, in the limit of infinite population size, this equation agrees with Haldane's result to a first-order approximation:

$$\begin{aligned}\lim_{N\to\infty} P_{\text{Fix}}(X_0 = \frac{1}{2N}) &= 1 - e^{-2s} \\ &= 2s - 2s^2 + \frac{4s^3}{s} + O(s^4) \\ &\approx 2s\end{aligned}$$

for $s > 0$ (where the series expansion was again taken around $s = 0$). This result supports the standard interpretation that Haldane's $2s$ represents the fixation (not establishment) probability.

Using equation 1 as a starting point, Gillespie (2004, sec. 3.9, eq. 3.25) later derived an expression for the establishment frequency, f^*, that makes the probability of fixation close to 1 (within a prescribed margin of error, $\epsilon \approx 0$, so that $1 - \epsilon \approx 1$). Gillespie first assumed that $2Ns$ is large enough that the denominator of equation 1 approaches 1 and can be ignored. If we define heterozygote fitness as $1 + s$ and homozygote fitness as $1 + 2s$ (rather than the $1 + \frac{1}{2}s$ and $1 + s$ that Gillespie used), we obtain

$$\begin{aligned}1 - \epsilon = P_{\text{Fix}}(X_0 = f^*) \\ \approx 1 - e^{-4Nsf^*}.\end{aligned}$$

Solving for f^*, we then obtain

$$f^* = \frac{-\log(\epsilon)}{4Ns},$$

which is consistent with Gillespie's reported result given our redefinition of heterozygote fitness. By setting $\epsilon = \frac{1}{2}$ so that exceeding the establishment frequency makes the probability of fixation greater than the probability of extinction, this yields an establishment frequency that is equivalent to Haldane's establishment count divided by $2N$. Note that in both Haldane and Gillespie's derivations, s was assumed positive and N was effectively assumed large so that the results do not necessarily apply to deleterious or neutral alleles, and do not account for finite population size effects.

We will now generalize the concept of allele establishment by combining the approaches of Haldane, Gillespie, and Kimura, and by relaxing assumptions about s and N. As we show, both the establishment frequency and probability have simple closed form approximations in the diffusion framework, even when the definition of establishment is varied according to the relative odds of eventual fixation to extinction.

Results

Diffusion approximations to the establishment frequency and probability are now developed. These will be contrasted with numerical results obtained by independent analyses of discrete Wright-Fisher models, which are based only on the definitions of establishment and the model itself. The method used to efficiently analyze the discrete models is described in the Appendix.

Establishment frequency in a Wright-Fisher diffusion (general case)

As above, we define the establishment frequency $f^* = \frac{c^*}{2N}$ for a given establishment count, c^*. For generality, we define establishment in terms of the odds ratio k, such that eventual fixation is k times more likely than extinction. Definitions based on quantities other than the odds ratio are possible, however, as we will see, the odds ratio produces a convenient simplification and allows existing arguments as special cases (*e.g.*, $k = 1$ for Haldane, $k = (1 - \epsilon)/\epsilon$ for Gillespie). Thus, the desired probability of fixation, $P_{\text{Fix}}(X_0 = f^*)$, can be written as:

$$\begin{aligned}P_{\text{Fix}}(X_0 = f^*) &= kP_{\text{Ext}}(X_0 = f^*) \\ &= k(1 - P_{\text{Fix}}(X_0 = f^*)) \\ &= \frac{k}{1+k}\end{aligned} \tag{2}$$

Following the approach used by Gillespie (2004), but without applying his approximations, equation 1 can be used to directly solve for the initial allele frequency $X_0 = f^*$ that satisfies equation 2:

$$\begin{aligned}\frac{1 - e^{-4Nsf^*}}{1 - e^{-4Ns}} &= \frac{k}{1+k} \\ f^* &= \frac{\log\left(1 - \left(\frac{k}{1+k}\right)(1 - e^{-4Ns})\right)}{-4Ns}\end{aligned} \tag{3}$$

Immediately, we can confirm the intuitive result that when fixation and extinction are equally likely ($k = 1$), the establishment frequency (equation 3) of a neutral variant is $\frac{1}{2}$:

$$\lim_{s\to 0} f^* = \frac{1}{2}$$

As a further check, multiplying the general solution in equation 3 by $2N$ and taking the limit as $N \to \infty$ gives:

$$\lim_{N\to\infty} 2Nf^* = \frac{\log(1+k)}{2s}$$

for $s > 0$. As expected, this result with $k = 1$ is equivalent to Haldane's establishment count when $s > 0$.

