Abstract. Many theoretical models of evolution assume that all competing individuals experience the same environment. Here, we consider the more realistic scenario of evolution in heterogeneous environments. The population is of finite size, \( N \), with two competing types and a fixed number of environmental sites. The types have different reproductive rates in each site. We introduce a novel formalism to approach any form of spatial fitness heterogeneity. We first calculate the condition for natural selection to favor the mutant type relative to the resident. In large populations, where stochastic effects are negligible, we find that the mutant is favored relative to the resident if and only if the arithmetic mean of the mutant’s possible fitness values exceeds that of the resident. More importantly, however, environmental heterogeneity elucidates an interesting asymmetry between the mutant and resident types. Mutant heterogeneity suppresses fixation probability, and if strong enough, it can completely offset the effects of natural selection. In contrast, resident heterogeneity can amplify a mutant’s fixation probability if \( N \) is small and has no effect on mutant fixation probability otherwise. Our results hold for any environmental heterogeneity and selection intensity as well as a wide range of population sizes.

1. Introduction

Evolutionary dynamics is concerned with the appearance and competition of traits over time. The success of a rare mutant in a population depends on a number of factors, including the population’s spatial structure and the mutant’s reproductive fitness relative to the resident. One quantitative measure of a mutant’s success is its fixation probability, which describes the chance that the mutant’s lineage will take over the population \([1–7]\). The effect of a particular property (such as a population’s spatial structure) on natural selection is often measured directly in terms of this probability of fixation. Among the components that affect evolutionary outcomes (and, in particular, fixation probability), environmental heterogeneity in reproductive fitness has received relatively little attention in the literature on evolutionary dynamics.

Population structure is one of the more well-studied properties that can result in interaction and migration heterogeneity. Lieberman et al. \([8]\) use graphs as a model for population structure and show that “isothermal” structures do not alter fixation probabilities, expanding upon a related observation for subdivided populations \([9]\). Non-isothermal graphs can alter this fixation probability and, in particular, act as amplifiers or suppressors of selection—a topic of considerable current interest \([8, 10–19]\). Recent work suggests that randomness in dispersal patterns yields either amplifiers or suppressors of selection \([17, 20–23]\). While spatial structure and frequency-dependent fitness have been incorporated into many evolutionary models, less is known about how environmental fluctuations in fitness affect evolutionary dynamics.

In population genetics, on the other hand, the role of environmental heterogeneity has received much more attention \([21–26]\). More than sixty years ago, Levene \([27]\) introduced a diploid model in which two alleles are favored in different ecological niches and showed that genetic equilibrium is possible even when there is no niche in which the heterozygote is favored over both homozygotes. Haldane and Jayakar \([28]\) treated a temporal analogue of this fitness asymmetry, which was then incorporated into a study of polymorphism under both spatial and temporal fitness heterogeneity \([29]\). Arnold and Anderson \([30]\) elaborated on the spatial model of Levene \([27]\), which they described as “the beginning of theoretical ecological genetics.”

Other studies of environmental heterogeneity have focused largely on metapopulation or island models under weak selective pressure, inspired by the evolution of habitat-specialist traits in heterogeneous environments \([31–38]\). These metapopulation models assume connected islands (habitats) where migration is allowed between islands, and environmental heterogeneity is parametrized by a variable fitness difference between two competing types and assumed to be small (i.e. weak selection). Notably, in the limit of strong
connectivity between islands, variations in fitness advantage do not affect fixation probability [34]. Others address the issue of fixation in two-island [35] and multi-habitat [36] models with variable fitness advantage.

A more fine-grained heterogeneity requires an extension of the stepping-stone models to evolutionary graphs [39, 42]. Manem et al. [42] demonstrated via death-birth simulations on a structured mesh that heterogeneity in fitness distribution can decrease the fixation probability of a beneficial mutant. Hauser et al. [40] showed that heterogeneity in background fitness almost always reduces this fixation probability. Using an interesting analytical approach, Masuda et al. [39] estimated the scaling behavior of the average consensus time in a voter model for random environments with uniform or power-law fitness distributions.

Taylor [33] distilled much of the research into heterogeneity with the remark that “[o]ne of the key insights to emerge from population genetics theory is that the effectiveness of natural selection is reduced by random variation in individual survival and reproduction.” However, beyond the fact that the Wright-Fisher model is the standard paradigm for many of these works in population genetics, results on environmental heterogeneity often rely on assumptions such as weak selection or restrictions on population structure or migration rates. Our purpose here is to study the effects of environmental heterogeneity in a haploid Moran model for any selection intensity. This heterogeneity could be deterministic, with fitness depending on an individual’s location in a fixed way, or it could be stochastic and distributed according to a random variable. We include both types of heterogeneity here, allowing for arbitrary fitness variations between locations.

Through explicit formulas for fixation probabilities in large populations, we show that selection favors the mutant type if and only if the expected fitness of a randomly-placed mutant exceeds that of a randomly-placed resident. In other words, the mutant type is neutral relative to the resident if and only if the arithmetic mean of all possible fitness values for the mutant is the same as that of the resident. We also study this selection condition in smaller populations, showing how a combination of heterogeneity and stochasticity results in a much more complicated criterion for the mutant to be favored over the resident.

More importantly, we show that mutant heterogeneity categorically suppresses selection; in particular, any such heterogeneity decreases the fixation probability of a beneficial mutant. In contrast, heterogeneity in resident fitness does not change a mutant’s probability of fixation when the population size is sufficiently large, and it acts as an amplifier in small populations. We are unaware of other studies that establish this asymmetry between the mutant and resident types in heterogeneous environments, focusing on the evolutionary fate of a mutant rather than just the condition for a mutant to be advantageous relative to the resident. Furthermore, since we impose no restrictions on selection intensity, our results highlight behavior that is difficult to see under weak heterogeneity.

Finally, when the environment is random, a mutant’s fitness is a random variable, and thus so is its fixation probability. We derive the asymptotic distribution of a mutant’s fixation probability and use it to establish an evolutionary-dynamical analogue of the central limit theorem, showing, in particular, that the uncertainty in fixation probability is inversely proportional to the square root of the population size.

2. Model and results

We consider a well-mixed population of size N in which each individual has one of two types, A (mutant) or B (resident). In location i, the mutant type has relative fitness a_i and the resident type has relative fitness b_i. At each time step, a player is chosen for reproduction with probability proportional to relative fitness; a random neighbor then dies and is replaced by the new offspring. This update rule defines a variation of the classical Moran process [43] that accounts for environmental fitness heterogeneity (see Fig. 1).

All possible fitness values of types A and B are then captured by vectors, $a := (a_1, \ldots, a_N)$ and $b := (b_1, \ldots, b_N)$, respectively. Letting $\overline{a} := \frac{1}{N} \sum_{i=1}^{N} a_i$ and $\overline{b} := \frac{1}{N} \sum_{i=1}^{N} b_i$ be the spatial averages of the fitness values, we can define the averaged fitness vectors $\overline{a} := (\overline{a}, \ldots, \overline{a})$ and $\overline{b} := (\overline{b}, \ldots, \overline{b})$. The classical Moran process is recovered when $a = \overline{a}$ and $b = \overline{b}$, in which case there exist a and b for which $a_i = a$ and $b_i = b$ for $i = 1, \ldots, N$, and the relative fitness of the mutant relative to the resident is simply $r := a/b$.

2.1. Neutrality of fixation probabilities. In the classical Moran process, the state of the population is completely determined by the number of individuals of type A. Let $\rho_A$ be the probability that a single mutant (A) fixates when starting in a population of $N-1$ residents (B). Similarly, let $\rho_B$ be the probability that a single B fixates in a population of $N-1$ individuals of type A. A standard way to measure the evolutionary success of A relative to B is to compare $\rho_A$ to $\rho_B$. Type A is said to be favored over B if
\[ \rho_A > \rho_B, \] disfavored over \( B \) if \( \rho_A < \rho_B \), and neutral relative to \( B \) if \( \rho_A = \rho_B \). This last equation, \( \rho_A = \rho_B \), defines what we refer to as the “neutrality condition” for fixation probability.

