Fixation times of *de novo* and standing beneficial variants in subdivided populations

Vitor Sudbrack[1] and Charles Mullon[1,*]

[1]Department of Ecology and Evolution, University of Lausanne, 1015 Lausanne, Switzerland

[*]Correspondence: charles.mullon@unil.ch

Short running title: Fixation times under limited dispersal

Keywords: Fixation time, Diffusion approximation, Inbreeding, Kin competition, Metapopulation, Selective sweeps

Abstract. The rate at which beneficial alleles fix in a population depends on the probability and time of fixation of such alleles. Both of these quantities can be significantly impacted by population subdivision and limited gene flow. Here, we investigate how limited dispersal influences the rate of fixation of beneficial *de novo* mutations, as well as fixation time from standing genetic variation. We investigate this for a population structured according to the island model of dispersal allowing us to use the diffusion approximation, which we complement with simulations. We find that fixation may take on average fewer generations under limited dispersal than under panmixia when selection is moderate. This is especially the case if adaptation occurs from *de novo* recessive mutations, and dispersal is not too limited (such that approximately $F_{\mathrm{ST}} < 0.25$). The reason is that mildly limited dispersal leads to only a moderate increase in effective population size (which slows down fixation), but is sufficient to cause a relative excess of homozygosity due to inbreeding, thereby exposing rare recessive alleles to selection (which accelerates fixation). We also explore the effect of meta-population dynamics through local extinction followed by recolonization, finding that such dynamics always accelerate fixation from standing genetic variation, while *de novo* mutations show faster fixation interspersed with longer waiting times. Finally, we discuss the implications of our results for the detection of sweeps, suggesting that limited dispersal mitigates the expected differences between the genetic signatures of sweeps involving recessive and dominant alleles.

1 Introduction

Populations can adapt to their environment via the fixation of beneficial alleles (Kimura and Ohta, 1971; Gillespie, 1994). Understanding the rate at which such fixation occurs has thus been a major goal for evolutionary biology (McCandlish and Stoltzfus, 2014), as well as more applied biosciences, such as population management and conservation (Wright et al., 2009). The rate of genetic adaptation of a diploid population is often quantified with $2N_{\text{tot}}\mu P_{\text{fix}}$, where N_{tot} is the population size, μ is the per-generation per-locus probability that a beneficial mutation occurs, and P_{fix} is the probability that it fixes (Kimura, 1962, 1968; Kimura and Ohta, 1971; Kryazhimskiy and Plotkin, 2008; Lanfear et al., 2014). This assumes that the mutation rate μ is small, such that the time taken for a beneficial allele to fix is negligible compared to the waiting time before such a mutation arises. Nevertheless, the time a beneficial allele takes to fix is in some cases relevant as it scales comparably to the number of generations it takes to arise, thus causing a slowdown in adaptation (Charlesworth, 2020, 2022; for more general considerations, see Hendry and Kinnison, 1999). Additionally, because whether an allele fixes quickly or slowly influences the genetic signatures of adaptation at surrounding loci, the time to fixation may be important in the detection of selected sites in the genomes of natural and experimental populations (Messer and Petrov, 2013; Charlesworth, 2020, 2022).

The probability that a beneficial mutation fixes and the time it takes to do so both depend on an interplay between selection and genetic drift. This interplay is especially well understood under panmixia, i.e. where individuals mate and compete at random (Crow and Kimura, 1970; Ewens, 2004). In particular, because a rare allele is found almost exclusively in heterozygotes under panmixia, the probability of fixation of a single-copy *de novo* mutation strongly depends on its genetic dominance (or penetrance). Dominant beneficial alleles are more likely to fix than recessive ones as their effects are more immediately exposed to positive selection (Haldane, 1927a). Population size, which scales inversely with genetic drift, increases fixation time (Kimura and Ohta, 1969), but tends to have limited effects on the probability that a newly arisen mutation will fix (Kimura, 1962). In fact, the probability of fixation of a *de novo* mutation becomes independent from population size in the limit of infinite population size such that invasion can be modelled as a branching process and invasion implies fixation (Haldane, 1927b; Otto and Whitlock, 2001).

Genetic adaptation is not restricted to the fixation of *de novo* mutations, but can also stem from standing genetic variation (Orr and Betancourt, 2001; Hermisson and Pennings, 2005; Pennings and Hermisson, 2006a,b; Barrett and Schluter, 2008; Hermisson and Pennings, 2017). This standing variation is thought to be neutral or mildly deleterious until an environmental change takes place such that it becomes beneficial. The probability that such a variant fixes and the time it takes to do so are especially sensitive to its initial frequency, with greater frequency increasing the probability of fixation (Kimura, 1962) and reducing the time taken to fix (Kimura and Ohta, 1969). The initial frequency, in turn, depends on how selection and genetic drift shaped variation before it turned beneficial, which has also been extensively studied in well-mixed populations (Kimura et al., 1963; Crow and Kimura, 1970; Orr and Betancourt, 2001; Ewens, 2004).

Many natural populations, however, are not well-mixed. The physical constraints on movement often cause dispersal to be limited, leading to genetic structure through limited gene flow (Clobert et al., 2001). Such genetic structure is extremely widespread though often mild with many estimates of among-populations genetic differentiation F_{ST} of the order of 0.1 (e.g. Ståhl, 1981; Glover et al., 2013; Benzie, 2000; Giles and Goudet, 1997; Potenko and Velikov, 1998; Tamaki et al., 2008; Irvin et al., 1998; Kumar and Singh, 2017; Forstmeier et al., 2007; pp. 302-303 in Hartl and Clark, 2007 for an overview). Genetic structure influences both drift and selection as it modulates effective population size N_{e} (Wang and Caballero, 1999; Rousset, 2004), and generates kin selection and inbreeding (Hamilton, 1964; Frank, 1998; Rousset, 2004; Charlesworth and Charlesworth, 2010). The influence of inbreeding on genetic adaptation can be investigated independently by considering populations where selfing (or assortative mating) takes place but that are otherwise well-mixed (so that there is no kin selection or competition, Glémin and Ronfort, 2013; Newberry et al., 2016; Hartfield and Bataillon, 2020; Charlesworth, 2020). These investigations show that selfing tends to speed up fixation as it causes both: (i) an increase in homozygosity that exposes rare mutations more readily to selection; as well as (ii) a decrease in effective population size that reduces the time to fixation.

How limited dispersal affects the probability of fixation through selection and drift is well-studied in the island model of dispersal, showing for instance that the fixation probability of beneficial alleles is increased by limited dispersal when recessive and decreased when

dominant (Roze and Rousset, 2003; Whitlock, 2003; Rousset, 2004). Assuming that dispersal between demes is so rare that segregation time within demes can be ignored, Slatkin (1981) show that limited dispersal always increases the time to fixation of *de novo* mutations. Using the diffusion approximation and thus considering segregation time within demes, Whitlock (2003) also reports from a numerical scan that limited dispersal makes the total time to fixation increase, through an increase in N_e as well as in kin competition (p. 778 in Whitlock, 2003). For adaptation from standing genetic variation, the implications of limited dispersal remain understudied (though see Paulose et al., 2019 for a discussion on this under isolation by distance).

Here, we contribute to this literature by investigating the rate of fixation of *de novo* and standing variants in subdivided populations. Firstly, we revisit the time to fixation of *de novo* mutations, complementing the analysis found in Whitlock (2003). We show that limited dispersal can in fact speed up fixation of non-additive alleles, as long as selection is not too weak and dispersal is mildly limited such that it generates $F_{ST} < 0.25$, which is typical of many natural populations. Secondly, we combine the waiting time and time to fixation to investigate the total rate of fixation from *de novo* mutations (as done in Glémin and Ronfort, 2013 for selfing). Thirdly, we investigate the impact of limited dispersal on the time for standing variation to fix. Finally, we consider the influence of meta-population dynamics whereby subpopulations can go extinct and be recolonized.

2 Model

2.1 Life cycle, genotype and fecundity

We consider a monoecious diploid population that is subdivided among N_d demes, each carrying a fixed number N_{ind} of adult individuals (so that the total population size is $N_{tot} = N_d \cdot N_{ind}$). Generations are discrete and non-overlapping with the following life cycle occurring at each generation: (i) Each adult produces a large number of gametes according to its fecundity and then dies. (ii) Each gamete either remains in its natal deme with probability $1-d$ or disperses with probability d. We assume that $d > 0$ so that demes are not completely isolated from one another. Dispersal is uniform among demes, following the island model

(Wright, 1931). (iii) Finally, gametes fuse randomly within each deme to form zygotes that in turn compete uniformly within each deme for N_{ind} breeding spots to become the adults of the next generation.

We are interested in evolution at an autosomal locus where two alleles segregate: a wild-type allele a and a beneficial mutant allele A. An individual's genotype determines its fecundity. As a baseline, aa individuals have a fecundity of 1, while relative to this, Aa and AA individuals have fecundity of $1+hs$ and $1+s$, respectively. The parameters $0 \leq h \leq 1$ and $s > 0$ thus capture the dominance and selective effects of A, respectively. For simplicity, we assume throughout the main text that selection is soft, i.e. that the same total number of gametes is produced in each deme. The case of hard selection is explored in our Appendix where we show that our main results are not affected by whether selection is soft or hard.

2.2 Diffusion approximation

The dynamics of the frequency p of the allele A in the whole population can be modelled as a diffusion process under the island model of dispersal (e.g. Barton, 1993; Roze and Rousset, 2003; Whitlock, 2003; Cherry and Wakeley, 2003; Cherry, 2003; Wakeley, 2003; Wakeley and Takahashi, 2004; Lessard, 2009; note that one cannot follow a single allele frequency when the population experiences isolation by distance, see Rousset, 2004 for general considerations). We follow the framework developed in Roze and Rousset (2003), which assumes that selection is weak and that the number N_{d} of demes is large (i.e. $s \sim \mathcal{O}(\delta)$ and $N_{\text{d}} \sim \mathcal{O}(1/\delta)$ where $\delta > 0$ is small). The diffusion approximation is based on the expectation and variance in the change in p, which we describe below.

2.2.1 Expected frequency change

We show in Appendix A.1 that the expected change in allele frequency p can be written as,

$$\mathrm{E}[\Delta p|p] = sp(1-p)\Bigg[\underbrace{p + r_0^{\text{D}}(1-p)}_{\text{direct effect in AA}} + \underbrace{h(1-r_0^{\text{D}})(1-2p)}_{\text{direct effect in Aa}} - \underbrace{\left(r_1^{\text{R}} + (r_1^{\text{R}} - a^{\text{R}})(2h-1)(1-2p)\right)}_{\text{kin competition}}\Bigg] + \mathcal{O}\left(\delta^2\right), \tag{1}$$

where r_0^{D} is the probability that the two homologous genes of an individual are identical-by-descent (IBD); r_1^{R} is the probability that two genes sampled from the same deme with replacement are IBD; a^{R} is the probability that two homologous genes of an individual coalesce with a third gene sampled from the same deme at random. All these coalescent probabilities are computed under neutrality (i.e. with $\delta = 0$) and their expression in terms of dispersal and population size can be found in Table 1 (Appendix A.3 for derivations). The part that is between squared brackets in eq. (1) is equivalent to eq. (23) of Roze and Rousset (2003) after plugging in their fitness eq. (36) (and additionally using our eq. A24 from our appendix). In the limit of low dispersal and large patches (i.e. $d \to 0$ and $N \to \infty$ such that Nd remains constant, $r_0^{\mathrm{D}} = r_1^{\mathrm{R}} = F_{\mathrm{ST}}$, and $a^{\mathrm{R}} = 2F_{\mathrm{ST}}^2/[1+F_{\mathrm{ST}}]$), eq. (1) reduces to eq. (12) of Whitlock (2002) (with his $\eta = 0$; we compare the time to fixation computed under this limit and more generally in Supplementary Figure SM1A-B, which shows overall good agreement).

Equation (1) decomposes selection on allele A among three effects. These can be understood by first considering that in a well-mixed population (so that $r_0^{\mathrm{D}} = r_1^{\mathrm{R}} = a^{\mathrm{R}} = 0$ in an infinite population), eq. (1) reduces to: $\mathrm{E}[\Delta p|p] = sp(1-p)\left[p + h(1-2p)\right] + \mathcal{O}\left(\delta^2\right)$ (Crow and Kimura, 1970, p. 408). In this baseline expression, the first term within square brackets, p, captures selection on A owing to the effects of the allele on the fitness of its bearer in homozygotes, while $h(1-2p)$ captures selection through heterozygotes. When dispersal is limited, direct effects increase to $p + r_0^{\mathrm{D}}(1-p)$ through homozygotes and decrease to $h(1-r_0^{\mathrm{D}})(1-2p)$ through heterozygotes in eq. (1). Selection through homozygotes is therefore more important under limited dispersal. This is because mating within demes leads to inbreeding and therefore a relative excess of homozygotes and a deficit of heterozygotes (according to r_0^{D}).

The remaining terms of eq. (1) – labelled "kin competition" – capture a second effect of limited dispersal: that competing individuals are more likely to carry identical gene copies than randomly sampled individuals. As shown by the negative sign in front of these terms in eq. (1), kin competition decreases selection on beneficial alleles. This is because kin competition results in an individual's reproductive success coming at the expense of genetic relatives. Equation (1) further shows that for additive alleles ($h = 1/2$), kin competition effects scale with genetic differentiation r_1^{R} only. For non-additive alleles, however, these effects also depend on allele frequency p, with kin competition effects being stronger when dom-

inant alleles are rare ($h > 1/2$ and $p < 1/2$) or when recessive alleles are common ($h < 1/2$ and $p > 1/2$).

2.2.2 Variance in frequency change and effective population size

The variance in allele frequency change can be written in the form

$$\mathrm{V}[\Delta p|p] = \frac{p(1-p)}{2N_\mathrm{e}} + \mathcal{O}\left(\delta^2\right), \tag{2}$$

where effective population size N_e for our model is given by eq. (28) of Roze and Rousset (2003),

$$N_\mathrm{e} = \frac{N_\mathrm{tot}}{1 - r_1^\mathrm{D}}, \tag{3}$$

with r_1^D is the probability that two different genes sampled from the same deme are IBD. See Appendix 2-A.2 for derivation of eq. (3), Table 1 for the probability r_1^D in terms of deme size and dispersal rate, and Fig. 1 for how limited dispersal increases the effective population size. Equation (3) reduces to the classical expression, $N_\mathrm{e} = N_\mathrm{tot}/(1 - F_\mathrm{ST})$ in the limit of low dispersal and large patches (such that $r_1^\mathrm{D} = r_1^\mathrm{R} = F_\mathrm{ST}$, Wright et al., 1939; Whitlock and Barton, 1997; Whitlock, 2003; Roze and Rousset, 2003).

2.3 Probability and time of fixation

From $\mathrm{E}[\Delta p|p]$ and $\mathrm{V}[\Delta p|p]$, the probability of fixation $P_\mathrm{fix}(p_0)$ of A with initial frequency p_0 is given by

$$P_\mathrm{fix}(p_0) = \frac{\int_0^{p_0} \exp\left(-2\int_0^x \frac{\mathrm{E}[\Delta p|p]}{\mathrm{V}[\Delta p|p]}\mathrm{d}p\right)\mathrm{d}x}{\int_0^1 \exp\left(-2\int_0^x \frac{\mathrm{E}[\Delta p|p]}{\mathrm{V}[\Delta p|p]}\mathrm{d}p\right)\mathrm{d}x} \tag{4}$$

(Crow and Kimura, 1970, p. 424). The expected number $T_\mathrm{fix}(p_0)$ of generations that an allele takes to fix (conditional on its fixation) given that its frequency is p_0 at generation $t = 0$ is,

$$T_\mathrm{fix}(p_0) = \int_{p_0}^1 \psi(x) P_\mathrm{fix}(x)\left[1 - P_\mathrm{fix}(x)\right]\mathrm{d}x + \frac{1 - P_\mathrm{fix}(p_0)}{P_\mathrm{fix}(p_0)} \int_0^{p_0} \psi(x) P_\mathrm{fix}^2(x)\mathrm{d}x\,, \tag{5}$$

with

$$\psi(x) = \frac{2\int_0^1 \exp\left(-2\int_0^z \frac{\mathrm{E}[\Delta p|p]}{\mathrm{V}[\Delta p|p]}\mathrm{d}p\right)\mathrm{d}z}{\mathrm{V}[\Delta x|x]\cdot\exp\left(-2\int_0^x \frac{\mathrm{E}[\Delta p|p]}{\mathrm{V}[\Delta p|p]}\mathrm{d}p\right)} \tag{6}$$

(Crow and Kimura, 1970, p. 430).

Equations (4)-(6) were numerically integrated with R using the built-in function `integrate` (with maximum number of subdivisions increased to 100′000). The time to fixation for non-additive alleles ($h \neq 1/2$) involves an integral with an integrand that spans many orders of magnitude, which can be prone to instability during numerical analysis. In the case of report of bad integrand behaviour, we translated integration limits by a small amount ϵ that we kept as low as possible ($\epsilon \leq 10^{-5}$ always).

To calculate the time to fixation more straightforwardly, Charlesworth (2020) suggests an approximation based on a decomposition of fixation dynamics between three phases: two stochastic phases when $p < p_1$ and $p > p_2$ connected by a deterministic phase for $p_1 \leq p \leq p_2$ (following semi-deterministic approaches, e.g. Martin and Lambert, 2015). This approximation can be summarized as

$$\begin{aligned} T^*_{\mathrm{fix}}(p_0) = {} & \underbrace{4N_\mathrm{e}\left[\frac{p_1}{1-p_1}\log(p_1)+1\right]}_{\text{initial stochastic phase}} + \underbrace{\frac{1-\exp(-1)}{s_0+(1-2h)s_1}}_{\text{final stochastic phase}} \\ & + \underbrace{\frac{1}{s_0}\log\left[\frac{s_0}{(1-p_1)(s_0+(1-2h)s_1)}\right] + \frac{1}{s_0+(1-2h)s_1}\log\left[\frac{s_0+(1-2h)s_1}{p_2 s_0}\right]}_{\text{deterministic phase}} \end{aligned} \tag{7}$$

("initial stochastic phase" is on p. 15, "final stochastic phase" in eq. A7, and "deterministic phase" in eq. 6a of Charlesworth, 2020), where

$$p_1 = 1 - \frac{1}{4N_\mathrm{e}s_0} \qquad \text{and} \qquad p_2 = \frac{1}{4N_\mathrm{e}\left[s_0+(1-2h)s_1\right]}\,, \tag{8}$$

and s_0 and s_1 are defined from decomposing the expected frequency change as $\mathrm{E}[\Delta p|p] =$

$p(1-p)\left[s_0 + (1-2h)s_1 p\right]$, which by comparison with our eq. (1) yields,

$$s_0 = s\left[r_0^{\mathrm{D}} - r_1^{\mathrm{R}} + h\left(1 - r_0^{\mathrm{D}}\right) - \left(r_1^{\mathrm{R}} - a^{\mathrm{R}}\right)(2h-1)\right] \tag{9a}$$

$$s_1 = s\left[1 - 2(r_1^{\mathrm{R}} - a^{\mathrm{R}}) - r_0^{\mathrm{D}}\right]. \tag{9b}$$

We will also use this approximation to compute fixation time.