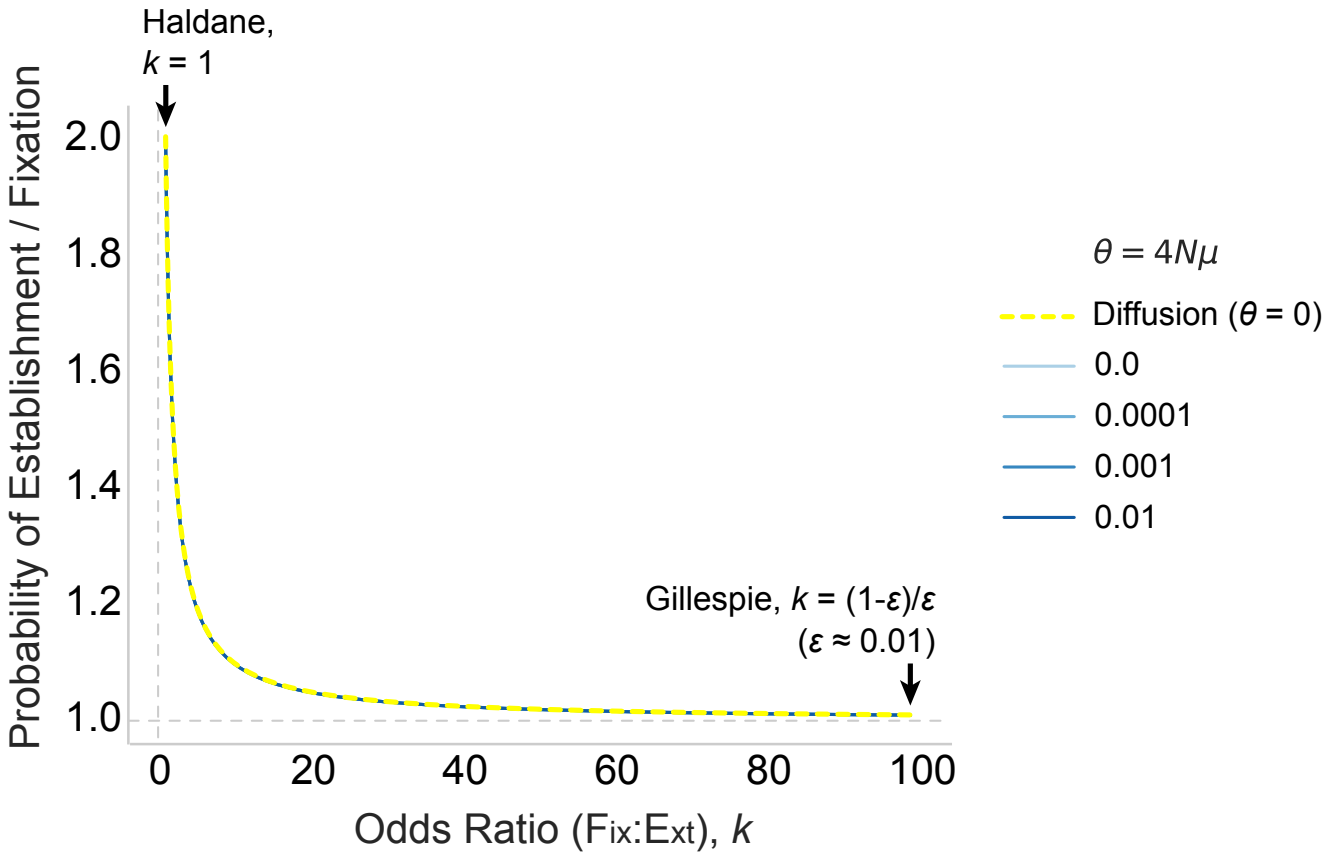

Figure 1 Establishment probability compared to fixation probability for a range of fixation-to-extinction odds ratios (k). The dashed yellow line shows results of calculations made assuming $p = 1/(2N)$ and using the diffusion method. Solid lines show results based on direct computation of the establishment count and its corresponding probability using the discrete-time Wright-Fisher model. Note, the methods agree and numerically correspond with $(1+k)/k$ as implied by equation 5. Accordingly, the results are invariant to N and s; these results were computed for $s = 0$ and $N = 10,000$. Values of $k = 1$ and $k = (1-\epsilon)/\epsilon$ capture the assumptions of Haldane and Gillespie, respectively. At Haldane's more liberal definition of establishment, the probability of establishment is twice the probability of fixation, whereas for Gillespie's definition, $P_{\text{Establish}} \to P_{\text{Fix}}$ approximately as $\epsilon \to 0$).

Establishment probability in a Wright-Fisher diffusion

Establishment probability can be defined as the probability of reaching the establishment frequency (or count) before going to extinction. Establishment may therefore be considered in a two absorbing-state Wright-Fisher model, with absorbing states defined at $X_t \in \{0, f^*\}$ (in the continuous state-space). An expression for the establishment probability can then be found using diffusion theory in a manner analogous to how Kimura derived the fixation probability, but where the fixation absorbing boundary is moved to the establishment frequency and its meaning redefined.