Suppose that \( a \) and \( b \) are the fitness values of \( A \) and \( B \), respectively, in the classical Moran process. Since there is no heterogeneity in the environment, one can think of \( \rho_A = \rho_A (a, b) \) and \( \rho_B = \rho_B (a, b) \) as functions of \( a \) and \( b \). Furthermore, \( \rho_B (a, b) = \rho_A (b, a) \) since \( A \) and \( B \) are distinguished by only their fitness. Therefore, \( A \) is neutral with respect to \( B \) if and only if \( \rho_A (a, b) = \rho_A (b, a) \). Since \( \rho_A (a, b) \to 1 - \frac{b}{a} \) as \( N \to \infty \), in large populations this condition holds if and only if \( a = b \), which makes sense intuitively because then \( A \) is neutral relative to \( B \) if and only if the two types are indistinguishable from a fitness standpoint.

In the Moran process with fitness heterogeneity, the fixation probability of a single \( A \)-individual might depend on its location, so it is important to account for this initial position when considering an analogue of the neutrality condition. Let \( \rho_A (a, b) \) be the fixation probability of an \( A \)-individual, averaged over all \( N \) initial locations of the mutant. A natural extension the comparison of \( \rho_A (a, b) \) to \( \rho_A (b, a) \) is the comparison of \( \rho_A (a, b) \) to \( \rho_A (b, a) \). If \( N \) is large, then the neutrality condition for the heterogeneous process is

\[ \rho_A (a, b) = \rho_A (b, a) \iff \bar{\sigma} = \bar{b} \tag{1} \]

(see Supporting Information). That is, \( A \) is neutral relative to \( B \) if and only if the expected fitness of a randomly-placed mutant is the same as the expected fitness of a randomly-placed resident. If there is no fitness heterogeneity and \( a_i = a \) and \( b_i = b \) for each \( i = 1, \ldots, N \), then this condition reduces to \( a = b \).

For smaller populations, the neutrality condition is not necessarily as simple as Eq. (1). When \( N = 2 \), we have \( \rho_A (a, b) = \rho_A (b, a) \) if and only if \( \sqrt{a_1 a_2} = \sqrt{b_1 b_2} \), i.e. if and only if the geometric mean of the fitness values for \( A \) equals the geometric mean of the fitness values for \( B \). For small \( N > 2 \), and actually already for \( N = 3 \), this condition is much more complicated. However, that Eq. (1) does not give the neutrality condition for smaller \( N \) is not surprising; even for the classical Moran process with just two parameters, \( a \) and \( b \), this condition is more complicated than the equation \( a = b \) (see Supporting Information). Fig. 2 illustrates the convergence of \( \rho_A (a, b) \) to \( \rho_A (b, a) \) as the population size gets large when \( \bar{\sigma} = \bar{b} \).

From the neutrality condition for large populations, we also obtain conditions for selection to favor (resp. disfavor) the mutant type: \( A \) is favored relative to \( B \) if and only if \( \bar{\sigma} > \bar{b} \), and \( A \) is disfavored relative to \( B \) if and only if \( \bar{\sigma} < \bar{b} \). Therefore, the performance of one type relative to another can be deduced from the classical (homogeneous) model by replacing each location’s fitness values, \( a_i \) and \( b_i \), by the spatial averages,
The difference in fixation probabilities in a random environment, \( \rho_A(a, b) - \rho_A(b, a) \), as a function of population size, \( N \). The fitness values for \( A \) and \( B \) at each location, \( a_i \) and \( b_i \), are chosen from uniform distributions with equal means. Each point corresponds to a fixed fitness configuration. For proper comparison between different population sizes, the fixation probability is normalized by \( N \). As \( N \) increases, \( \rho_A(a, b) \) converges to \( \rho_A(b, a) \), and the bounding envelope between these two values is inversely proportional to \( \sqrt{N} \).

\[ \bar{a} = \bar{b} = 1 \]

\[ N (\rho_A - \rho_B) \]

\[ \begin{array}{c}
\text{difference in fixation probabilities} \\
n \text{population size} \\
\pm 1 \\
p_N \bar{a} = \bar{b} = 1 \\
N \left( \rho_A (a, b) \right)
\end{array} \]

**Figure 2.** The difference in fixation probabilities in a random environment, \( \rho_A(a, b) - \rho_A(b, a) \), as a function of population size, \( N \). The fitness values for \( A \) and \( B \) at each location, \( a_i \) and \( b_i \), are chosen from uniform distributions with equal means. Each point corresponds to a fixed fitness configuration. For proper comparison between different population sizes, the fixation probability is normalized by \( N \). As \( N \) increases, \( \rho_A(a, b) \) converges to \( \rho_A(b, a) \), and the bounding envelope between these two values is inversely proportional to \( \sqrt{N} \).

\( \bar{a} \) and \( \bar{b} \). Although one can make a rough comparison of two types by looking at their mean fitness values, we show in the next section that mutant heterogeneity furthermore acts as a suppressor of selection.

### 2.2. Mutant heterogeneity suppresses selection.

If \( a \) and \( b \) are the fitness vectors for a finite population, then the fraction of each fitness value present in these vectors defines density functions, \( f(a) \) and \( g(b) \). Thus, for any population size, one can speak of the fixation probability of a randomly-placed mutant when the mutant and resident fitness values are distributed according to \( f \) and \( g \), respectively. Let \( \bar{a} = \sum_{a} a f(a) \) and \( \bar{b} = \sum_{b} b g(b) \) be the mean fitness values of the mutant type and the resident type, respectively. The limiting (or “infinite”) fixation probability, \( \rho^\infty_A(f, g) \), satisfies the equation(s)

\[ \rho^\infty_A(f, g) = 0 \quad \text{if} \quad \bar{a} \leq \bar{b}; \quad (2a) \]

\[ \int_{0}^{\infty} \frac{a f(a)}{\bar{b} + a \rho^\infty_A(f, g)} \, da = 1 \quad \text{if} \quad \bar{a} > \bar{b}. \quad (2b) \]

Moreover, provided \( N \) is large, we can approximate \( \rho_A \) when \( \bar{a} > \bar{b} \) with a finite-population correction,

\[ \rho_A(a, b) \approx \frac{\rho^\infty_A(f, g)}{1 - \prod_{i=1}^{N} \left( 1 + \frac{a_i}{\bar{b}} \rho^\infty_A(f, g) \right)^{-1}}. \quad (3) \]

Together with Eq. 2a, this correction gives an (approximate) expression for the fixation probability in large, heterogeneous populations. The proofs of these equations are in Supporting Information.

The results can be expanded in moments of the fitness distribution, \( a \). If \( \rho_A(\bar{a}, \bar{b}) \) is the fixation probability in the uniform (homogeneous) system, and if \( E_f \) denotes the expectation with respect to the distribution.
Figure 3. Fixation probability of the mutant type, $A$, as a function of the width of the mutant fitness distribution, $\Delta_a$. The fitness values for the mutant and resident are uniformly distributed on $[\bar{\alpha} - \Delta_a, \bar{\alpha} + \Delta_a]$ and $[\bar{\beta} - \Delta_b, \bar{\beta} + \Delta_b]$, respectively (solid line). Similarly, for a bi-modal distribution, the fitness values for the mutant and resident are $\bar{\alpha} - \Delta_a$ or $\bar{\alpha} + \Delta_a$ and $\bar{\beta} - \Delta_b$ and $\bar{\beta} + \Delta_b$, respectively, with each option being chosen with probability $1/2$ (dashed lines). These widths, $\Delta_a$ and $\Delta_b$, are measures of mutant and resident heterogeneity, respectively. The population size is $N = 50$ and the solid/dashed lines indicate the analytical predictions from Eq. 4. As $\Delta_a$ grows, a beneficial mutant’s fixation probability decreases. However, this fixation probability does not change as $\Delta_b$ varies (not shown).