3 Results

3.1 The antagonistic effects of limited dispersal on the time to fixation of *de novo* mutations

We first complement Whitlock (2003)'s analyses on the effects of limited dispersal on the time it takes for A to fix as a *de novo* mutation (i.e., when arising as a single copy, $P_{\mathrm{fix}}(p_0)$ with $p_0 = 1/(2N_{\mathrm{tot}})$). Whitlock (2003) considers a fecundity advantage $s = 0.0002$, and shows that in this case, limited dispersal always increases the time to fixation by increasing effective population size (see also our Supplementary Fig. SM1B). We consider the case where selection is stronger, though still weak, with $s = 0.01$ so that carrying two copies of A causes a 1% increase in fecundity (we consider other selection strengths later). We show that by modulating the interaction between selection and drift, limited dispersal can decrease the time to fixation in this case.

We integrated eq. (5) with $p_0 = 1/(2N_{\mathrm{tot}})$ for a range of dispersal d and dominance h values. Results of these calculations and of individual based simulations are shown in Fig. 2. We find that the effect of limited dispersal on the time to fixation depends on the dominance of the beneficial allele A. Where dominance is incomplete (approximately $0.1 \leq h \leq 0.9$ Fig. 2B), the expected time to fixation always increases as dispersal becomes more limited (Fig. 2A, blue). In contrast, the time to fixation of a partially dominant ($h > 0.9$) or partially recessive ($h < 0.1$) allele initially decreases as dispersal becomes limited and only increases once past below a dispersal threshold (Fig. 2A, green and purple). The reason the time to fixation eventually increases when dispersal becomes severely limited ($d \lesssim 0.01$ so that $F_{\mathrm{ST}} \gtrsim 0.2$ in Fig. 2A) is because N_{e} increases hyperbolically as dispersal decreases (eq. A40, Fig. 1). As

a result, the effects of N_e on time to fixation overwhelms any other effects when dispersal becomes small. These results so far are consistent with those in Whitlock (2003), where selection is sufficiently weak such that the effects of limited dispersal on the time to fixation are mostly through its effects on N_e.

Under mild dispersal limitation ($0.01 \lesssim d < 1$), however, our results show that the increase in effective population size N_e can be outweighed by an increase in selection, resulting in partially recessive and partially dominant alleles fixing more rapidly than under panmixia (Fig. 2A, dark gray line in B). The reason selection reduces the time to fixation here is because limited dispersal leads to inbreeding and thus a relative excess of homozygotes. How this excess boosts selection depends on whether the allele is recessive or dominant, as revealed by considering an increase in r_0^D in eq. (1). For a recessive beneficial allele A ($h < 1/2$), selection on A is greater when A is relatively rare (i.e. $p < 1/2$) because in this case, inbreeding creates a relative excess of AA homozygotes through which the recessive allele A can be picked up by selection. For a dominant beneficial allele A ($h > 1/2$), the excess of homozygosity boosts selection at high frequency (i.e. $p > 1/2$) as it allows to purge more efficiently the deleterious (and recessive) a allele through an excess of aa individuals. These frequency-dependent effects are amplified by kin competition (the fact that under limited dispersal, competing individuals are more likely to carry identical gene copies than randomly sampled individuals, Rousset, 2004). Kin competition weakens selection as it results in an individual's reproductive success coming at the expense of genetic relatives. The strength of kin competition on a non-additive allele depends on its frequency (see eq. 1). Specifically, kin competition is weaker and thus selection is stronger when a recessive allele is rare ($h < 1/2$ and $p < 1/2$) and a dominant allele is common ($h > 1/2$ and $p > 1/2$).

The frequency-dependent effects of limited dispersal on selection are reflected in the trajectory profiles of recessive and dominant alleles that fix (Fig. 2C). In a panmictic population, a recessive beneficial allele tends to spend longer periods at low frequency (for enough homozygotes to appear) and a dominant allele at high frequency (for heterozygotes to be purged, Fig. 2C, top). Under limited dispersal, however, these differences are mitigated as selection is increased at low frequency for recessive alleles and at high frequency for dominant alleles. As a result, the trajectory profiles of beneficial alleles that eventually fix become independent of their dominance as dispersal becomes limited (Fig. 2C, bottom). This

can also be seen from the decomposition of the time to fixation into three relevant phases according to the semi-deterministic approximation eq. (7). As shown in Fig. 3, limited dispersal decreases the share of the time that recessive alleles spend in the initial stochastic phase (lower shaded region in top row of Fig. 3) and the share that dominant alleles spend in the final stochastic phase (top shaded region in bottom row of Fig. 3). In fact, owing to an excess homozygosity that reduces boundary effects at low and high frequency, the semi-deterministic approximation eq. (7) performs even better under limited dispersal when alleles are recessive or dominant (compare black full and dashed gray lines in Fig. 3).

The above shows that limited dispersal can reduce the time to fixation of partially recessive and dominant alleles, provided the selection coefficient s is above some threshold. We investigate this threshold numerically in Supplementary Fig. SM1C, which suggests that it is close to $s = 0.0025$, i.e. such that carrying two copies of the beneficial allele causes a 0.25% increase in fecundity (with the parameters of our model). This value sits well within empirically estimated distribution of fitness effects across several taxa (Eyre-Walker and Keightley, 2007).

Although the time to fixation is useful for multiple reasons (e.g. Whitlock, 2003; Glémin and Ronfort, 2013; Charlesworth, 2020, 2022), it may not always provide a good reflection of the time for a population to show high mean fecundity, especially when beneficial alleles are dominant. To see this, we computed the expected time taken for the genetic load to drop to 10% in the whole population (τ_G), as well as within each deme (τ_L), using individual based simulations (Fig. 4). We observe $\tau_G < \tau_L < T_{fix}$ throughout, with their differences larger the greater h is. This is because dominant alleles confer high fitness in heterozygotic form and thus allow the population to show low load at lower frequency. Therefore, though limited dispersal reduces the time to fixation of both recessive and dominant beneficial mutations, the time taken for the population to show high fecundity is reduced only for recessive alleles.

3.2 The impact of limited dispersal on the total time for *de novo* mutations to arise and fix

In addition to the number of generations taken for a beneficial allele to fix, another relevant consideration is how long it takes for such an allele to emerge. To capture this, we follow

Glémin and Ronfort (2013) and quantify the total expected number T_{new} of generations for an adaptive *de novo* mutation to fix in the population by the sum of two terms:

$$T_{\text{new}} = \underbrace{\frac{1}{2N_{\text{tot}}\mu P_{\text{fix}}(p_0)}}_{\text{waiting time}} + \underbrace{T_{\text{fix}}(p_0)}_{\text{fixation time}}, \tag{10}$$

where μ is the mutation rate from a to A, and P_{fix} is the fixation probability of A when it arises as a single copy, i.e. when $p_0 = 1/(2N_{\text{tot}})$ (eq. 4 for the diffusion approximation to this probability). The first term is the expected number of generations for the emergence of a beneficial mutation that fixes, and the second is the expected number of generations taken by such fixation. The underlying assumption behind using eq. (10) is that beneficial mutations appear at a per-site per-generation rate μ that is such that new mutations segregate independently (as in e.g. Gillespie, 1994's strong-selection weak-mutation regime).

The waiting time is inversely proportional to the fixation probability P_{fix}, whose dependence on limited dispersal is well known: while dispersal limitation has no influence on the probability of fixation of additive alleles, it increases (respectively, decreases) the probability that a recessive (dominant) beneficial allele fixes (Roze and Rousset, 2003; Whitlock, 2003). Hence, the waiting time for a fixing additive allele is not affected by dispersal limitation, but is reduced for a recessive or partially recessive allele and increased for a dominant or partially dominant allele (Fig. 5A). Accordingly, the total number of generations T_{new} for fixation always increases with dispersal limitation when such an allele is additive (as T_{fix} increases, section 3.1, Fig. 5A central column). Recessive alleles, meanwhile, benefit from dispersal limitation in two ways, as limited dispersal not only reduces the time to fixation (provided that dispersal is not too limited) but also the waiting time for such a fixing allele to appear. This results in a significant drop in the total time for fixation as dispersal becomes limited, before eventually increasing under severe dispersal limitation (Fig. 5A left column).

The case of dominant alleles is more complicated as on one hand dispersal limitation increases the waiting time, but on the other reduces the time to fixation (section 3.1 and Fig. 5A right column). The balance between these opposing effects depends on the mutation rate μ. When this rate is very small, the waiting time dominates eq. (10) so that dispersal limitation always increases the total time T_{new} for dominant alleles to fix (Fig. 5A top-right). As the mutation rate increases, however, the time to fixation becomes more relevant in eq. (10)

so that dispersal limitation may reduce total time T_{new}, though less than for recessive alleles (Fig. 5A bottom-right).

Overall, we thus find that the rate of fixation of adaptive alleles depends on the interaction between the dominance h of these alleles and dispersal d. To see this more definitively, we compare T_{new} across levels of dispersal and dominance with T_{new} under panmixia in Fig. 5B for a mutation rate of $\mu = 10^{-6}$. This figure shows that, for this set of parameters, the total time for the fixation of *de novo* mutations can be up to four times more rapid under limited dispersal compared to panmixia when beneficial alleles are recessive (Fig. 5B, dark purple region, dotted black lines). The effect for dominant alleles, although weaker, is still non-negligible with fixation up to 30% faster under limited dispersal (Fig. 5B, light purple region where $h > 0.5$). Below a dispersal threshold, however, fixation is slower whatever the dominance of beneficial alleles (Fig. 5B, green region).

3.3 Fixation from standing variation: limited dispersal and dominance reversal

To investigate fixation from standing genetic variation, we now let the initial frequency of allele A in the whole global population p_0 be a random variable, whose distribution is determined by assuming that A is initially deleterious, maintained at a mutation-selection-drift equilibrium until an environmental change takes place that causes A to become beneficial (following Orr and Betancourt, 2001; Hermisson and Pennings, 2005, 2017; Orr and Unckless, 2008; Glémin and Ronfort, 2013). Given a realisation p_0, the initial frequency in each deme when environment changes is thus on average p_0 but there is variation among demes, i.e. there is genetic differentiation among demes due to local sampling effects. The expected number of generations taken for fixation is now computed with

$$T_{\text{sgv}} = \int_0^1 T_{\text{fix}}(p_0)\phi(p_0)\,\mathrm{d}p_0, \tag{11}$$

where $\phi(p_0)$ is the probability density function for the frequency p_0 of allele A in the whole population at the time of the environmental change when allele A becomes beneficial. This distribution $\phi(p_0)$ can be calculated at mutation-selection-drift equilibrium using the dif-

fusion approximation (eq. A49 in Appendix). Prior to the environmental change, allele A is deleterious such that aa, Aa and AA individuals have fecundity of 1, $1 - h_D s_D$, and $1 - s_D$, respectively, while mutations from a to A, and from A to a, occur at rate μ. Population structure and dispersal rate are assumed to be the same before and after the environmental change. We computed eq. (11) numerically under different values of genetic dominance before (h_D) and after (h) the environmental change and various dispersal rates (d).

Let us first consider scenarios where the dominance of A is preserved before and after the environmental change (i.e. $h_D = h$). We find that under mildly limited dispersal, fixation takes longer time when A is additive and shorter time when A is dominant (Fig. 6A blue and green). Thus, limited dispersal has the same effect on the time for A to fix as a standing genetic variant than on the total time T_{new} for A to fix as a *de novo* mutation (provided mutation is strong enough so that waiting time does not dominate T_{new} when A is dominant). By contrast, while limited dispersal can lead to shorter time T_{new} for A to fix as a *de novo* mutation when recessive, it always increases the time T_{sgv} for A to fix as a standing genetic variant (Fig. 6A purple). To understand this, recall that when the frequency of a beneficial recessive allele A is low, dispersal limitation speeds up the segregation of that allele by producing an excess of homozygotes AA. If such an allele is initially deleterious and recessive, however, its initial frequency p_0 tends to be higher at the moment of environmental change. Consequently, the allele is likely to already exist in the homozygous form when it becomes beneficial, and thus is easily picked up by selection regardless of dispersal limitation (Fig. 6B purple for the distribution $\phi(p_0)$). Similarly to *de novo* mutations, the sweeping trajectories of standing genetic variants with different levels of dominance also become more similar as dispersal is reduced, in fact converging to the trajectories of recessive alleles under panmixia (compare top and bottom of Fig. 6C).

One scenario that has been argued to be particularly relevant in the context of fixation from standing genetic variation is that of dominance reversal, whereby an initially recessive deleterious allele ($h_D = 0$) becomes beneficial and dominant ($h = 1$) in the new environment (Muralidhar and Veller, 2022). This facilitates fixation because at mutation-selection-drift equilibrium, a recessive deleterious allele can be maintained at significant frequency, such that it can be readily picked up by selection when it turns beneficial, especially if it simultaneously becomes dominant. Comparing the case where A is additive before and after the

environmental change, with the case where it shifts from being recessive to dominant, we see that limited dispersal reduces the effects of dominance reversal (Fig. 7A). This is because limited dispersal, through an excess of homozygosity, diminishes the effects of dominance on both: (i) the expected frequency p_0 at which the deleterious allele is maintained before environmental change; and (ii) selection when beneficial. In fact, as dispersal becomes increasingly limited, the trajectory profiles of alleles that experience a dominance shift become almost indistinguishable from the profiles of alleles that did not (Fig. 7B).

3.4 Longer waiting but faster fixation under extinction-recolonization

We have assumed that demes are of fixed and constant size. But deme extinctions, whereby entire local populations vanish and their habitat is made available for recolonization, are ecologically and evolutionarily relevant as they modulate the consequences of dispersal (Pannell and Charlesworth, 1999; Rousset, 2004). To explore how the interplay between extinction-recolonization and dispersal limitation influences fixation times, we assume that before step (i) of the life cycle (see section 2.1), each deme independently goes extinct with a probability $0 \leq e < 1$, in which case all individuals present in that deme die before producing any gamete (Appendix C for details). Each extinct deme is then available for recolonization by $2N_{\text{ind}}$ gametes from extant demes during dispersal. We sample these $2N_{\text{ind}}$ gametes in two ways to examine contrasting scenarios of recolonization (as in Slatkin, 1977; Whitlock and McCauley, 1990): (i) in the propagule model, gametes are sampled from the gametic pool of a single extant deme, which is chosen at random among all extant demes; while (ii) in the migrant pool model, gametes are sampled from the joint gametic pool of all extant demes.

We look at the total time of taken for a *de novo* beneficial mutation, T_{new} (eq. 10), which depends on the waiting time for a fixing allele to arise and on the time to fixation (Fig. 8A). We find that deme extinctions tend to have limited effects on T_{new}, unless the mode of recolonization follows the propagule model and dispersal is strongly limited (third row of Fig. 8A). In this case, T_{new} is greater under deme extinctions mostly due to an inflation in waiting time. This is because by increasing the covariance in allele frequency among demes, extinction-recolonization dynamics reduces effective population size relative to census size

(Fig. 8B; Slatkin, 1977; Whitlock and McCauley, 1990; Barton, 1993; Barton and Whitlock, 1997). This reduction is especially significant under the propagule model because in this case, a recolonized deme and the deme of origin for the propagule have the same allele frequency on average, which boosts the covariance among demes (Fig. 8B bottom). The resulting increase in genetic drift reduces the fixation probability of beneficial alleles, and thus in turn, increases the waiting time (first term of eq. 10; Fig. 8C bottom). The increase in genetic drift also causes a reduction in fixation time (second term of eq. 10), but this does not compensate for the inflated waiting time in the case of propagule recolonization (Fig. 8A). Altogether, our results indicate that adaptation from *de novo* mutations is characterized by faster fixations separated by longer waiting times under extinction-recolonization dynamics.

In contrast to fixation of *de novo* mutations, extinctions almost always reduce the expected number of generations a beneficial allele from standing genetic variation takes to fix, T_{sgv} (eq. 11, Fig. 8D). This is because the waiting time is no longer relevant when the allele is already present in the population. In fact, the reduction in N_{e} owing to extinction-recolonization dynamics accelerates adaptation as it both: (i) reduces fixation time; and (ii) leads to on average a greater frequency p_0 of A at the time of environmental change. As a result, fixation of standing genetic variants can be significantly quicker when extinctions are common and recolonization follows the propagule model (Fig. 8D bottom).

4 Discussion

Our analyses indicate that dispersal limitation can accelerate the fixation of beneficial *de novo* alleles when: (i) dispersal is mildly limited; and (ii) allelic effects on fitness are not too weak and far from additive (e.g. $h < 0.1$ or $h > 0.9$ in Fig. 2B). This may be relevant to natural populations as the dispersal rates under which we found that recessive and dominant mutations fix quicker than under panmixia lead to F_{ST} levels that agree with estimates from a wide range of taxa (roughly $d > 10^{-2}$ so on average 1 or more immigrants per generation, leading to $F_{\text{ST}} < 0.25$, figs. 2A and B; e.g. fish, Ståhl, 1981; Glover et al., 2013; crustaceans, Benzie, 2000; plants, Giles and Goudet, 1997; Potenko and Velikov, 1998; Tamaki et al., 2008; insects, Irvin et al., 1998; Kumar and Singh, 2017; birds, Forstmeier et al., 2007; pp. 302-303

in Hartl and Clark, 2007 for an overview). Additionally, the notion that alleles have non-additive fitness effects is supported by multiple lines of evidence, both theoretical (Fisher, 1928; Wright, 1934; Kacser and Burns, 1981; Manna et al., 2011; Billiard et al., 2021) and empirical, with mutations thought to be often at least partially recessive, with an average dominance coefficient h close to 0.2 (Mukai et al., 1972; Agrawal and Whitlock, 2011; Huber et al., 2018; reviewed in Orr, 2010 and Li and Bank, 2023). Further, the selection coefficient that is required to observe a decrease in the time to fixation under limited dispersal (s greater than 0.0025) sits well within empirically estimated fitness effects (Eyre-Walker and Keightley, 2007). We considered specifically the case where $s = 0.01$ in our main text figures, which corresponds to a 1% increase in fecundity due to a single substitution. We tested the effect of stronger selection with simulations whose results are shown in Supplementary Fig. SM2. These show similar patterns to our baseline model, i.e. mild dispersal limitation speeds up fixation when alleles are recessive or dominant. In fact, strong selection tends to amplify this effect (Supplementary Fig. SM2).

