Fixation Probability in a Haploid-Diploid Population

Kazuhiro Bessho*,1 and Sarah P. Otto†,1

*JSPS Research Fellow, Department of Evolutionary Studies of Biosystems, School of Advanced Sciences, SOKENDAI (The Graduate University for Advanced Studies), Hayama, 240-0193, Japan, and †Department of Zoology, University of British Columbia, Vancouver, V6T 1Z4, Canada

ORCID ID: 0000-0003-3042-0818 (S.P.O.)

ABSTRACT Classical population genetic theory generally assumes either a fully haploid or fully diploid life cycle. However, many organisms exhibit more complex life cycles, with both free-living haploid and diploid stages. Here we ask what the probability of fixation is for selected alleles in organisms with haploid-diploid life cycles. We develop a genetic model that considers the population dynamics using both the Moran model and Wright–Fisher model. Applying a branching process approximation, we obtain an accurate fixation probability assuming that the population is large and the net effect of the mutation is beneficial. We also find the diffusion approximation for the fixation probability, which is accurate even in small populations and for deleterious alleles, as long as selection is weak. These fixation probabilities from branching process and diffusion approximations are similar when selection is weak for beneficial mutations that are not fully recessive. In many cases, particularly when one phase predominates, the fixation probability differs substantially for haploid-diploid organisms compared to either fully haploid or diploid species.

KEYWORDS fixation probability; Moran model; Wright–Fisher model; haploid-diploid life cycle; variance effective population size

CLASSICAL genetic theories generally assume either a fully haploid (haplont) or fully diploid (diplont) life cycle (*e.g.*, Crow and Kimura 1970). Organisms often exhibit more complex sexual life cycles, however, with both free-living haploid and diploid stages (haploid-diploid life cycle) (Mable and Otto 1998). For example, marine macroalgae generally exhibit an alternation of generations between a haploid-gametophyte stage and a diploid-sporophyte stage (Bell 1994, 1997). Furthermore, haploid and diploid life stages can differ substantially in morphology. For example, diploid-dominant species (*e.g.*, Laminariales) have large diploid sporophytes and small haploid gametophytes, with the reverse found in haploid-dominant species (*e.g.*, *Cutleriales multifida*) (Hirose 1975; Bell 1994, 1997).

Despite haploid-diploid life cycles being common in several groups of organisms, classical population genetic quantities,

such as the fixation probability and the effective population size, have not been determined for haploid-diploid organisms. The fixation probability is fundamental to predicting the rate of adaptation of a species with an alternation between free-living haploid and diploid generations. It is also important for predicting the ability of organisms to persist in a changing environment and to predict how patterns of molecular evolution depend on the life cycle. Indeed, a total of 5480 articles mention the probability of fixation in a genetic context [Google Scholar search, September 10, 2016: ("fixation probability" ‖ "probability of fixation") (gene‖genetic‖allele‖mutation‖mutant)], in contexts ranging from the interpretation of sequence data, to the prediction of quantitative trait variation, to the spread of disease resistance, *etc.*

In addition, the fixation probability is needed to model the evolution of the life cycle itself (*i.e.*, the relative proportion of haploidy *vs.* diploidy), when one accounts for several of the complexities seen in haploid-diploid species such as coexistence of haploids and diploids in a population (*e.g.*, Destombe *et al.* 1989; Hiraoka and Yoshida 2010) and differences in spatial distribution and dispersal ability (*e.g.*, Bell 1997). In a related article, we model the dynamics of a modifier allele controlling the life cycle (Mable and Otto 1998), explicitly tracking the demographic

and spatial dynamics of haploids and diploids (K. Bessho and S. P. Otto, unpublished data). This related work requires the fixation probability in haploid-diploid species, as derived in the current article.

Given the fundamental importance of the fixation probability in evolutionary theory, there is a surprising lack of stochastic analyses describing the fate of selected alleles with haploid-diploid life cycles, compared to the many analyses on exclusively haploid or diploid life cycles (*e.g.*, Crow and Kimura 1970; Ewens 2004). To fill this gap, we develop genetic models that consider the population dynamics for haploid and diploid generations, deriving the fixation probability for a selected allele in such species. These results are further evaluated by comparison to numerical simulations.

Model

To calculate the fixation probability, we first extended the classical population genetic models, the Moran model and the Wright–Fisher (WF) model, to account for free-living haploid and diploid phases. In these models, we track the haploid and diploid population dynamics and changes in allele frequencies under the assumption of a one-locus genetic model with two alleles (resident allele R and mutant allele M). We assume the following: (1) that the total population size N is fixed, (2) that there is an obligate alternation of generations between haploid and diploid phases, (3) that the species is monoecious with equal investment in male and female gametes, and (4) that mating is random.

Because the number of individuals with genotype (GT) is a Markov process over time in both Moran and WF models, we let $X_{(GT)}(t)$ represent the number at time t as a random variable, and we let $x_{(GT)}$ represent the realized number in a given generation. The index (GT) indicates the genotype of individuals or cells [(GT) = R or M for haploids and (GT) = RR, RM, or MM for diploids]. Summed over all genotypes, the population size is N ($x_R + x_M + x_{RR} + x_{RM} + x_{MM} = N$). Parameters and variables in our model are represented in Table 1.

Moran model

In the Moran model, at every time step, death, reproduction, fertilization, and recruitment occur. First, one individual is randomly eliminated from the population in proportion to its mortality rate $d_{(GT)}$. This individual is then replaced by a birth. At reproduction, diploids produce haploid spores at rate $\beta_{(GT)}$ [$0 < \beta_{(GT)}$], and haploids produce female haploid gametes at rate $\beta_{(GT)}/2$ [and male gametes at rate $\beta_{(GT)}/2$]. The female gametes can be fertilized with male gametes with a fertilization success of $f_{(GT)}$ [$0 < f_{(GT)} < 1$], where we assume that male gametes are not limiting. All haploid spores and fertilized diploid zygotes are candidate reproductive cells to replace the dead individual and are equally likely to be recruited.

Wright-Fisher (WF) model

We next consider a different type of population dynamics using the WF model. In this model, all diploids produce a number $w_{(GT)}$ [$0 < w_{(GT)}$] of spores, and haploids produce $w_{(GT)}/2$ female gametes [and $w_{(GT)}/2$ male gametes]. As before, the female gametes are fertilized with male gametes with a fertilization success rate of $f_{(GT)}$ [$0 < f_{(GT)} < 1$]. After the end of each time step, all parents are eliminated from the population, and the next generation is constructed by randomly sampling N individuals in proportion to the number of haploid spores and diploid zygotes produced by the previous generation (see Equation D1).

Individual-based simulations

To check the accuracy of the analytical results, we simulated the dynamics numerically. We started the system fixed on the R allele, assuming that the system has reached ecological and demographic equilibrium (*i.e.*, in the absence of selection, see Appendix A and Appendix D for details), with $\hat{H} = \hat{\rho}_H^{\text{Model}} N$ haploids and $\hat{D} = \hat{\rho}_D^{\text{Model}} N$ diploids (rounded to have integer numbers of individuals) and with $\hat{\rho}_H^{\text{Model}} = 1 - \hat{\rho}_D^{\text{Model}}$ giving the frequency of haploids in the population (see Equation A8b for the Moran model and Equation D2 for the WF model). Next, we mutated one resident allele at random and simulated until this allele was fixed or lost in the population. To estimate the fixation probability, we repeated the process and estimated the probability from the resulting frequency of fixation.

Data availability

The analysis, numerical simulations, and scripts to generate the original figures were coded in Wolfram Mathematica 10 (Supplemental Material, File S1), available for download from the Dryad Digital Repository (http://doi.org/10.5061/dryad.n4k76).

Results

Fixation probability

Fixation probabilities in the Moran model using a branching process approximation: We first develop a branching process approximation (*e.g.*, Haldane 1927; Yeaman and Otto 2011) for haploid-diploid life cycle species. Specifically, we calculate the fixation probability of a single allele that appears in the haploid phase (genotype M), π_H^{Moran}, and the fixation probability for a single allele that appears in a heterozygous diploid (genotype RM), π_D^{Moran}, after which we calculate the average probability of fixation for an allele that appears randomly in any individual. When using a branching process approach, the fates of mutations in different individuals are assumed to be independent of one another, which implicitly requires that the population size is large and that mutant alleles are lost or established while rare. Under these assumptions,

Table 1 The variables and parameters in the Moran and WF models

Symbol	Interpretation
π_H^{Model}, (π_D^{Model})	Fixation probability in the branching process approximation when the single mutant arises in haploids (diploids); (1) for the Moran model and (D4) for the WF model
$u_{\text{branch}}^{\text{Model}}$	Average fixation probability from the branching process approximation; (2) for the Moran model and (5a) for the WF model
$u_{\text{diffusion}}^{\text{Model}}$	Fixation probability from the diffusion approximation; (4) with (3) for the Moran model and (7) for the WF model
p_0	Initial allele frequency; $p_0 = 1/\left\{ N\left(\hat{\rho}_H^{\text{Model}} + 2\hat{\rho}_D^{\text{Model}}\right)\right\}$
N	Total population size
$\hat{H}$, $(\hat{D})$	Number of haploids (diploids) at the population equilibrium; $\hat{H} = \hat{\rho}_H^{\text{Model}} N$ and $\hat{D} = \hat{\rho}_D^{\text{Model}} N$
$\hat{\rho}_H^{\text{Model}}$, $(\hat{\rho}_D^{\text{Model}})$	Equilibrium fraction of haploids (diploids) in the resident population; (A8) for the Moran model and (D2) for the WF model
$d_{(GT)}$	Mortality rate of an individual with genotype (GT) in the Moran model
$l_{(GT)}$	Expected life span of an individual in the Moran model; $l_{(GT)} = 1/d_{(GT)}$
$\beta_{(GT)}$	Fertility rate of an individual with genotype (GT) in the Moran model
$w_{(GT)}$	Fitness of an individual with genotype (GT) in the WF model
$f_{(GT)}$	Fertilization success rate of an individual with genotype (GT)
$\bar{d}$	Arithmetic mean mortality rate; $\bar{d} = (d_R + d_{RR})/2$
$\hat{d}$	Harmonic mean mortality rate; $\hat{d} = 2/\left(\frac{1}{d_R} + \frac{1}{d_{RR}}\right)$
$\check{d}$	Weighted mean mortality rate; $\check{d} = d_R \hat{\rho}_H + d_{RR} \hat{\rho}_D$
$\hat{\beta}$, $(\hat{w})$	Weighted mean fertility rate in the Moran model, $\frac{f_R \beta_R}{2}\hat{\rho}_H^{\text{Moran}} + \beta_{RR}\hat{\rho}_D^{\text{Moran}}$ $\left(\hat{w} = \frac{f_R w_R}{2}\hat{\rho}_H^{\text{WF}} + w_{RR}\hat{\rho}_D^{\text{WF}}\right.$ in the WF model$\left.\right)$
$\bar{f}$	Average fertilization success rate for rare mutant gametes; $\bar{f} = (f_R + f_M)/2$
$\phi_{(GT)}$	Expected reproductive rate of an individual with genotype (GT); $\phi_{(GT)} = \beta_{(GT)}/d_{(GT)}$
$\tilde{\phi}_{(GT)}$	Expected reproductive rate of a haploid individual with genotype (GT) considering the cost of sex; $\tilde{\phi}_M = \bar{f}\phi_M/2$ and $\tilde{\phi}_R = f_R\phi_R/2$
$s_{(GT)}^\theta$	Selection coefficient of the parameters $\theta \in \{f, \beta, d, w\}$
h^θ	Dominance coefficient of the parameters $\theta \in \{f, \beta, d, w\}$
$s_{\text{ave}}^{\text{Model}}$	Average selection coefficient; $s_{\text{ave}}^{\text{Moran}} = \frac{1}{2}\left(\frac{s_M^f}{2} + s_M^\beta + s_M^d\right) + \frac{1}{2}(s_{RM}^\beta + s_{RM}^d)$ for the Moran model, and $s_{\text{ave}}^{\text{WF}} = \frac{1}{2}\left(\frac{s_M^f}{2} + s_M^w\right) + \frac{1}{2}s_{RM}^w$ for the WF model
N_e^{Model}	Effective population size relative to the diploid-only WF model; (8)
s_e^{Model}	Effective selection coefficient; (9a) for the Moran model and (10a) for the WF model
h_e^{Model}	Effective dominance coefficient; (9b) for the Moran model and (10b) for the WF model
$x_{(GT)}$	Number of individual with genotype (GT)
$c_{(GT)}$	Number of haploid spores (A2) and diploid zygotes (A3) with genotype (GT)
$\tilde{d}_{(GT)}$	Probability that an individual with genotype (GT) is eliminated from the population; (A1)
$q_{(GT)}$	Probability that genotype (GT) is recruited into the empty site; (A4)
m^{Model}, (v^{Model})	First (second) moment of change in allele frequency; (3) for the Moran model and (7) for the WF model
$X_{(GT)}(t)$	Random variable describing the number of individual with genotype (GT) at time t
p_M	Average allele frequency weighted appropriately so that the dynamics of weak selection depend only on the allele frequency; $p_M = \frac{l_R}{l_R + l_{RR}}p_H + \frac{l_{RR}}{l_R + l_{RR}}p_D$ for the Moran model, and $p_M = \frac{1}{2}p_H + \frac{1}{2}p_D$ for the WF model
p_H, (p_D)	Allele frequency in haploids (diploids); $p_H = \frac{x_M}{x_R + x_M}$ and $p_D = \frac{(x_{RM}/2) + x_{MM}}{x_{RR} + x_{RM} + x_{MM}}$
δ_p	Difference in allele frequencies between haploids and diploids; $\delta_p = p_D - p_H$
ρ_H	Frequency of haploids in the entire population; $\rho_H = (x_R + x_M)/N$
F_D	Departure from Hardy–Weinberg in diploids; $F_D = 1 - \frac{1}{2p_D(1-p_D)}\frac{x_{RM}}{x_{RR} + x_{RM} + x_{MM}}$

we obtain approximations for the fixation probabilities (Appendix B):