For a uniform probability distribution, $\bar{\alpha} - \Delta_a \leq a \leq \bar{\alpha} + \Delta_a$, the second and fourth moments are $\mathbb{E}_f [(a - \bar{\alpha})^2] = \Delta_a^2 / 3$ and $\mathbb{E}_f [(a - \bar{\alpha})^4] = \Delta_a^4 / 5$, and the third moment (skewness) vanishes. The results are plotted in Fig. 3 which shows that our analytical results are in excellent agreement with simulation data for any selection strength and fitness heterogeneity.
An upshot of Eqs. 2a–3 is that \( \rho_A(a, b) = \rho_A(\pi, b) \), so environmental heterogeneity of the resident \((B)\) does not affect the fixation probability of the mutant \((A)\) when \(N\) is sufficiently large. On the other hand, we can also deduce from Eqs. 2a–3 (see Supporting Information) that \( \rho_A(a, b) \leq \rho_A(\pi, b) \) when \( \pi \geq b \) and \( \rho_A(a, b) \geq \rho_A(\pi, b) \) when \( \pi < b \); in both cases, equality holds if and only if \( a = \pi \) (there is no mutant heterogeneity) or \( \pi = b \) \((A\) is neutral relative to \(B)\), so heterogeneity suppresses selection (Fig. 3).

2.3. Resident heterogeneity can amplify selection. While environmental heterogeneity of the resident is irrelevant when the population size is sufficiently large, it can have an effect on fixation probability for small population sizes. In most cases, this effect (which is of order \(1/N\)) can be ignored, but we observe that for small population sizes, and in particular near neutrality \((\pi = \bar{b})\), heterogeneity in resident fitness values can amplify a mutant’s fixation probability. One example of this amplification effect is presented in Fig. 4 where \(\pi\) is close to 1 and \(\mathbf{b}\) is distributed uniformly on \(\bar{b} - \Delta_b \leq b \leq \bar{b} + \Delta_b\) with \(\bar{b} = 1\). A second, bi-modal distribution is also tested, with fitness values randomly chosen from two values, \(\bar{b} - \Delta_b\) or \(\bar{b} + \Delta_b\). In both cases, we observe that fixation probability is increased for near neutral mutants. However, fixation probability is increased for both on-average beneficial and on-average deleterious mutations, which indicates that the mechanism of amplification is somewhat different from that of an amplifier of selection on evolutionary graphs (for example, a star graph).

2.4. Random environments. In the previous section, we observed that fitness vectors, \(\mathbf{a}\) and \(\mathbf{b}\), define probability density functions, \(f\) and \(g\), each concentrated on a finite number of fitness values. Here, we reverse this process and assume that the fitness vectors themselves are uncertain and obtained from probability distributions given by density functions \(f\) and \(g\). In other words, for each location, \(i\), the probability that \(a_i \leq a_i \leq A\) is \(\int_a^A f(a)\, da\) and the probability that \(b_i \leq b_i \leq B\) is \(\int_b^B g(b)\, db\) (independent of \(i\)).

For brevity, we assume that \(b_i\) is constant—and, in fact, equal to 1 for each \(i = 1, \ldots, N\). Perhaps the most natural question is how the stochasticity of (mutant) fitness values affects fixation probability. Since the probability of fixation depends on fitness, which is stochastic, fixation probability is itself a random variable. A natural question, then, arises: what is the distribution of this random variable?

In fact, as \(N\) gets large, the probability that \(x \leq \rho \leq y\) approaches \(\int_x^y d(\rho)\, d\rho\), where \(d\) is the density

\[
d(\rho) = \frac{1}{\sqrt{2\pi \text{Var}_f}} \exp \left\{ -\frac{N \left( 1 - \text{E}_f \left[ \frac{a}{1+\alpha_\rho} \right] \right)^2}{2 \text{Var}_f \left[ \frac{a}{1+\alpha_\rho} \right]} \right\}
\]

(see Supporting Information). By Eq. 2a, \(d(\rho)\) is maximized at \(\rho = \rho_A^\ast\). From Eq. 5 we can relate the uncertainty (standard deviation) in mutant fitness, \(\sigma_a\), to the uncertainty in fixation probability, \(\sigma_\rho\). Specifically, if \(\text{Var}_f[a]\) is the variance of the fitness distribution of the mutant type, then

\[
\sigma_\rho := \sqrt{\text{Var}_d[\rho]} \propto \frac{1}{\sqrt{N}} \sqrt{\text{Var}_f[a]} = \frac{1}{\sqrt{N}} \sigma_a.
\]

The above result may be thought of as an evolutionary-dynamical analogue of the central limit theorem. In particular, the uncertainty in fixation probability is inversely proportional to \(\sqrt{N}\) (see Fig. 2).

3. Discussion

In heterogeneous environments, there is a notable asymmetry between the mutant and resident types. Any variation in mutant fitness acts as a suppressor of selection. In particular, mutant heterogeneity decreases the fixation probability of beneficial mutants and increases the fixation probability of deleterious mutants. Resident heterogeneity, on the other hand, has no effect on a mutant’s fixation probability in large populations and can even amplify it in small populations. Our finding differs from what is seen in processes with dispersal heterogeneity, which can amplify or suppress selection but need not do either [19, 23, 46].

While the neutrality condition of Eq. 1 namely \(\pi = \bar{b}\) admits a simple interpretation when the population is large (i.e. the types have the same expected fitness), we do not expect this condition to be quite as intuitive for smaller population sizes. This issue arises already in populations without heterogeneity and is compounded further when there is spatial variation in resources (see Supporting Information). For
$\rho_A(a, b)$

**Figure 4.** Fixation probability of the mutant type, $A$, as a function of the width of the resident fitness distribution width, $\Delta_b$. The fitness values for the resident are uniformly distributed on $[\bar{b} - \Delta_b, \bar{b} + \Delta_b]$ (solid line). Similarly, for a bi-modal distribution, the fitness values for the resident are either $\bar{b} - \Delta_b$ or $\bar{b} + \Delta_b$, each chosen with probability 1/2 (dashed lines). The population size is $N = 10$ and $\bar{a} = 0.8, 0.9, 1.0$ and 1.1. The results are obtained from exact solutions of the Kolmogorov equation for the fixation probability (see Supporting Information). As $\Delta_b$ grows, a near-neutral mutant’s fixation probability increases.

smaller $N$, stochastic effects are stronger and the neutrality condition is complicated by the interplay between natural selection and random drift; in large populations, selection becomes the primary effect.

Other studies of environmental heterogeneity have also produced conditions in terms of fitness “means.” In a two-allele model with ecological variation, the condition for the maintenance of a protected polymorphism is stated in terms of the harmonic mean of the fitness values [27]. If fitness heterogeneity is temporal rather than environmental [47], then the mean in this condition is geometric [28]. The approach we take here is somewhat different from these studies since we are focused on the contrast between two types under environmental heterogeneity. Furthermore, we treat a haploid Moran model rather than a diploid model with random mating, which is a biologically meaningful framework that has not been studied as extensively as the Wright-Fisher model in population genetics. Nonetheless, it is intriguing that the condition for a mutant to be favored over the resident also involves a mean (arithmetic) of the fitness values.

Heterogeneity, in its many and varied forms, is commonplace in evolving populations. Our focus here is on environmental fitness heterogeneity that can arise, for example, from spatial fluctuations in the availability of resources. Although mutant heterogeneity always suppresses selection and resident heterogeneity can amplify selection, it would be interesting to understand its interaction with other asymmetries such as those induced
by spatial structure. In particular, how a combination of fitness and migration heterogeneity influences selection is poorly understood and represents an interesting topic for future research.