In addition to the time to fixation, the pace of adaptation also depends on the waiting time for a fixing mutation to appear (eq. 10; Glémin and Ronfort, 2013). Because limited dispersal reduces most significantly the waiting time for a fixing recessive allele to appear, the total time for a *de novo* mutation to fix is most shortened when beneficial alleles are recessive (purple region in Fig. 5B). Overall, our results thus suggest that with all else being equal, a subdivided population should be better adapted and show greater mean fecundity than a panmictic population, provided dispersal is only mildly limited and adaptation is driven by recessive *de novo* mutations.

In contrast, limited dispersal always slows down fixation of standing genetic variants that are recessive before and after they become beneficial due to an environmental change (Fig. 6A, purple line). Rather, mild dispersal limitation tends to accelerate the fixation of dominant alleles here (Fig. 6A, green line). This is because limited dispersal leads to a greater boost in frequency of a deleterious allele when it is dominant compared to when it is recessive (Fig. 6B, compare top to bottom). Nevertheless, the time taken for dominant genetic variants to fix in response to changes in selective pressures is typically greater than recessive variants, although limited dispersal tends to reduce this difference (Fig. 6A).

More broadly, dispersal limitation diminishes the importance of genetic dominance on the

time taken by alleles to fix. This is in part because limited dispersal leads to inbreeding, which causes an excess of homozygotes whose fecundity does not depend on genetic dominance. Models involving partial selfing (or assortative mating), which also causes elevated homozygosity, similarly found lesser importance of dominance for fixation (Roze and Rousset, 2004; Glémin and Ronfort, 2013; Newberry et al., 2016; Hartfield and Bataillon, 2020; Charlesworth, 2020). Our model, however, contrasts with these scenarios because dispersal limitation also: (i) leads to kin competition, which reduces the overall strength of selection; and (ii) increases effective population size N_e whereas selfing alone reduces N_e. These two effects explain why strong dispersal limitation always delays fixation, whereas selfing generally speeds up fixation (Roze and Rousset, 2004; Glémin and Ronfort, 2013; Newberry et al., 2016; Hartfield and Bataillon, 2020; Charlesworth, 2020). In fact, our results under extinction-recolonization dynamics align more closely with those under selfing as limited dispersal reduces N_e when extinctions are sufficiently common (Fig. 8B).

Through its effects on time to fixation, dispersal limitation may have implications for the signature of selective sweeps. The idea behind this is that when a selected allele goes to fixation rapidly, there are fewer opportunities for recombination so that genetic diversity at nearby neutral sites tends to be reduced, leading to what is referred to as a *hard sweep*; conversely, when fixation is slow, recombination is more likely to break the association between an adaptive allele and its original background before fixation, leading to a *soft sweep* (Hermisson and Pennings, 2005, 2017; Messer and Petrov, 2013; Jensen, 2014). More specifically, the linkage disequilibrium between a new beneficial allele and a linked neutral allele decreases at a rate given by their recombination rate r in a large well-mixed population, i.e. linkage disequilibrium decays as $\exp(-rt)$, where t is the number of generations that the fixing allele takes to rise to high frequency (Smith and Haigh, 1974; Barton, 2000). Accordingly, the probability of observing a hard sweep is lower under limited dispersal if limited dispersal increases t (Barton, 2000; Pennings and Hermisson, 2006d; Kim and Maruki, 2011). However, modelling studies have found contrasting effects of limited dispersal on the signature of sweeps, which is typically quantified by F_{ST} at linked neutral loci. In fact, the fixation of a beneficial allele can increase or decrease F_{ST}, depending on initial conditions and on dominance (Slatkin and Wiehe, 1998; Santiago and Caballero, 2005; Kim and Maruki, 2011; Teshima and Przeworski, 2006; Roze and Rousset, 2008; Ewing et al., 2011). In particular, F_{ST} at linked neutral loci is expected to increase when a recessive allele fixes, whereas

F_{ST} is expected to decrease when a dominant allele fixes (eq. 79 in Roze and Rousset, 2008). We performed simulations of evolution at two linked loci where one is neutral and initially polymorphic, and the other is under positive selection (Appendix D for details). The results we find align with those of Roze and Rousset (2008). When beneficial alleles are additive, dispersal has no effect on the probability of observing a hard sweep, i.e. on the probability that the polymorphism at the neutral locus is lost with fixation of the beneficial allele (Supplementary Fig. SM3A blue). This is because although the time to fixation is greater, and so are recombination opportunities under limited dispersal, most of the new haplotypes created by recombination are lost due to local drift within patches (Supplementary Fig. SM3B). Meanwhile, limited dispersal increases the probability of observing a hard sweep for a recessive allele and decreases it for a dominant allele so that these probabilities converge to that of an additive allele (Supplementary Fig. SM3A purple and green). This is because limited dispersal decreases (respectively, increases) the proportion of time that a recessive (dominant) allele spends at low frequency (Fig. 3), thus affecting the recombination opportunities with new backgrounds for these alleles.

The association between a selected allele and its original background can also be broken when recurrent mutations create beneficial mutations that are identical-by-state and that fix with different backgrounds (Pennings and Hermisson, 2006c; Ralph and Coop, 2010; Paulose et al., 2019). How likely this is to happen can be inferred from comparing the waiting and fixation time (eq. 10). If the fixation time is long compared to waiting time, then recurrent mutations should be more likely to lead to a soft sweep. Inspection of Fig. 5 reveals that the time to fixation can become longer than the waiting time as dispersal becomes more limited, especially if mutations are common ($\mu = 10^{-6}$). This suggests that limited dispersal may favour soft sweeps through recurrent mutations. To investigate more definitively how limited dispersal affects the signature of sweeps, it would be interesting to extend our model to consider multiple linked loci (e.g., extending Roze and Rousset, 2008 to finite number of demes or Lehmann and Rousset, 2009 to limited dispersal).

Finally, our results are based on several assumptions. First, we assumed that dispersal is gametic, which is relevant for plant and marine taxa but less so for terrestrial animals where it is often zygotes that disperse. But provided that mating is random within demes and demes are large enough, alellic segregation is similar under gametic and zygotic dispersal (Roze

and Rousset, 2003). Second, we assumed that selection is soft, i.e. that each deme produces the same number of gametes. We explore the case of hard selection in Appendix E such that demes showing greater frequency of allele A produce more gametes. Hard selection reduces by a small margin the time to fixation, but does not affect our results otherwise (Fig. E1). Third, the diffusion approximation also relies on the assumption that demes are homogeneous and that dispersal is uniform among them (i.e. no isolation-by-distance). Isolation by distance in principle delays fixation (Rousset, 2006), unless demes show specific patterns of connectivity that create sources and sinks that may facilitate fixation (e.g. Marrec et al., 2021). Fourth, we focused on the expectation of the number of generations taken for fixation, which may be misleading if the underlying distribution is fat tailed and skewed towards large values. To check for this, we computed the median time to fixation using individual based simulations. We found that the mean and the median are close, indicating that the distribution of times to fixation is fairly symmetrical around the mean (Supplementary Fig. SM4).

Acknowledgements

We thank Ehouarn Le Faou, Claudia Bank, Denis Roze and Michael Whitlock for discussions, and the Swiss National Science Foundation for funding (PCEFP3181243 to CM).

Data Availability Statement

The code used for our simulations is available here: http://sudbrack.me/LimitedDispersalSM. Otherwise, the authors affirm that all data necessary for confirming the conclusions of the article are present within the article and figures.

References

Agrawal, A. F. and Whitlock, M. C. (2011). Inferences about the distribution of dominance drawn from yeast gene knockout data. *Genetics*, 187(2):553–566.

Barrett, R. D. and Schluter, D. (2008). Adaptation from standing genetic variation. *Trends in ecology & evolution*, 23(1):38–44.

Barton, N. H. (1993). The probability of fixation of a favoured allele in a subdivided population. *Genetics Research*, 62(2):149–157.

Barton, N. H. (2000). Genetic hitchhiking. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 355(1403):1553–1562.

Barton, N. H. and Whitlock, M. C. (1997). The evolution of metapopulations. In *Metapopulation biology*, pages 183–210. Elsevier.

Benzie, J. H. (2000). Population genetic structure in penaeid prawns. *Aquaculture Research*, 31(1):95–119.

Billiard, S., Castric, V., and Llaurens, V. (2021). The integrative biology of genetic dominance. *Biological Reviews*, 96(6):2925–2942.

Charlesworth, B. (2020). How long does it take to fix a favorable mutation, and why should we care? *The American Naturalist*, 195(5):753–771.

Charlesworth, B. (2022). The effects of weak selection on neutral diversity at linked sites. *Genetics*, 221(1):iyac027.

Charlesworth, B. and Charlesworth, D. (2010). *Elements of evolutionary genetics*. Roberts and Company.

Cherry, J. L. (2003). Selection in a subdivided population with dominance or local frequency dependence. *Genetics*, 163(4):1511–1518.

Cherry, J. L. and Wakeley, J. (2003). A diffusion approximation for selection and drift in a subdivided population. *Genetics*, 163(1):421–428.

Clobert, J., Danchin, E., Dhondt, A. A., and Nichols, J. (2001). *Dispersal*. Oxford University Press.

Crow, J. and Kimura, M. (1970). *An Introduction to Population Genetics Theory*. Burgess Publishing Company.

Ewens, W. J. (2004). *Mathematical population genetics: theoretical introduction*, volume 1. Springer.

Ewing, G., Hermisson, J., Pfaffelhuber, P., and Rudolf, J. (2011). Selective sweeps for recessive alleles and for other modes of dominance. *Journal of mathematical biology*, 63(3):399–431.

Eyre-Walker, A. and Keightley, P. D. (2007). The distribution of fitness effects of new mutations. *Nature Reviews Genetics*, 8(8):610–618.

Fisher, R. A. (1928). The possible modification of the response of the wild type to recurrent mutations. *The American Naturalist*, 62(679):115–126.

Flintham, E. O., Savolainen, V., and Mullon, C. (2021). Dispersal alters the nature and scope of sexually antagonistic variation. *The American Naturalist*, 197(5):543–559.

Forstmeier, W., Segelbacher, G., Mueller, J. C., and Kempenaers, B. (2007). Genetic variation and differentiation in captive and wild zebra finches (taeniopygia guttata). *Molecular ecology*, 16(19):4039–4050.

Frank, S. A. (1998). *Foundations of social evolution*. Princeton University Press.

Giles, B. E. and Goudet, J. (1997). A case study of genetic structure in a plant metapopulation. In *Metapopulation biology*, pages 429–454. Elsevier.

Gillespie, J. H. (1994). *The causes of molecular evolution*, volume 2. Oxford University Press On Demand.

Glémin, S. and Ronfort, J. (2013). Adaptation and maladaptation in selfing and outcrossing species: new mutations versus standing variation. *Evolution: International Journal of Organic Evolution*, 67(1):225–240.

Glover, K. A., Pertoldi, C., Besnier, F., Wennevik, V., Kent, M., and Skaala, Ø. (2013). Atlantic salmon populations invaded by farmed escapees: quantifying genetic introgression with a bayesian approach and snps. *BMC genetics*, 14(1):1–19.

Haldane, J. B. S. (1927a). A mathematical theory of natural and artificial selection, part v: selection and mutation. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 23, pages 838–844. Cambridge University Press.

Haldane, J. B. S. (1927b). A mathematical theory of natural and artificial selection, part v: selection and mutation. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 23, pages 838–844. Cambridge University Press.

Hamilton, W. D. (1964). The genetical evolution of social behaviour. ii. *Journal of theoretical biology*, 7(1):17–52.

Hartfield, M. and Bataillon, T. (2020). Selective sweeps under dominance and inbreeding. *G3: Genes, Genomes, Genetics*, 10(3):1063–1075.

Hartl, D. and Clark, A. (2007). Principles of population genetics. 4th. *Sunderland, MA: Sinauer*.

Hendry, A. P. and Kinnison, M. T. (1999). Perspective: the pace of modern life: measuring rates of contemporary microevolution. *Evolution*, 53(6):1637–1653.

Hermisson, J. and Pennings, P. S. (2005). Soft sweeps: molecular population genetics of adaptation from standing genetic variation. *Genetics*, 169(4):2335–2352.

Hermisson, J. and Pennings, P. S. (2017). Soft sweeps and beyond: understanding the patterns and probabilities of selection footprints under rapid adaptation. *Methods in Ecology and Evolution*, 8(6):700–716.

Huber, C. D., Durvasula, A., Hancock, A. M., and Lohmueller, K. E. (2018). Gene expression drives the evolution of dominance. *Nature communications*, 9(1):2750.

Irvin, S. D., Wetterstrand, K. A., Hutter, C. M., and Aquadro, C. F. (1998). Genetic variation and differentiation at microsatellite loci in drosophila simulans: evidence for founder effects in new world populations. *Genetics*, 150(2):777–790.

Jensen, J. D. (2014). On the unfounded enthusiasm for soft selective sweeps. *Nature communications*, 5(1):5281.

Kacser, H. and Burns, J. A. (1981). The molecular basis of dominance. *Genetics*, 97(3-4):639–666.

Karlin, S. (1968). Equilibrium behavior of population genetic models with non-random mating. part ii: Pedigrees, homozygosity and stochastic models. *Journal of Applied Probability*, 5(3):487–566.

Kim, Y. and Maruki, T. (2011). Hitchhiking effect of a beneficial mutation spreading in a subdivided population. *Genetics*, 189(1):213–226.

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

Kimura, M. (1968). Evolutionary rate at the molecular level. *Nature*, 217:624–626.

Kimura, M., Maruyama, T., and Crow, J. F. (1963). The mutation load in small populations. *Genetics*, 48(10):1303.

Kimura, M. and Ohta, T. (1969). The average number of generations until fixation of a mutant gene in a finite population. *Genetics*, 61(3):763.

Kimura, M. and Ohta, T. (1971). On the rate of molecular evolution. *Journal of Molecular Evolution*, 1:1–17.

Kryazhimskiy, S. and Plotkin, J. B. (2008). The population genetics of dn/ds. *PLoS genetics*, 4(12):e1000304.

Kumar, S. and Singh, A. K. (2017). Population genetics of drosophila: genetic variation and differentiation among indian natural populations of drosophila ananassae. *Zoological studies*, 56.

Lanfear, R., Kokko, H., and Eyre-Walker, A. (2014). Population size and the rate of evolution. *Trends in ecology & evolution*, 29(1):33–41.

Lehmann, L. and Rousset, F. (2009). Perturbation expansions of multilocus fixation probabilities for frequency-dependent selection with applications to the hill–robertson effect and to the joint evolution of helping and punishment. *Theoretical Population Biology*, 76(1):35–51.

Lessard, S. (2009). Diffusion approximations for one-locus multi-allele kin selection, mutation and random drift in group-structured populations: a unifying approach to selection models in population genetics. *Journal of mathematical biology*, 59:659–696.

Li, J. and Bank, C. (2023). Dominance and multi-locus interaction. *arXiv preprint arXiv:2310.04730*.

Manna, F., Martin, G., and Lenormand, T. (2011). Fitness landscapes: an alternative theory for the dominance of mutation. *Genetics*, 189(3):923–937.

Marrec, L., Lamberti, I., and Bitbol, A.-F. (2021). Toward a universal model for spatially structured populations. *Physical Review Letters*, 127(21):218102.

Martin, G. and Lambert, A. (2015). A simple, semi-deterministic approximation to the distribution of selective sweeps in large populations. *Theoretical Population Biology*, 101:40–46.

McCandlish, D. M. and Stoltzfus, A. (2014). Modeling evolution using the probability of fixation: history and implications. *The Quarterly review of biology*, 89(3):225–252.

Messer, P. W. and Petrov, D. A. (2013). Population genomics of rapid adaptation by soft selective sweeps. *Trends in ecology & evolution*, 28(11):659–669.

Mukai, T., Chigusa, S. I., Mettler, L., and Crow, J. F. (1972). Mutation rate and dominance of genes affecting viability in drosophila melanogaster. *Genetics*, 72(2):335–355.

Muralidhar, P. and Veller, C. (2022). Dominance shifts increase the likelihood of soft selective sweeps. *Evolution*, 76(5):966–984.

Newberry, M. G., McCandlish, D. M., and Plotkin, J. B. (2016). Assortative mating can impede or facilitate fixation of underdominant alleles. *Theoretical population biology*, 112:14–21.

Orr, H. A. (2010). The population genetics of beneficial mutations. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1544):1195–1201.

Orr, H. A. and Betancourt, A. J. (2001). Haldane's sieve and adaptation from the standing genetic variation. *Genetics*, 157(2):875–884.

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

Otto, S. P. and Whitlock, M. C. (2001). Fixation probabilities and times. *e LS*.

Pannell, J. R. and Charlesworth, B. (1999). Neutral genetic diversity in a metapopulation with recurrent local extinction and recolonization. *Evolution*, 53(3):664–676.

Paulose, J., Hermisson, J., and Hallatschek, O. (2019). Spatial soft sweeps: patterns of adaptation in populations with long-range dispersal. *PLoS genetics*, 15(2):e1007936.

Pennings, P. S. and Hermisson, J. (2006a). Soft sweeps ii — molecular population genetics of adaptation from recurrent mutation or migration. *Molecular biology and evolution*, 23(5):1076–1084.

Pennings, P. S. and Hermisson, J. (2006b). Soft sweeps iii: the signature of positive selection from recurrent mutation. *PLoS genetics*, 2(12):e186.

Pennings, P. S. and Hermisson, J. (2006c). Soft sweeps iii: the signature of positive selection from recurrent mutation. *PLoS genetics*, 2(12):e186.

Pennings, P. S. and Hermisson, J. (2006d). Soft sweeps ii—molecular population genetics of adaptation from recurrent mutation or migration. *Molecular biology and evolution*, 23(5):1076–1084.

Potenko, V. and Velikov, A. (1998). Genetic diversity and differentiation of natural populations of *Pinus koraiensis* (sieb. et zucc.) in russia. *Silvae Genetica*, 47(4):202–207.

Ralph, P. and Coop, G. (2010). Parallel adaptation: one or many waves of advance of an advantageous allele? *Genetics*, 186(2):647–668.

Rousset, F. (2004). *Genetic structure and selection in subdivided populations*, volume 40. Princeton University Press.