$$\pi_H^{\text{Moran}} = \frac{\tilde{\phi}_M \phi_{RM} - \tilde{\phi}_R \phi_{RR}}{\tilde{\phi}_M \phi_{RM} + \frac{\phi_{RM}}{2}\sqrt{\tilde{\phi}_R \phi_{RR}}}, \quad (1a)$$

$$\pi_D^{\text{Moran}} = \frac{\tilde{\phi}_M \phi_{RM} - \tilde{\phi}_R \phi_{RR}}{\tilde{\phi}_M \phi_{RM} + 2\tilde{\phi}_M \sqrt{\tilde{\phi}_R \phi_{RR}}}, \quad (1b)$$

where we define the expected reproductive rate $\phi_{(GT)} = \beta_{(GT)}/d_{(GT)}$, the average fertilization rate for mutant gametes $\bar{f} = (f_R + f_M)/2$ (averaged over cases where the mu-

tant is carried by the sperm and when it is carried by the egg), the expected reproductive rate of a haploid mutant $\tilde{\phi}_M = \bar{f}\phi_M/2$, and the expected reproductive rate of a haploid resident $\tilde{\phi}_R = f_R\phi_R/2$. Observe that the expected reproductive rate of haploids is scaled by the average value of $f_{(GT)}$ for the combining gametes divided by two, which accounts for the costs of sex (specifically for the possibility that not all female gametes are fertilized and for the fact that haploids must invest half of their reproductive resources in male gametes, which are unable to reproduce on their own). Equations 1a and 1b apply as long as the fixation probabilities π_H and π_D are positive, which occurs when the average mutant fitness is greater than the resident, where the relevant average across the

haploid-diploid life cycle requires $\bar{f}\phi_M\phi_{RM} > f_R\phi_R\phi_{RR}$. This condition agrees with the invasion condition in a deterministic model (K. Bessho and S. P. Otto, unpublished data). When haploids and diploids have similar reproductive rates ($\tilde{\phi}_M \approx \phi_{RM}$), the fixation probability of a mutation appearing first in a haploid tends to be larger than a mutation appearing first in a diploid ($\pi_H^{\text{Moran}} > \pi_D^{\text{Moran}}$), because all of the haploid mutant offspring bear the mutant allele, whereas only half of the offspring of an RM diploid do. As the cost of sex rises, however ($\tilde{\phi}_M$ declines relative to ϕ_{RM}), the probability of fixation becomes increasingly more likely when the mutant first appears in a diploid, thereby delaying this cost.

The average fixation probability of a new mutation that arises uniformly at random among the individuals present in the resident population, $u_{\text{branch}}^{\text{Moran}}$, can then be defined as:

$$u_{\text{branch}}^{\text{Moran}} = \frac{\hat{\rho}_H^{\text{Moran}}}{\hat{\rho}_H^{\text{Moran}} + 2\hat{\rho}_D^{\text{Moran}}}\pi_H^{\text{Moran}} + \frac{2\hat{\rho}_D^{\text{Moran}}}{\hat{\rho}_H^{\text{Moran}} + 2\hat{\rho}_D^{\text{Moran}}}\pi_D^{\text{Moran}},$$
$$(2)$$

where $\hat{\rho}_H^{\text{Moran}} = \frac{\sqrt{\beta_{RR}d_{RR}}}{\sqrt{f_R\beta_R d_R/2} + \sqrt{\beta_{RR}d_{RR}}}$ is the equilibrium proportion of haploids in the resident population (A8b) and where the "2" accounts for the fact that diploids carry two alleles that could have potentially mutated. Equation 2 allows the fixation probability to be calculated for arbitrarily strong selection, but it will be accurate only as long as the fixation probabilities (1) are not negative.

Fixation probabilities in the Moran model using a diffusion approximation: Next, we derive the fixation probability of a selected mutant allele in a haploid-diploid population using Kolmogorov's backward equation, given the first and second moments for the fluctuations in allele frequencies. This diffusion equation provides an adequate approximation for the fixation probability even for small populations and deleterious mutations (*e.g.*, Kimura 1957), unlike a branching process approximation. Rather than analyzing the genotype frequencies, $x_{(GT)}$, we transform the model into a new set of variables that allow us to perform a separation of timescales, assuming that selection occurs on a slower timescale and that ecological and demographic processes rapidly bring the system to a "quasi-equilibrium" (see Appendix C). This approach allows the dynamics of the mutant allele to be described by moments that depend only on the parameters and the current allele frequency (Appendix C).

Specifically, we let the allele frequencies in haploids and diploids equal $p_H = x_M/(x_R + x_M)$ and $p_D = (x_{RM}/2 + x_{MM})/(x_{RR} + x_{RM} + x_{MM})$, respectively, and focus on the following new variables: the difference in allele frequencies between haploids and diploids $\delta_p = p_D - p_H$, the frequency of haploids in the population $\rho_H = (x_R + x_M)/N$, the departure from the Hardy–Weinberg principle in diploids $F_D = 1 - \frac{1}{2p_D(1-p_D)}\frac{x_{RM}}{x_{RR} + x_{RM} + x_{MM}}$, and the allele frequency averaged across ploidy levels and weighted by the expected

life span ($l_{(GT)} = 1/d_{(GT)}$) of haploids and diploids $p_M = \frac{l_R}{l_R + l_{RR}}p_H + \frac{l_{RR}}{l_R + l_{RR}}p_D$. Only with this weighted allele frequency is it possible to separate the timescales so that the leading-order change in p_M is independent of the other variables (Appendix C). This averaging also corresponds to weighting by the reproductive value of each class (Taylor 1990).

To perform the separation of timescales, we assume selection is weak by first defining the following selection [$s_{(GT)}$] and dominance (h) coefficients: $f_M = f_R(1 + s_M^f)$, $\beta_M = \beta_R(1 + s_M^\beta)$, $\beta_{RM} = \beta_{RR}(1 + s_{RM}^\beta)$, $\beta_{MM} = \beta_{RR}(1 + s_{MM}^\beta)$, $d_M = d_R(1 - s_M^d)$, $d_{RM} = d_{RR}(1 - s_{RM}^d)$, $d_{MM} = d_{RR}(1 - s_{MM}^d)$, $h^\beta = s_{RM}^\beta/s_{MM}^\beta$, and $h^d = s_{RM}^d/s_{MM}^d$. We then assume weak selection by setting $s_{(GT)}^\theta = \varepsilon\tilde{s}_{(GT)}^\theta$ ($\theta \in \{f, \beta, d\}$) for each life history parameter, where ε is a small term.

The fast-scale dynamics can be tracked by determining changes in the system to leading order ($\varepsilon \approx 0$), assuming that selection alters these dynamics only over a longer period of time. Doing so, we find that the system rapidly approaches a quasi-equilibrium state where there is a similar allele frequency in haploid and diploid populations ($\delta_p \approx 0$), the fraction of haploids is approximately $\rho_H \approx \hat{\rho}_H^{\text{Moran}}$, the population is nearly at Hardy–Weinberg ($F_D \approx 0$), and the allele frequency does not change ($\Delta p_M \approx 0$) (see Appendix C and File S1). We then calculate the average change in allele frequency to leading order, which requires that we keep linear-order terms in ε and which accounts for selection (Appendix C and File S1).

Furthermore to leading order, the second and higher moments of the change in allele frequency are the same as for the neutral case (*i.e.*, when R and M are equivalent, dropping all terms involving ε). Consequently, we derive the second moment for the change in allele frequency using the diffusion limit ($N \to \infty$) when selection is absent (Appendix C and File S1), which allows us to determine the variance effective population size for this model. We also confirmed that the third moment goes to zero in the diffusion limit for the Moran model, justifying the use of a diffusion approximation (Karlin and Taylor 1981, p. 165).

Using the first and second moments for the expected change in the frequency of mutant alleles, m^{Moran} and v^{Moran}, we can calculate the backward equation for the fixation probability with an initial allele frequency p_0, $u_{\text{diffusion}}^{\text{Moran}}(p_0)$ (Appendix C):

$$v^{\text{Moran}}(p_0)\frac{d^2 u_{\text{diffusion}}^{\text{Moran}}}{dp_0^2} + m^{\text{Moran}}(p_0)\frac{du_{\text{diffusion}}^{\text{Moran}}}{dp_0} = 0, \quad (3a)$$

where

$$m^{\text{Moran}}(p_M) = \frac{Np_M(1-p_M)\left\{2s_{ave}^{\text{Moran}} + p_M\left[\left(1 - 2h^\beta\right)s_{MM}^\beta + \left(1 - 2h^d\right)s_{MM}^d\right]\right\}}{2},$$
$$(3b)$$

$$v^{\text{Moran}}(p_M) = \frac{p_M(1-p_M)\left[(d_R/\bar{d})\hat{\rho}_H^{\text{Moran}} + 2(d_{RR}/\bar{d})\hat{\rho}_D^{\text{Moran}}\right]}{4\hat{\rho}_H^{\text{Moran}}\hat{\rho}_D^{\text{Moran}}},$$
$$(3c)$$

and $s_{ave}^{Moran} = \frac{1}{2}\left(\frac{s_M^f}{2} + s_M^\beta + s_M^d\right) + \frac{1}{2}(s_{RM}^\beta + s_{RM}^d)$ is the average selection experienced by a rare allele across haploid and diploid stages. Here, we scale time as described in Appendix C and define the arithmetic mean mortality rate as $\bar{d} = (d_R + d_{RR})/2$. Under the diffusion approximation, selection is assumed to be weak and of order $O(1/N)$, in which case both m^{Moran} and v^{Moran} are of $O(1)$ when we take the diffusion limit ($N \to \infty$). Solving Equations 3a–3c with the boundary conditions, $u_{diffusion}^{Model}(0) = 0$ and $u_{diffusion}^{Model}(1) = 1$, we then obtain the fixation probability for any diffusion process (Crow and Kimura 1970, p. 424):

$$u_{diffusion}^{Model}(p_0) = \frac{\int_0^{p_0} \exp\left[-2Q(p')\right]dp'}{\int_0^1 \exp\left[-2Q(p')\right]dp'}, \quad (4a)$$

with $Q(p) = \int \frac{m^{Model}(p)}{v^{Model}(p)} dp$. When a mutation arises at random on a single chromosome within the population, the expected initial allele frequency is:

$$
\begin{aligned}
p_0 &= \frac{l_R}{l_R + l_{RR}} \frac{1}{\hat{H}} \frac{\hat{\rho}_H^{Moran}}{\hat{\rho}_H^{Moran} + 2\hat{\rho}_D^{Moran}} \\
&\quad + \frac{l_{RR}}{l_R + l_{RR}} \frac{1}{2\hat{D}} \frac{2\hat{\rho}_D^{Moran}}{\hat{\rho}_H^{Moran} + 2\hat{\rho}_D^{Moran}} \\
&= \frac{1}{N(\hat{\rho}_H^{Moran} + 2\hat{\rho}_D^{Moran})}, \quad (4b)
\end{aligned}
$$

again weighting the allele frequency in each ploidy phase by its longevity.

Fixation probabilities in the WF model using a branching process approximation: We next develop the fixation probabilities u_{branch}^{WF} and $u_{diffusion}^{WF}$ using the WF model, again applying both branching process and diffusion methods (Appendix D).

The branching process approximation yields a pair of coupled nonlinear equations in π_H^{WF} and π_D^{WF} (Equations D4a and D4b), which can be solved numerically to obtain the fixation probabilities for an allele appearing at random within the population when selection is strong:

$$u_{branch}^{WF} = \frac{\hat{\rho}_H^{WF}}{\hat{\rho}_H^{WF} + 2\hat{\rho}_D^{WF}} \pi_H^{WF} + \frac{2\hat{\rho}_D^{WF}}{\hat{\rho}_H^{WF} + 2\hat{\rho}_D^{WF}} \pi_D^{WF}, \quad (5a)$$

where now the fraction of haploids in the resident population is given by $\hat{\rho}_H^{WF} = \frac{\sqrt{w_{RR}}}{\sqrt{\tilde{w}_R} + \sqrt{w_{RR}}}$ (see Equation D2). For compactness, we have written the fitness of a resident haploid considering the cost of fertilization and sex as $\tilde{w}_R = f_R w_R/2$.

We can also approximate the fixation probabilities π_H^{WF} and π_D^{WF} explicitly when selection is weak. To do so, we define $w_M = w_R(1 + s_M^w)$, $w_{RM} = w_{RR}(1 + s_{RM}^w)$, $w_{MM} = w_{RR}(1 + s_{MM}^w)$, and $h^w = s_{RM}^w/s_{MM}^w$, where again we assume weak selection by setting $s_{(GT)}^w = \varepsilon \tilde{s}_{(GT)}^w$ for small ε. To leading order in ε, the fixation probabilities for an allele initially in a haploid individual or in a diploid individual become:

$$\pi_H^{WF} \approx \frac{2\tilde{w}_R w_{RR}}{\tilde{w}_R w_{RR} + \frac{w_{RR}}{2}\sqrt{\tilde{w}_R w_{RR}}} 2s_{ave}^{WF} = \frac{4\hat{\rho}_D^{WF}}{\hat{\rho}_H^{WF} + 2\hat{\rho}_D^{WF}} 2s_{ave}^{WF}, \quad (5b)$$

$$\pi_D^{WF} \approx \frac{2\tilde{w}_R w_{RR}}{\tilde{w}_R w_{RR} + 2\tilde{w}_R \sqrt{\tilde{w}_R w_{RR}}} 2s_{ave}^{WF} = \frac{2\hat{\rho}_H^{WF}}{\hat{\rho}_H^{WF} + 2\hat{\rho}_D^{WF}} 2s_{ave}^{WF}, \quad (5c)$$

where the average selective effect for a rare allele in the WF model is defined as $s_{ave}^{WF} = \frac{1}{2}\left(\frac{s_M^f}{2} + s_M^w\right) + \frac{1}{2}s_{RM}^w$. Plugging Equation 5b and (5c) into (5a), we have

$$u_{branch}^{WF} \approx \frac{8\hat{\rho}_H^{WF}\hat{\rho}_D^{WF}}{\left(\hat{\rho}_H^{WF} + 2\hat{\rho}_D^{WF}\right)^2} 2s_{ave}^{WF}. \quad (6)$$

This branching process approximation is valid only for mutations whose average effect is beneficial ($s_{ave}^{WF} > 0$).