ACKNOWLEDGMENTS

This work was supported by the Office of Naval Research (grant N00014-16-1-2914). The Program for Evolutionary Dynamics is supported in part by a gift from B Wu and Eric Larson.

REFERENCES

[1] T. Nagylaki. *Introduction to Theoretical Population Genetics*. Springer Berlin Heidelberg, 1992. doi: 10.1007/978-3-642-76214-7.

[2] R. Durrett. *Probability Models for DNA Sequence Evolution*. Springer New York, 2002. doi: 10.1007/978-1-4757-6285-3.

[3] M. A. Nowak, A. Sasaki, C. Taylor, and D. Fudenberg. Emergence of cooperation and evolutionary stability in finite populations. *Nature*, 428(6983):646–650, Apr 2004. doi: 10.1038/nature02414.

[4] M. A. Nowak. *Evolutionary Dynamics: Exploring the Equations of Life*. Belknap Press, 2006.

[5] M. A. Nowak, C. E. Tarnita, and T. Antal. Evolutionary dynamics in structured populations. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1537):19–30, Nov 2009. doi: 10.1098/rstb.2009.0215.

[6] A. Traulsen and C. Hauert. Stochastic Evolutionary Game Dynamics. In *Reviews of Nonlinear Dynamics and Complexity*, pages 25–61. Wiley-Blackwell, Jun 2010. doi: 10.1002/9783527628001.ch2.

[7] M. Broom and J. Rychtář. *Game-Theoretical Models in Biology*. Taylor & Francis, 2013.

[8] E. Lieberman, C. Hauert, and M. A. Nowak. Evolutionary dynamics on graphs. *Nature*, 433(7023):312–316, Jan 2005. doi: 10.1038/nature03204.

[9] T. Maruyama. A simple proof that certain quantities are independent of the geographical structure of population. *Theoretical Population Biology*, 5(2):148–154, Apr 1974. doi: 10.1016/0040-5809(74)90037-9.

[10] H. Ohtsuki and M. A. Nowak. Evolutionary games on cycles. *Proceedings of the Royal Society B: Biological Sciences*, 273(1598):2249–2256, Sep 2006. doi: 10.1098/rspb.2006.3576.

[11] M. Broom and J. Rychtář. An analysis of the fixation probability of a mutant on special classes of non-directed graphs. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 464(2098):2609–2627, Oct 2008. doi: 10.1098/rspa.2008.0058.

[12] Z. Patwa and L. M. Wahl. The fixation probability of beneficial mutations. *Journal of The Royal Society Interface*, 5(28):1279–1289, Nov 2008. doi: 10.1098/rsif.2008.0248.

[13] M. Broom, J. Rychtář, and B. Stadler. Evolutionary dynamics on small-order graphs. *Journal of Interdisciplinary Mathematics*, 12(2):129–140, Apr 2009. doi: 10.1080/09720502.2009.10700618.

[14] B. Houchmandzadeh and M. Vallade. The fixation probability of a beneficial mutation in a geographically structured population. *New Journal of Physics*, 13(7):073020, Jul 2011. doi: 10.1088/1367-2630/13/7/073020.

[15] G. B. Mertzios, S. Nikoletseas, C. Raptopoulos, and P. G. Spirakis. Natural models for evolution on networks. *Theoretical Computer Science*, 477:76–95, Mar 2013. doi: 10.1016/j.tcs.2012.11.032.

[16] T. Monk, P. Green, and M. Paulin. Martingales and fixation probabilities of evolutionary graphs. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 470(2165):20130730–20130730, Mar 2014. doi: 10.1098/rspa.2013.0730.

[17] B. Adlam and M. A. Nowak. Universality of fixation probabilities in randomly structured populations. *Scientific Reports*, 4:6692, Oct 2014. doi: 10.1038/srep06692.

[18] B. Allen, C. Sample, Y. Dementieva, R. C. Medeiros, C. Paoletti, and M. A. Nowak. The Molecular Clock of Neutral Evolution Can Be Accelerated or Slowed by Asymmetric Spatial Structure. *PLOS Computational Biology*, 11(2):e1004108, Feb 2015. doi: 10.1371/journal.pcbi.1004108.

[19] B. Adlam, K. Chatterjee, and M. A. Nowak. Amplifiers of selection. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science*, 471(2181):20150114, Sep 2015. doi: 10.1098/rspa.2015.0114.
[20] T. Antal, S. Redner, and V. Sood. Evolutionary Dynamics on Degree-Heterogeneous Graphs. *Physical Review Letters*, 96(18), May 2006. doi: 10.1103/physrevlett.96.188104.

[21] V. Sood, T. Antal, and S. Redner. Voter models on heterogeneous networks. *Physical Review E*, 77(4), Apr 2008. doi: 10.1103/physreve.77.041121.

[22] V. S. K. Manem, M. Kohandel, N. L. Komarova, and S. Sivaloganathan. Spatial invasion dynamics on random and unstructured meshes: Implications for heterogeneous tumor populations. *Journal of Theoretical Biology*, 349:66–73, May 2014. doi: 10.1016/j.jtbi.2014.01.009.

[23] L. Hindersin and A. Traulsen. Most Undirected Random Graphs Are Amplifiers of Selection for Birth-Death Dynamics, but Suppressors of Selection for Death-Birth Dynamics. *PLOS Computational Biology*, 11(11):e1004437, Nov 2015. doi: 10.1371/journal.pcbi.1004437.

[24] J. H. Gillespie. *The Causes of Molecular Evolution*. Oxford University Press, 1991.

[25] N. H. Barton. The probability of fixation of a favoured allele in a subdivided population. *Genetical Research*, 62(02):149, Oct 1993. doi: 10.1017/s0016672300031748.

[26] N. H. Barton. Linkage and the limits to natural selection. *Genetics*, 140(2):821–841, Jun 1995.

[27] H. Levene. Genetic Equilibrium When More Than One Ecological Niche is Available. *The American Naturalist*, 87(836):331–333, Sep 1953. doi: 10.1086/281792.

[28] J. B. S. Haldane and S. D. Jayakar. Polymorphism due to selection of varying direction. *Journal of Genetics*, 58(2):237–242, Mar 1963. doi: 10.1007/bf02986143.

[29] E. P. Ewing. Genetic Variation in a Heterogeneous Environment VII. Temporal and Spatial Heterogeneity in Infinite Populations. *The American Naturalist*, 114(2):197–212, Aug 1979. doi: 10.1086/283468.

[30] J. Arnold and W. W. Anderson. Density-Regulated Selection in a Heterogeneous Environment. *The American Naturalist*, 121(5):656–668, May 1983. doi: 10.1086/284093.

[31] R. Levins. Theory of Fitness in a Heterogeneous Environment. I. The Fitness Set and Adaptive Function. *The American Naturalist*, 96(891):361–373, Nov 1962. doi: 10.1086/282245.

[32] R. Levins. Theory of Fitness in a Heterogeneous Environment. II. Developmental Flexibility and Niche Selection. *The American Naturalist*, 97(893):75–90, Mar 1963. doi: 10.1086/282258.

[33] E. Pollak. On the Survival of a Gene in a Subdivided Population. *Journal of Applied Probability*, 3(1):142, Jun 1966. doi: 10.2307/3212043.

[34] T. Nagylaki. The strong-migration limit in geographically structured populations. *Journal of Mathematical Biology*, 9(2):101–114, Apr 1980. doi: 10.1007/bf0275916.

[35] S. Gavrilets and N. Gibson. Fixation probabilities in a spatially heterogeneous environment. *Population Ecology*, 44(2):51–58, Aug 2002. doi: 10.1007/s101440200007.

[36] M. C. Whitlock and R. Gomulkiewicz. Probability of Fixation in a Heterogeneous Environment. *Genetics*, 171(3):1407–1417, Aug 2005. doi: 10.1534/genetics.104.040089.