Following Kimura (Kimura 1962), let $u(f_0, t)$ be the probability that the mutant allele will absorb at the upper absorbing boundary during time interval t, given initial frequency f_0. We are interested in a solution to $\lim_{t\to\infty} u(f_0, t) = u(f_0)$, with the boundary conditions $u(0) = 0$ and $u(f^*) = 1$. We solve for $u(f_0)$ by setting up the appropriate Kolmogorov backward equation, ignoring higher-order terms, and solving the ordinary differential equation:

$$M\frac{du(f_0)}{df_0} + \frac{V}{2}\frac{d^2u(f_0)}{df_0{}^2} = 0$$

Under a diploid Wright-Fisher model accounting for drift and selection, the effective per generation mean and variance of the diffusion are $M = sp(1-p)$ and $V = p(1/p)/(2N)$, respectively (when heterozygote fitness is $1+s$)). With the appropriate boundary conditions defined above, the solution is then

$$u(f_0) = \frac{1 - e^{-4Nsf_0}}{1 - e^{-4Nsf^*}}, \tag{4}$$

which has an obvious similarity to the fixation probability in equation 1. Substituting in the establishment frequency from equation 3, we obtain the simple result:

$$P_{\text{Establish}}(f_0) = \left(\frac{1+k}{k}\right)\frac{1 - e^{-4Nsf_0}}{1 - e^{-4Ns}} \tag{5}$$

or simply $(1+k)/k \cdot P_{\text{Fix}}(f_0)$. While f_0 will usually be assumed to be $1/(2N)$, it need not be so and could reasonably take any value between $1/(2N)$ and $f^* - 1/(2N)$. This result also holds for dominant and recessive variants (not shown for brevity).