Fixation probabilities in the WF model using a diffusion approximation: We also derive the fixation probability using a diffusion approach, $u_{diffusion}^{WF}$, finding the appropriate drift and diffusion coefficients for use in Equation 4a. Again, we sought a transformation of variables that allowed a separation of timescales, such that the leading-order term in selection depended only on the parameters and the allele frequency. For the WF model, where all parents die each generation, the appropriate definition of the average mutant allele frequency is $p_M = (p_H + p_D)/2$ (Appendix D). Applying the separation of timescales, we obtain the drift and diffusion coefficients:

$$m^{WF}(p_M) = \frac{Np_M(1 - p_M)\left[2s_{ave}^{WF} + p_M(1 - 2h^w)s_{MM}^w\right]}{2}, \quad (7a)$$

$$v^{WF}(p_M) = \frac{p_M(1 - p_M)\left(\hat{\rho}_H^{WF} + 2\hat{\rho}_D^{WF}\right)}{8\hat{\rho}_H^{WF}\hat{\rho}_D^{WF}}, \quad (7b)$$

where time is measured in units of N generations (see Appendix D for the derivation). Equation 4a then gives the fixation probability, with the initial allele frequency $p_0 = 1/[N(\rho_H^{WF} + 2\rho_D^{WF})]$ for a mutation appearing at random on a single chromosome within the population. In contrast to the branching process approximation, this fixation probability can be calculated for both beneficial ($s_{ave}^{WF} > 0$) and deleterious ($s_{ave}^{WF} < 0$) mutations, but assumes that this selection is weak.

Effective genetic parameters: To compare the fixation probabilities derived above to classical results for fully haploid or fully diploid populations, we define the variance effective population size (N_e^{Model}), selection coefficient (s_e^{Model}), and dominance coefficient (h_e^{Model}) that would give the same fixation probability obtained by a
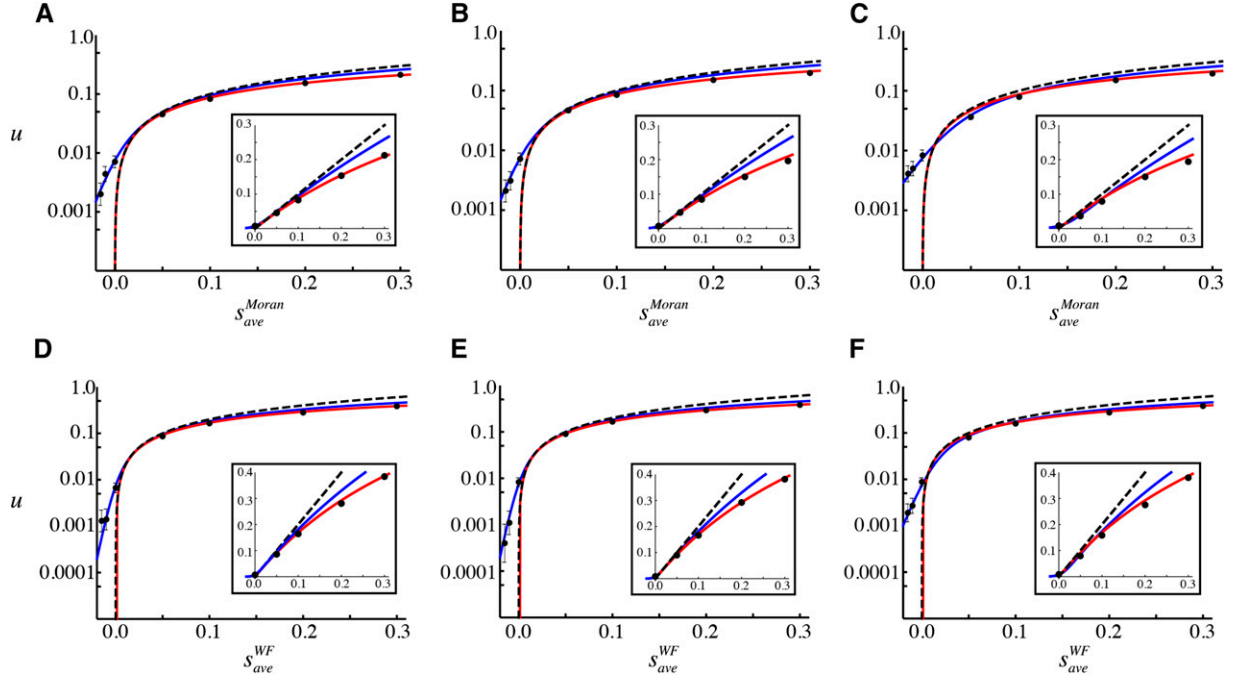
Figure 1 The fixation probability on a log scale for different values of s_{ave}^{Model} from the (A–C) Moran model and (D–F) WF model, allowing for different selection coefficients in (A and D) haploids or selection and dominance coefficients in diploids (remaining panels). Inset panels present the same results on a linear scale when $s_{ave}^{Model} \geq 0$. Curves give the branching process results allowing for strong selection [red; Moran from (2) with (1); WF from (5a) with (D4)], the weak-selection approximation to the branching process [black dashed; Moran from (13b); WF from (13a)], and the diffusion approximation [blue; Moran from (4a) with (3); WF from (4a) with (7)]. ● indicates the fixation probability estimated from 10,000 numerical simulations, with intervals indicating the Wilson score 95% C.I. Parameters: (A) $s_M^\beta = 2s_{ave}^{Moran}$ and $s_{RM}^\beta = s_{MM}^\beta = 0$; (B) $s_M^\beta = 0$, $s_{RM}^\beta = 2s_{ave}^{Moran}$, and $h^\beta = 0.5$; (C) $s_M^\beta = 0$, $s_{RM}^\beta = 2s_{ave}^{Moran}$, and $h^\beta = 1$; (D) $s_M^w = 2s_{ave}^{WF}$, and $s_{RM}^w = 0$; (E) $s_M^w = 0$, $s_{RM}^w = 2s_{ave}^{WF}$, and $h^w = 0.5$; (F) $s_M^w = 0$, $s_{RM}^w = 2s_{ave}^{WF}$, and $h^w = 1$. Other parameters: $N = 100$, $f_R = 0.5$, $\beta_R = \beta_{RR} = 1000$, $d_R = d_{RR} = 0.005$, $w_R = w_{RR} = 1000$, and $s_M^f = s_M^d = s_{RM}^d = s_{MM}^d = 0$, such that the fraction of haploids in the resident population is $\hat{\rho}_H^{Model} = 2/3$.

diffusion approximation in the diploid WF model, a classical standard for comparison.

We first define the variance effective population size using the number of diploid individuals in a WF model that would exhibit the same allele frequency variation for a haploid-diploid species by defining $N_e^{Model}/N = p_M(1 - p_M)/(2v_p^{Model})$ (Ewens 2004). Note that N_e would be the census size N in the diploid-only WF model but $N/2$ in the haploid-only WF model, because we use the diploid WF model to define the effective parameters. From the diffusion approximations above, we find:

$$N_e^{Moran} = \frac{2\hat{\rho}_H^{Moran}\hat{\rho}_D^{Moran}}{(d_R/\bar{d})\hat{\rho}_H^{Moran} + 2(d_{RR}/\bar{d})\hat{\rho}_D^{Moran}} N, \quad (8a)$$

$$N_e^{WF} = \frac{4\hat{\rho}_H^{WF}\hat{\rho}_D^{WF}}{\hat{\rho}_H^{WF} + 2\hat{\rho}_D^{WF}} N. \quad (8b)$$

Note that the variance effective population size is substantially reduced in populations where one of the phases is rare ($\hat{\rho}_H^{Model} \approx 0$ or $\hat{\rho}_D^{Model} \approx 0$), reflecting the increased amount of drift that occurs in that phase in species with a strict alternation of generations. When the death rates are equal ($d_R = d_{RR} = \bar{d}$), N_e^{Moran} is half N_e^{WF}, reflecting the higher variance in reproductive success

that is typically observed in the Moran model compared to the WF model (*e.g.*, Otto and Day 2007, p. 583 and p. 643).

We next consider the effective selection and dominance coefficients. To do so, we define s_e and h_e as the selection coefficient and the degree of dominance in the classic diploid WF model, in which the drift term for the change in allele frequency is $\Delta p_M = s_e p_M(1 - p_M)[h_e + (1 - 2h_e)p_M]$ (see Crow and Kimura 1970). Note that the drift term obtained for haploid-diploid populations per generation (m^{Model}/N) has the same functional form with respect to p_M (see Equation 3b and Equation 7a). Consequently, we can find s_e and h_e from the pair of equations describing the changes in allele frequency when $p_M = 0$ ($\Delta p_M/[p_M(1 - p_M)] = s_e h_e$) and when $p_M = 1$ ($\Delta p_M/[p_M(1 - p_M)] = s_e(1 - h_e)$), obtaining:

$$s_e^{Moran} = 2s_{ave}^{Moran} + \frac{\left[\left(1 - 2h^\beta\right)s_{MM}^\beta + \left(1 - 2h^d\right)s_{MM}^d\right]}{2}, \quad (9a)$$

$$h_e^{Moran} = \frac{2s_{ave}^{Moran}}{4s_{ave}^{Moran} + \left[\left(1 - 2h^\beta\right)s_{MM}^\beta + \left(1 - 2h^d\right)s_{MM}^d\right]}, \quad (9b)$$
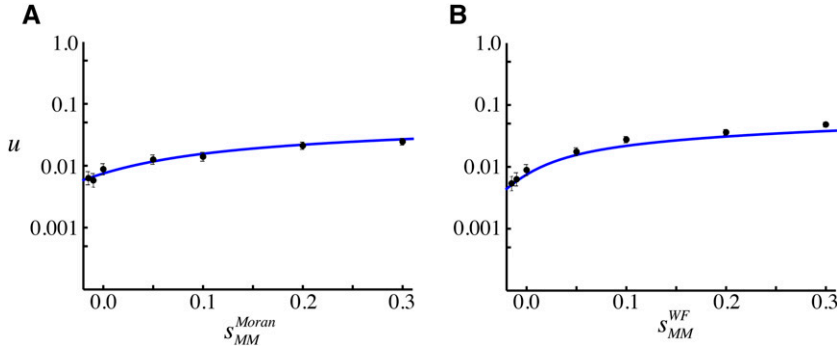
for the Moran model, and

Figure 2 The fixation probability of a fully recessive allele in the (A) Moran model and (B) WF model for different values of s_{MM}^{Model}. Other parameters: $N = 100$, $f_R = 0.5$, $\beta_R = \beta_{RR} = 1000$, $d_R = d_{RR} = 0.005$, $w_R = w_{RR} = 1000$, and $s_M^\beta = s_{RM}^\beta = s_M^f = s_M^d = s_{RM}^d = s_{MM}^d = 0$. See Figure 1 for further information.

$$s_e^{\text{WF}} = 2s_{\text{ave}}^{\text{WF}} + \frac{(1 - 2h^w)s_{MM}^w}{2}, \tag{10a}$$

$$h_e^{\text{WF}} = \frac{2s_{\text{ave}}^{\text{WF}}}{4s_{\text{ave}}^{\text{WF}} + (1 - 2h^w)s_{MM}^w}, \tag{10b}$$

for the WF model. For an additive beneficial allele ($h^\beta = h^d = h^w = 1/2$), we have $s_e^{\text{Moran}} = 2s_{\text{ave}}^{\text{Moran}}$, $s_e^{\text{WF}} = 2s_{\text{ave}}^{\text{WF}}$, and $h_e^{\text{Moran}} = h_e^{\text{WF}} = 1/2$.

Given the above definitions for the variance effective population size (N_e), the effective selection coefficient (s_e), and the effective dominance coefficient (h_e), we regain the fixation probabilities derived above, $u_{\text{diffusion}}^{\text{Model}}$, from the fixation probability in the classic diploid WF model (Kimura 1957, 1962; Crow and Kimura 1970, p. 427):

$$u^*(p_0) = \frac{\int_0^{p_0} exp[-2N_e s_e \{(2h_e - 1)p'(1 - p') + p'\}]dp'}{\int_0^1 exp[-2N_e s_e \{(2h_e - 1)p'(1 - p') + p'\}]dp'}. \tag{11}$$

For arbitrary dominance, the result must be numerically integrated, but for additive mutations ($h^\beta = h^d = h^w = 1/2$), we have:

$$u_{\text{diffusion}}^{\text{Model}}(p_0) = \frac{\int_0^{p_0} exp[-2N_e^{\text{Model}} s_e^{\text{Model}} p']dp'}{\int_0^1 exp[-2N_e^{\text{Model}} s_e^{\text{Model}} p']dp'}$$
$$= \frac{1 - exp[-2s_e^{\text{Model}} N_e^{\text{Model}} p_0]}{1 - exp[-2s_e^{\text{Model}} N_e^{\text{Model}}]}. \tag{12}$$

Assuming weak positive selection on an initially rare mutation ($s_e^{\text{Model}} N_e^{\text{Model}} p_0 \approx 0$) in a large population ($s_e^{\text{Model}} N_e^{\text{Model}} \gg 1$), we obtain the classic approximation to the fixation probability, $u_{\text{diffusion}}^{\text{Model}}(p_0) \approx 2s_e^{\text{Model}} N_e^{\text{Model}} p_0$. The result is the same as that obtained using branching processes if we assume weak selection, as given by Equation 6 for the WF model:

$$u_{\text{branch}}^{\text{WF}} \approx u_{\text{diffusion}}^{\text{WF}} \approx \frac{8\hat\rho_H^{\text{WF}} \hat\rho_D^{\text{WF}}}{\left(\hat\rho_H^{\text{WF}} + 2\hat\rho_D^{\text{WF}}\right)^2} 2s_{\text{ave}}^{\text{WF}}, \tag{13a}$$

and by taking the Taylor series of Equation 2 to leading order in selection (ε) for the Moran model:

$$u_{\text{branch}}^{\text{Moran}} \approx u_{\text{diffusion}}^{\text{Moran}}$$

$$\approx \frac{8\hat\rho_H^{\text{Moran}} \hat\rho_D^{\text{Moran}}}{\left(\hat\rho_H^{\text{Moran}} + 2\hat\rho_D^{\text{Moran}}\right)\left[(d_R/\bar d)\hat\rho_H^{\text{Moran}} + 2(d_{RR}/\bar d)\hat\rho_D^{\text{Moran}}\right]} s_{\text{ave}}^{\text{Moran}}. \tag{13b}$$

We next turn to numerical analyses of these equations to better understand how the fixation probability depends on life history parameters of haploids and diploids.