[37] S. J. Schreiber and J. O. Lloyd-Smith. Invasion Dynamics in Spatially Heterogeneous Environments. *The American Naturalist*, 174(4):490–505, Oct 2009. doi: 10.1086/605405.

[38] S. Vuilleumier, J. Goudet, and N. Perrin. Evolution in heterogeneous populations: From migration models to fixation probabilities. *Theoretical Population Biology*, 78(4):250–258, Dec 2010. doi: 10.1016/j.tpb.2010.08.004.

[39] N. Masuda, N. Gibert, and S. Redner. Heterogeneous voter models. *Physical Review E*, 82(1), Jul 2010. doi: 10.1103/physreve.82.010103.

[40] O. P. Hauser, A. Traulsen, and M. A. Nowak. Heterogeneity in background fitness acts as a suppressor of selection. *Journal of Theoretical Biology*, 343:178–185, Feb 2014. doi: 10.1016/j.jtbi.2013.10.013.

[41] W. Maciejewski and G. J. Puleo. Environmental evolutionary graph theory. *Journal of Theoretical Biology*, 360:117–128, Nov 2014. doi: 10.1016/j.jtbi.2014.06.040.

[42] V. S. K. Manem, K. Kaveh, M. Kohandel, and S. Sivaloganathan. Modeling Invasion Dynamics with Spatial Random-Fitness Due to Micro-Environment. *PLOS ONE*, 10(10):e0140234, Oct 2015. doi: 10.1371/journal.pone.0140234.

[43] J. Taylor. The Common Ancestor Process for a Wright-Fisher Diffusion. *Electronic Journal of Probability*, 12(0):808–847, 2007. doi: 10.1214/ejp.v12-418.

[44] P. A. P. Moran. Random processes in genetics. *Mathematical Proceedings of the Cambridge Philosophical Society*, 54(01):60, Jan 1958. doi: 10.1017/s0305004100033193.
[45] C. E. Tarnita, H. Ohtsuki, T. Antal, F. Fu, and M. A. Nowak. Strategy selection in structured populations. *Journal of Theoretical Biology*, 259(3):570–581, Aug 2009. doi: 10.1016/j.jtbi.2009.03.035.

[46] A. Pavlogiannis, J. Tkadlec, K. Chatterjee, and M. A. Nowak. Amplification on Undirected Population Structures: Comets Beat Stars. *Scientific Reports*, 7(1), Mar 2017. doi: 10.1038/s41598-017-00107-w.

[47] I. Cvijović, B. H. Good, E. R. Jerison, and M. M. Desai. Fate of a mutation in a fluctuating environment. *Proceedings of the National Academy of Sciences*, 112(36):E5021–E5028, Aug 2015. doi: 10.1073/pnas.1505406112.
Supporting Information

SI.1. Fixation probabilities in heterogeneous environments

Here, we use what could be called an “approximate Martingale” technique to derive fixation probabilities in heterogeneous environments. Suppose that the population size is \( N \) and that the fitness values for \( A \) and \( B \) at each location are given by \( a = (a_1, \ldots, a_N) \) and \( b = (b_1, \ldots, b_N) \), respectively. If there are \( M \) distinct environmental conditions (or “sites”) of size \( N_1, \ldots, N_M \), then the state of the population is given by an \( M \)-tuple, \( \mathbf{n} = (n_1, \ldots, n_M) \), where \( n_i \in [0, N_i] \) is the number of mutants in site \( i \). Clearly, \( N_1 + \cdots + N_M = N \). When \( M = 2 \), there are only two different environmental conditions (also known as “patches”).

Since we discuss what happens to fixation probabilities as the population size grows, we need a way to parametrize the population by only its size, \( N \). Therefore, we assume that there are \( M \) patches with \( A \)-fitness given by \( a_1, \ldots, a_M \) and \( B \)-fitness given by \( b_1, \ldots, b_M \). The number of sites of type \( i \), \( N_i \), is a function of \( N \) so that \( N_1 (N) + \cdots + N_M (N) = N \) for each \( N \geq 1 \). Since the dispersal structure is the same as that of a well-mixed population, \( N \) then completely specifies the population structure and fitness heterogeneity.

For \( i = 1, \ldots, M \), let \( P_i^+ (\mathbf{n}) \) (resp. \( P_i^- (\mathbf{n}) \)) be the probability that the number of mutants in site \( i \) goes up (resp. down) after being in state \( \mathbf{n} \). In other words, \( P_i^+ (\mathbf{n}) \) is the probability that \( n_i \) becomes \( n_i + 1 \) and \( P_i^- (\mathbf{n}) \) is the probability that \( n_i \) becomes \( n_i - 1 \). By the definition of the Moran process,

\[
P_i^- (\mathbf{n}) = \left( \frac{\sum_{j=1}^{M} b_j (N_j - n_j)}{\sum_{j=1}^{M} (a_j n_j + b_j (N_j - n_j))} \right) \left( \frac{n_i}{N} \right); \quad (SI.1a)
\]

\[
P_i^+ (\mathbf{n}) = \left( \frac{\sum_{j=1}^{M} a_j n_j}{\sum_{j=1}^{M} (a_j n_j + b_j (N_j - n_j))} \right) \left( \frac{N_i - n_i}{N} \right). \quad (SI.1b)
\]

Denote by \( S = [0, N_1] \times \cdots \times [0, N_M] \) the state space of the process and suppose that \( M \) is the transition matrix for the corresponding Markov chain. Let \( S^\infty := S - \{ \mathbf{A}, \mathbf{B} \} \), where \( \mathbf{A} \) and \( \mathbf{B} \) denote states where all individuals are of type \( A \) and \( B \), respectively. Let \( M = \begin{pmatrix} I & 0 \\ R & Q \end{pmatrix} \) be a partition of the Markov chain, where \( Q \) is a substochastic transition matrix on \( S^\infty \). Since all non-absorbing states are transient, we have

\[
M^\infty := \lim_{i \to \infty} M^i = \begin{pmatrix} I & 0 \\ FR & 0 \end{pmatrix}, \quad (SI.2)
\]

where \( F := (1 - Q)^{-1} \) is the so-called “fundamental matrix” of the chain.

**Lemma 1.** For each \( \mathbf{n}, \mathbf{n}' \in S^\infty \), we have \( \lim_{N \to \infty} \frac{1}{N} F_{\mathbf{n}, \mathbf{n}'} < \infty \).

**Proof.** For fixed \( \mathbf{n} \in S^\infty \) and \( i \in \{1, \ldots, N\} \), we have

\[
F_{\mathbf{n}, \mathbf{e}_i, \mathbf{B}} = \sum_{\mathbf{n}' \in S^\infty} F_{\mathbf{n}, \mathbf{n}'} Q_{\mathbf{n}', \mathbf{B}} = \rho_{\mathbf{n}, \mathbf{B}}. \quad (SI.3)
\]

Since \( Q_{\mathbf{e}_i, \mathbf{B}} = P_i^- (\mathbf{e}_i) \), it follows at once that \( \lim_{N \to \infty} \frac{1}{N} F_{\mathbf{n}, \mathbf{e}_i, \mathbf{B}} < \infty \). Suppose now that \( \lim_{N \to \infty} \frac{1}{N} F_{\mathbf{n}, \mathbf{n}'} < \infty \) for some \( \mathbf{n}' \) and let \( \mathbf{n}'' \neq \mathbf{n}' \) be such that \( Q_{\mathbf{n}'', \mathbf{n}'} \neq 0 \) (i.e. \( |\mathbf{n}'' - \mathbf{n}'| = 1 \)). Since \( N (I - Q) = I \), we have

\[
F_{\mathbf{n}, \mathbf{n}''} Q_{\mathbf{n}'', \mathbf{n}'} \leq (1 - Q_{\mathbf{n}'', \mathbf{n}'}) F_{\mathbf{n}, \mathbf{n}'} - \delta_{\mathbf{n}, \mathbf{n}'}. \quad (SI.4)
\]