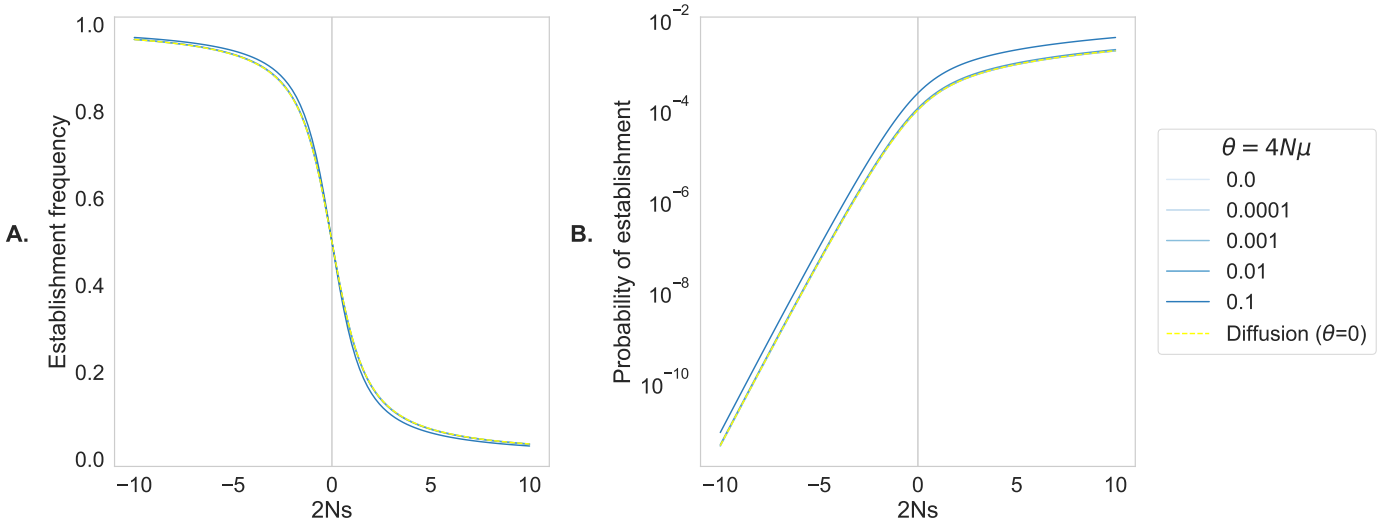

Figure 2 Comparison of establishment frequencies and probabilities for $k = 1$. As before, dashed yellow lines show analytic diffusion solutions, solid blue lines show discrete solutions. The diffusion approximation agrees with the discrete solution in this parameter range. (A) Establishment frequency as a function of selection. At neutrality, establishment frequency is $\frac{1}{2}$, increasing for negatively selected, and decreasing for positively selected alleles. (B) Probability of establishment as a function of selection (Y-axis on log scale). Establishment probability of a neutral variant is $\frac{1}{N}$, increasing for positively selected, and sharply decreasing for negatively selected alleles. Population size $N = 10,000$, heterozygote fitness $(1 + s)$

Establishment and fixation probabilities differ according to the definition of establishment

Establishment and fixation probabilities differ according to how certain we wish to be that establishment implies likely fixation (Figure 1). On the left, establishment is defined as exceeding the frequency at which eventual fixation and extinction are equally likely ($k = 1$). As explained in the introduction, this roughly corresponds to the assumptions of Haldane (although he considered only the case of $s > 0$ in a branching-process model). With $k = 1$, the probability of establishment is twice the probability of fixation, while with $k = 2$ (where fixation is twice as likely as extinction), the establishment probability is $1.5\times$ larger than that for fixation. On the right of Figure 1, establishment is defined such that fixation is extremely likely (100 times more likely than extinction, $k = 100$). This corresponds to Gillespie's definition of establishment with ϵ set to ~ 0.01 ($1/101$ to be precise). Comparison of a direct analysis of the discrete Wright-Fisher model with the diffusion results of the previous sections confirms the results in equation 5. Since the direct analysis of the discrete model is agnostic to details of the model itself, results for several values of the population mutation rate parameter, θ, were computed where both forward and backward recurrent mutation were allowed. The close correspondence of these results with the diffusion results (which assumed no new mutations) suggests that the relationship between P_{Fix} and $P_{\text{Establish}}$ is robust as defined in equation 5.

In Figure 2, establishment frequency and probability are contrasted under Haldane's definition of establishment ($k = 1$), using both the general diffusion solutions presented above, and the direct analysis of the discrete Wright-Fisher model. Establishment frequency is large and close to one for deleterious alleles, illustrating that strongly deleterious alleles are only expected to establish in an equilibrium population when they start at frequencies that are already very high. Contrariwise, beneficial alleles have comparatively low establishment frequencies, which decrease as the strength of positive selection is increased. For neutral mutations, the establishment frequency is 0.5, and the shape of the curve is symmetric with respect to the neutral establishment frequency. Only minor variations are observed when bidirectional recurrent mutation is introduced (where faster mutation generally exaggerates the effects of selection).