Rousset, F. (2006). Separation of time scales, fixation probabilities and convergence to evolutionarily stable states under isolation by distance. *Theoretical Population Biology*, 69(2):165–179.

Roze, D. and Rousset, F. (2003). Selection and drift in subdivided populations: a straightforward method for deriving diffusion approximations and applications involving dominance, selfing and local extinctions. *Genetics*, 165(4):2153–2166.

Roze, D. and Rousset, F. (2004). The robustness of hamilton's rule with inbreeding and dominance: kin selection and fixation probabilities under partial sib mating. *The American Naturalist*, 164(2):214–231.

Roze, D. and Rousset, F. (2008). Multilocus models in the infinite island model of population structure. *Theoretical Population Biology*, 73(4):529–542.

Santiago, E. and Caballero, A. (2005). Variation after a selective sweep in a subdivided population. *Genetics*, 169(1):475–483.

Slatkin, M. (1977). Gene flow and genetic drift in a species subject to frequent local extinctions. *Theoretical population biology*, 12(3):253–262.

Slatkin, M. (1981). Fixation probabilities and fixation times in a subdivided population. *Evolution*, pages 477–488.

Slatkin, M. and Wiehe, T. (1998). Genetic hitch-hiking in a subdivided population. *Genetics Research*, 71(2):155–160.

Smith, J. M. and Haigh, J. (1974). The hitch-hiking effect of a favourable gene. *Genetics Research*, 23(1):23–35.

Ståhl, G. (1981). Genetic differentiation among natural populations of atlantic salmon (*Salmo salar*) in northern sweden. *Ecological Bulletins*, pages 95–105.

Tamaki, I., Setsuko, S., and Tomaru, N. (2008). Genetic variation and differentiation in populations of a threatened tree, magnolia stellata: factors influencing the level of within-population genetic variation. *Heredity*, 100(4):415–423.

Teshima, K. M. and Przeworski, M. (2006). Directional positive selection on an allele of arbitrary dominance. *Genetics*, 172(1):713–718.

Wakeley, J. (2003). Polymorphism and divergence for island-model species. *Genetics*, 163(1):411–420.

Wakeley, J. and Takahashi, T. (2004). The many-demes limit for selection and drift in a subdivided population. *Theoretical population biology*, 66(2):83–91.

Wang, J. (1997). Effective size and f-statistics of subdivided populations. ii. dioecious species. *Genetics*, 146(4):1465–1474.

Wang, J. and Caballero, A. (1999). Developments in predicting the effective size of subdivided populations. *Heredity*, 82(2):212–226.

Whitlock, M. C. (2002). Selection, load and inbreeding depression in a large metapopulation. *Genetics*, 160(3):1191–1202.

Whitlock, M. C. (2003). Fixation probability and time in subdivided populations. *Genetics*, 164(2):767–779.

Whitlock, M. C. and Barton, N. H. (1997). The effective size of a subdivided population. *Genetics*, 146(1):427–441.

Whitlock, M. C. and McCauley, D. E. (1990). Some population genetic consequences of colony formation and extinction: genetic correlations within founding groups. *Evolution*, 44(7):1717–1724.

Wright, S. (1931). Evolution in mendelian populations. *Genetics*, 16(2):97.

Wright, S. (1934). Physiological and evolutionary theories of dominance. *The American Naturalist*, 68(714):24–53.

Wright, S., Teissier, G., et al. (1939). Statistical genetics in relation to evolution. *Actualites Scientifiques et Industrielles 802. Exposes de Biometrie et de la Statistique Biologique XIII.*

Wright, S. D., Gillman, L. N., Ross, H. A., and Keeling, D. J. (2009). Slower tempo of microevolution in island birds: implications for conservation biology. *Evolution*, 63(9):2275–2287.

Symbol	Expression	Limit
r_0^{D}	$\dfrac{(1-d)^2}{2N_{\mathrm{ind}}-(1-d)^2(2N_{\mathrm{ind}}-1)}$	$\dfrac{1}{1+4N_{\mathrm{ind}}d}$
r_1^{D}	"	"
r_1^{R}	$\dfrac{1}{2N_{\mathrm{ind}}}+\left(1-\dfrac{1}{2N_{\mathrm{ind}}}\right)r_1^{\mathrm{D}}$	"
a^{R}	$\dfrac{1}{N_{\mathrm{ind}}}r_0^{\mathrm{D}}+\left(1-\dfrac{1}{N_{\mathrm{ind}}}\right)\dfrac{\left[1+3\left(2N_{\mathrm{ind}}-1\right)r_1^{\mathrm{D}}\right](1-d)^3}{\left(2N_{\mathrm{ind}}\right)^2-\left(2N_{\mathrm{ind}}-1\right)\left(2N_{\mathrm{ind}}-2\right)(1-d)^3}$	$\dfrac{1}{(1+4N_{\mathrm{ind}}d)\,(1+2N_{\mathrm{ind}}d)}$

Table 1: Probabilities of coalescence in the island model. Expressions for the various probabilities of coalescence that are relevant to the analysis, including their values in the limit of low dispersal and large patches ($d \to 0$ and $N \to \infty$ such that Nd remains constant, see Appendix A.3 for derivations). r_0^{D} is the probability that the two homologous genes of the same individual are identical-by-descent (IBD), which is equivalent to F_{IT} in the island model of dispersal; r_1^{D} is the probability that two different genes sampled from the same deme are IBD, which is equivalent to F_{ST} in the island model of dispersal; r_1^{R} is the probability that two genes sampled from the same deme with replacement are IBD; and a^{R} is the probability that two homologous genes of an individual coalesce with a third gene sampled from the same deme at random (Roze and Rousset, 2003; Rousset, 2004).

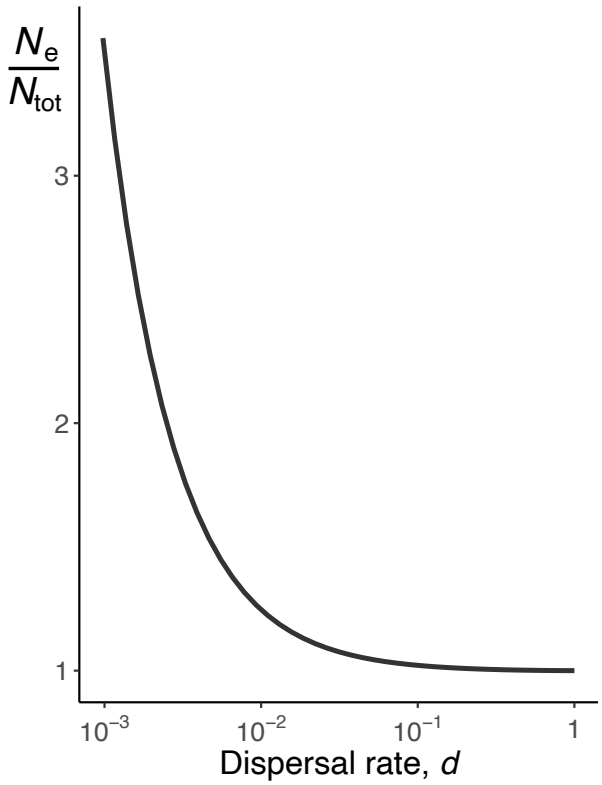


Figure 1: Effective population size relative to census size, N_e/N_{tot} (from eq. 3 with r_1^D in Table 1). In a well-mixed population ($d = 1$), $N_e = N_{tot}$, otherwise limited dispersal increases N_e relative to N_{tot}. Parameters: $N_d = 200, N_{ind} = 100$.

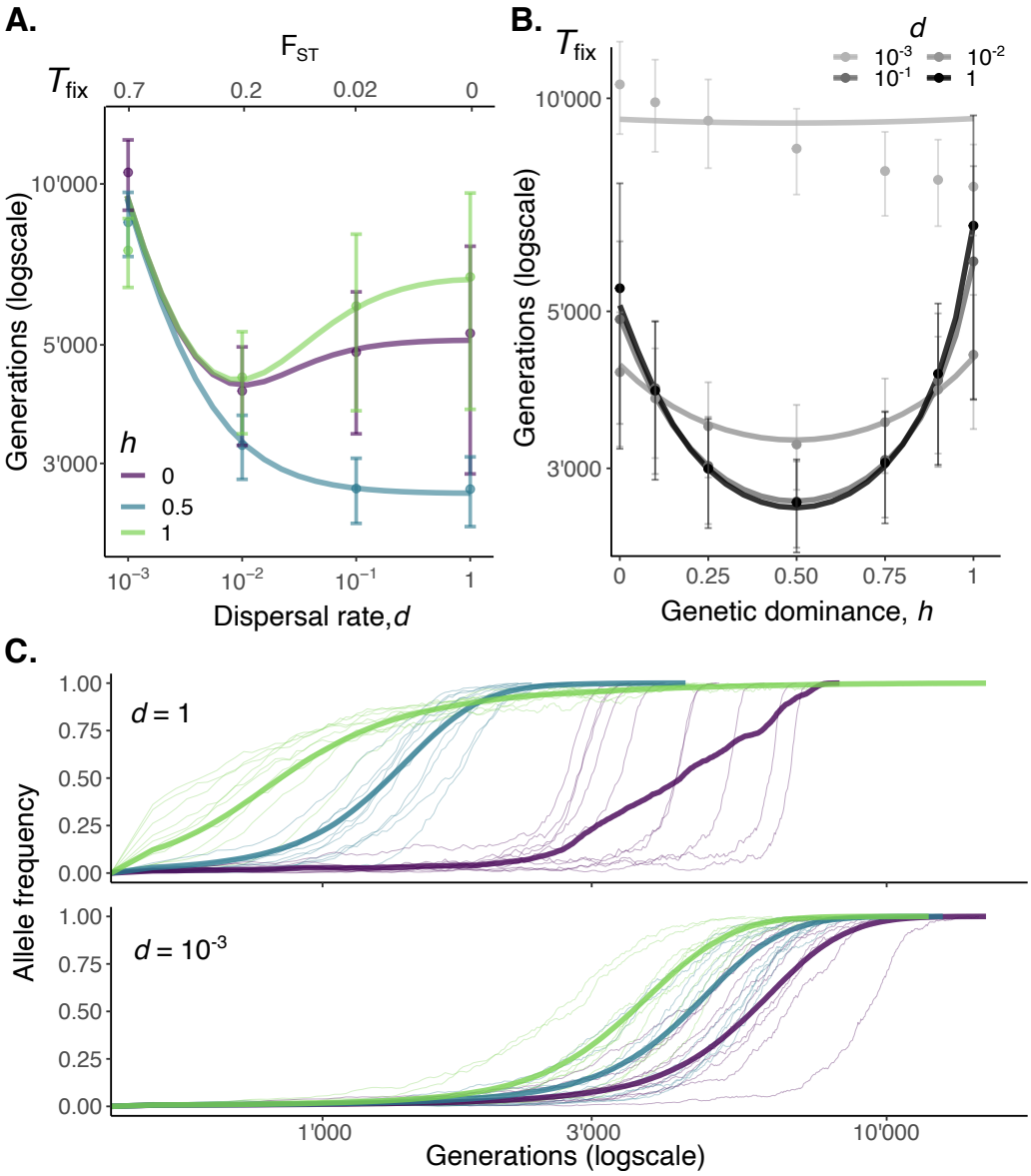


Figure 2: Time and trajectory to fixation of beneficial mutations according to dispersal. **A.** Expected time to fixation $T_{\text{fix}}(p_0)$ of recessive ($h = 0$), additive ($h = 0.5$) and dominant ($h = 1$) alleles (in different colors, see legend) arising as single copies $p_0 = 1/(2N_{\text{tot}})$, with solid lines from diffusion approximation (i.e. numerical integration of eq. 5) and dots as averages from simulations (fixation events among 40'000 replicates for each set of parameters, error bars show standard deviation, Appendix B for details). Parameters: $N_{\text{d}} = 200$, $N_{\text{ind}} = 100$, $s = 0.01$. Under strong dispersal limitation, $d \lesssim 10^{-3}$, the time to fixation asymptotically approaches the neutral expectation $T_{\text{fix}} \sim 4N_{\text{e}}$ (Kimura and Ohta, 1969), regardless of genetic dominance. **B.** Expected time to fixation $T_{\text{fix}}(p_0)$, i.e. same as A, but plotted against dominance for different levels of dispersal (see legend). **C.** Fixation trajectories of beneficial mutations in a well-mixed (top, $d = 1$) and dispersal-limited (bottom, $d = 10^{-3}$) population. For each level of dominance (in different colours, see A for legend), thin lines show ten randomly sampled trajectories, thick lines show the mean trajectory among all trajectories. Parameters: same as A.

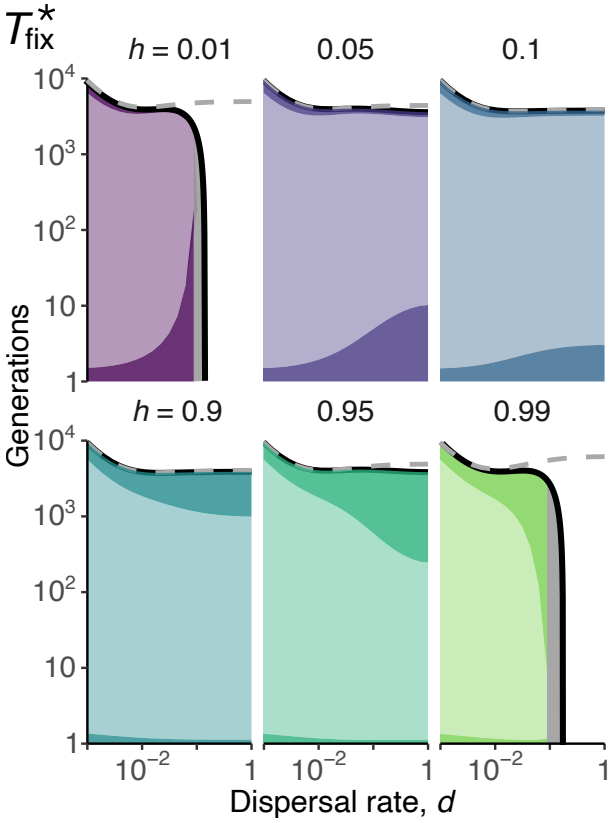


Figure 3: Semi deterministic approximation to fixation time under limited dispersal. Solid black lines show Charlesworth (2020)'s approximation to the expected time to fixation T^*_{fix} (eq. 7) of partially recessive (h = 0.01, 0.05, 0.1; top row) and partially dominant (h = 0.9, 0.95, 0.99; bottom row) alleles arising as single copies $p_0 = 1/(2N_{\text{tot}})$; and dashed gray lines show numerical integration of eq. (5). Shaded regions below curve represent the proportion of time spend in each phase of the approximation of eq. (7), from bottom to top: initial stochastic phase (dark shade), deterministic phase (light shade), final stochastic phase (dark shade). The shaded gray areas in top left ($h = 0.01$) and bottom right ($h = 0.99$) graphs indicate where eq. (7) diverges. Other parameters: same as Fig. 2.

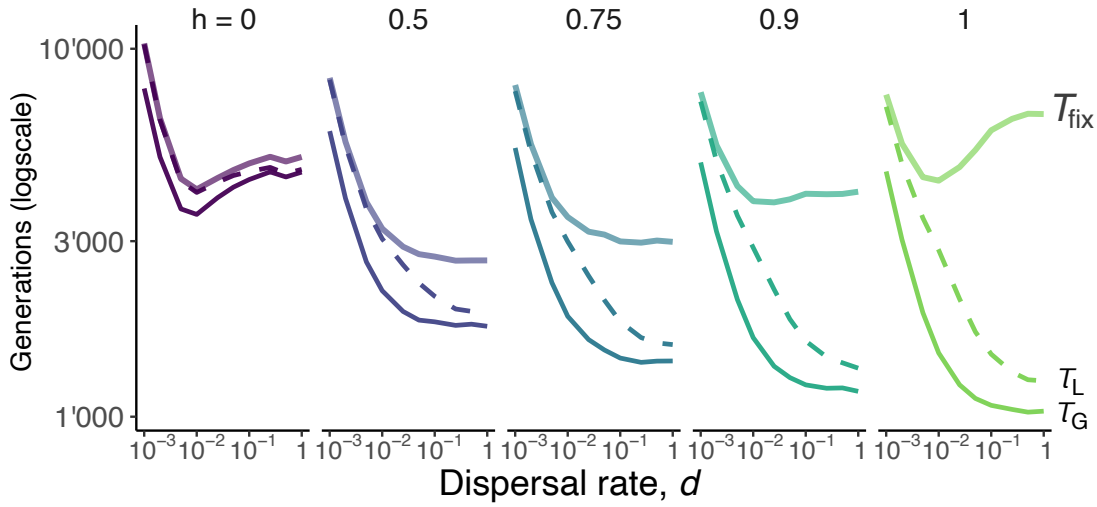


Figure 4: Expected time to purge 90% of the genetic load. We define the genetic load as $L = (1+s-z)/s$, i.e. as the difference between the population mean fecundity, z, and the maximum fecundity, $1+s$, normalized such that a population monomorphic for the wild-type allele, aa ($z = 1$), has $L = 1$, while a population where the beneficial mutation A has fixed has no genetic load, $L = 0$. We also define a local genetic load L_i in each deme i as $L_i = (1+s-z_i)/s$, where z_i is the mean fecundity at deme i. Plots show the average time to fixation T_{fix} (thick line), average time to purge 90% of the load τ_{G}, i.e. average time for $L = 0.1$ (thin line), and average time to purge 90% of genetic load in every deme τ_{L}, i.e. average time for $\max_i L_i = 0.1$ (dashed line), for recessive ($h = 0$), additive ($h = 0.5$) and (partially) dominant ($h = 0.75, 0.9$ and 1) alleles (in different columns) arising as single copies $p_0 = 1/(2N_{\text{tot}})$ in individual-based simulations (Appendix B for details). Other parameters: same as Fig. 2.