By the hypothesis and the transition probabilities for the Moran process, \( \lim_{N \to \infty} \frac{1}{N} F_{\mathbf{n}, \mathbf{n}'} < \infty \). Therefore, by induction, it follows that \( \lim_{N \to \infty} \frac{1}{N} F_{\mathbf{n}, \mathbf{n}'} < \infty \) for any fixed \( \mathbf{n}, \mathbf{n}' \in S^\infty \), which completes the proof. \( \square \)

For \( z_1, \ldots, z_M \in \mathbb{R} \), let \( \mathbf{z} \) be the vector defined by \( \mathbf{z}_n := z_1^n \cdots z_M^n \). It is clear that the two absorbing states, \( \mathbf{A} \) and \( \mathbf{B} \), satisfy \( (M \mathbf{z})_A = \mathbf{z}_A \) and \( (M \mathbf{z})_B = \mathbf{z}_B \), respectively. Let \( v = \left( \frac{\mathbf{z}_A}{\mathbf{z}_B} \right) \) and let \( w \) consist of the remaining entries of \( \mathbf{z} \) indexed by \( S^\infty \). From the equations \( (M \mathbf{z})_A = \mathbf{z}_A \) and \( (M \mathbf{z})_B = \mathbf{z}_B \), we have

\[
e_i^T M^\infty \mathbf{z} = z_i + e_i^T F (R v - (I - Q) w). \quad (SI.5)
\]
With \( f_i (N) := e_i^T F (Rv) - (I - Q) w \), our goal is to show that \( \lim_{N \to \infty} f_i (N) = 0 \) for appropriately chosen \( z_1, \ldots, z_M \). It is in this sense that we consider an “approximate Martingale.” For \( i = 1, \ldots, N \), let
\[
z_i^* = -\left( \frac{a_i}{N} \right) (1 - \frac{a_i}{N}) + \sqrt{\left( \frac{b_i}{N} \right) (1 - \frac{a_i}{N})^2 - \frac{4}{N} \frac{a_i}{N} \frac{1}{N} \frac{b_i}{N} - \frac{b_i}{N}}.
\]
where \( \frac{a_i}{N} \neq 1 \) satisfies the equation
\[
z_i^* = \frac{1}{N} \sum_{i=1}^{N} \left( \frac{a_i}{N} \right) (1 - \frac{a_i}{N}) + \sqrt{\left( \frac{b_i}{N} \right) (1 - \frac{a_i}{N})^2 - \frac{4}{N} \frac{a_i}{N} \frac{1}{N} \frac{b_i}{N} - \frac{b_i}{N}}.
\]

**Lemma 2.** For this choice of \( z_1^*, \ldots, z_M^* \) and for \( n \in S \), we have \( \lim_{N \to \infty} N (Rv^* - (I - Q) w^*)_n = 0 \).

**Proof.** By Eq. [SI.1] we see that \( \lim_{N \to \infty} N P_i^- (n) \) exists for each \( n \) and that
\[
\lim_{N \to \infty} N P_i^- (n) = \left( \lim_{N \to \infty} N P_i^- (e_i) \right) n_i;
\]
\[
\lim_{N \to \infty} N P_i^+ (n) = \sum_{k=1}^{M} \left( \lim_{N \to \infty} N P_i^+ (e_k) \right) n_k.
\]

Since the Moran process proceeds one birth event at a time,
\[
N (Rv^* - (I - Q) w^*)_n = \sum_{i=1}^{M} N P_i^- (n) (z_1^*)^{n_1} \cdots (z_i^*)^{n_i - 1} \cdots (z_M^*)^{n_M}
+ \sum_{i=1}^{M} N P_i^+ (n) (z_1^*)^{n_1} \cdots (z_i^*)^{n_i + 1} \cdots (z_M^*)^{n_M}
- \left( \sum_{i=1}^{M} N P_i^- (n) + \sum_{i=1}^{M} N P_i^+ (n) \right) (z_1^*)^{n_1} \cdots (z_M^*)^{n_M}.
\]

On the other hand, \( z_1^*, \ldots, z_M^* \) are defined precisely to ensure that for \( k = 1, \ldots, M \),
\[
N (Rv^* - (I - Q) w^*)_e_k = NP_i^- (e_k) + \sum_{i=1}^{M} N P_i^- (e_k) z_i^* - \left( \sum_{i=1}^{M} N P_i^- (e_k) \right) z_k^* = 0.
\]

Combining the limits of Eq. [SI.8] with Eqs. [SI.9][SI.10] gives \( \lim_{N \to \infty} N (Rv^* - (I - Q) w^*)_n = 0 \).

**Lemma 3.** For this choice of \( z_1^*, \ldots, z_M^* \), we have \( \lim_{N \to \infty} e_i^T F (Rv^* - (I - Q) w^*) = 0 \).

**Proof.** Since \((Mz)_e_j = z_j\) for each \( j = 1, \ldots, N \), we have
\[
e_i^T F (Rv^* - (I - Q) w^*) = \sum_{n} F_{e_i, n} (Rv^* - (I - Q) w^*)_n
= \sum_{n \neq e_1, \ldots, e_N} F_{e_i, n} (Rv^* - (I - Q) w^*)_n
= \sum_{n \neq e_1, \ldots, e_N} \left( \frac{1}{N} F_{e_i, n} \right) (N (Rv^* - (I - Q) w^*)_n).
\]

By the previous lemma and the fact that \((z_i^*)^{n_1} \cdots (z_M^*)^{n_M} \to 0\) as \( n_1 + \cdots + n_M \to \infty \), we are done.

Since all non-absorbing states are transient, we also know that
\[
e_i^T M^\infty z = \rho_{e_i, A} z_A + \rho_{e_i, B} z_B
= \rho_{e_i, A} z_A + (1 - \rho_{e_i, A}) z_B.
\]
Therefore, using the equation $e_z^i M^N z = z_i + f_i(N)$ and solving for $\rho_{e_i,A}$, we find that

$$ \rho_{e_i,A}(N) = \frac{1 - z_i - f_i(N)}{z_B - z_A} - \frac{1 - z_i^*}{1 - (z_i^*)^{N_1} \cdots (z_M^*)^{N_M}} - \frac{f_i(N)}{1 - (z_i^*)^{N_1} \cdots (z_M^*)^{N_M}}. \quad (SI.13) $$

In particular, since $\lim_{N \to \infty} f_i(N) = 0$, we have

$$ \rho_{e_i,A}(N) \approx \frac{1 - z_i^*}{1 - (z_i^*)^{N_1} \cdots (z_M^*)^{N_M}} \quad (SI.14) $$

given whenever $N$ is sufficiently large. As $N \to \infty$, we obtain $\rho_{e_i,A}^* := \lim_{N \to \infty} \rho_{e_i,A}(N) = 1 - z_i^*$.

**SI.2. Neutrality conditions**

Consider the expected fixation probability of a randomly-placed mutant,

$$ \rho_A(a,b) := \frac{1}{N} \sum_{i=1}^{M} N_i \rho_{e_i,A}(a,b). \quad (SI.15) $$

The neutrality condition, which says when $A$ is neutral relative to $B$, is defined by the equation $\rho_A(a,b) = \rho_A(b,a)$. In what follows, we illustrate two extremes: (i) extremely small populations, where drift plays a greater role in the dynamics, and (ii) the large population limit, where selection dominates.

**SI.2.1. Small populations.** When $N$ is small, the terms $(z_i^*)^{N_1} \cdots (z_M^*)^{N_M}$ and $f_i(N)$ both contribute to fixation probability expressed in Eq. [SI.13]. Consequently, we cannot expect the neutrality condition to be nearly as simple as it is when $N$ is large and these two terms are negligible and can be safely ignored.