In the discrete framework, we can also easily find the expected time to establishment (for those variants that establish; Figure 3A), the expected time of segregation after establishment (Figure 3B), and the expected time to fixation after establishment (for those variants that go on to be fixed; Figure 3C). These quantities could also be found using diffusion theory for particular model parameterizations, which we do not pursue here for brevity. As expected, the establishment time for advantageous alleles is shortest (Figure 3A). However, it is interesting to note that mildly deleterious alleles take a significantly longer time to establish than do either neutral or strongly deleterious alleles. Similar effects have been observed in computations of expected allele age and times to absorption, and have been attributed to 'stochastic slowdowns' under weak selection in the presence of dominance (Mafessoni and Lachmann 2015; de Sanctis *et al.* 2017) and mutation (de Sanctis *et al.* 2017). Notably, we see these effects here even when the mutation rate is zero. Similarly, we observe that once established, weakly adaptive alleles take longer to go to fixation than do neutral alleles (Figure 3C). Contrariwise, the average time to absorption (at either boundary) after establishment is symmetric with respect to selection about $s = 0$ (Figure 3B).

Discussion

We have shown that establishment (survival until escaping loss by drift) and fixation (taking over the population) are easily differentiated in a Wright-Fisher model, and that clearly defining the establishment frequency and probability clarifies the meaning of Haldane's important result ($P_{\text{Fix}} \approx 2s, s > 0$). Despite that Haldane's fixation probability is widely interpreted as if it were

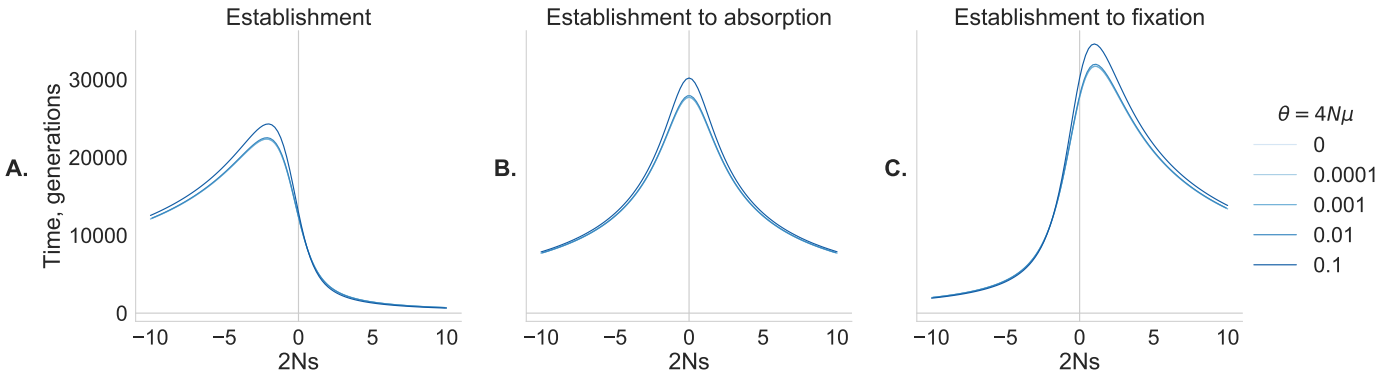

Figure 3 Expected times to establishment and absorption in the discrete Wright-Fisher model with establishment defined by $k = 1$. (A) Expected time to establishment (from $f_0 = 1/(2N)$); (B) Expected time to absorption, post-establishment (at either absorbing boundary); and (C) Expected time to fixation, post-establishment. Solid blue lines correspond to different values of the bidirectional mutation rate. Population size $N = 10,000$, heterozygote fitness $(1 + s)$.

the establishment probability, we have argued that the probability of reaching the establishment frequency does not appear to have been addressed until now (except in the trivial case where k is large and establishment therefore implies certain fixation). Accordingly, we do not believe Haldane intended his probability of fixation to be interpreted as the probability of reaching the establishment count. Indeed, it is plausible that his result has been misinterpreted owing to the use of terminology from the theory of branching processes that may be ambiguous when taken out of context. For example, Haldane called $2s$ the 'probability of survival', alluding to the so-called ultimate survival probability in branching process theory. While this sounds very much like survival to escape of loss by drift, which he also discussed, it really means fixation since it refers to the ultimate fate of an allele in the limit of infinite (or very long) time.

We recognize that there could be reasonable disagreement amongst geneticists about the definition of establishment in comparison to fixation. After all, historical usage of the term 'establishment' has been inconsistent, with some authors using it as a synonym for complete fixation (*e.g.*, Kimura and Ohta 1969) and others using it specifically to mean survival to escape from loss by drift (*e.g.*, Haldane 1927). Nevertheless, the modern usage is more consistent with the latter definition, which we have assumed throughout this work. Due to this ambiguous legacy, it would probably be helpful to avoid future colloquial uses of the term 'establishment', *sensu lato*, without explicitly defining what is intended.