Comparisons to numerical simulations

Similar haploid and diploid life histories: We first consider the accuracy of these analytical approximations by comparing the fixation probability to numerical simulations when the life history parameters are equivalent between haploid and diploid residents ($\beta_R = \beta_{RR}$ and $d_R = d_{RR}$ for the Moran model, $w_R = w_{RR}$ for the WF model). This scenario may be appropriate for isomorphic species with morphologically similar haploid and diploid stages (*e.g.*, the green alga *Ulva*, the brown alga *Dictyota*, the red alga *Chondrus, etc.*).

In Figure 1, we illustrate the fixation probability for the Moran model (A–C) and WF model (D–F) as a function of the average selection experienced by a rare allele, $s_{\text{ave}}^{\text{Model}}$, where selection acts only in the haploid phase (Figure 1, A and D; s_M^β) or in the diploid phase (remaining panels; s_{MM}^β and h^β), assuming for simplicity that selection acts only through fertility ($s_M^f = s_M^d = s_{RM}^d = s_{MM}^d = 0$). Figure 2 represents the fixation probability for different values of s_{MM}^{Model}, when selection does not act on the allele while rare (*i.e.*, the allele is fully recessive in diploids and has no selective benefit in haploids, $s_{ave}^{\text{Model}} = 0$) in the Moran (A) and WF model (B). In Figure 3, we illustrate how the fixation probability varies for a given total number of individuals, N.

The branching process approximation allowing for strong selection (solving numerically for $u_{\text{branch}}^{\text{Moran}}$ from Equation 2 and $u_{\text{branch}}^{\text{WF}}$ from Equation 5a with Equation D4) accurately predicts the fixation probability from numerical simulations regardless of the strength of selection (red curve, see inset panels of Figure 1), but this method is valid only when the population size N is large and the average effect of a rare mutation is beneficial, $\tilde\phi_M \phi_{RM} > \tilde\phi_R \phi_{RR}$, failing when there is no selection on a rare allele (Figure 2). As expected, approximating the branching process results for weak selection
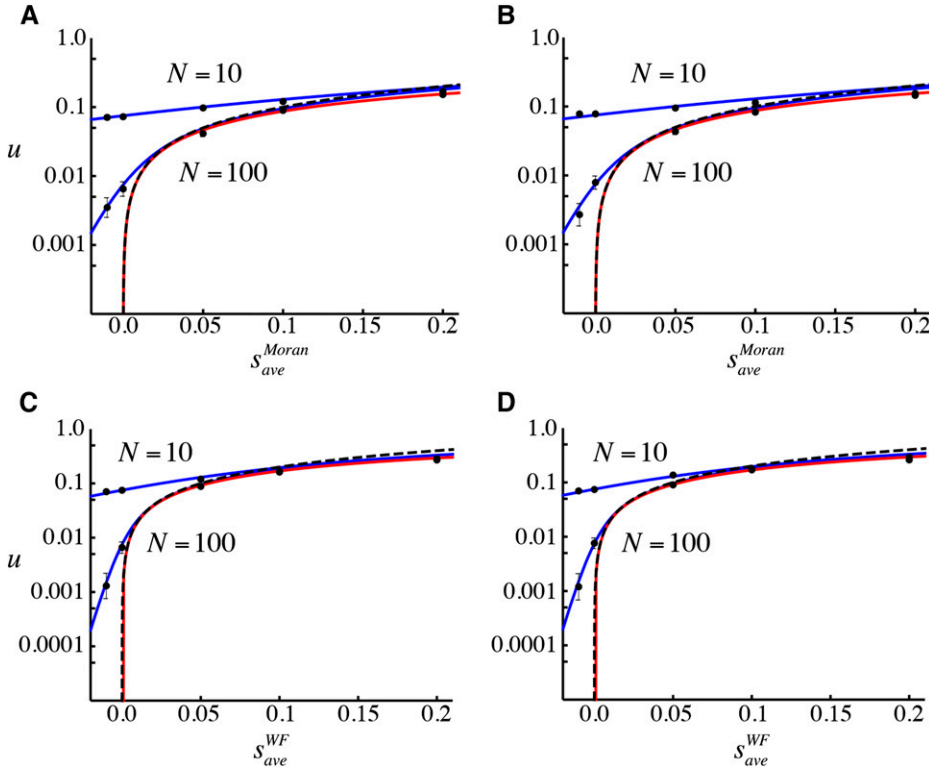
Figure 3 The fixation probability for different population sizes (N) from the (A and B) Moran model and (C and D) WF model, plotted as a function of (A and C) selection in haploids, and (B and D) selection in diploids with $N = 10$ or 100. Parameters: (A) $s_M^\beta = 2s_{ave}^{Moran}$ and $s_{RM}^\beta = s_{MM}^\beta = 0$; (B) $s_M^\beta = 0$, $s_{RM}^\beta = 2s_{ave}^{Moran}$, and $h^\beta = 0.5$; (C) $s_M^w = 2s_{ave}^{WF}$ and $s_{RM}^w = s_{MM}^w = 0$; (D) $s_M^w = 0$, $s_{RM}^w = 2s_{ave}^{WF}$, and $h^w = 0.5$. Other parameters: $f_R = 0.5$, $\beta_R = \beta_{RR} = 1000$, $d_R = d_{RR} = 0.005$, and $s_M^f = s_M^d = s_{RM}^d = s_{MM}^d = 0$. See Figure 1 for further information.

(Equation 13) is accurate only when selection is weak (black dashed line in Figure 1).

By contrast, the diffusion equation (4a) with drift and diffusion coefficients (3) for the Moran model and (7) for the WF model provides a good approximation as long as selection is weak (Figure 1, Figure 2, and Figure 3, blue curves), even when the average effect of a rare mutation is not beneficial and/or the population size is small. Also, the diffusion well approximates the fixation probability when the mutation is fully recessive and has no effect in haploids (Figure 2), which is a case that cannot be handled by the branching process approximation because the fate of the allele is not determined while it is rare.

Because we have assumed equal death rates for haploids and diploids, the fixation probabilities behave similarly for the Moran (Figure 1, A–C) and WF models (Figure 1, D–F), except for the fact that the fixation probability is approximately halved in the Moran model, $u_{diffusion}^{WF} \approx 2u_{diffusion}^{Moran}$, due to the difference in N_e (see Equation 8).

Different haploid and diploid life histories: We next consider the fixation probability in organisms whose life histories differ between haploid and diploid stages ($\beta_R \neq \beta_{RR}, d_R \neq d_{RR}$ for the Moran model, and $w_R \neq w_{RR}$ for the WF model), as would be appropriate for heteromorphic species (*e.g.*, the green alga *Derbesia*, the brown alga *Macrocystis*, the red alga *Porphyra*, etc.). Again, both the branching process and diffusion approximations work well under the conditions assumed during their derivation (Figure 5). This figure also compares the results to classical haploid-only or diploid-only models,

showing that the fixation probability in a haploid-diploid population differs dramatically as the frequency of haploids *vs.* diploids in the population varies.
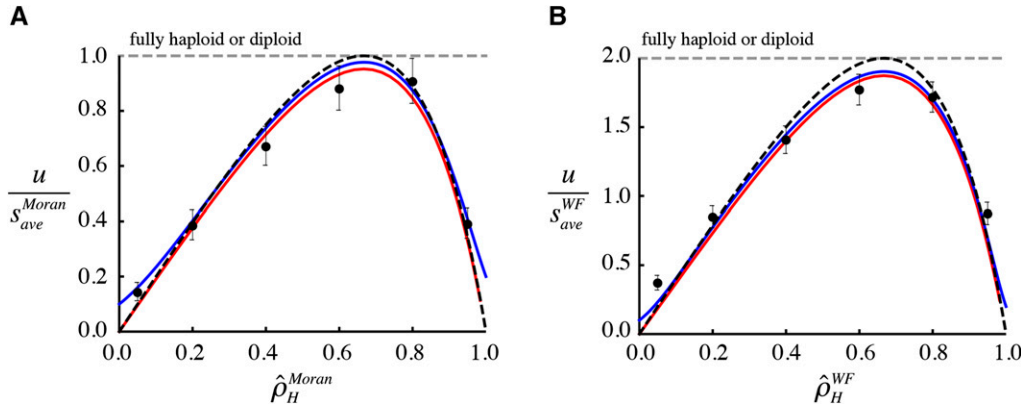
We first discuss the case where the mortality rates of haploids and diploids are equivalent ($d_R = d_{RR}$). With equal death rates, haploids and diploids have the same expected life span in the Moran model, as is assumed in the WF model where all individuals die at each time step. In this case, the fixation probability of beneficial mutations with weak selection and a large population becomes $u^{Moran} = \frac{8\hat{\rho}_H^{Moran}\hat{\rho}_D^{Moran}}{(\hat{\rho}_H^{Moran}+2\hat{\rho}_D^{Moran})^2}s_{ave}^{Moran}$ for the Moran model and $u^{WF} = \frac{8\hat{\rho}_H^{WF}\hat{\rho}_D^{WF}}{(\hat{\rho}_H^{WF}+2\hat{\rho}_D^{WF})^2}2s_{ave}^{WF}$ for the WF model, using either branching processes or a diffusion approximation (Equation 13). We next compare these fixation probabilities for haploid-diploid populations to the classic results for selection only in haploids (Moran: $s_H = s_M^f/2 + s_M^\beta + s_M^d$; WF: $s_H = s_M^f/2 + s_M^w$) *vs.* selection only in diploids (Moran: $s_D = s_{RM}^\beta + s_{RM}^d$; WF: $s_D = s_{RM}^w$), assuming selection is additive.

The fixation probability (13) reaches a maximum of $u^{Moran} = (s_H + s_D)/2$ for the Moran model and $u^{WF} = (s_H + s_D)$ for the WF model when $\hat{\rho}_H^{Model} = 2/3$ (Figure 4). When haploids comprise two-thirds of a haploid-diploid population, there is an equal number of alleles present in haploids and in diploids, which maximizes $N_e^{Model}p_0$ (minimizing drift for a newly arisen mutation), thereby maximizing the fixation probability for a given total number of individuals. At this maximum, the fixation probability (Equation 13, black dashed curve in Figure 4) is the average of the fixation probability in a fully haploid population (s_H for the Moran model

Figure 4 The fixation probability as a function of the fraction of haploids in the population ($\hat{\rho}_H^{Model}$). The fixation probability is measured relative to the average selection coefficient experienced by a rare allele ($u^{Model}/s_{ave}^{Model}$) in (A) the Moran model and (B) the WF model. The horizontal gray dashed lines give the fixation probability in haploid-only and diploid-only populations. The values of $\hat{\rho}_H^{Model}$ of {0.05, 0.2, 0.4, 0.6, 0.8, 0.95} are obtained by setting $\beta_R = w_R$ equal to {1.444 × 10⁶, 64000, 9000, 1777.78, 250, 11.08}. Selection parameters are fixed at $s_M^{\beta} = s_{RM}^{\beta} = s_M^w = s_{RM}^w = 0.1$, $s_M^f = s_M^d = s_{RM}^d = s_{MM}^d = 0$, and $h^{\beta} = h^w = 0.5$, with $f_R = 0.5$, $\beta_{RR} = w_{RR} = 1000$, and $d_R = d_{RR} = 0.005$. See Figure 1 for further information.

and $2s_H$ for the WF model) and in a fully diploid population (s_D for the Moran model and $2s_D$ for the classic diploid WF model). In Figure 4, we explore a case where the fixation probability is the same in haplont and diplont populations (grey dashed curve), in which case the fixation probability in a haploid-diploid population under weak selection (13) is only the same if two-thirds of the population is haploid (black dashed curve). The simulation results tend to fall slightly below this expectation, however, due to the small population size and relatively large selection coefficients assumed (black dots in Figure 4).

When the fraction of haploids departs from two-thirds, however, the fixation probability decreases dramatically (Figure 4), reaching zero when most individuals are haploids ($\hat{\rho}_H^{Model} \approx 1$) or diploids ($\hat{\rho}_D^{Model} \approx 1$). This sensitivity to the relative proportions of haploids and diploids occurs because the new allele must be carried by both haploid and diploid individuals, in alternating generations, in a haploid-diploid population. Hence, when the ratio of haploids to diploids in the population is very skewed, the strength of random genetic drift is increased in whichever ploidy phase has fewer individuals.

We next allow the mortality rates of haploids and diploids to differ as well ($d_R \neq d_{RR}$). As seen in Figure 5A, we again find that the fixation probability approaches zero whenever the ratio of haploids to diploids is extremely skewed ($\hat{\rho}_H^{Model}$ near zero or one). With mortality rates varying between the ploidy levels, however, the peak fixation probability no longer occurs at $\hat{\rho}_H^{Moran} = 2/3$. As a consequence, increasing the mortality rate of haploids relative to diploids may increase or decrease the fixation probability of a beneficial mutation, depending on the overall population frequency of haploids. In Figure 5B, we illustrate the parameter space within which the fixation probability in a haploid-diploid population is larger than in a fully haploid and fully diploid population (red), or smaller than the two (blue), as a function of the frequency of haploids (x-axis: $\hat{\rho}_H^{Moran}$) and the relative mortality rate of haploids (y-axis: $d_R/(d_R + d_{RR})$).

At an intuitive level, beneficial alleles are better preserved and more likely to fix when the more common ploidy phase also has the lower mortality rate, allowing alleles to remain in this phase for longer.

Ploidally antagonistic selection: In this section, we consider cases where selection acts in opposite directions in haploid and diploid phases (ploidally antagonistic selection). The above derivations have not made assumptions about the sign of selection in each phase, although the branching process approximation assumes that the fixation probabilities in (1) are both positive. To assess the robustness of these approximations when selection acts in opposite directions in haploids and diploids, we set $s_M^{\theta} = 0.2 + z$ and $s_{RM}^{\theta} = -z$ ($s_{ave}^{Model} = 0.1$) for $\theta \in \{\beta, w\}$ and varied the degree of antagonism (z) in Figure 6.