Already when $N = 2$, we can see that the neutrality condition is significantly different from $\bar{a} = \bar{b}$. In this case, there is environmental heterogeneity if there are $M = 2$ sites (otherwise the process is equivalent to the classical Moran process in a population of size $N = 2$). For such a small population, it is simple to directly solve the standard recurrence equations for fixation probabilities (i.e. the equation $Mz = z$) to get

$$ \rho_{e_{i,A}} = \frac{a_1}{a_1 + b_2}; \quad (SI.16a) $$
$$ \rho_{e_{i,B}} = \frac{a_2}{a_2 + b_1}. \quad (SI.16b) $$

It follows that the average fixation probabilities of the mutant and resident types, respectively, are

$$ \rho_A(a,b) = \frac{1}{2} \left( \frac{a_1}{a_1 + b_2} + \frac{a_2}{a_2 + b_1} \right); \quad (SI.17a) $$
$$ \rho_A(b,a) = \frac{1}{2} \left( \frac{b_1}{b_1 + a_2} + \frac{b_2}{b_2 + a_1} \right), \quad (SI.17b) $$

and the neutrality condition, $\rho_A(a,b) = \rho_A(b,a)$, simplifies to $a_1 a_2 = b_1 b_2$ (i.e. $\sqrt{a_1 a_2} = \sqrt{b_1 b_2}$).

On the other hand, even $N = 3$ demonstrates how the neutrality condition quickly gets complicated for small values of $N$ greater than 2. Again, we can solve the equation $Mz = z$ for fixation probabilities, $z$, but their expressions are complicated and not especially easy to interpret. Consider the averages

$$ \|a\|_1 = \frac{a_1 + a_2 + a_3}{3}; \quad (SI.18a) $$
$$ \|a\|_0 = \sqrt{a_1 a_2 a_3}; \quad (SI.18b) $$
$$ \|a\|_{-1} = \frac{3}{a_1 + a_2 + a_3}, \quad (SI.18c) $$

which are the arithmetic, geometric, and harmonic means of $a = (a_1, a_2, a_3)$, respectively. Under the simplifying assumption $b_1 = b_2 = b_3 = 0$, the neutrality condition, $\rho_A(a,b) = \rho_A(b,a)$ is equivalent to

$$ 90 \|a\|_1^2 \|a\|_0^3 \|a\|_{-1}^{-1} + 324 \|a\|_1^2 \|a\|_0^3 + 8 \|a\|_1 \|a\|_0^6 \|a\|_{-1}^{-2} \|a\|_1 \|a\|_0^6 \|a\|_{-1}^{-1} - 864 \|a\|_1^2 \|a\|_0^3 + 144 \|a\|_1 \|a\|_0^6 \|a\|_{-1}^{-2} \|a\|_0 \|a\|_{-1}^{-1} = 0. \quad (SI.19) $$
For larger (but still finite \( N \)), the condition \( \rho_A (a, b) = \rho_A (b, a) \) gets only more complicated.

**SI.2.2. Large population limit.** As \( N \to \infty \), \( z_i^* \) (as defined by Eq. \[ \text{SI.6} \]) approaches

\[
  z_i^* = \frac{\bar{b}}{b + a_i \left( 1 - z^* \right)},
\]  

(SI.20)

where, if \( f \) is the density of \( A \)-fitness values, \( z^* \) is a solution to the integral equation

\[
  1 - z^* = (1 - z^*) \int_0^\infty \frac{af (a)}{b + a \left( 1 - z^* \right)} da. 
\]  

(SI.21)

If \( g \) is the density of \( B \)-fitness values, then, by Eq. \[ \text{SI.13} \], the limiting value of the fixation probability of a randomly-placed mutant as \( N \to \infty \), \( \rho^\infty_A (f, g) \), therefore satisfies the integral equation

\[
  \rho^\infty_A (f, g) = \rho^\infty_A (f, g) \int_0^\infty \frac{af (a)}{b + a \rho^\infty_A (f, g)} da. 
\]  

(SI.22)

Moreover, since \( \bar{\pi} = \int_0^\infty af (a) da \) and \( \bar{b} = \int_0^\infty bg (b) db \), it follows from this equation that

\[
  \rho^\infty_A (f, g) = 0 \iff \bar{\pi} \leq \bar{b}. 
\]  

(SI.23)

Therefore, \( \rho^\infty_A (f, g) = \rho^\infty_A (g, f) \) if and only if \( \bar{\pi} = \bar{b} \), which gives the neutrality condition for large populations.

**SI.3. Heterogeneity in mutant fitness**

In this section, we prove the statement that heterogeneity in the fitness of an advantageous mutant or invading strategy always decreases its fixation probability, meaning that if \( \bar{\pi} \geq \bar{b} \), then

\[
  \rho_A (a, b) \leq \rho_A (\bar{\pi}, b). 
\]  

(SI.24)

First of all, suppose that \( f \) and \( g \) are fitness density functions for the mutant and resident type, respectively. Let \( \delta_\pi \) be the Dirac measure centered at \( \bar{\pi} = \int_0^\infty af (a) da \). Consider the function

\[
  \psi : [0, \infty) \longrightarrow [0, \infty) \quad : a \longmapsto \frac{a}{\bar{b} + a \rho^\infty_A (f, g)}. 
\]  

(SI.25)

Since \( \psi \) is concave, it follows from Jensen’s inequality that

\[
  1 = \int_0^\infty \psi (a) f (a) da \leq \psi (\bar{\pi}) = \frac{\bar{a}}{\bar{b} + \bar{\pi} \rho^\infty_A (f, g)}. 
\]  

(SI.26)

Since \( \rho^\infty_A (\delta_\pi, g) = 1 - \frac{\bar{b}}{\bar{a}} \), it follows that \( \rho^\infty_A (f, g) \leq 1 - \frac{\bar{b}}{\bar{a}} = \rho^\infty_A (\delta_\pi, g) \).

Furthermore, by the arithmetic-geometric mean inequality, we have

\[
  (z_i^*)^N \cdots (z_M^*)^N \leq \frac{1}{N} \sum_{i=1}^M N_i z_i^* = \frac{\bar{N}}{\bar{z}^N} \leq 1 - \rho^\infty_A (f, g). 
\]  

(SI.27)

It follows from this inequality that

\[
  \rho_A (a, b) \approx \frac{\rho^\infty_A (f, g)}{1 - (z_i^*)^N \cdots (z_M^*)^N} \leq \frac{\rho^\infty_A (f, g)}{1 - (1 - \rho^\infty_A (f, g))^N} \leq \frac{\rho^\infty_A (\delta_\pi, g)}{1 - (1 - \rho^\infty_A (\delta_\pi, g))^N} \approx \rho_A (\bar{\pi}, b). 
\]  

(SI.28)

Since \( \rho^\infty_A (f, g) \leq \rho^\infty_A (\delta_\pi, g) \) and \( \frac{x}{1 - (1-x)^N} \) is an increasing function of \( x \) for \( x \in [0, 1] \). Consequently, heterogeneity in mutant fitness suppresses the fixation probability of an invading mutant. Furthermore, provided \( N \) is large, these calculations show that resident heterogeneity has no effect on a mutant’s fixation probability.
SI.3.1. Perturbative expansion of fixation probability. Here, we discuss limiting expansions for the fixation probability in various limiting cases. We first derive the weak-heterogeneity limit that was presented in the main text. Suppose that the fitness distributions, $\mathbf{a}$ and $\mathbf{b}$, are given by $\mathbf{a} = \mathbf{a}_0 + \varepsilon \mathbf{a}_1$ and $\mathbf{b} = \mathbf{b}_0 + \varepsilon \mathbf{b}_1$, where $\mathbf{a}_0 = (a_0, \ldots, a_0)$ and $\mathbf{b}_0 = (b_0, \ldots, b_0)$ denote constant average fitness and $\mathbf{a}_1$ and $\mathbf{b}_1$ are the deviations from uniform fitness and are tuned by the (small) parameter $\varepsilon$. The averages of both $\mathbf{a}_1$ and $\mathbf{b}_1$ are zero.