Defining establishment with respect to the odds ratio of eventual fixation to extinction provides a flexible definition that generalizes previous approaches, and which allows the establishment probability to be simply computed as $(1 + k)/k$ times P_{Fix}. Under Haldane's assumptions, this corresponds to an establishment probability of $4s$ (when $P_{\text{Fix}} = 2s$; $k = 1$), and $3s$ when establishment is defined by fixation being twice as likely as extinction ($k = 2$). When heterozygote fitness is $1 + sh$ and $h = 1/2$, the establishment and fixation probabilities are half of these values (not shown). For larger values of k (*e.g.*, $k > 10$), $P_{\text{Establish}}$ is quite close to P_{Fix} and the concept itself becomes equivalent to fixation. We therefore suggest that $k = 1$ and $k = 2$ are reasonable choices for defining establishment in practice.

Acknowledgments

We thank Bianca de Sanctis for helpful comments and discussion. This work was supported by a Discovery Grant to APJdK from the Natural Sciences and Engineering Research Council of Canada (NSERC), and by an Alberta Innovates Technology Futures (AITF) Doctoral Fellowship to IK. The authors gratefully acknowledge infrastructure support from the Canada Foundation for Innovation, Alberta Advanced Education, and the Alberta Children's Hospital Research Institute.

Literature Cited

de Koning, A. P. J. and B. D. de Sanctis, 2018 The Rate of Molecular Evolution When Mutation May Not Be Weak. bioRxiv https://doi.org/10.1101/259507 .

de Sanctis, B., I. Krukov, and A. P. J. de Koning, 2017 Allele Age Under Non-Classical Assumptions is Clarified by an Exact Computational Markov Chain Approach. Scientific Reports **7**: 11869.

Ewens, W. J., 2004 *Mathematical Population Genetics: I. Theoretical Introduction.*, volume 27 of *Interdisciplinary Applied Mathematics*. Springer New York, New York, second edition, OCLC: 958522782.

Fisher, R. A., 1922 On the Dominance Ratio. Proceedings of the Royal Society of Edinburgh **42**: 321–241.

Gerrish, P. J. and R. E. Lenski, 1998 The fate of competing beneficial mutations in an asexual population. Genetica **102**: 127.

Gillespie, J. H., 2004 *Population Genetics, a Concise Guide*. John Hopkins University Press, Baltimore and London, second edition.

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

Kimura, M., 1962 On the Probability of Fixation of Mutant Genes in a Population. Genetics **47**: 713–719.

Kimura, M., 1964 Diffusion Models in Population Genetics. Journal of Applied Probability **1**: 177.

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.

Krukov, I., B. de Sanctis, and A. P. J. de Koning, 2017 Wright–Fisher exact solver (WFES): Scalable analysis of population genetic models without simulation or diffusion theory. Bioinformatics **33**: 1416–1417.

Mafessoni, F. and M. Lachmann, 2015 Selective Strolls: Fixation and Extinction in Diploids Are Slower for Weakly Selected Mutations Than for Neutral Ones. Genetics **201**: 1581–1589.

Messer, P. W. and D. A. Petrov, 2013 Population genomics of rapid adaptation by soft selective sweeps. Trends in Ecology & Evolution **28**: 659–669.

Otto, S. and M. Whitlock, 1997 The Probability of Fixation in Populations of Changing Size. Genetics **146**: 723–733.

Peck, J. R., 1994 A ruby in the rubbish: Beneficial mutations, deleterious mutations and the evolution of sex. Genetics **137**: 597–606.

Appendix

Establishment count in a Wright-Fisher Markov model (general case)

Here we derive establishment properties for arbitrary discrete-time Wright-Fisher models. Detailed discussion of efficient computational solutions to the discrete-time model can be found in Krukov *et al.* (2017); de Sanctis *et al.* (2017); de Koning and de Sanctis (2018).