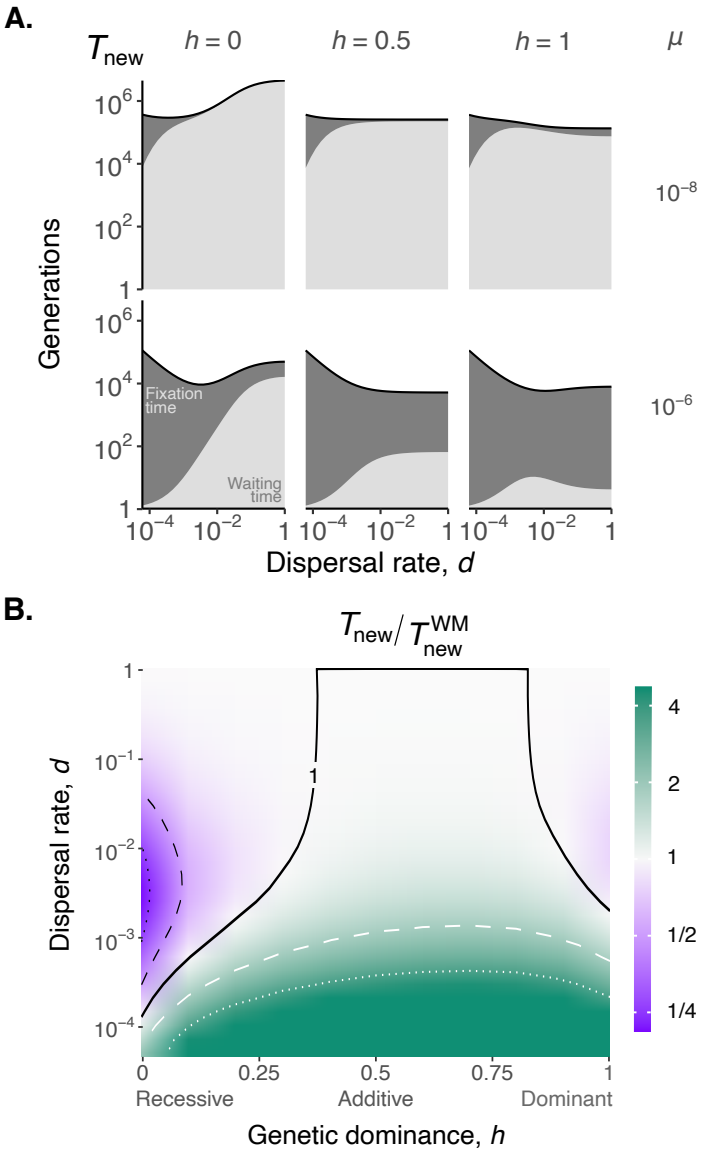


Figure 5: Total time before a *de novo* mutations arises and fixes under limited dispersal. A. Expected total time T_{new} of fixation (on a log scale) of a recessive ($h = 0$, left), additive ($h = 0.5$, middle) and dominant ($h = 1$, right) *de novo* mutations $p_0 = 1/(2N_{tot})$ for different mutation rates ($\mu = 10^{-8}$, top; $\mu = 10^{-6}$, bottom), with solid black lines from eq. (10) (with eqs. 4-5). Dark and light gray shades underneath curves represent the proportion of time spent in each component of T_{new} (on a linear scale). Parameters: same as Fig. 2. **B.** Effect of population subdivision on the total time of fixation according to dispersal rate d and genetic dominance h. Ratio between the expected time to fixation of *de novo* mutations under limited dispersal T_{new} and panmixia (T_{new}^{WM}, i.e. where $d = 1$, both from eq. 10 with eqs. 4-5). Full contour shows $T_{new}/T_{new}^{WM} = 1$; $T_{new}/T_{new}^{WM} = 2$ and 4 in black dashed and dotted, respectively; $T_{new}/T_{new}^{WM} = 1/2$ and 1/4 in white dashed and dotted, respectively. Parameters: $\mu = 10^{-6}$, other parameters: same as Fig. 2.

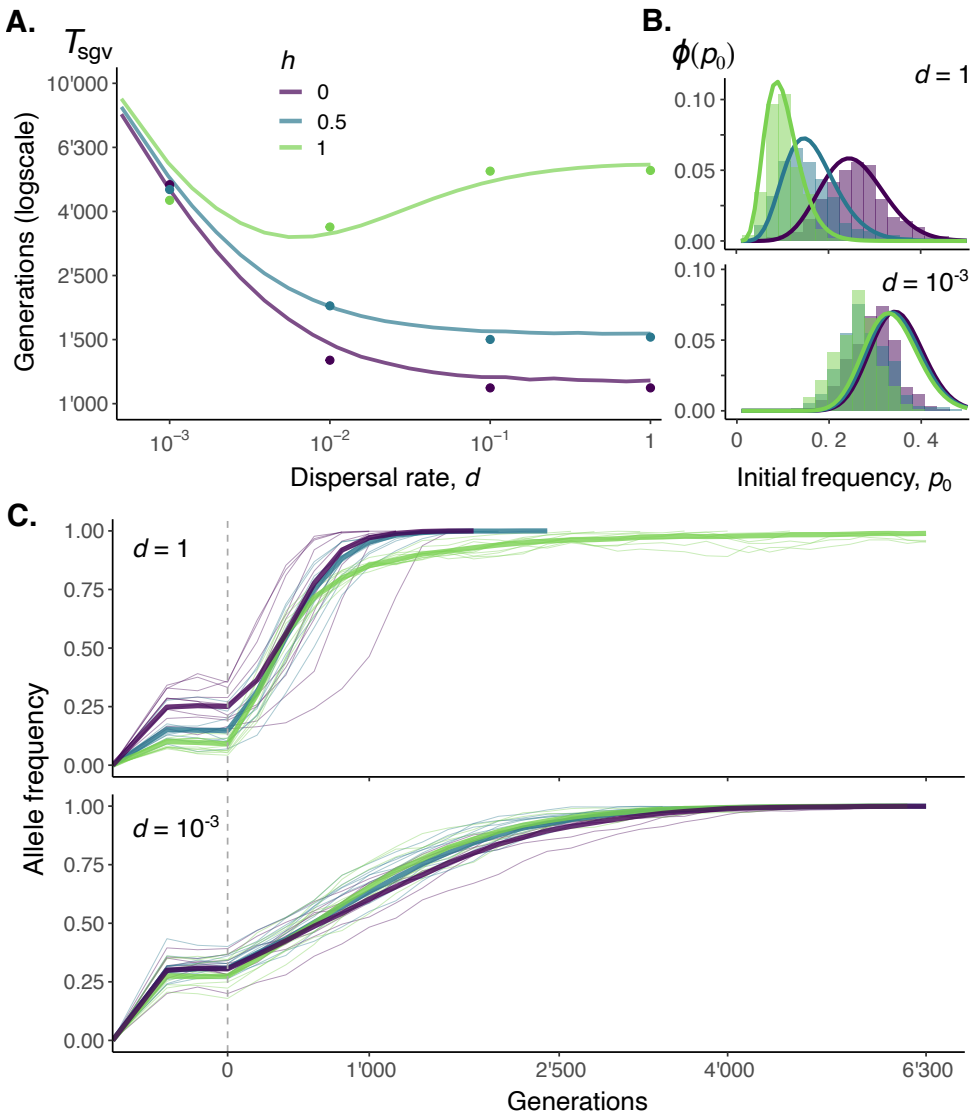


Figure 6: Fixation times of standing genetic variants. **A.** Expected time T_{sgv} of fixation (on a log scale) of a standing recessive ($h = 0$), additive ($h = 0.5$) and dominant ($h = 1$) variants (in different colors, see legend) with solid lines from diffusion approximation (eq. 11 with eqs. 5 and A49), and dots for the average from individual based simulations (300 replicates for each set of parameters, Appendix B for details; we do not show standard deviations of simulations here as they are typically large and therefore lead to an overcrowded figure; this is because T_{sgv} is affected by two sources of variance: variance in p_0 and in the time to fixation). Parameters: $s_D = 10^{-3}$, $\mu = 10^{-4}$, $h_D = h$, other parameters: same as Fig. 2. **B.** Distribution of initial frequencies $\phi(p_0)$ at the moment of environmental change in a well-mixed (top, $d = 1$) and dispersal-limited (bottom, $d = 10^{-3}$) population. Vertical bars represent histograms of simulations and lines from diffusion approximation (eq. A49). Note that the diffusion approximation fares less well when $d = 10^{-3}$ as dispersal is much weaker than selection (Roze and Rousset, 2003; Wakeley, 2003). Parameters: same as A. **C.** Fixation of standing genetic variants in a well-mixed (top, $d = 1$) and a dispersal-limited (bottom, $d = 10^{-3}$) population. Environmental change takes place at $t = 0$ (dashed vertical line). For each level of dominance (in different colours, see A for legend), thin lines show ten randomly sampled trajectories, thick lines show the mean trajectory among all trajectories. Parameters: same as A.

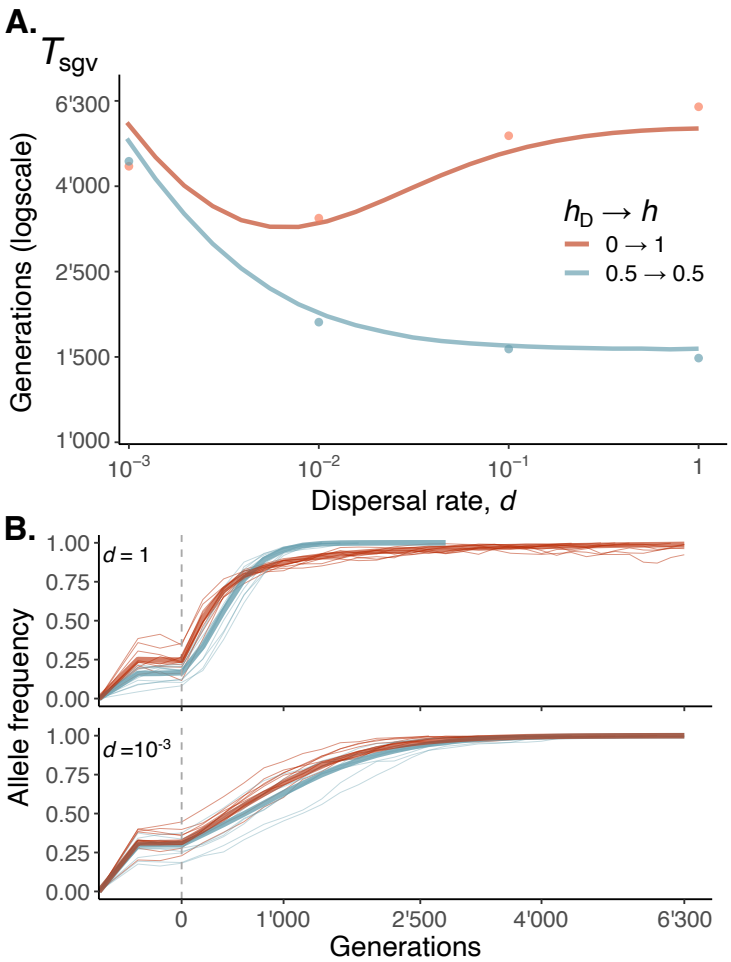


Figure 7: The effects of dominance reversal under limited dispersal. **A.** Expected time T_{sgv} of fixation (on a log scale) of recessive deleterious alleles that become beneficial dominant ($h_D = 0$ and $h = 1$, in red), and of additive alleles ($h_D = h = 0.5$ in blue) with solid lines from diffusion approximation (eq. 11 with eqs. 5 and A49), and dots for the average from individual based simulations (300 replicates for each set of parameters, Appendix B for details). Parameters: same as Fig. 6. **B.** Fixation trajectories of alleles showing dominance reversal ($h_D = 0$ and $h = 1$, in red) and additive alleles ($h_D = h = 0.5$ in blue) in a well-mixed (top, $d = 1$) and dispersal-limited (bottom, $d = 10^{-3}$) population. Environmental change takes place at $t = 0$ (dashed vertical line). For each scenario (in different colours, see A for legend), thin lines show ten randomly sampled trajectories, and thick lines show the mean trajectory among all trajectories. Parameters: same as A.

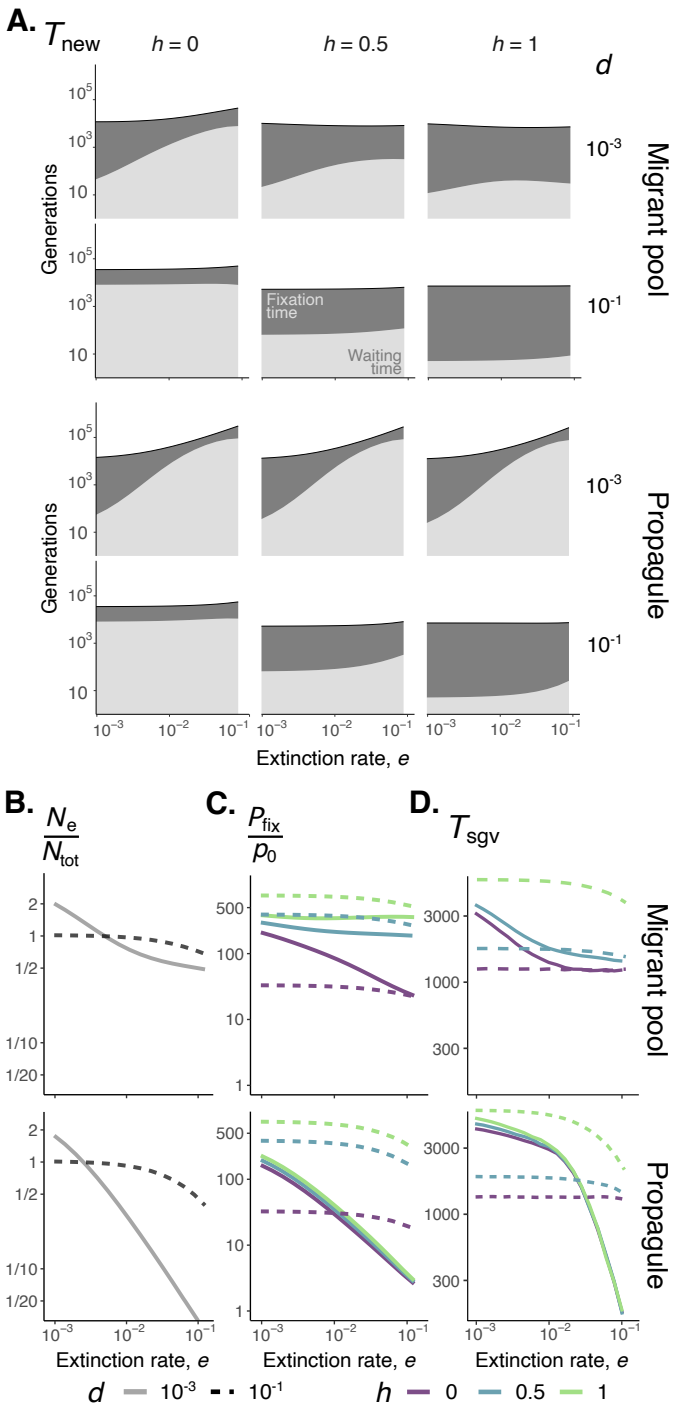


Figure 8: The effects of local extinctions and recolonization dynamics. **A.** Expected time T_{new} of fixation (on a log scale) of recessive ($h = 0$, left), additive ($h = 0.5$, middle) and dominant ($h = 1$, right) *de novo* mutations, $p_0 = 1/(2N_{\text{tot}})$, for different dispersal rates ($d = 10^{-1}$ in bottom, and 10^{-3} in top) and recolonization models (migrant pool model in top row and propagule model in bottom, also for panels B-D), from eq. (10) (with eqs. 4 and 5). Parameters: same as Fig. 2. **B.** Effective population size relative to census size, N_e/N_{tot} (from eq. C4 in Appendix). Dashed line for $d = 10^{-1}$ and full line for $d = 10^{-3}$ (also for panels C-D). Parameters: same as A. **C.** Fixation probabilities normalised to initial frequency p_0 of recessive ($h = 0$), additive ($h = 0.5$) and dominant ($h = 1$) alleles (in different colours, see legend) arising as single copies $p_0 = 1/(2N_{\text{tot}})$, from eq. (4) with $d = 10^{-1}$ (dashed) and 10^{-3} (full), and under different recolonization models (top and bottom rows). Parameters: same as A. **D.** Expected time T_{sgv} of fixation (on a log scale) of standing recessive ($h = 0$), additive ($h = 1/2$) and dominant ($h = 1$) variants (in different colors) from eq. 11 (with eqs. 5 and A49). The case $h = 1$ with $d = 10^{-3}$ is omitted as comparisons with simulations showed a poor fit (Fig. C1 in Appendix). Parameters: same as Fig. 6.

Appendices

for

Fixation times of *de novo* and standing beneficial variants in subdivided populations

Vitor Sudbrack and Charles Mullon [1]

List of appendices

A Diffusion approximation for baseline model

B Computer simulations

C Extinction-recolonization dynamics

D Hard and soft sweeps

E Hard selection

[1] Correspondence: charles.mullon@unil.ch

A Diffusion approximation for baseline model

Here, we briefly describe how we used the diffusion approximation to model the segregation of allele A following Roze and Rousset (2003) for the island model of dispersal (see also Barton, 1993; Whitlock, 2003; Cherry and Wakeley, 2003; Cherry, 2003; Wakeley, 2003; Wakeley and Takahashi, 2004). Readers familiar with this approach may skip this appendix, which we present for the sake of completeness.

The diffusion approximation relies on two quantities: the expectation of, and the variance in, the change in frequency over one generation (Crow and Kimura, 1970; Ewens, 2004). To specify this frequency change, consider a given generation t. Let $p_{ij1} \in \{0,1\}$ and $p_{ij2} \in \{0,1\}$ be the frequency of A at each homologous chromosomes of individual $j \in \{1,\ldots,N_{\text{ind}}\}$ in deme $i \in \{1,\ldots,N_{\text{d}}\}$ at this generation t. The allele frequency then is

$$p_{ij} = \frac{p_{ij1} + p_{ij2}}{2} \tag{A1}$$

in that individual, and

$$p = \frac{1}{N_{\text{tot}}} \sum_{i=1}^{N_{\text{d}}} \sum_{j=1}^{N_{\text{ind}}} p_{ij} \tag{A2}$$

in the whole population (recall $N_{\text{tot}} = N_{\text{ind}} \cdot N_{\text{d}}$). We use a prime symbol to denote quantities in the next generation $t+1$, so that p' refers to the frequency of A in the next generation, and thus

$$\Delta p = p' - p \tag{A3}$$

is the allele frequency change.