The diffusion approximation (blue curves) applies only while selection in both phases is weak. The diffusion approximation depends only on the average strength of selection (see Equation 3b and Equation 7a) and fails to capture the large oscillations in allele frequency that occur between haploid and diploid phases when z is large.

The branching process does, however, capture the decline in the fixation probability that occurs as the oscillations in selection across haploid and diploid phases increase in magnitude. When ploidally antagonistic selection becomes very extreme, however, selection no longer favors allele M when it becomes common. We can measure selection on M once common as the difference in fitness of M-bearing *vs.* R-bearing individuals relative to the now common genotypes, MM and M: $\frac{1}{2}\left[\frac{(1+s_{RM}^{\theta})}{(1+s_{MM}^{\theta})} - 1\right] + \frac{1}{2}\left(\frac{1}{1-s_M^{\theta}} - 1\right)$, averaging across the two phases. Selection only remains positive in Figure 6, favoring M when it becomes common, if z falls between -0.377 and 0.177. Beyond these points, the branching process begins to break down because the loss or fixation of M is no longer decided while it is rare. Indeed, once z extends beyond -0.432 and 0.232, both alleles M and R can invade when rare according to the branching process approximation in both the Moran and WF models. Beyond these values, we
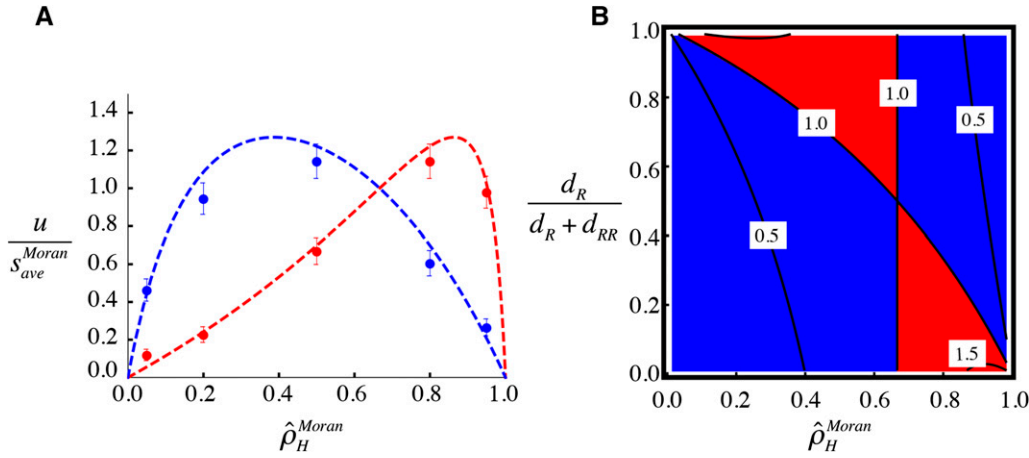
Figure 5 The dependence of the fixation probability on the mortality rates in the Moran model. (A) Panel is equivalent to the weak-selection approximation shown in Figure 3A with different mortality rates: $d_R = 0.0005$ and $d_{RR} = 0.005$ [$d_R/(d_R + d_{RR}) \approx 0.09$] (red), and $d_R = 0.005$ and $d_{RR} = 0.0005$ [$d_R/(d_R + d_{RR}) \approx 0.91$] (blue). Selection parameters are fixed so that s_{ave}^{Moran} is the same in a haploid-diploid, haplont, or diplont population, with $s_M^{\beta} = s_{RM}^{\beta} = 0.05$, $s_M^f = s_M^d = s_{RM}^d = s_{MM}^d = 0$, $h^{\beta} = 0.5$, $f_R = 0.5$, and $\beta_R = \beta_{RR} = 1000$. (B) The fixation probability scaled to selection ($u^{Moran}/s_{ave}^{Moran}$) is shown as a function of $\hat{\rho}_H^{Moran}$ (x-axis) and $d_R/(d_R + d_{RR})$ (y-axis). The blue region indicates when the fixation probability is less than that for both a haploid-only and diploid-only population ($u^{Moran}/s_{ave}^{Moran} < 1.0$). The red region indicates the region when the fixation probability is higher than both a haploid-only and diploid-only population ($1.0 < u^{Moran}/s_{ave}^{Moran}$). See Figure 1 for further information.

expect ploidally antagonistic selection to maintain polymorphism for extended periods of time [see bottom of Figure 6, A and B, in File S1; the stochastic analog of the protected polymorphism identified by Immler *et al.* 2012].

In the WF model, a further complexity arises because of the fact that a mutation that appears in a haploid will only be in diploid offspring, haploid grand-offspring, *etc.*, because of the assumption of completely nonoverlapping generations. This alternation can lead to the situation where the entire population is composed of M haploids and RR diploids, followed in the next generation by R haploids and MM diploids; continuing to alternate until the population happens to lose one genotype or the other. At that point, allele M tends to fix only when its geometric mean fitness is higher: $\sqrt{\frac{f_M}{2} w_M w_{MM}} > \sqrt{\frac{f_R}{2} w_R w_{RR}}$, which requires $z < 0.122$. Because M is now in a different context, finding itself in MM diploids that are disfavored when z is positive, the branching process fails to capture the probability of fixation for large values of z.

Overall, we find that our results continue to apply with ploidally antagonistic selection as long as selection in each phase is weak for the diffusion approximation and the fate of an allele is decided while it remains rare for the branching process approximation.

Discussion

Sexual organisms often exhibit complex life cycles alternating between different ploidy phases, with free-living haploid and diploid individuals. We develop a genetic model explicitly considering the population dynamics of haploid and diploid individuals by generalizing classical genetic models: the Moran model and the WF model. Using a branching process and a diffusion approximation, our work is the first to derive the fixation probability for organisms with haploid-diploid life cycles.

The fixation probability obtained by using branching processes (Equation 2), u_{branch}^{Model}, provides a good approximation to our simulation results when the average effect of a mutation is beneficial ($s_{ave}^{Model} > 0$) and the total population size N is sufficiently large, as long as ploidally antagonistic selection is not so strong that the allele becomes deleterious as it rises in frequency. On the other hand, the fixation probability from the diffusion approximation (Equation 4a), $u_{diffusion}^{Model}$, is accurate for both deleterious and beneficial mutations for both small and large populations, but only when selection is weak in each phase.

When average selection is weak, additive, and positive, and when starting from a single mutation, the branching process and diffusion approximations converge to Equation 13a for the WF model and Equation 13b for the Moran model. When the mortality rates are equal in haploids and diploids, the fixation probability reduces in both models to:

$$u^{Model} = \frac{8\hat{\rho}_H^{Model}\hat{\rho}_D^{Model}}{\left(\hat{\rho}_H^{Model} + 2\hat{\rho}_D^{Model}\right)^2} C^{Model} s_{ave}^{Model}, \tag{14}$$

where $C^{WF} = 2$ for the WF model and $C^{Moran} = 1$ for the Moran model; reflecting the increased stochasticity in the Moran model, which samples both individuals to die and to give birth (all parents die in the WF model). Equation 14 demonstrates an important aspect of the fixation probability in organisms with free-living haploid and diploid phases: beneficial alleles are less likely to fix in populations with skewed ratios of haploids to diploids ($\hat{\rho}_H^{Model}$ near zero or one). This contrasts with classical results for haplont or diplont populations (Crow and Kimura 1970), where the other ploidy phase is assumed to be effectively infinite (no drift).

The "average effect" of selection, which averages selection across reproduction, fertilization, and mortality for both haploid and diploid phases, plays a key role in the fixation probability in haploid-diploid populations. Mutations can
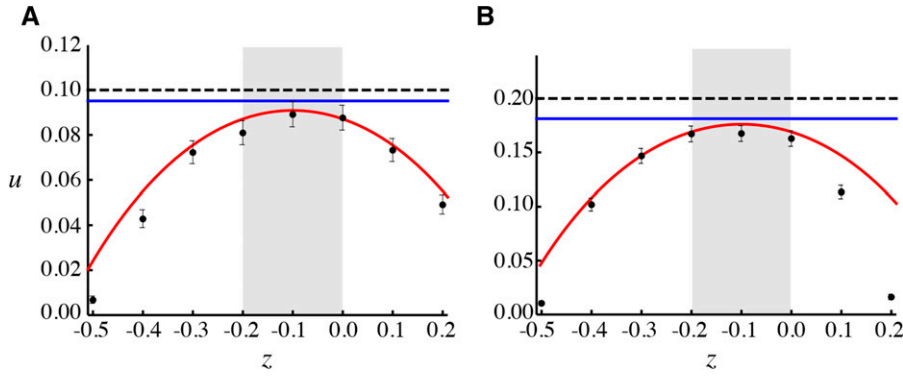
Figure 6 The fixation probability with ploidally antagonistic selection in the (A) Moran model and (B) WF model, plotted as a function of the degree of ploidally antagonistic selection, z, where $s_M^{\text{Model}} = s_{\text{ave}}^{\text{Model}} + z$ and $s_{RM}^{\text{Model}} = -z$. For small values of z, selection remains directional (gray box) but becomes ploidally antagonistic outside this range. Parameters: (A) $s_M^\beta = 0.2 + z$, $s_{RM}^\beta = -z$, $h^\beta = 0.5$; and (B) $s_M^w = 0.2 + z$, $s_{RM}^w = -z$, and $h^w = 0.5$. Other parameters: $f_R = 0.5$, $\beta_R = \beta_{RR} = w_R = w_{RR} = 1000$, $d_R = d_{RR} = 0.005$, and $s_M^f = s_{RM}^d = s_{MM}^d = 0$. See Figure 1 for further information.

have deleterious effects at one ploidy level, but not the other, and still be beneficial overall. Alternatively, such ploidally antagonistic selection can maintain polymorphism (Immler *et al.* 2012). In that case, the diffusion model can also be used with the mean and variance components derived here to obtain the time to loss of one or the other allele, or the stationary distribution, as well as the fixation probability.

Our results in a haploid-diploid population offer an interesting comparison to models with separate sexes (*e.g.*, Crow and Kimura 1970; Hill 1979; Pollak 1990; Caballero 1995) or two patches (*e.g.*, Maruyama 1970; Whitlock and Barton 1997; Whitlock 2003; Yeaman and Otto 2011). Consider first the case with separate sexes. To see the connection it is instructive to recalculate the variance effective population size using the method described in Crow and Kimura (1970), rather than the more formal derivation in Appendix D. In a population with two sexes consisting of M adult males and F adult females (both diploid) with allele frequencies $p_\male$ and $p_\female$, the average allele frequency among offspring is $(p_\male + p_\female)/2$ and the variance in allele frequency is $[\text{Var}(p_\male) + \text{Var}(p_\female)]/4$. This in turn equals $\frac{1}{4}\frac{p(1-p)}{2M} + \frac{1}{4}\frac{p(1-p)}{2F}$, assuming similar allele frequencies in the two sexes ($p_\male = p_\female = p$). Setting this to the variance expected in the classic diploid WF model without separate sexes, $p(1-p)/(2N_e)$, we obtain the variance effective population size $N_e = 4MF/(M+F)$ (Crow and Kimura 1970). For X-linked genes, the average allele frequency within the population is instead $(p_\male + 2p_\female)/3$, so that the expected variance is $\frac{1}{9}\frac{p(1-p)}{M} + \frac{4}{9}\frac{p(1-p)}{2F}$ (note there is only one allele in males), yielding $N_e = 9MF/(4M+2F)$. One might initially assume that a haploid-diploid population would have a similar variance effective population size to an X-linked gene, but this is not the case. With a strict alternation of generations, a gene spends half of the generations in haploids and half in diploids, so the average allele frequency is $(p_H + p_D)/2$ (now averaged over the alternation of generations), giving a variance in allele frequency of $\frac{1}{4}\frac{p(1-p)}{\hat{H}} + \frac{1}{4}\frac{p(1-p)}{2\hat{D}}$ and a variance effective population size of $N_e = 4\hat{H}\hat{D}/(\hat{H} + 2\hat{D})$ (Equation 8b). This is most similar to the two sex, autosomal model but with $M = \hat{H}/2$ and $F = \hat{D}$. In both cases, descendants have an evolutionary history spent half in one type (males or haploids) and half in the other (females or diploids),

and the variance effective population size is sensitive to low numbers of either type. As a consequence, N_e is maximized in the haploid-diploid model when the phase with fewer genes (haploids) is relatively common ($\hat{H}/(\hat{H} + \hat{D}) \approx 0.586$), whereas N_e is maximized in the sex-linked model when the phase with fewer genes (males) is relatively rare ($M/(F + M) \approx 0.414$).

The haploid-diploid model can also be thought of as a two-patch model, sharing similarities with models of subdivided populations (*e.g.*, Tachida and Iizuka 1991; Gavrilets and Gibson 2002; Yeaman and Otto 2011). In the classical island and stepping stone models, the fixation probability in a subdivided population is the same as that in an undivided population of the same total population size (Maruyama 1970). This assumes, however, "conservative migration," where each population contributes to the dynamics of a metapopulation according to the local population size and where migration has no effect on the local population size (Whitlock and Barton 1997; Whitlock 2003). The haploid-diploid model, however, does not exhibit conservative migration. Rather, all haploids become diploids and vice versa (equivalent to a migration rate of 100% in the WF model). When one population is smaller (say haploids), that population has a much larger impact on the dynamics of the whole population because the larger population (say diploids) descends entirely from it. Consequently, as we have shown here, the fixation probability of a haploid-diploid population does not obey Maruyama's result for a single population, either haploid or diploid, with the same total number of individuals (Figure 4 and Figure 5).

By deriving the fixation probability for haploid-diploid populations, this work has extended classical population genetic results to organisms with an alternation of generations, allowing a better understanding of the role of stochasticity and drift in such species. Given that a large number of genes are expressed during both the haploid and diploid generations (Coelho *et al.* 2007), the theoretical framework developed here also allows for the integration of selection of these "shared" (pleiotropic) genes across the life cycle.

Acknowledgments

Literature Cited

Bell, G., 1994   The comparative biology of the alternation of generations. Lect. Math Life Sci. 25: 1–26.

Bell, G., 1997   The evolution of the life cycle of brown seaweeds. Biol. J. Linn. Soc. Lond. 60: 21–38.