Consider a discrete-time Wright-Fisher model with the transition probability matrix:

$$\mathbf{P}_{(i,j)} = \binom{2N}{j} \psi_i^j (1-\psi_i)^{2N-j} \tag{S1}$$

States $0, 2N \in \mathcal{A}$ are absorbing states, corresponding to extinction and fixation, respectively. The rest of the states are transient, $i \in \overline{\mathcal{A}}$. The transition probability matrix (equation S1) can be partitioned into transient-to-transient transition probabilities $\mathbf{Q}_{(i,j)} | i, j \in \overline{\mathcal{A}}$, and transient-to-absorbing transition probabilities $\mathbf{R}_{(i,k)} | i \in \overline{\mathcal{A}}, k \in \mathcal{A}$:

$$\mathbf{P} = \begin{pmatrix} \mathbf{Q} & \mathbf{R} \\ \mathbf{0} & \mathbf{I}_2 \end{pmatrix} \tag{S2}$$

In this case, $\mathbf{R} = [\mathbf{R}_{\text{Ext}}, \quad \mathbf{R}_{\text{Fix}}]$ has two columns, corresponding to extinction and fixation. The full model has $2N+1$ states, corresponding to allele counts from 0 to $2N$, inclusive. The fundamental matrix of the Markov chain is:

$$\mathbf{N}_{(i,j)} = \sum_{n=0}^{\infty} \mathbf{Q}^n = (\mathbf{I} - \mathbf{Q})^{-1} \tag{S3}$$

The entries $\mathbf{N}_{(i,j)}$ express the expected number of generations spent in each state j prior to absorption, conditional on starting in state i. The probability of absorbing in each absorbing state, conditional on starting in state i, is given by the matrix:

$$\mathbf{B} = \mathbf{N}\mathbf{R}, \tag{S4}$$

which has one column of probabilities for each of the absorbing states.

We can find the establishment count c^* directly by scanning $\mathbf{B}$ for increasing values of c_0 (the initial state), until we find the first entry of $\mathbf{B}$'s second column such that $B_{c_0,2} \geq k B_{c_0,1}$. Note that this does not require solving for every row of $\mathbf{N}$, since we can rearrange

$$\begin{aligned} \mathbf{B} &= \mathbf{N}\mathbf{R} \\ &= (\mathbf{I} - \mathbf{Q})^{-1}\mathbf{R} \\ \mathbf{R} &= (\mathbf{I} - \mathbf{Q})\mathbf{B}, \end{aligned} \tag{S5}$$

which can be easily computed using an LU decomposition of the matrix $\mathbf{I} - \mathbf{Q}$.

Probability of establishment in a Wright-Fisher Markov model (general case)

Given the establishment count c^*, we can define a new Markov model with c^* as the upper absorbing boundary, yielding a model with $c^* - 1 \times c^* - 1$ transient-to-transient state transitions. This reduced model is constructed by truncation of $\mathbf{Q}$, and by setting the second column of R to $\mathbf{R}_{2,i} = 1 - \mathbf{R}_{1,i} - \sum_{j=1}^{c^*-1} Q_{(i,j)}$.

With the full transition matrix $\mathbf{P}'$ thus defined, where we call the corresponding transient-to-transient submatrix $\mathbf{Q}'$, and the transient-to-absorbing submatrix $\mathbf{R}'$, we can find the properties of interest by using $\mathbf{N}' = (\mathbf{I} - \mathbf{Q}')$, as in equation S3. We can then use $\mathbf{N}'$ to derive properties such as probabilities and expected times using standard definitions (Krukov *et al.* 2017; de Sanctis *et al.* 2017). The matrix $\mathbf{Q}'$ (just as matrix $\mathbf{Q}$) can be based on any parameterization of the underlying model, including with arbitrary mutation, dominance, and selection.

To integrate quantities of interest over the likely distribution of starting states $\mathbf{c_0}$, which can become important when the population mutation rate is not small, we integrate over each state according to the probability of mutation creating $1, 2, 3, ...$ copies in a single generation, starting from zero mutant copies (*i.e.*, $P_{0,1}, P_{0,2}, ...$). As in de Sanctis *et al.* (2017) and de Koning and de Sanctis (2018), the summation can be truncated at x terms for x when $P_{0,x}$ falls below some small threshold (*e.g.*, 10^{-7}).