A.1 Expected allele frequency change

Here, we derive eq. (1) of the main text, i.e. the expected frequency change over one generation, $\mathrm{E}[\Delta p|p]$, to the first order of a small parameter δ such that $s \sim \mathcal{O}(\delta)$ and $N_{\text{d}} \sim \mathcal{O}(1/\delta)$. Under random chromosomal segregation and in the absence of mutation, this expectation

can be written as

$$\mathrm{E}[\Delta p|p] = \mathrm{E}\left[\underbrace{\frac{1}{2N_{\mathrm{tot}}}\sum_{i=1}^{N_{\mathrm{d}}}\sum_{j=1}^{N_{\mathrm{ind}}}\mathrm{E}[w_{ij}p_{ij}|\mathbf{P}]-p}_{\mathrm{E}[\Delta p|\mathbf{P}]}\;\middle|\;p\right], \tag{A4}$$

where $\mathbf{P}$ is the set of realised genotypes of all individuals, i.e. $\mathbf{P} = \left\{p_{ij\lambda} : \lambda \in \{1,2\}, j \in \{1,\ldots,N_{\mathrm{ind}}\}, i \in \{1,\ldots,N_{\mathrm{d}}\}\right\}$, and w_{ij} in eq. (A4) is the expected number of successful gametes (i.e. that are recruited) produced by individual j from deme i, given the genotype of every individual in its generation. We refer to w_{ij} as the fitness of individual j from deme i.

A.1.1 Fitness

To specify the fitness w_{ij} of a focal individual, we first connect relative fecundity z_{ij} of this individual with its genotype through the expression

$$z_{ij} = 1 + s\left[2hp_{ij} + (1-2h)p_{ij1}p_{ij2}\right], \tag{A5}$$

such that aa, aA, and AA individuals have relative fecundity 1, $1+hs$, and $1+s$, respectively (as per the model described in section 2.1 of the main text; eq. 11 in Roze and Rousset, 2003 with their $z_{\mathrm{aa}} = 1$). In the limit where absolute fecundity is large, fitness for our baseline model can be written as

$$w_{ij} = \underbrace{2N_{\mathrm{ind}}\cdot(1-d)\cdot\frac{z_{ij}}{\sum_{k=1}^{N_{\mathrm{ind}}} z_{ik}}}_{\text{philopatric fitness}} + \underbrace{\sum_{\substack{r=1\\ r\neq i}}^{N_{\mathrm{d}}} 2N_{\mathrm{ind}}\cdot\frac{d}{N_{\mathrm{d}}-1}\cdot\frac{z_{ij}}{\sum_{k=1}^{N_{\mathrm{ind}}} z_{ik}}}_{\text{dispersal fitness}}, \tag{A6}$$

which consists in the sum of philopatric fitness (expected number of gametes recruited in their deme of origin) and dispersal fitness (recruited in all other demes). Equation (A6) simplifies to

$$w_{ij} = 2\cdot\frac{z_{ij}}{z_i}, \tag{A7}$$

where

$$z_i = \frac{1}{N_{\mathrm{ind}}}\sum_{j=1}^{N_{\mathrm{ind}}} z_{ij} \tag{A8}$$

is the average fecundity in deme i.

More generally, we consider in the next two sections models where individual fitness can be written to the first order of δ as a function of three variables,

$$w_{ij} = w(z_{ij}, z_i, z) + \mathcal{O}\left(\delta^2\right), \tag{A9}$$

where

$$z = \frac{1}{N_\mathrm{d}} \sum_{i=1}^{N_\mathrm{d}} z_i \tag{A10}$$

is the population average fecundity. One useful property of fitness is that since the total population size N_tot is constant, the average population fitness is also fixed at $\sum_{i=1}^{N_\mathrm{d}} \sum_{j=1}^{N_\mathrm{ind}} w_{ij}/N_\mathrm{tot} = 2$. As a consequence, we also have

$$\left(\frac{\partial w_{ij}}{\partial z_{ij}} + \frac{\partial w_{ij}}{\partial z_i} + \frac{\partial w_{ij}}{\partial z}\right)_{s=0} = 0 \tag{A11}$$

(p. 96 in Rousset, 2004).

A.1.2 Conditional expected allele frequency change

To compute eq. (A4), we follow Roze and Rousset (2003) and first characterise the expected allele frequency change conditioned on all genotypes, $\mathrm{E}[\Delta p|\mathbf{P}]$. Plugging eq. (A9) into eq. (A4) and performing a Taylor expansion around $s = 0$ obtains

$$\mathrm{E}[\Delta p|\mathbf{P}] = \frac{s}{2N_\mathrm{tot}} \sum_{i=1}^{N_\mathrm{d}} \sum_{j=1}^{N_\mathrm{ind}} \left(\frac{\partial w_{ij}}{\partial z_{ij}} \frac{dz_{ij}}{ds} + \frac{\partial w_{ij}}{\partial z_i} \frac{dz_i}{ds} + \frac{\partial w_{ij}}{\partial z} \frac{dz}{ds}\right) p_{ij} + \mathcal{O}\left(\delta^2\right), \tag{A12}$$

where here and hereafter all derivatives are estimated under neutrality (i.e., $s = \delta = 0$ and thereby $z_{ij} = z_i = z = 1$). Using eqs. (A5), (A8) and (A10), eq. (A12) then becomes

$$\begin{aligned}\mathrm{E}[\Delta p|\mathbf{P}] = \frac{s}{2N_{\mathrm{tot}}}\sum_{i=1}^{N_{\mathrm{d}}}\sum_{j=1}^{N_{\mathrm{ind}}}\Bigg(&\frac{\partial w_{ij}}{\partial z_{ij}}\left[2hp_{ij} + (1-2h)p_{ij1}p_{ij2}\right] \\ &+ \frac{1}{N_{\mathrm{ind}}}\sum_{k=1}^{N_{\mathrm{ind}}}\frac{\partial w_{ij}}{\partial z_i}\left[2hp_{ik} + (1-2h)p_{ik1}p_{ik2}\right] \\ &+ \frac{1}{N_{\mathrm{tot}}}\sum_{r=1}^{N_{\mathrm{d}}}\sum_{k=1}^{N_{\mathrm{ind}}}\frac{\partial w_{ij}}{\partial z}\left[2hp_{rk} + (1-2h)p_{rk1}p_{rk2}\right]\Bigg)p_{ij} + \mathcal{O}\left(\delta^2\right) .\end{aligned} \tag{A13}$$

This can be written after some straightforward rearrangements as

$$\begin{aligned}\mathrm{E}[\Delta p|\mathbf{P}] = hs\Bigg(&\frac{\partial w_{ij}}{\partial z_{ij}}\left\langle p_{ij}^2\right\rangle_{ij} + \frac{\partial w_{ij}}{\partial z_i}\left\langle p_i^2\right\rangle_i + \frac{\partial w_{ij}}{\partial z}p^2\Bigg) \\ &+ (1-2h)\frac{s}{2}\Bigg(\frac{\partial w_{ij}}{\partial z_{ij}}\left\langle p_{ij1}p_{ij2}\right\rangle_{ij} + \frac{\partial w_{ij}}{\partial z_i}\left\langle p_{ij1}p_{ij2}p_i\right\rangle_{ij} + \frac{\partial w_{ij}}{\partial z}p\left\langle p_{ij1}p_{ij2}\right\rangle_{ij}\Bigg) \\ &+ \mathcal{O}\left(\delta^2\right) ,\end{aligned} \tag{A14}$$

where we used the convention

$$\left\langle f_{ij}\right\rangle_{ij} = \frac{1}{N_{\mathrm{tot}}}\sum_{i=1}^{N_{\mathrm{d}}}\sum_{j=1}^{N_{\mathrm{ind}}} f_{ij} , \qquad \left\langle f_{ij}\right\rangle_{j} = \frac{1}{N_{\mathrm{ind}}}\sum_{j=1}^{N_{\mathrm{ind}}} f_{ij} , \qquad \left\langle f_{i}\right\rangle_{i} = \frac{1}{N_{\mathrm{d}}}\sum_{i=1}^{N_{\mathrm{d}}} f_{i} , \tag{A15}$$

for any quantity f_{ij} that can be ascribed to an individual ij, and the property that

$$\left\langle p_{ij1}p_{ij2}p_{ij}\right\rangle_{ij} = \left\langle p_{ij1}p_{ij2}\right\rangle_{ij} \tag{A16}$$

since p_{ij1} and p_{ij2} are indicator variables. Equation (A14) is equivalent to eq. (14) in Roze and Rousset (2003). Using eq. (A11) to substitute for $\frac{\partial w_{ij}}{\partial z}$, eq. (A14) becomes

$$\begin{aligned}\mathrm{E}[\Delta p|\mathbf{P}] = hs\Bigg[&\frac{\partial w_{ij}}{\partial z_{ij}}\left(\left\langle p_{ij}^2\right\rangle_{ij} - p^2\right) + \frac{\partial w_{ij}}{\partial z_i}\left(\left\langle p_i^2\right\rangle_i - p^2\right)\Bigg] \\ &+ (1-2h)\frac{s}{2}\left[(1-p)\frac{\partial w_{ij}}{\partial z_{ij}}\left\langle p_{ij1}p_{ij2}\right\rangle_{ij} + \frac{\partial w_{ij}}{\partial z_i}\left\langle p_{ij1}p_{ij2}(p_i - p)\right\rangle_{ij}\right] + \mathcal{O}\left(\delta^2\right) .\end{aligned} \tag{A17}$$

A.1.3 Unconditional expected allele frequency change

We then plug eq. (A17) into eq. (A4) to obtain

$$\begin{aligned}\mathrm{E}[\Delta p|p] = hs\left(\frac{\partial w_{ij}}{\partial z_{ij}}\left(\mathrm{E}\left[\left\langle p_{ij}^2\right\rangle_{ij}\middle|p\right]-p^2\right)+\frac{\partial w_{ij}}{\partial z_i}\left(\mathrm{E}\left[\left\langle p_i^2\right\rangle_i\middle|p\right]-p^2\right)\right)\\ +(1-2h)\frac{s}{2}\left((1-p)\frac{\partial w_{ij}}{\partial z_{ij}}\mathrm{E}\left[\left\langle p_{ij1}p_{ij2}\right\rangle_{ij}\middle|p\right]+\frac{\partial w_{ij}}{\partial z_i}\mathrm{E}\left[\left\langle p_{ij1}p_{ij2}(p_i-p)\right\rangle_{ij}\middle|p\right]\right)\\ +\mathcal{O}\left(\delta^2\right),\end{aligned} \tag{A18}$$

which depends on various moments of the genetic distribution in the population. These only need to be evaluated under neutrality as they are already weighed by s in eq. (A18). In that case, they can be connected to probabilities of identity-by-descent under neutrality (Roze and Rousset, 2003, for more details). In fact,

$$\mathrm{E}\left[\left\langle p_{ij1}p_{ij2}\right\rangle_{ij}\middle|p\right] = r_0^{\mathrm{D}}p+\left(1-r_0^{\mathrm{D}}\right)p^2+\mathcal{O}(\delta) \tag{A19}$$

$$\mathrm{E}\left[\left\langle p_{ij}^2\right\rangle_{ij}\middle|p\right] = r_0^{\mathrm{R}}p+\left(1-r_0^{\mathrm{R}}\right)p^2+\mathcal{O}(\delta) \tag{A20}$$

$$\mathrm{E}\left[\left\langle p_i^2\right\rangle_i\middle|p\right] = r_1^{\mathrm{R}}p+\left(1-r_1^{\mathrm{R}}\right)p^2+\mathcal{O}(\delta) \tag{A21}$$

$$\mathrm{E}\left[\left\langle p_{ij1}p_{ij2}p_i\right\rangle_{ij}\middle|p\right] = a^{\mathrm{R}}p+\left(1-a^{\mathrm{R}}-c^{\mathrm{R}}\right)p^2+c^{\mathrm{R}}p^3+\mathcal{O}(\delta) \tag{A22}$$

where r_0^{D} is the probability that the two homologous genes of the same individual are identical-by-descent (IBD);

$$r_0^{\mathrm{R}} = \frac{1}{2}+\frac{1}{2}r_0^{\mathrm{D}} \tag{A23}$$

is the probability that two genes sampled with replacement in the same individual are IBD; r_1^{R} is the probability that two genes sampled from the same deme with replacement are IBD; a^{R} is the probability that two homologous genes of an individual coalesce with a third gene (sampled from the same deme at random), and

$$c^{\mathrm{R}} = 1-a^{\mathrm{R}}-\left(r_0^{\mathrm{D}}-a^{\mathrm{R}}\right)-2\left(r_1^{\mathrm{R}}-a^{\mathrm{R}}\right) \tag{A24}$$

is the probability that none of these three genes are IBD (see eq. A-23 in Flintham et al., 2021 for details on eq. A24).

Plugging eqs. (A19) through (A22) into eq. (A18) yields

$$\begin{aligned}\mathrm{E}[\Delta p|p] = hs\left(\frac{\partial w_{ij}}{\partial z_{ij}} r_0^{\mathrm{R}} + \frac{\partial w_{ij}}{\partial z_i} r_1^{\mathrm{R}}\right) p(1-p) \\ + (1-2h)\,\frac{s}{2}\left(\frac{\partial w_{ij}}{\partial z_{ij}}\left[r_0^{\mathrm{D}} + (1-r_0^{\mathrm{D}})p\right] + \frac{\partial w_{ij}}{\partial z_i}\left[a^{\mathrm{R}} + 2(r_1^{\mathrm{R}} - a^{\mathrm{R}})p\right]\right) p(1-p) + \mathcal{O}\left(\delta^2\right),\end{aligned} \tag{A25}$$

which is equivalent to eq. (23) in Roze and Rousset (2003) with our eq. (A24) substituted into it. We compute probabilities of identity-by-descent in Appendix A.3 in terms of dispersal and deme size.

From eq. (A7), we have

$$\frac{\partial w_{ij}}{\partial z_{ij}} = 2 \qquad \text{and} \qquad \frac{\partial w_{ij}}{\partial z_i} = -2\,, \tag{A26}$$

which substituted into eq. (A25) and after some re-arrangements using eq. (A23) yields eq. (1) of the main text.

A.2 Variance in frequency change

Here, we derive eqs. (2) and (3) of the main text, i.e. the variance in frequency change over one generation, $\mathrm{V}[\Delta p|p]$, to the leading order of δ. Following eq. A8 in Roze and Rousset (2003), we decompose this change as

$$\mathrm{V}[\Delta p|p] = \frac{1}{{N_\mathrm{d}}^2}\sum_{i=1}^{N_\mathrm{d}} \mathrm{V}\left[p_i'\middle|p\right] + \frac{1}{{N_\mathrm{d}}^2}\sum_{i=1}^{N_\mathrm{d}}\sum_{\substack{r=1\\ r\neq i}}^{N_\mathrm{d}} \mathrm{C}\left[p_i', p_r'\middle|p\right]\,, \tag{A27}$$

where recall $p_i' = \sum_{j=1}^{N_\mathrm{ind}} p_{ij}' \big/ N_\mathrm{ind}$ is the frequency of A in deme i at generation $t+1$. Since sampling of gametes occurs independently in each deme in our baseline model, the covariance term vanishes, leaving us with

$$\mathrm{V}[\Delta p|p] = \frac{1}{{N_\mathrm{d}}^2}\sum_{i=1}^{N_\mathrm{d}} \mathrm{V}\left[p_i'\middle|p\right] = \frac{1}{N_\mathrm{d}}\left\langle \mathrm{V}\left[p_i'\middle|p\right]\right\rangle_i \tag{A28}$$

as in eq. (A1) in Roze and Rousset (2003). Since $\left\langle \mathrm{V}\left[p_i'\middle|p\right]\right\rangle_i$ is multiplied by $1/N_\mathrm{d}$ in eq. (A28), we can compute it assuming the number of demes is infinite and there is no selection ($s=0$). To do so, we take a slight detour from the derivation of Roze and Rousset (2003) (though we

eventually get the same result). Let us first condition on all genetic information $\mathbf{P}$ at the previous generation (in particular, in knowing all deme frequencies) using the law of total variance,

$$\mathrm{V}[\Delta p|p] = \frac{1}{N_\mathrm{d}}\left\langle \mathrm{E}\left[\mathrm{V}[p_i'|\mathbf{P}]\middle|p\right] + \mathrm{V}\left[\mathrm{E}[p_i'|\mathbf{P}]\middle|p\right]\right\rangle_i . \tag{A29}$$

To characterise this, let $N_{\mathrm{M},i}$ be the realized number of migrant gametes received in deme i at generation $t+1$ from all the other demes. Under neutrality, $N_{\mathrm{M},i}$ is distributed binomially with parameters,

$$N_{\mathrm{M},i} \overset{\mathrm{law}}{\sim} \mathrm{Binom}(2N_\mathrm{ind}, d) . \tag{A30}$$

Conditional on $N_{\mathrm{M},i}$, the allele frequency at focal deme i has mean and variance given by,

$$\mathrm{E}\left[p_i'\middle|\mathbf{P}, N_{\mathrm{M},i}\right] = \left(\frac{2N_\mathrm{ind} - N_{\mathrm{M},i}}{2N_\mathrm{ind}}\right)p_i + \left(\frac{N_{\mathrm{M},i}}{2N_\mathrm{ind}}\right)p + \mathcal{O}(\delta) \tag{A31}$$

$$\mathrm{V}\left[p_i'\middle|\mathbf{P}, N_{\mathrm{M},i}\right] = \frac{N_{\mathrm{M},i}}{(2N_\mathrm{ind})^2}p(1-p) + \frac{2N_\mathrm{ind} - N_{\mathrm{M},i}}{(2N_\mathrm{ind})^2}p_i(1-p_i) + \mathcal{O}(\delta) . \tag{A32}$$

Marginalizing those over $N_{\mathrm{M},i}$ (eq. A30), we obtain

$$\mathrm{E}\left[p_i'\middle|\mathbf{P}\right] = (1-d)p_i + d\,p + \mathcal{O}(\delta) \tag{A33}$$

$$\begin{aligned}\mathrm{V}\left[p_i'\middle|\mathbf{P}\right] &= \mathrm{V}\left[\mathrm{E}\left[p_i'\middle|\mathbf{P}, N_{\mathrm{M},i}\right]\middle|\mathbf{P}\right] + \mathrm{E}\left[\mathrm{V}\left[p_i'\middle|\mathbf{P}, N_{\mathrm{M},i}\right]\middle|\mathbf{P}\right] \\ &= \frac{d(1-d)(p_i-p)^2 + dp(1-p) + (1-d)p_i(1-p_i)}{2N_\mathrm{ind}} + \mathcal{O}(\delta) .\end{aligned} \tag{A34}$$

Plugging eqs. (A33) and (A34) into eq. (A29), one obtains

$$\mathrm{V}[\Delta p|p] = \frac{1}{N_\mathrm{d}}\left(\frac{p - [1-(1-d)^2]p^2 - (1-d)^2\left\langle \mathrm{E}\left[{p_i}^2\middle|p\right]\right\rangle_i}{2N_\mathrm{ind}} + (1-d)^2\left\langle \mathrm{V}[p_i|p]\right\rangle_i\right) + \mathcal{O}\left(\delta^2\right) \tag{A35}$$

after re-arrangements and using the fact that $\mathrm{E}\left[p_i\middle|p\right] = p$.