Consider the power series expansion of a mutant’s fixation probability in $\varepsilon$,

$$
\rho_A^\varepsilon(f, g) = c_0 + c_1 \varepsilon + c_2 \varepsilon^2 + c_3 \varepsilon^3 + c_4 \varepsilon^4 + O(\varepsilon^5).
$$

We can solve for $c_0, c_1, \ldots, c_4$ using a perturbative expansion of Eq. SI.22

$$
1 = \int_0^\infty \frac{(a_0 + \varepsilon a_1)f_1(a_1)}{\rho + (a_0 + \varepsilon a_1)\rho_A^\varepsilon(f, g)}da_1,
$$

and matching the coefficients for different powers of $\varepsilon$ up to $\varepsilon^4$. Doing so gives

$$
c_0 = 1 - \frac{b_0}{a_0}, \quad \text{(SI.31a)}
$$

$$
c_1 = 0; \quad \text{(SI.31b)}
$$

$$
c_2 = -\left(1 - \frac{b_0}{a_0}\right)\left(\frac{b_0}{a_0}^2\right)E_{f_1}[a_1^2]; \quad \text{(SI.31c)}
$$

$$
c_3 = \left(1 - \frac{b_0}{a_0}\right)\left(\frac{b_0(a_0 - b_0)}{a_0^2}\right)E_{f_1}[a_1^3]; \quad \text{(SI.31d)}
$$

$$
c_4 = -\left(1 - \frac{b_0}{a_0}\right)\left\{\left(\frac{b_0^2(a_0 - 2b_0)}{a_0^3}\right)E_{f_1}[a_1^4]\right\} - \left(\frac{b_0(a_0 - b_0)^2}{a_0^3}\right)E_{f_1}[a_1^4]; \quad \text{(SI.31e)}
$$

Since $E_{f_1}[a_1] := \int_0^\infty a_1f_1(a_1)da_1 = 0$. Therefore, using the fact that $\rho_A^\varepsilon(\delta_{a_0}, \delta_{b_0}) = 1 - \frac{b_0}{a_0}$, we have

$$
\rho_A^\varepsilon(f, g) \approx \rho_A^\varepsilon(\delta_{a_0}, \delta_{b_0}) \left\{1 - \left(\frac{b_0}{a_0}^2\right)E_{f_1}[a_1^2]\varepsilon^2 + \left(\frac{b_0(a_0 - b_0)}{a_0^2}\right)E_{f_1}[a_1^3]\varepsilon^3 - \left(\frac{b_0^2(a_0 - 2b_0)}{a_0^3}\right)E_{f_1}[a_1^4]\varepsilon^4\right\}. \quad \text{(SI.32)}
$$

In addition to the means of $\mathbf{a}$ and $\mathbf{b}$, we can explicitly express fixation probability in terms of higher moments of the fitness distribution, including quantities such as standard deviation, skewness, and kurtosis. For symmetric distributions, the odd moments cancel and this expansion can be simplified even further.

It follows that the fixation probability in a finite population can be approximated as

$$
\rho_A(\mathbf{a}, \mathbf{b}) \approx \rho_A(\bar{\mathbf{a}}, \bar{\mathbf{b}}) \left\{1 - \left(\frac{\bar{b}}{\bar{a}}\right)E_f[(a - \bar{a})^2] + \left(\frac{\bar{b}(\pi - \bar{b})}{\bar{a}^2}\right)E_f[(a - \bar{a})^3] - \left(\frac{\bar{b}^2(\pi - 2\bar{b})}{\bar{a}^3}\right)E_f[(a - \bar{a})^4]\right\}. \quad \text{(SI.33)}
$$

where $\rho_A(\bar{\mathbf{a}}, \bar{\mathbf{b}})$ is the Moran fixation probability for a population of finite size, $N$. The results are plotted in Fig. 3 for both uniform and bi-modal distributions, which is in excellent agreement with our closed-form result.

SI.4. HETEROGENEITY IN RESIDENT FITNESS

In Eq. SI.32 only the mean of the resident fitness distribution appears in the approximation of $\rho_A^\varepsilon$. Therefore, the fixation probability of a single mutant does not depend on whether there is heterogeneity in resident fitness, i.e. $\rho_A(\mathbf{a}, \mathbf{b}) = \rho_A(\bar{\mathbf{a}}, \bar{\mathbf{b}})$. However, more interesting behavior is seen in smaller population sizes, where there can be deviations from this equation.

We tested the above results for resident fitness values that are uniformly distributed in the interval $[\bar{b} - \Delta_b, \bar{b} + \Delta_b]$ and a fixed value for mutant fitness, $a_i = \bar{a}$. The results from solutions to the Kolmogorov
equation for $N = 10$ and several values of resident heterogeneity, $\Delta b$, and mean, $\bar{b}$, are shown in Fig. 4. For such a small population size, the fixation probability near the neutral limit is increased as we increase resident heterogeneity. We also varied both mutant and resident fitnesses, and the results are plotted in Fig. 1.

**Figure 1.** Heat map for the fixation probability of the mutant type, $A$, as a function of the width of the mutant fitness distribution, $\Delta a$, and resident fitness distribution, $\Delta b$. The fitness values for the mutant and resident are uniformly distributed on $[\bar{a} - \Delta a, \bar{a} + \Delta a]$ and $[\bar{b} - \Delta b, \bar{b} + \Delta b]$, respectively. The population size is $N = 10$ and the results are obtained from numerical solutions to the Kolmogorov equation.

### SI.5. Random Environments

Here, we derive the limiting distribution for the fixation probability in a random environment. Without a loss of generality, we may assume that the resident fitness is constant and equal to 1 in all locations. The mutant fitness at each location is a random variable whose probability distribution function is $f(a)$.

When $N$ is sufficiently large, the fixation probability, $\rho_A = \rho_A(a, b)$, satisfies the equation

$$1 = \frac{1}{N} \sum_{i=1}^{N} \frac{a_i}{1 + a_i \rho_A}.$$  \hspace{1cm} (SI.34)

We can now ask a different question: given a fixed $\rho_A$, what is the probability that the environment is such that Eq. (SI.34) holds? This probability is precisely the probability that the fixation probability of a mutant is equal to $\rho_A$ given the random environment whose distribution is governed by the density function $f$. 


Let $\mu(\rho_A) = \mathbb{E}_f\left[\frac{a}{1+a\rho_A}\right]$ and $\sigma^2(\rho_A) = \text{Var}_f\left[\frac{a}{1+a\rho_A}\right]$ be the mean and variance of the function $\frac{a}{1+a\rho_A}$ when $a$ is distributed according to $f$. Since locations are independent, by the central limit theorem $\sqrt{N} \left( \frac{1}{N} \sum_{i=1}^{N} \frac{a_i}{1+a_i\rho_A} - \mu \right)$ converges in distribution to a Gaussian random variable with density

$$p(x) := \frac{1}{\sqrt{2\pi}\sigma^2} \exp\left\{ -\frac{x^2}{2\sigma^2} \right\}$$

as $N$ grows. Therefore, for $\rho_A$ fixed, it follows that the density function for the distribution of $\rho$ is

$$d(\rho) = p\left( \sqrt{N} (1 - \mu(\rho)) \right)$$

$$= \frac{1}{\sqrt{2\pi}\sigma^2(\rho)} \exp\left\{ -\frac{(\sqrt{N} (1 - \mu(\rho)))^2}{2\sigma^2(\rho)} \right\}$$

$$= \frac{1}{\sqrt{2\pi}\sigma^2(\rho)} \exp\left\{ -N \frac{(1 - \mu(\rho))^2}{2\sigma^2(\rho)} \right\}.$$  \hspace{1cm} (SI.36)