Caballero, A., 1995   On the effective size of populations with separate sexes, with particular reference to sex-linked genes. Genetics 139: 1007–1011.

Coelho, S. M., A. F. Peters, B. Charrier, D. Roze, C. Destombe *et al.*, 2007   Complex life cycles of multicellular eukaryotes: new approaches based on the use of model organisms. Gene 406: 152–170.

Crow, J. F., and M. Kimura, 1970   *An Introduction to Population Genetics Theory*. Blackburn Press, Caldwell, NJ.

Destombe, C., M. Valero, P. Vernet, and D. Couvet, 1989   What controls haploid—diploid ratio in the red alga, *Gracilaria verrucosa?* J. Evol. Biol. 2: 317–338.

Ewens, W. J., 2004   *Mathematical Population Genetics 1: Theoretical Introduction*. Springer Science & Business Media, New York.

Gavrilets, S., and N. Gibson, 2002   Fixation probabilities in a spatially heterogeneous environment. Popul. Ecol. 44: 51–58.

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

Hill, W. G., 1979   A note on effective population size with overlapping generations. Genetics 92: 317–322.

Hiraoka, M., and G. Yoshida, 2010   Temporal variation in isomorphic phase and sex ratios of a natural population of *Ulva pertusa* (Chlorophyta). J. Phycol. 46: 882–888.

Hirose, H., 1975   *General Phycology*, revised Ed. 2, Uchida Rokakuho, Tokyo (in Japanese).

Immler, S., G. Arnqvist, and S. P. Otto, 2012   Ploidally antagonistic selection maintains stable genetic polymorphism. Evolution 66: 55–65.

Karlin, S., and H. E. Taylor, 1981   *A Second Course in Stochastic Processes*. Academic Press, New York.

Kimura, M., 1957   Some problems of stochastic processes in genetics. Ann. Math. Stat. 28: 882–901.

Kimura, M., 1962   On the probability of fixation of mutant genes in a population. Genetics 47: 713.

Mable, K. B., and P. S. Otto, 1998   The evolution of life cycles with haploid and diploid phases. BioEssays 20: 453–462.

Maruyama, T., 1970   Effective number of alleles in a subdivided population. Theor. Popul. Biol. 1: 273–306.

Nagylaki, T., 1976   The evolution of one-and two-locus systems. Genetics 83: 583–600.

Otto, S. P., and T. Day, 2007   *A Biologist's Guide to Mathematical Modeling in Ecology and Evolution*. Princeton University Press, Princeton, NJ.

Pollak, E., 1990   The effective population size of an age-structured population with a sex-linked locus. Math. Biosci. 101: 121–130.

Tachida, H., and M. Iizuka, 1991   Fixation probability in spatially changing environments. Genet. Res. 58: 243–251

Taylor, P. D., 1990   Allele-frequency change in a class-structured population. Am. Nat. 135: 95–106.

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

Whitlock, M. C., and N. Barton, 1997   The effective size of a subdivided population. Genetics 146: 427–441.

Yeaman, S., and S. P. Otto, 2011   Establishment and maintenance of adaptive genetic divergence under migration, selection, and drift. Evolution 65: 2123–2129.

Communicating editor: L. M. Wahl

Appendix A: Change in Allele Frequency Within a Haploid-Diploid Moran Model

The Moran model is a Markov process where the state of the population at time $t + 1$ can be treated probabilistically given the frequencies of all genotypes, $x_{(GT)}$, in the previous time step. At each time step, a single individual is chosen to die and one is chosen to give birth, so that the population size remains constant.

We first choose an individual to die, weighting the chance that each type is eliminated by its mortality rate. Thus, the probability that an individual with genotype (GT) is eliminated from the population is

$$\tilde{d}_{(GT)} = \frac{d_{(GT)}x_{(GT)}}{d_R x_R + d_M x_M + d_{RR}x_{RR} + d_{RM}x_{RM} + d_{MM}x_{MM}}. \tag{A1}$$

We then replace that individual with a birth chosen randomly from all of the spores and zygotes that could be produced by the remaining population at time t. Because we assume a large population size, we assume that the change in genotype frequency due to the death in this time step is negligible for common genotypes (this assumption is not made in the simulations or for rare genotypes in our analysis). The probability that a haploid spore is recruited is then proportional to the rate at which diploids reproduce by meiosis to produce either R- or M-bearing spores:

$$c_R = \beta_{RR}x_{RR} + \frac{\beta_{RM}x_{RM}}{2}, \tag{A2a}$$

$$c_M = \beta_{MM}x_{MM} + \frac{\beta_{RM}x_{RM}}{2}. \tag{A2b}$$

Given that resident and mutant gametes are produced at rates $g_R = \beta_R x_R/2$ and $g_M = \beta_M x_M/2$, respectively, and accounting for the probability of fertilization, the probability that a diploid zygote is recruited is:

$$c_{RR} = \frac{f_R}{2}\frac{\beta_R^2 x_R^2}{\beta_R x_R + \beta_M x_M}, \tag{A3a}$$

$$c_{RM} = \frac{f_R + f_M}{2}\frac{\beta_R \beta_M x_R x_M}{\beta_R x_R + \beta_M x_M}, \tag{A3b}$$

$$c_{MM} = \frac{f_M}{2}\frac{\beta_M^2 x_M^2}{\beta_R x_R + \beta_M x_M}. \tag{A3c}$$

Overall, the probability that genotype (GT) is recruited into the empty site is:

$$q_{(GT)} = \frac{c_{(GT)}}{c_R + c_M + c_{RR} + c_{RM} + c_{MM}}. \tag{A4}$$

To derive the first moment of mutant allele frequency fluctuations, we calculate the conditional expected value for the change in any function F of the random variable as

$$E\left[\Delta F(\overrightarrow{X}(t))|\overrightarrow{X}(t) = \overrightarrow{x}\right] = E\left[F(\overrightarrow{X}(t+1)) - F(\overrightarrow{X}(t))|\overrightarrow{X}(t) = \overrightarrow{x}\right], \tag{A5}$$

where $\overrightarrow{X}(t) = [X_R(t), X_M(t), X_{RR}(t), X_{RM}(t), X_{MM}(t)]^T$ and $\overrightarrow{x} = (x_R, x_M, x_{RR}, x_{RM}, x_{MM})^T$.

Using Equation A1 and (A4), the expected change in the number of individuals with genotype (GT), $\Delta X_{(GT)}(t) = X_{(GT)}(t+1) - X_{(GT)}(t)$ is:

$$E\left[\Delta X_{(GT)}(t)\middle|\overrightarrow{X}(t) = \overrightarrow{x}\right] = \left(1 - \tilde{d}_{(GT)}\right)q_{(GT)} - \tilde{d}_{(GT)}\left(1 - q_{(GT)}\right) = q_{(GT)} - \tilde{d}_{(GT)}, \tag{A6}$$

Note that Equation A6 can be rewritten as

$$E\left[\Delta X_{(GT)}(t)\middle|\overrightarrow{X}(t) = \overrightarrow{x}\right] = \frac{\gamma c_{(GT)} - d_{(GT)}x_{(GT)}}{N d_{\mathrm{mean}}}, \tag{A7}$$

where $\gamma = (d_R x_R + d_M x_M + d_{RR} x_{RR} + d_{RM} x_{RM} + d_{MM} x_{MM})/(c_R + c_M + c_{RR} + c_{RM} + c_{MM})$ and $d_{mean} = (d_R x_R + d_M x_M + d_{RR} x_{RR} + d_{RM} x_{RM} + d_{MM} x_{MM})/N$.

These expected changes in genotype frequencies can be used to track the dynamics of the system and determine its equilibrium properties.

Equilibrium

When the mutant allele is absent, the resident population approaches an equilibrium with $\hat{H}$ haploids of genotype R (x_R) and $\hat{D}$ diploids of genotype RR (x_{RR}), with stochastic fluctuations around this point. Assuming the population size is large (dropping terms of order $1/N$), we can solve for the equilibrium number of haploids and diploids,

$$\hat{H} = \hat{\rho}_H^{\text{Moran}} N \text{ and } \hat{D} = \hat{\rho}_D^{\text{Moran}} N, \tag{A8a}$$

using (A7), which yields $\hat{\gamma} \beta_{RR} \hat{D} - d_R \hat{H} = 0$ and $\hat{\gamma} = (d_R \hat{H} + d_{RR} \hat{D})/[(f_R \beta_R \hat{H}/2) + \beta_{RR} \hat{D}]$, to obtain the frequencies:

$$\hat{\rho}_H^{\text{Moran}} = \frac{\sqrt{\beta_{RR} d_{RR}}}{\sqrt{f_R \beta_R d_R/2} + \sqrt{\beta_{RR} d_{RR}}} \text{ and } \hat{\rho}_D^{\text{Moran}} = \frac{\sqrt{f_R \beta_R d_R/2}}{\sqrt{f_R \beta_R d_R/2} + \sqrt{\beta_{RR} d_{RR}}}. \tag{A8b}$$

Equation A8b emphasizes the fact that haploids are most common when diploids reproduce and/or die at a higher rate, and vice versa.

Appendix B: Fixation Probability in the Haploid-Diploid Moran Model Using a Branching Process Approximation

To derive the fixation probability of a mutant allele, we first apply the branching process approximation for a two-patch model (e.g., Yeaman and Otto 2011). Because the fate of a mutant allele is assumed to be determined while it is rare in a branching process approximation, we ignore homozygous MM mutants and consider only the change in number of haploid mutants X_M and diploid mutants X_{RM}. As in Appendix A, we assume that the life cycle consists of the death of a single individual, followed by the birth of an individual from the pool of haploid spores and zygotes that could be produced at that time step. In this birth step, we again assume that the population size is large enough that we can ignore the small change in $q_{(GT)}$ that occurs following the death of an individual with one of the common genotypes.

When a single haploid mutant appears in the resident population, the numbers of individuals are $x_M = 1$, $x_R = \hat{H} - 1$, $x_{RR} = \hat{D}$, and $x_{RM} = x_{MM} = 0$. When a single diploid mutant appears in the resident population, we have $x_R = \hat{H}$, $x_M = 0$, $x_{RR} = \hat{D} - 1$, $x_{RM} = 1$, and $x_{MM} = 0$. Plugging these into Equation A1, the probabilities that a mutant individual is chosen to die are:

$$\tilde{d}_{M|H} = \frac{d_M}{\hat{d}N} \text{ and } \tilde{d}_{RM|D} = \frac{d_{RM}}{\hat{d}N}, \tag{B1}$$

where $\hat{d} = d_R \hat{\rho}_H + d_{RR} \hat{\rho}_D$ (assuming the population is large) and otherwise residents are sampled to die (i.e., $\tilde{d}_{M|H} = 1 - \tilde{d}_{R|H} - \tilde{d}_{RR|H}$ and $\tilde{d}_{RM|D} = 1 - \tilde{d}_{R|D} - \tilde{d}_{RR|D}$). We add "$|H$" or "$|D$" to the subscript to clarify that we start with one haploid or one diploid mutant individual.

When a single haploid mutant appears in the resident population, the expected number of zygotes is: $c_{RR|H} = \frac{f_R \beta_R (\hat{H}-1)}{2} \frac{\beta_R (\hat{H}-1)}{\beta_R (\hat{H}-1)+\beta_M}$, $c_{RM|H} = \frac{f_R + f_M}{2} \frac{\beta_R \beta_M (\hat{H}-1)}{\beta_R (\hat{H}-1)+\beta_M}$, and $c_{MM|H} = \frac{f_M \beta_M}{2} \frac{\beta_M}{\beta_R (\hat{H}-1)+\beta_M}$. Because the diploid population consists solely of RR individuals, the rate of production of haploid spores is: $c_{R|H} = \beta_{RR} \hat{D}$ and $c_{M|H} = 0$. Because N is assumed to be very large in a branching process approximation and homozygotes are rare, the probability that the sampled offspring will be of genotype (GT) given by Equation A4 is approximately:

$$q_{RM|H} \approx \frac{(f_R + f_M)\beta_M}{2\hat{\beta}N} \text{ and } q_{MM|H} \approx 0, \tag{B2}$$

where $\hat{\beta} = \frac{f_R \beta_R}{2} \hat{\rho}_H^{\text{Moran}} + \beta_{RR} \hat{\rho}_D^{\text{Moran}}$.