We can simplify eq. (A35) further by noticing first that $\mathrm{V}[\langle p_i\rangle_i|p] = \langle \mathrm{V}[p_i|p]\rangle_i$ since the covariance between deme frequencies is zero (when the number of demes is infinite). We also have that by definition, $\mathrm{V}[\langle p_i\rangle_i|p] = \mathrm{V}[p|p] = 0$. These relationships then entail that $\langle \mathrm{V}[p_i|p]\rangle_i = 0$. Second, from eq. (A21) we have

$$\left\langle \mathrm{E}\left[{p_i}^2\middle|p\right]\right\rangle_i = \mathrm{E}\left[\left\langle {p_i}^2\right\rangle_i\middle|p\right] = r_1^\mathrm{R}p + \left(1 - r_1^\mathrm{R}\right)p^2 + \mathcal{O}(\delta) . \tag{A36}$$

Substituting for these into eq. (A35) yields,

$$\mathrm{V}[\Delta p|p] = p(1-p)\frac{1-(1-d)^2 r_1^{\mathrm{R}}}{2N_{\mathrm{d}}N_{\mathrm{ind}}} + \mathcal{O}\left(\delta^2\right) \tag{A37}$$

which is equal to eq. (A4) of Roze and Rousset (2003) using the fact that $r_1^{\mathrm{D}} = (1-d)^2 r_1^{\mathrm{R}}$.

Following standard theory, we define the (variance) effective population size N_{e} such that

$$\mathrm{V}[\Delta p|p] = \frac{p(1-p)}{2N_{\mathrm{e}}} + \mathcal{O}\left(\delta^2\right), \tag{A38}$$

which by comparison to eq. (A37) leads to

$$N_{\mathrm{e}} \quad = \quad \frac{N_{\mathrm{d}}N_{\mathrm{ind}}}{1-(1-d)^2 r_1^{\mathrm{R}}} \tag{A39}$$

and using $r_1^{\mathrm{D}} = (1-d)^2 r_1^{\mathrm{R}}$ obtains eq. (3) of the main text. Note also that because $r_1^{\mathrm{D}} = F_{\mathrm{ST}}$ in the island model of dispersal, eq. (3) is equal to eq. (9.48) in Rousset (2004). We compute r_1^{D} in eq. (A44). Plugging for this into eq. (3) then gives

$$N_{\mathrm{e}} = \left(1 + \frac{1}{2N_{\mathrm{ind}}} \cdot \frac{(1-d)^2}{1-(1-d)^2}\right) N_{\mathrm{d}}N_{\mathrm{ind}} \tag{A40}$$

in terms of the dispersal rate d.

A.3 Probabilities of coalescence

Here, we briefly explain how relevant probabilities of IBD are derived using standard coalescent arguments (e.g. Karlin, 1968; Wang, 1997; Rousset, 2004). Since these probabilities already weigh on terms of order δ in the expected and variance in allelic frequency change, they are derived under the assumption that there is no selection and the number of demes is effectively infinite. This latter assumption means in particular that the probability that two genes sampled in different demes coalesce is zero.

A.3.1 Pairwise coalescence probability

Recall that r_0^{D} denotes the probability of coalescence between the two homologous genes of a focal individual, r_0^{R} the probability of coalescence between two genes sampled with replacement in the same individual (eq. A23), and r_1^{R} the probability of coalescence between two genes sampled with replacement from the same deme. These last two probabilities are connected by

$$r_1^{\mathrm{R}} = \frac{1}{N_{\mathrm{ind}}} r_0^{\mathrm{R}} + \left(1 - \frac{1}{N_{\mathrm{ind}}}\right) r_1^{\mathrm{D}} \tag{A41}$$

where r_1^{D} is the probability of coalescence of two different genes sampled in the same deme (superscripts R and D generally denote samplings of genes with and without replacement, respectively, while subscripts 0 and 1 denote samplings of genes within the same individual and within the same deme, respectively). The first term of eq. (A41) corresponds to the probability of sampling the same individual, and the second, of sampling different individuals.

The probability of coalescence of two different genes in a given deme after one generation can be expressed as

$$r_1^{\mathrm{D}\prime} = (1-d)^2 \left[\underbrace{\frac{1}{N_{\mathrm{ind}}}\left(\frac{1}{2} + \frac{1}{2} r_0^{\mathrm{D}}\right)}_{\text{(i)}} + \underbrace{\left(1 - \frac{1}{N_{\mathrm{ind}}}\right) r_1^{\mathrm{D}}}_{\text{(ii)}} \right], \tag{A42}$$

where $(1-d)^2$ is the probability of sampling two philopatric genes, and we have decomposed their coalescence as to whether they have the same parent (i) or not (ii). However, since we assume that gametes fuse randomly in each deme following dispersal, we have

$$r_0^{\mathrm{D}} = r_1^{\mathrm{D}}, \tag{A43}$$

or equivalently, $F_{\mathrm{IT}} = F_{\mathrm{ST}}$ (and $F_{\mathrm{IS}} = 0$). Substituting for r_0^{D} into eq. (A42) and solving the recursion for $r_1^{\mathrm{D}\prime} = r_1^{\mathrm{D}}$ yields

$$r_0^{\mathrm{D}} = r_1^{\mathrm{D}} = \frac{(1-d)^2}{2N_{\mathrm{ind}} - (1-d)^2(2N_{\mathrm{ind}} - 1)}. \tag{A44}$$

This together with eqs. (A23) and (A41) allows us to compute r_1^{R}, which gives all the required components for the selection gradient when $h = 1/2$ (eq. 1). Equation (A44) is equivalent to eq. (3.14) for F_{ST} in Rousset (2004) considering diploids (i.e., their N as $2N_{\mathrm{ind}}$).

A.3.2 Threeway coalescence probability

When $h \neq 1/2$, the selection gradient further depends on the probability a^{R} that two homologous genes of an individual coalesce with a third gene sampled with replacement from the same deme. This probability can be expressed as

$$a^{\mathrm{R}} = \frac{1}{N_{\mathrm{ind}}} r_0^{\mathrm{D}} + \left(1 - \frac{1}{N_{\mathrm{ind}}}\right) a^{\mathrm{D}}, \tag{A45}$$

where a^{D} is the probability that two homologous genes of an individual coalesce with a third gene of a different individual from the same deme (i.e. sampled without replacement). This probability satisfies the recursion,

$$a^{\mathrm{D}\prime} = \underbrace{(1-d)^3}_{\substack{\text{three philopatric}\\ \text{genes}}} \left[\underbrace{\left(\frac{1}{2N_{\mathrm{ind}}}\right)^2}_{\text{(i)}} + \underbrace{3\frac{1}{2N_{\mathrm{ind}}}\left(1 - \frac{1}{2N_{\mathrm{ind}}}\right) r_1^{\mathrm{D}}}_{\text{(ii)}} + \underbrace{\left(1 - \frac{1}{2N_{\mathrm{ind}}}\right)\left(1 - \frac{2}{2N_{\mathrm{ind}}}\right) a^{\mathrm{D}}}_{\text{(iii)}} \right] \tag{A46}$$

where (i) captures immediate coalescence of these three copies in the same gene; (ii) captures immediate coalescence of only two alleles, followed by coalescence with the third gene further in the past with probability r_1^{D}; and finally, (iii) captures the case where none of the three genes coalesce in the previous generation. Solving eq. (A46) for $a^{\mathrm{D}\prime} = a^{\mathrm{D}}$ yields

$$a^{\mathrm{D}} = \frac{\left[1 + 3\left(2N_{\mathrm{ind}} - 1\right) r_1^{\mathrm{D}}\right](1-d)^3}{\left(2N_{\mathrm{ind}}\right)^2 - \left(2N_{\mathrm{ind}} - 1\right)\left(2N_{\mathrm{ind}} - 2\right)(1-d)^3}, \tag{A47}$$

which substituted into eq. (A45) with eq. (A44) gives a^{R} as required for the selection gradient (eq. 1).

A.4 Stationary frequency distribution

In section 3.3 of the main text, we consider the expected number of generations taken by alleles to fix whose initial frequencies are sampled from a stationary frequency distribution $\phi(p)$. Here, we detail how we computed this distribution $\phi(p)$. According to the assump-

tions of this model (section 3.3 for details), the expected change in frequency is given by

$$\begin{aligned}\mathrm{E}[\Delta p|p] = \mu(1-2p) - s_{\mathrm{D}}\, p(1-p)\Big[& p + r_0^{\mathrm{D}}(1-p) + h_{\mathrm{D}}(1-r_0^{\mathrm{D}})(1-2p) \\ & - \Big(r_1^{\mathrm{R}} + (r_1^{\mathrm{R}} - a^{\mathrm{R}})(2h_{\mathrm{D}}-1)(1-2p)\Big)\Big] + \mathcal{O}\big(\delta^2\big)\,,\end{aligned} \tag{A48}$$

where the first term captures mutation with $\mu \sim \mathcal{O}(\delta)$, and the second negative selection with parameters s_{D} and h_{D}. The expression for the variance in allele frequency change is still given by eq. (A38) with eq. (A40). We can then compute the stationary allele frequency distribution under the diffusion approximation by solving the Fokker-Planck equation (or Kolmogorov-forward equation) with eqs. (A48) and (A38) (Crow and Kimura, 1970, p. 372). Specifically, the probability density function $\phi(p)$ is given by

$$\phi(p) = \frac{c_0}{\mathrm{V}[\Delta p|p]} \exp\left(2\int_0^p \frac{\mathrm{E}[\Delta x|x]}{\mathrm{V}[\Delta x|x]}\,dx\right)\,, \tag{A49}$$

where c_0 is a constant that guarantees $\int_0^1 \phi(p)\,dp = 1$ (Crow and Kimura, 1970, p. 425).

B Computer simulations

B.1 Basic algorithm

The Python code used for our simulations is available here: http://sudbrack.me/LimitedDispersalSM. These simulations track the number of adults at each deme $i = 1,\ldots,N_{\mathrm{d}}$ according to their genotypes: $n_i^{aa}, n_i^{aA}, n_i^{AA}$ (with $n_i^{aa} + n_i^{aA} + n_i^{AA} = N_{\mathrm{ind}}$) over time according to the following algorithm.

Given $n_i^{aa}, n_i^{aA}, n_i^{AA}$, we first calculate the proportion of A gametes produced in each deme i with

$$p_i^{(\mathrm{S})} = \underbrace{\frac{1}{2N_{\mathrm{ind}} + 2hsn_i^{aA} + 2sn_i^{AA}}}_{\text{total number of gametes}} \underbrace{\Big[(1+hs)n_i^{aA} + 2(1+s)n_i^{AA}\Big]}_{\text{total A gametes}}, \tag{B1}$$

where the superscript (S) is for selection. Second, mutations occur, so that this proportion

becomes,

$$p_i^{(\mathrm{M})} = p_i^{(\mathrm{S})} + \underbrace{\mu\left(1 - p_i^{(\mathrm{S})}\right)}_{a \to A} - \underbrace{\mu\, p_i^{(\mathrm{S})}}_{A \to a}, \tag{B2}$$

in each deme i. Third, dispersal alters the proportion of A gametes in each deme i to,

$$p_i^{(\mathrm{D})} = (1-d)\, p_i^{(\mathrm{M})} + d\; p_{(-i)}^{(\mathrm{M})}\,, \tag{B3}$$

where $p_{(-i)}^{(\mathrm{M})} = \frac{N_\mathrm{d}}{N_\mathrm{d}-1} p^{(\mathrm{M})} - \frac{1}{N_\mathrm{d}-1} p_i^{(\mathrm{M})}$ (with $p^{(\mathrm{M})}$ as the mean across all demes) is the average of $p_j^{(\mathrm{M})}$ among all demes excluding i. Finally, gametes fuse randomly within demes and density-dependent regulation occurs, so that the number of adults of each genotype after one iteration of the life cycle is thus given by

$$\left[n_i^{aa}, n_i^{aA}, n_i^{AA}\right]' \overset{\text{law}}{\sim} M_3\left(N_\mathrm{ind};\, \left(1 - p_i^{(\mathrm{D})}\right)^2,\, 2p_i^{(\mathrm{D})}\left(1 - p_i^{(\mathrm{D})}\right),\, \left(p_i^{(\mathrm{D})}\right)^2\right), \tag{B4}$$

where M_3 is a Multinomial distribution with three classes.

B.2 Analyses

B.2.1 Time to fixation of *de novo* mutations

To study the time to fixation of *de novo* mutations (conditioned on fixation), we iterate the life cycle starting with all aa adults but one heterozygous aA. We let 40'000 simulations run with no mutation – i.e., $\mu = 0$ in eq. (B2) so that eventually, A is fixed or lost. We recorded the number of generations until fixation whenever A fixed.

B.2.2 Fixation of standing genetic variants

To study fixation of standing genetic variants, we first let the population evolve for 10'000 generations under selection, mutation and drift. We inspected that simulations reached equilibrium, which they did very rapidly (e.g. fig. 6C). During this time, mutation rate is set to $\mu = 10^{-4}$, selection and dominance coefficients in eq. (B1) is set to $s = -s_\mathrm{D} = -10^{-3}$ and $h = h_\mathrm{D}$ (see Figures for values). At $t = 10'001$, selection on A becomes positive $s = 10^{-2} > 0$ and to study the time to fixation from standing variation, we then set $\mu = 0$. We record the

number of generations until the eventual fixation of allele A.

C Extinction-recolonization dynamics

Here we describe our analyses to investigate the impact of extinction-recolonization dynamics, whose results are described in section 3.4. We also present the algorithm to simulate extinction and recolonization dynamics.

C.1 Expected allele frequency change

When selection is soft, fitness eq. (A7) still holds with extinction-recolonization occurring as described in section 3.4 (eq. 35 of Roze and Rousset, 2003 for rationale; section E.3 here for the case of hard selection). Consequently, the expected change in allele frequency is still given by eq. (1), i.e. by

$$\begin{aligned}\mathrm{E}[\Delta p|p] = sp(1-p)\Big[p + r_0^{\mathrm{D}}(1-p) + h(1-r_0^{\mathrm{D}})(1-2p) - \Big(r_1^{\mathrm{R}} + (r_1^{\mathrm{R}} - a^{\mathrm{R}})(2h-1)(1-2p)\Big)\Big] \\ + \mathcal{O}\big(\delta^2\big).\end{aligned} \tag{C1}$$

What does change due to extinction-recolonization are the probabilities of coalescence that appear in this equation. We derive those below (eqs. C6 and C8).

C.2 Variance in allele frequency change

Due to recolonization, the covariance in gene frequency between demes can no longer be ignored in eq. (A27). As shown by Roze and Rousset (2003), the variance in allele frequency change can still be written as

$$\mathrm{V}[\Delta p|p] = \frac{p(1-p)}{2N_{\mathrm{e}}} + \mathcal{O}\big(\delta^2\big), \tag{C2}$$

but now the effective population size is given by

$$N_e = \left(\frac{1-e}{2r_1^R\left[1-(1-e)^2(1-d)^2\right]}\right)N_d \tag{C3}$$

$$= \left(\frac{1-e}{2\left(r_0^R+(N_{ind}-1)r_1^D\right)\left[1-(1-e)^2(1-d)^2\right]}\right)N_d N_{ind} \tag{C4}$$

(eq. 29 in Roze and Rousset, 2003; then using eq. A41). We compute r_1^R for our extinction-recolonization model below.

C.3 Probabilities of coalescence under local extinctions

C.3.1 Pairwise coalescence probability

After one generation, the probability of coalescence of two alleles sampled without replacement from the same deme can be expressed as

$$r_1^{D'} = (1-e)\underbrace{(1-d)^2\left[\frac{1}{N_{ind}}\left(\frac{1}{2}+\frac{1}{2}r_0^D\right)+\left(1-\frac{1}{N_{ind}}\right)r_1^D\right]}_{\text{no extinction, eq. (A42)}}+e\varphi\underbrace{\left[\frac{1}{N_{ind}}\left(\frac{1}{2}+\frac{1}{2}r_0^D\right)+\left(1-\frac{1}{N_{ind}}\right)r_1^D\right]}_{\text{extinction}}, \tag{C5}$$

where φ is the probability that two randomly sampled gametes among the $2N_{ind}$ colonizers come from the same deme. We model recolonization in two ways: (i) with a migrant pool model, in which each gamete comes from different demes so $\varphi = 0$; and (ii) with a propagule model, wherein all gametes come from the same randomly sampled deme so $\varphi = 1$. As the labels indicate, the first term of eq. (C5) captures coalescence in the case where the deme did not go extinct, and the second term where it did. Equation (C5) differs, although only marginally, to eq. (B20) of Roze and Rousset (2003) because they assume two rounds of reproduction within recolonization dynamics. As before, $r_0^D = r_1^D$. Substituting for this into eq. (C5) and solving for $r_1^{D'} = r_1^D$, yields,

$$r_0^D = r_1^D = \frac{(1-e)(1-d)^2+e\varphi}{2N_{ind}-[(1-e)(1-d)^2+e\varphi](2N_{ind}-1)}\,. \tag{C6}$$

The probability r_1^R is found by plugging eq. (C6) into eq. (A41).

C.3.2 Threeway coalescence probability

The probability a^{D} that the two homologous genes of an individual plus a third gene randomly sampled from the remaining genes in that same deme, are all IBD is given by solving the recursion

$$a^{\mathrm{D}\prime} = \left((1-e)(1-d)^3 + e\varphi^2\right)\left[\left(\frac{1}{2N_{\mathrm{ind}}}\right)^2 + 3\frac{1}{2N_{\mathrm{ind}}}\left(1-\frac{1}{2N_{\mathrm{ind}}}\right)r_1^{\mathrm{D}} + \left(1-\frac{1}{2N_{\mathrm{ind}}}\right)\left(1-\frac{2}{2N_{\mathrm{ind}}}\right)a^{\mathrm{D}}\right]. \tag{C7}$$

This is similar in structure but slightly different to eq. (B21) of Roze and Rousset (2003) owing to our different assumptions about recolonization. Solving eq. (C7) for $a^{\mathrm{D}\prime} = a^{\mathrm{D}}$ yields,

$$a^{\mathrm{D}} = \frac{\left(1+3(2N_{\mathrm{ind}}-1)r_1^{\mathrm{D}}\right)\left[(1-d)^3(1-e)+e\varphi^2\right]}{(2N_{\mathrm{ind}})^2-(2N_{\mathrm{ind}}-1)(2N_{\mathrm{ind}}-2)\left[(1-d)^3(1-e)+e\varphi^2\right]}, \tag{C8}$$

where r_1^{D} is given by eq. (C6).