When a single diploid mutant appears in the resident population, we have $c_{RR|D} = f_R \beta_R \hat{H}/2$, $c_{RM|D} = c_{MM|D} = 0$, $c_{R|D} = \beta_{RR}(\hat{D}-1) + (\beta_{RM}/2)$, and $c_{M|D} = \beta_{RM}/2$. Because N is large and homozygous mutants rare, the probability that a mutant offspring is sampled is:

$$q_{M|D} \approx \frac{\beta_{RM}}{2\hat{\beta}N}. \tag{B3}$$

Let $\pi_{0|H}$ ($\pi_{0|D}$) equal the probability that a mutant allele introduced into a single haploid individual with genotype M (single diploid individual with genotype RM) is lost from the descendant population at some point in the future. Assuming that its fate is determined before the homozygote mutant MM appears, that the population is at demographic equilibrium, and that the probabilities of loss are independent of each other; these probabilities satisfy the law of total probability, accounting for all possible death-birth transitions:

$$\pi_{0|H} = \tilde{d}_{M|H} + \tilde{d}_{R|H}\left(1 - q_{RM|H}\right)\pi_{0|H} + \tilde{d}_{R|H}q_{RM|H}\pi_{0|H}\pi_{0|D} + \tilde{d}_{RR|H}\left(1 - q_{RM|H}\right)\pi_{0|H} + \tilde{d}_{RR|H}q_{RM|H}\pi_{0|H}\pi_{0|D} \tag{B4}$$

when starting with one initial M haploid, and

$$\pi_{0|D} = \tilde{d}_{RM|D} + \tilde{d}_{RR|D}\left(1 - q_{M|D}\right)\pi_{0|D} + \tilde{d}_{RR|D}q_{M|D}\pi_{0|H}\pi_{0|D} + \tilde{d}_{R|D}\left(1 - q_{M|D}\right)\pi_{0|D} + \tilde{d}_{R|D}q_{M|D}\pi_{0|H}\pi_{0|D} \tag{B5}$$

when starting with one initial RM diploid. The probabilities of fixation are then defined as the chance that loss does not occur: $\pi_{1|H} = 1 - \pi_{0|H}$ and $\pi_{1|D} = 1 - \pi_{0|D}$. Recalling the branching process assumption that the fate of the allele is determined while the mutation is rare, the probability that a single mutant allele is chosen to die $\left(\tilde{d}_{(GT)}\right)$ or reproduce $(q_{(GT)})$ remains roughly constant during this period. The pair in Equation B5 can thus be solved to give the fixation probabilities:

$$\pi_{1|H} = \frac{q_{RM|H}\left(1 - \tilde{d}_{M|H}\right) - \frac{\tilde{d}_{M|H}\tilde{d}_{RM|D}}{q_{M|D}\left(1 - \tilde{d}_{RM|D}\right)}}{q_{RM|H}\left(1 - \tilde{d}_{M|H}\right) + \tilde{d}_{M|H}}, \tag{B6a}$$

$$\pi_{1|D} = \frac{q_{M|D}\left(1 - \tilde{d}_{RM|D}\right) - \frac{\tilde{d}_{M|H}\tilde{d}_{RM|D}}{q_{RM|H}\left(1 - \tilde{d}_{M|H}\right)}}{q_{M|D}\left(1 - \tilde{d}_{RM|D}\right) + \tilde{d}_{RM|D}}. \tag{B6b}$$

Because the population size is very large, the probabilities of death or recruitment of a mutant $\tilde{d}_{M/H}, \tilde{d}_{RM/D}, q_{M/D}$, and $q_{RM/H}$ are very small. Hence, the fixation probabilities are well approximated by:

$$\pi_H^{\text{Moran}} = \frac{q_{M|D}q_{RM|H} - \tilde{d}_{M|H}\tilde{d}_{RM|D}}{q_{M|D}\left(q_{RM|H} + \tilde{d}_{M|H}\right)}, \tag{B7a}$$

$$\pi_D^{\text{Moran}} = \frac{q_{RM|H}q_{M|D} - \tilde{d}_{M|H}\tilde{d}_{RM|D}}{q_{RM|H}\left(q_{M|D} + \tilde{d}_{RM|D}\right)}. \tag{B7b}$$

Plugging (B1), (B2), and (B3) into (B7) along with the frequency of haploids among residents, (A8b), we obtain (1).

Furthermore, we can apply a Taylor expansion in ε to Equation 1, assuming weak selection. Specifically, we plug (A8b) into (1) and rewrite $\tilde{\phi}_{(GT)}$ and $\phi_{(GT)}$ in terms of the life history parameters and the mutant's effect on those parameters $\left(e.g., f_M = f_R\left(1 + s_M^f\right) = f_R\left(1 + \varepsilon\tilde{s}_M^f\right)\right)$. Carrying out the Taylor series, the fixation probability is then approximately

$$\pi_H^{\text{Moran}} \approx \frac{4d_{RR}\hat{\rho}_D^{\text{Moran}}}{d_R\hat{\rho}_H^{\text{Moran}} + 2d_{RR}\hat{\rho}_D^{\text{Moran}}}s_{\text{ave}}^{\text{Moran}}, \tag{B8a}$$

$$\pi_D^{\text{Moran}} \approx \frac{2d_R\hat{\rho}_H^{\text{Moran}}}{d_R\hat{\rho}_H^{\text{Moran}} + 2d_{RR}\hat{\rho}_D^{\text{Moran}}}s_{\text{ave}}^{\text{Moran}}. \tag{B8b}$$

Plugging (B8) into (2), we have (13) for the weak-selection approximation to the fixation probability based on branching processes in the Moran model.

Appendix C: Fixation Probability in the Haploid-Diploid Moran Model Using a Diffusion Approximation

To derive the fixation probability from a diffusion approximation, we begin by deriving the first and second moments for the expected change in mutant allele frequency.

Separation of Timescales

Because the total number of individuals is held constant at N, dynamics for $X_{(GT)}$ are described by four variables. Transforming the original system of Equation A7 into the new variables, p_M, δ_p, ρ_H, and F_D, we have equivalent dynamics for the expected changes in the new variables. Applying a separation of timescales (*e.g.*, Nagylaki 1976; Otto and Day 2007), we then calculate the slow dynamics for the change in frequency of mutant alleles p_M. The details of the calculations are given in a supplementary Mathematica file (File S1).

Let f_η represent the expected change in each variable ($\eta \in \{p_M, \delta_p, \rho_H, F_D\}$) across one generation given the current state of the population:

$$E[\Delta\eta(t)|\overrightarrow{X}(t) = \overrightarrow{x}] = f_\eta(\varepsilon, p_M, \delta_p, \rho_H, F_D). \tag{C1}$$

The exact functional form for each variable is provided in File S1.

To approximate the above with weak selection, we apply a Taylor series expansion in ε to Equation C1. To constant order $[O(1)]$, we have:

$$f_{p_M}(0, p_M, \delta_p, \rho_H, F_D) = g_{p_M}(\rho_H)\delta_p, \tag{C2a}$$

$$f_{\delta_p}(0, p_M, \delta_p, \rho_H, F_D) = g_{\delta_p}(\rho_H)\delta_p, \tag{C2b}$$

$$f_{\rho_H}(0, p_M, \delta_p, \rho_H, F_D) = \frac{2\beta_{RR}d_{RR}(1-\rho_H)^2 - f_R\beta_R d_R\rho_H^2}{N\{d_R\rho_H + d_{RR}(1-\rho_H)\}\{f_R\beta_R\rho_H + 2\beta_{RR}(1-\rho_H)\}}, \tag{C2c}$$

$$f_{F_D}(0, p_M, \delta_p, \rho_H, F_D) = g_{F_D}(p_M, \delta_p, \rho_H, F_D), \tag{C2d}$$

(see detailed form of the g_η functions in the File S1). We assume that the population rapidly reaches a quasi-equilibrium according to these equations, which are independent of selection ($\varepsilon = 0$) and reflect ecological dynamics and the approach to Hardy–Weinberg. Solving for this quasi-equilibrium by setting Equations C2 to zero, we have the steady state approximations: $\overline{\delta}_p = \overline{F}_D = 0$ and $\overline{\rho}_H = \hat{\rho}_H$, which gives $f_{p_M}(0, p_M, \overline{\delta}_p, \overline{\rho}_H, \overline{F}_D) = 0$ for any p_M. Consequently, the population quickly approaches a state where there is a similar allele frequency between the haploid and diploid population ($\delta_p \approx 0$), a Hardy–Weinberg ratio in the diploid population ($F_D \approx 0$), and a haploid-to-diploid ratio close to that in the resident population given by Equation (A8b) ($\rho_H \approx \hat{\rho}_H$). We refer to this as the quasi-equilibrium, to emphasize the point that the dynamics will still change slowly due to selection.

After the state of the population approaches this steady state, the variables change over a longer timescale due to selection $[\delta_p = F_D = O(\varepsilon)$ and $\rho_H = \hat{\rho}_H + O(\varepsilon)]$. To derive these slow dynamics, we set $\delta_p = \tilde{\delta}_p\varepsilon$, $F_D = \tilde{F}_D\varepsilon$, and $\rho_H = \hat{\rho}_H + \tilde{\rho}_H\varepsilon$ in Equation C1 and perform a Taylor series expansion around $\varepsilon = 0$, now keeping $O(\varepsilon)$ order terms. Doing so, we find that the change in allele frequency represents a closed variable system that does not depend on the dynamics of the other variables to this order:

$$E[\Delta p_M(t)|\overrightarrow{X}(t) = \overrightarrow{x}] \approx M_p^{\text{Moran}}(p_M) = \frac{p_M(1-p_M)\left\{2s_{\text{ave}}^{\text{Moran}} + p_M\left[\left(1-2h^\beta\right)s_{MM}^\beta + \left(1-2h^d\right)s_{MM}^d\right]\right\}}{2N}\frac{\hat{d}}{\check{d}}, \tag{C3}$$

where the weighted mean mortality rate is $\hat{d} = \hat{\rho}_H d_R + \hat{\rho}_D d_{RR}$ and the harmonic mean mortality rate is $\check{d} = 2\left/\left(\frac{1}{d_R} + \frac{1}{d_{RR}}\right)\right.$.

Appropriate Weight for the Average Allele Frequency p_M

When we transform the variables using different allele frequency definitions, the dynamics of the allele frequency did depend on the other variables to $O(\varepsilon)$. That is, defining the average allele frequency as $p_M = ap_H + (1-a)p_D$ with an arbitrary weighting,

a, only $a = l_R/(l_R + l_{RR})$ (as used above) allows the allele frequency dynamics to be independent of the other variables to $O(\varepsilon)$ (see File S1).

Deriving the Second and Third Moments

To derive the backward diffusion equation, we need both the second and third moments of change in mutant allele frequencies. Because we assume weak selection, we derive the second and third moments from the leading-order dynamics where the allele R and M are equivalent ($d_R = d_M, d_{RR} = d_{RM} = d_{MM}, \beta_R = \beta_M, \beta_{RR} = \beta_{RM} = \beta_{MM}$, and $f_R = f_M$). As described above, to leading order, the system rapidly approaches a state where the allele frequencies are similar between haploids and diploids, diploids are at Hardy–Weinberg, and the frequency of haploids is $\hat{\rho}_H^{\text{Moran}}$. In the following, we assume that these steady state relationships hold.

We define the change in mutant allele frequency when an individual with genotype (a) dies and is replaced by an individual with genotype (b) as

$$\Delta p_{a \to b} = \frac{l_R}{l_R + l_{RR}} \frac{x'_M}{x'_R + x'_M} + \frac{l_{RR}}{l_R + l_{RR}} \frac{(x'_{RM}/2) + x'_{MM}}{x'_{RR} + x'_{RM} + x'_{MM}} - p_M, \tag{C4}$$

where the number of the genotype that died decreases by one $[x'_{(a)} = x_{(a)} - 1]$, the number of the genotype that was born increases by one $[x'_{(b)} = x_{(b)} + 1]$, and the number of all other genotypes remains the same $[x'_{(GT)} = x_{(GT)}$ when $(GT) \notin \{a, b\}]$. We can ignore the case where an individual is replaced by an individual of the same genotype because that causes no change in allele frequency. For example, when a haploid resident is replaced by a haploid mutant $[(a) = R$ and $(b) = M]$, Δp can be described as $\Delta p_{R \to M} = \frac{l_R}{l_R + l_{RR}} \frac{(x_M + 1)}{(x_R - 1) + (x_M + 1)} + \frac{l_{RR}}{l_R + l_{RR}} \frac{(x_{RM}/2) + x_{MM}}{x_{RR} + x_{RM} + x_{MM}} - p_M$, and this event occurs with probability $\tilde{d}_R q_M$. Using Equation C3, we can define the expected moments of change in mutant allele frequency through a series of twenty terms:

$$E\left[(\Delta p_M)^k \middle| \overrightarrow{X}(t) = \overrightarrow{x}\right] = \tilde{d}_R q_M \Delta p_{R \to M}^k + \tilde{d}_R q_{RR} \Delta p_{R \to RR}^k + \cdots + \tilde{d}_{MM} q_{RM} \Delta p_{MM \to RM}^k, \tag{C5}$$

Transforming variables $x_{(GT)}$ into p_M, δ_p, ρ_H, and F_D, taking $\delta_p = F_D = 0$ and $\rho_H = \hat{\rho}_H^{\text{Moran}}$; transforming the timescale by a new time variable $\tau = t/\{N^2(\hat{d}/\bar{d})\}$ and $P_M(\tau) = p_M[N^2(\hat{d}/\bar{d})t]$, where the change in this transformed time variable is $\Delta\tau = \left(\frac{t+1}{N^2} - \frac{t}{N^2}\right)\frac{\hat{d}}{\bar{d}} = \frac{1}{N^2}\frac{\hat{d}}{\bar{d}}$; and taking the limit as $N \to \infty$, we have the expected second moment for the instantaneous change in mutant allele frequency

$$\nu^{\text{Moran}}(p_M) = \lim_{N \to \infty} E_{\Delta\tau}\left[\frac{(P_M(\tau + \Delta\tau) - P_M(\tau))^2}{\Delta\tau}\right], \tag{C6a}$$

$$= p_M(1 - p_M) \frac{d_R \hat{\rho}_H^{\text{Moran}} + 2d_{RR}\hat{\rho}_D^{\text{Moran}}}{\hat{\rho}_H^{\text{Moran}} \hat{\rho}_D^{\text{Moran}}} \frac{1}{4\bar{d}},$$

and the expected third moment:

$$\lim_{N \to \infty} E_{\Delta\tau}\left[\frac{(P_M(\tau + \Delta\tau) - P_M(\tau))^3}{\Delta\tau}\right] = 0. \tag{C6b}$$

The diffusion equation is then described by the backward Equation 3a with the first moment from Equation C3 $[m^{\text{Moran}} = N^2(\hat{d}/\bar{d})M^{\text{Moran}}]$ and the second moment from Equation C6a. The details of these calculations are described in File S1.

Appendix D: Fixation Probability in the Haploid-Diploid WF Model

Here we apply the branching process and diffusion approximation to obtain the fixation probability of a mutant allele in the WF model.

Equilibrium

In the WF model, the offspring generation is formed by sampling randomly from the spores $\left(c_R = w_{RR}x_{RR} + \frac{w_{RM}x_{RM}}{2},\right.$ $\left. c_M = w_{MM}x_{MM} + \frac{w_{RM}x_{RM}}{2}\right)$ and zygotes $\left(c_{RR} = \frac{f_R}{2}\frac{w_R^2 x_R^2}{w_R x_R + w_M x_M}, c_{RM} = \frac{f_R + f_M}{2}\frac{w_R w_M x_R x_M}{w_R x_R + w_M x_M}, c_{MM} = \frac{f_M}{2}\frac{w_M^2 x_M^2}{w_R x_R + w_M x_M}\right)$ produced by the previous

generation of adults, who all die. Letting $q_{(GT)} = \frac{c_{(GT)}}{c_R + c_M + c_{RR} + c_{RM} + c_{MM}}$ be the probability that a reproductive cell of genotype (GT) is sampled to occupy an empty site at the recruitment step, the expectation for change in the number of individuals of genotype (GT) is:

$$E\left[\Delta X_{(GT)}(t)\right] = Nq_{(GT)} - x_{(GT)}. \tag{D1}$$

Equation D1 is based on the multinomial distribution and assumes that the number of spores and zygotes is large relative to N, so that sampling is nearly with replacement.

In a large population, the equilibrium fraction of haploids can be found deterministically by solving $Nq_{RR} - \hat{D} = 0$ and $Nq_R - \hat{H} = 0$ (recalling that $x_R = \hat{H}$, $x_{RR} = \hat{D}$, and $\hat{H} + \hat{D} = N$). The equilibrium in a resident population then satisfies $\hat{H} = \hat{\rho}_H^{WF} N$ and $\hat{D} = \hat{\rho}_D^{WF} N$, where:

$$\hat{\rho}_H^{WF} = \frac{\sqrt{w_{RR}}}{\sqrt{f_R w_R/2} + \sqrt{w_{RR}}} \text{ and } \hat{\rho}_D^{WF} = \frac{\sqrt{f_R w_R/2}}{\sqrt{f_R w_R/2} + \sqrt{w_{RR}}}. \tag{D2}$$

When we derive the fixation probability, we assume that the frequency of haploids remains near this equilibrium. Despite ignoring fluctuations in the frequency of haploids, we find that the fixation probability remains accurate (*e.g.*, Figure 1).

Branching Process Approximation

Because the branching process approximation focuses on mutant alleles while they are rare, we again only consider the change in the number of haploid mutants X_M and diploid heterozygotes X_{RM}.

When a single haploid mutant appears in the resident population, the expected fraction of offspring produced by this parent is $q_{RM|H} = 2\frac{(\bar{f}/2)w_M}{\hat{w}N}$, where the first "2" accounts for the fact that RM offspring can be formed when the mutant serves as the mother or as the father; where $\bar{f} = (f_R + f_M)/2$ is the average fertility, which gets divided by two to account for the cost of sex; and where we define $\hat{w} = \frac{f_R w_R}{2}\hat{\rho}_H^{WF} + w_{RR}\hat{\rho}_D^{WF} = \sqrt{\frac{f_R w_R w_{RR}}{2}}$. Similarly, when a single diploid mutant arises in the resident population, the expected fraction of M haploid offspring it produces is $q_{M|D} = \frac{w_{RM}}{2\hat{w}N}$, where the "2" here accounts for the segregation of M alleles into only half of the haploid spores produced by an RM diploid.

Although sampling in the WF model follows a multinomial distribution, the sampling distribution for a rare mutant is approximately Poisson in a large population (as assumed in a branching process approximation). The probabilities for loss would then satisfy:

$$1 - \pi_H = P_0\left(Nq_{RM|H}\right) + P_1\left(Nq_{RM|H}\right)(1 - \pi_D) + P_2\left(Nq_{RM|H}\right)(1 - \pi_D)^2 + \cdots, \tag{D3a}$$

$$1 - \pi_D = P_0\left(Nq_{M|D}\right) + P_1\left(Nq_{M|D}\right)(1 - \pi_H) + P_2\left(Nq_{M|D}\right)(1 - \pi_H)^2 + \cdots, \tag{D3b}$$

where $P_k(\lambda) = (\lambda^k e^{-\lambda}/k!)$ is the Poisson probability of producing k offspring. Using the Taylor series expansion for the exponential function $\left\{\exp[Nq(1 - \pi)] = \sum_{k=0}^{\infty} \frac{[Nq(1-\pi)]^k}{k!}\right\}$, we can rewrite (D3) as:

$$1 - \pi_H = \exp\left(-\pi_D Nq_{RM|H}\right) = \exp\left(-\pi_D \frac{\bar{f}w_M}{\hat{w}}\right), \tag{D4a}$$

$$1 - \pi_D = \exp\left(-\pi_H Nq_{M|D}\right) = \exp\left(-\pi_H \frac{w_{RM}}{2\hat{w}}\right). \tag{D4b}$$

Assuming that the fixation probabilities $\pi_H = \tilde{\pi}_H \varepsilon$ and $\pi_D = \tilde{\pi}_D \varepsilon$ are small under weak selection, applying the Taylor series expansion, and ignoring the third order terms, $O(\varepsilon^3)$, we have

$$\pi_H \approx \frac{f_R w_R}{\hat{w}}\pi_D + \frac{2f_R w_R s_M^w + f_R w_R s_M^f}{2\hat{w}}\pi_D - \frac{f_R^2 w_R^2}{2\hat{w}^2}\pi_D^2, \tag{D5a}$$

$$\pi_D \approx \frac{w_{RR}}{2\hat{w}}\pi_H + \frac{w_{RR}s_{RM}^w}{2\hat{w}}\pi_H - \frac{w_{RR}^2}{8\hat{w}^2}\pi_H^2. \tag{D5b}$$

Solving (D5) and dropping $O(\varepsilon^3)$ terms in the solution gives Equation 5 for the fixation probability in the WF model under weak selection using a branching process approximation.

First Moment of Change in Average Allele Frequency

As with the Moran model, we next apply a separation of timescales to describe the first moment for the change in mutant allele frequency. First, we transform the variables, using the same definitions for the Moran model $(p_M, \delta_p, \rho_H, F_D)$, except that the average allele frequency is defined as $p_M = (p_H + p_D)/2$. [Again, we set $p_M = ap_H + (1-a)p_D$ and found that the only value of the weighting term a that caused the change in allele frequency to be independent of the other variables to leading order in selection was $a = 1/2$.]

We then calculate the expected change in each variable, $f_\eta(\varepsilon, p_M, \delta_p, \rho_H, F_D)$, assuming weak selection. To constant order (no selection), we find:

$$f_{p_M}(0, p_M, \delta_p, \rho_H, F_D) = 0, \tag{D6a}$$

$$f_{\delta_p}(0, p_M, \delta_p, \rho_H, F_D) = -2\delta_p, \tag{D6b}$$

$$f_{\rho_H}(0, p_M, \delta_p, \rho_H, F_D) = \frac{2w_{RR}(1-\rho_H)^2 - f_R w_R \rho_H^2}{f_R w_R \rho_H + 2w_{RR}(1-\rho_H)}, \tag{D6c}$$

$$f_{F_D}(0, p_M, \delta_p, \rho_H, F_D) = -F_D. \tag{D6d}$$

Consequently, the system rapidly approaches the quasi-equilibrium, $\overline{\delta}_p = \overline{F}_D = 0$ and $\overline{\rho}_H = \hat{\rho}_H$ (see Equation D2), after which the system slowly changes according to selection. Taking the next term in the Taylor series $O(\varepsilon)$, we find that the change in mutant allele frequency is:

$$E[\Delta p_M(t)|\overrightarrow{X}(t) = \overrightarrow{x}] \approx M^{\mathrm{WF}}(p_M) = \frac{p_M(1-p_M)\left[2s_{\mathrm{ave}}^{\mathrm{WF}} + p_M(1-2h^w)s_{MM}^w\right]}{2}, \tag{D7}$$

which does not depend on any of the other variables, to this order. That is, the system of equations has been reduced to one variable describing selection. Equation D7 provides the drift coefficient for use in the diffusion Equation 4a. The drift in allele frequency is driven both by the average selection acting on allele M and the nonadditive effects of selection on the homozygous MM genotype once the mutant allele frequency becomes common.

Second Moment of Change in Average Allele Frequency

We next derive the diffusion coefficient for use in Equation 4a, which can be derived to leading order by ignoring selection, as in Appendix C (see supporting information in the File S1).

We assume that the numbers of both haploids and diploids are large and focus on the sampling properties of mutations within each of these two pools, ignoring the impact of the small fluctuations in the frequency of haploids around the quasi-equilibrium $\hat{H}$. Focusing first on the haploid pool, haploid offspring are sampled according to a binomial distribution, $X_M(t+1) \sim \mathrm{Bin}(q_M, \hat{H})$ with $q_M = c_M/(c_R + c_M)$. Ignoring terms of order ε, the first and second moment of change in the number of mutant individuals among the population of haploids are then:

$$E[\Delta X_M|\overrightarrow{X}(t) = \overrightarrow{x}] = E[X_M(t+1)|\overrightarrow{X}(t) = \overrightarrow{x}] - x_M = \hat{H}\delta_p, \tag{D8a}$$

$$E\left[(\Delta X_M)^2\Big|\overrightarrow{X}(t) = \overrightarrow{x}\right] = E\left[X_M(t+1)^2\Big|\overrightarrow{X}(t) = \overrightarrow{x}\right] - 2E[X_M(t+1)|\overrightarrow{X}(t) = \overrightarrow{x}]x_M + x_M^2 = \hat{H}p_D(1-p_D) + \hat{H}^2\delta_p^2, \tag{D8b}$$

where $p_H = x_M/\hat{H}$, $p_D = (x_{RM}/2 + x_{MM})/\hat{D}$, and $\delta_p = p_D - p_H$.

We next consider the pool of diploid offspring. Because the number of individuals in the next generation depends on a trinomial distribution, we have $E[X_{RM}(t+1)|\overrightarrow{X}(t) = \overrightarrow{x}] = \hat{D}q_{RM}$, $E[X_{MM}(t+1)|\overrightarrow{X}(t) = \overrightarrow{x}] = \hat{D}q_{MM}$, $E[X_{RM}(t+1)^2|\overrightarrow{X}(t) = \overrightarrow{x}] = \hat{D}q_{RM}(\hat{D}q_{RM} + 1 - q_{RM})$, and $E[X_{MM}(t+1)^2|\overrightarrow{X}(t) = \overrightarrow{x}] = \hat{D}q_{MM}(\hat{D}q_{MM} + 1 - q_{MM})$; where $q_{RM} = c_{RM}/(c_{RR} + c_{RM} + c_{MM})$ and $q_{MM} = c_{MM}/(c_{RR} + c_{RM} + c_{MM})$. Furthermore, because the random variables are correlated, we have $E[X_{RM}(t+1)X_{MM}(t+1)|\overrightarrow{X}(t) = \overrightarrow{x}] = \hat{D}(\hat{D} - 1)q_{RM}q_{MM}$.

Using these and ignoring terms of order ε, the expected values for the first and second moments of change in the number of mutant individuals among the diploid offspring pool are:

$$E[\Delta X_{RM}(t)|\overrightarrow{X}(t) = \overrightarrow{x}] = \hat{D}\{2p_H(1 - p_H) - p_{RM}\}, \tag{D9a}$$

$$E[\Delta X_{MM}(t)|\overrightarrow{X}(t) = \overrightarrow{x}] = \hat{D}(p_H^2 - p_{MM}), \tag{D9b}$$

$$E\left[(\Delta X_{RM}(t))^2\Big|\overrightarrow{X}(t) = \overrightarrow{x}\right] = \hat{D}^2(q_{RM} - p_{RM})^2 + \hat{D}q_{RM}(1 - q_{RM}), \tag{D9c}$$

$$E\left[(\Delta X_{MM}(t))^2\Big|\overrightarrow{X}(t) = \overrightarrow{x}\right] = \hat{D}^2(q_{MM} - p_{MM})^2 + \hat{D}q_{MM}(1 - q_{MM}), \tag{D9d}$$

$$E\left[(\Delta X_{RM}(t))(\Delta X_{MM}(t))|\overrightarrow{X}(t) = \overrightarrow{x}\right] = \hat{D}^2(q_{RM} - p_{RM})(q_{MM} - p_{MM}) - \hat{D}q_{RM}q_{MM}. \tag{D9e}$$

These genotypic frequencies can be used to obtain the change in mutant allele frequency among diploids by calculating the change in the random variable $Y_M(t) = X_{RM}(t)/2 + X_{MM}(t)$. From (D9), we have

$$E[\Delta Y_M(t)|\overrightarrow{X}(t) = \overrightarrow{x}] = -\hat{D}\delta_p, \tag{D10a}$$

$$E\left[(\Delta Y_M(t))^2\Big|\overrightarrow{X}(t) = \overrightarrow{x}\right] = \hat{D}\frac{p_H(1 - p_H)}{2} + \hat{D}^2\delta_p^2. \tag{D10b}$$

To consider the change in mean mutant allele frequency across the entire population (combining haploid and diploid pools), we calculate the first and second moment of change in allele frequency: $Z_M(t) = \frac{1}{2}\frac{X_M(t)}{\hat{H}} + \frac{1}{2}\frac{Y_M(t)}{\hat{D}}$ [recall that we defined the allele frequency as $(p_H + p_D)/2$ to isolate the dynamics of the allele frequency from the other variables]. Assuming the changes in mutant allele frequency in haploid and diploid populations are nearly independent in a large population $[\Delta X_M(t) \perp \Delta Y_M(t)]$, we have the first and second moments for the change in allele frequency to constant order in ε:

$$E[\Delta Z_M(t)|\overrightarrow{X}(t) = \overrightarrow{x}] \approx 0, \tag{D11a}$$

$$E\left[(\Delta Z_M(t))^2\Big|\overrightarrow{X}(t) = \overrightarrow{x}\right] \approx \frac{1}{4N}\left(\frac{p_D(1 - p_D)}{\rho_H^{WF}} + \frac{p_H(1 - p_H)}{2\rho_D^{WF}}\right). \tag{D11b}$$

We transform timescales using the time variable $\tau = t/N$ and the allele variable $P_M(\tau) = Z_M(N\tau)$ (Otto and Day 2007). Taking the limit as $N \to \infty$, we then obtain the diffusion coefficient for a neutral allele from Equation D11b, $v^{WF}(p_M) = \lim_{N \to \infty} E_{\Delta\tau}\left[\frac{(P_M(\tau + \Delta\tau) - P_M(\tau))^2}{\Delta\tau}\right]$ We also have the drift coefficient from Equation D7 $(m^{WF} = NM^{WF})$. Furthermore, because the difference in frequency of a mutant allele between the haploid and diploid pools is very small ($\delta_p \approx 0$ so that $p_M = p_D = p_H$), we have the diffusion coefficient, Equation 7b. With the drift and diffusion coefficients in hand, the fixation probability for the WF model can be readily calculated using standard methods (Karlin and Taylor 1981).