C.4 Simulations

To simulate evolution with extinctions and recolonization dynamics, we extend our basic algorithm described in section B.1. Before the production of gametes, we sample a random number n_{e} of demes that go extinct from a binomial distribution with parameters,

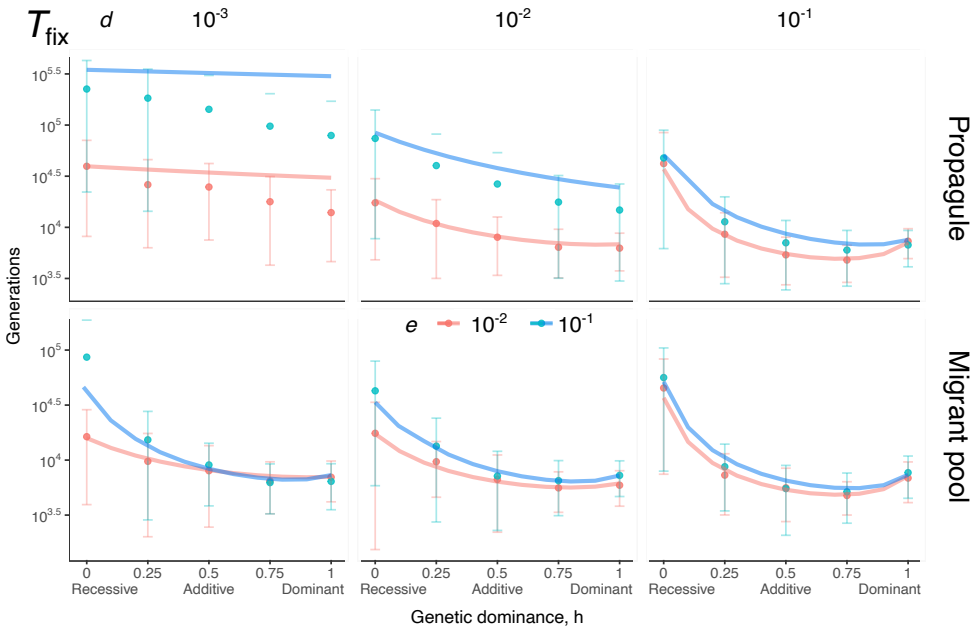
$$n_{\mathrm{e}} \overset{\mathrm{law}}{\sim} \mathrm{Binom}(N_{\mathrm{d}}, e). \tag{C9}$$

Those n_{e} demes are sampled uniformly without replacement from all demes.

Extinct demes do not produce any gametes and therefore do not participate of the dispersal phase (i.e., $p^{(\mathrm{M})}_{(-i)}$ in eq. B3 is averaged only over extant demes).

Extinct demes are available for recolonization and this can occur in two ways depending on the model. In the migrant pool model of recolonization, each deme i among these n_{e} demes is recolonized by sampling gametes from a multinomial distribution

$$\left[n_i^{aa}, n_i^{aA}, n_i^{AA}\right]' \overset{\mathrm{law}}{\sim} M_3\left(N_{\mathrm{ind}};\ (1-p_{\mathrm{e}})^2\ ,\ 2p_{\mathrm{e}}(1-p_{\mathrm{e}})\ ,\ (p_{\mathrm{e}})^2\right), \tag{C10}$$



Appendix C, Figure C1: Comparison with simulations under local extinctions. Expected time to fixation $T_{\text{fix}}(p_0)$ (on a log scale) of alleles arising as single copies $p_0 = 1/(2N_{\text{tot}})$ when $e = 10^{-1}$ and 10^{-2}, for different dispersal rates (columns) and models of recolonization (rows). Solid lines are the numerical integration of eq. (5), and dots are the average of simulations (180 replicates for each set of parameters, error bars show standard deviation). Parameters: same as fig. 2.

wherein p_{e} is the average of gamete frequencies $p_i^{(\text{M})}$ before dispersal among all $N_{\text{d}} - n_{\text{e}}$ extant demes.

In the propagule model, one deme r among all $N_{\text{d}} - n_{\text{e}}$ extant demes is chosen at random to recolonize a given extinct deme i. The multinomial sampling to establish the adults in the next adult generation is

$$\left[n_i^{aa}, n_i^{aA}, n_i^{AA}\right]' \overset{\text{law}}{\sim} M_3\left(N_{\text{ind}};\ \left(1-p_r^{(\text{M})}\right)^2,\ 2p_r^{(\text{M})}\left(1-p_r^{(\text{M})}\right),\ \left(p_r^{(\text{M})}\right)^2\right). \tag{C11}$$

Comparisons between these simulations and predictions from the diffusion approximation reveal that that the diffusion approximation typically works well under extinction-recolonization dynamics (fig. C1), except with propagule recolonization and very weak dispersal, $d = 10^{-3}$. In this case, the diffusion approximation tends to overestimate the time taken by alleles to fix, especially if the allele is dominant and the extinction rate is high.

D Hard and soft sweeps

To investigate the probability of hard and soft sweeps (and produce Supplementary Figure SM3), we simulated evolution at two linked loci, one under selection and another neutral. Each locus is diallelic with, as before, an adapted allele A and a wild-type allele a at the locus under selection, and neutral alleles B and b at the neutral locus.

The algorithm has the same broad structure as in Appendix B.1, with modifications to accommodate two linked loci. Briefly, we track the number of adults $n_i^{X|Y}$ that carry genotypes made up of haplotypes X and Y (with $X, Y \in \{AB, Ab, aB, ab\}$) in each deme i. Given these numbers, we can straightforwardly calculate gametic frequencies after selection and recombination in each deme i ($p_{X,i}^{(S)}$). For instance, the frequency of gametes with haplotype AB before dispersal is

$$p_{AB,i}^{(S)} \propto (1+s)n_i^{AB|AB} + \frac{1}{2}(1+s)n_i^{AB|Ab} + \frac{1}{2}(1+hs)(1-r)n_i^{AB|ab} + \frac{1}{2}(1+hs)rn_i^{Ab|aB}, \quad \text{(D1)}$$

where r is the recombination rate and the proportionality is such that gametic frequencies sum to one (i.e. such that $p_{AB,i}^{(S)} + p_{Ab,i}^{(S)} + p_{aB,i}^{(S)} + p_{ab,i}^{(S)} = 1$). Gametic dispersal is implemented the same way as in eq. (B3) for each of the four gametic types. Finally, gametes fuse randomly within each deme and density-dependent regulation occurs. This means the number of adults of each genotype is given by a Multinomial sampling, with probabilities built from the product of relevant gametic frequencies after dispersal (i.e., similar to eq. B4).

Initially, haplotypes are either aB or ab with equal probability, so that wild-type a is fixed and, neutral B and b are equally frequent on average. To reflect how dispersal and drift shaped variation among demes before the appearance of beneficial allele A, we use eq. 2 of Cherry and Wakeley, 2003 (with their $\overline{x} = 0.5$) and sample the frequency in each deme i of haplotype aB from an appropriate Beta distribution (the frequency of haplotype ab is given by the reciprocal of the frequency of aB). We form the first simulated generation of adults from these frequencies (i.e. obtain $n_i^{X|Y}$ for each X, Y) by assuming Hardy-Weinberg equilibrium in each deme (though this is very quickly lost owing limited dispersal). One individual of this first generation is chosen at random in the whole population and mutates at its selected locus, from wild-type a to beneficial mutant A.

We let the population evolve until the adaptive allele A is either fixed or lost. Among all the simulations where A fixes, we record the frequency of B at the neutral locus, and refer to the sweep as a *hard sweep* if either B or b is also fixed at the time A fixes. Otherwise, we refer to the sweep as a *soft sweep*.

For each set of parameters, we perform 10'000 replicates and calculate the proportion π_{HS} of hard sweeps among those replicates where A fixes. Results are shown in Supplementary Figure SM3.

E Hard selection

As a baseline, we have assumed that all demes produce an equal number of gametes, i.e. that selection is soft. Here, we consider the case of hard selection, whereby demes in which the frequency of A is greater produce more gametes. We find overall that fixations occur marginally faster under hard selection than under soft selection, but otherwise our results under hard selection are qualitatively similar to those under soft selection.

E.1 Expected and variance in allele frequency change

Under hard selection, the fitness of an individual j in deme i can be written as

$$w(z_{ij}, z_i, z) = \frac{2(1-d)z_{ij}}{(1-d)z_i + dz} + \frac{2dz_{ij}}{z}, \tag{E1}$$

which yields

$$\frac{\partial w_{ij}}{\partial z_{ij}} = 2 \qquad \text{and} \qquad \frac{\partial w_{ij}}{\partial z_i} = -2(1-d)^2. \tag{E2}$$

(evaluated at $s = 0$, i.e. $z_{ij} = z_i = z = 1$). Plugging these derivatives into eq. (A25), we obtain that the expected change in allele frequency is

$$\begin{aligned}\mathrm{E}[\Delta p|p] =& sp(1-p)\left[p + r_0^{\mathrm{D}}(1-p) + h(1-r_0^{\mathrm{D}})(1-2p) - (1-d)^2\left(r_1^{\mathrm{R}} + (r_1^{\mathrm{R}} - a^{\mathrm{R}})(2h-1)(1-2p)\right)\right] \\ &+ \mathcal{O}\left(\delta^2\right).\end{aligned} \tag{E3}$$

Comparing eq. (E3) with the case of soft selection (eq. 1) reveals that under hard selection, kin competition is reduced by $(1-d)^2$. This means that effective selection is stronger under hard selection.

The relevant probabilities of coalescence are not affected by whether selection is soft or hard since we only need these probabilities under neutrality. We can thus use those computed in sections A.3 (eqs. A44 and A47). This also entails that the variance in allele frequency change is still written as

$$\mathrm{V}[\Delta p|p] = \frac{p(1-p)}{2N_\mathrm{e}} + \mathcal{O}\left(\delta^2\right), \tag{E4}$$

with the effective population size given by eq. (A40), i.e.

$$N_\mathrm{e} = \left(1 + \frac{1}{2N_\mathrm{ind}} \cdot \frac{(1-d)^2}{1-(1-d)^2}\right) N_\mathrm{d} N_\mathrm{ind}\,. \tag{E5}$$

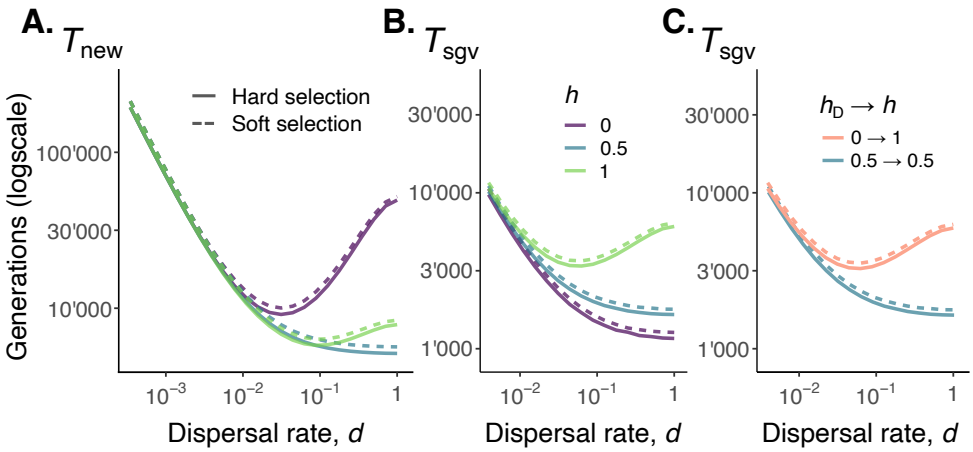
E.2 Analyses

We used the above to compute the total time of fixation of *de novo* mutations (i.e. the sum of waiting and fixation times, eq. 10 with eqs. 4 and 5, fig. E1A) and the time of fixation of standing genetic variants (eq. 11 with eqs. 5 and A49) where dominance is preserved (fig. E1B) and reversed (fig. E1C). We focused on the case where demes are small, $N_\mathrm{ind} = 10$, because this is where differences among hard and soft selection are the greatest (Roze and Rousset, 2003). Figure E1 shows that hard selection speeds up adaptation regardless if it stems from new *de novo* mutations or standing genetic variation. This is because kin competition is weaker and so effective selection is stronger under hard selection. The quantitative consequences of this for fixation times, however, are small (compare dashed and solid lines in fig. E1).

E.3 Hard selection and local extinction

We also investigated the case of hard selection under extinction-recolonization dynamics. In that case, fitness is given by,

$$w(z_{ij}, z_i, z) = (1-e)\underbrace{\left[\frac{2(1-d)z_{ij}}{(1-d)z_i + dz} + \frac{2dz_{ij}}{z}\right]}_{\text{as in eq. (E1) if not extinct}} + (1-e)\cdot\underbrace{2N_\mathrm{ind}(eN_\mathrm{d})\cdot\frac{z_i}{(1-e)N_\mathrm{d}z}\cdot\frac{z_{ij}}{N_\mathrm{ind}z_i}}_{\text{descendants via recolonization}}, \tag{E6}$$



Appendix E, Figure E1: Fixation times under hard selection. A. Expected total time of fixation of *de novo* mutants T_{new} (on a log scale) that are recessive ($h = 0$), additive ($h = 0.5$) and dominant ($h = 1$, in different colors) with $p_0 = 1/(2N_{\text{tot}})$ under hard (full line) and soft (dashed line) selection. Parameters: $N_{\text{d}} = 2000$, $N_{\text{ind}} = 10$, $\mu = 10^{-6}$, $s = 10^{-2}$, so that N_{tot} is the same as in the main text. **B.** Expected time of fixation of standing genetic variants T_{sgv} (on a log scale) with $h_{\text{D}} = h$ for recessive ($h = 0$), additive ($h = 0.5$) and dominant ($h = 1$) variants under hard (full line) and soft (dashed line) selection. Parameters: $s_{\text{D}} = 10^{-3}$, $\mu = 10^{-4}$, other parameters same as A. **C.** Expected time of fixation of standing genetic variants T_{sgv} (on a log scale) of recessive deleterious variant that becomes beneficial dominant ($h_{\text{D}} = 0$ and $h = 1$, orange), and an additive allele ($h_{\text{D}} = h = 1/2$, blue) under hard (full line) and soft (dashed line) selection. Parameters: same as B.

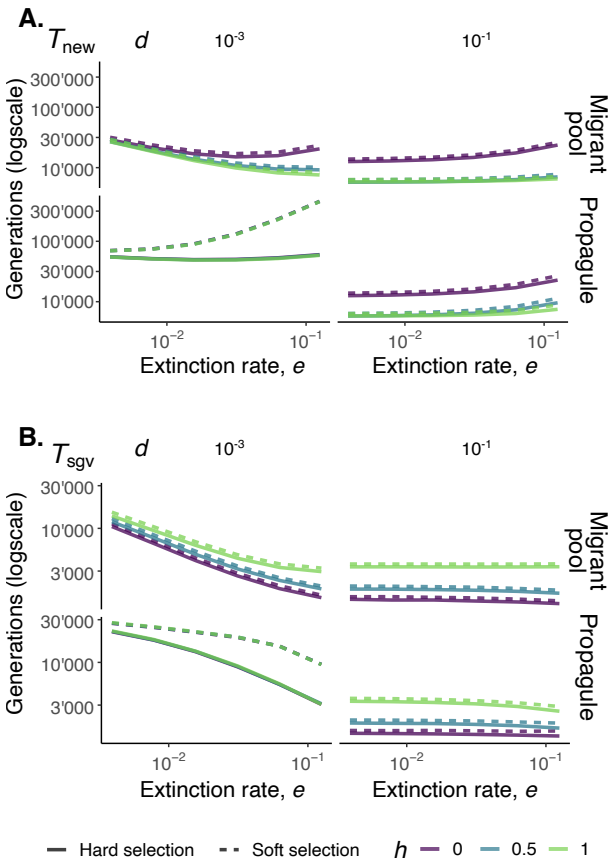
such that

$$\frac{\partial w_{ij}}{\partial z_{ij}} = 2 \qquad \text{and} \qquad \frac{\partial w_{ij}}{\partial z_i} = -2(1-e)(1-d)^2\,, \tag{E7}$$

which plugged into eq. (A25) gives

$$\begin{aligned}\mathrm{E}[\Delta p|p] &= sp(1-p)\\ &\quad\times\left[p + r_0^{\text{D}}(1-p) + h(1-r_0^{\text{D}})(1-2p) - (1-e)(1-d)^2\Big(r_1^{\text{R}} + (r_1^{\text{R}} - a^{\text{R}})(2h-1)(1-2p)\Big)\right]\\ &\quad + \mathcal{O}\big(\delta^2\big)\,.\end{aligned} \tag{E8}$$

Note that fitness is not influenced as to whether recolonization occurs according to the migrant pool or propagule model. Probabilities of coalescence are not influenced by hard selection and we thus use those computed in section C.3. Effective population size, meanwhile, is given by eq. (C4). Our results for time of fixation of *de novo* mutations and standing genetic variants are shown in fig. E2. This shows that hard selection accelerates adaptation, especially under the propagule model of recolonization where dispersal is weak (fig. E2A and B, bottom left).



Appendix E, Figure E2: The effects of extinction and recolonization dynamics under hard selection. **A.** Expected total time of fixation of *de novo* mutations T_{new} (on a log scale) with recessive ($h = 0$), additive ($h = 0.5$) and dominant ($h = 1$) effects (in different colors) arising as single copies $p_0 = 1/(2N_{\text{tot}})$ under hard (full line) and soft (dashed line) selection for $d = 10^{-3}$ (left) and $d = 10^{-1}$ (right), and under the migrant pool (top) and propagule (bottom) model of recolonization. Other parameters: same as fig. E1A. **B.** Expected time of fixation of standing genetic variants T_{sgv} (on a log scale) with $h_{\text{D}} = h$ for recessive ($h = 0$), additive ($h = 0.5$) and dominant ($h = 1$) alleles under hard (full line) and soft (dashed line) selection for $d = 10^{-3}$ (left) and $d = 10^{-1}$ (right), and under the migrant pool (top) and propagule (bottom) model of recolonization. $h_{\text{D}} = h$. Parameters: same as fig. E1B.

Supplementary Figures

for

Fixation times of *de novo* and standing beneficial variants in subdivided populations

Vitor Sudbrack and Charles Mullon [2]

List of supplementary figures

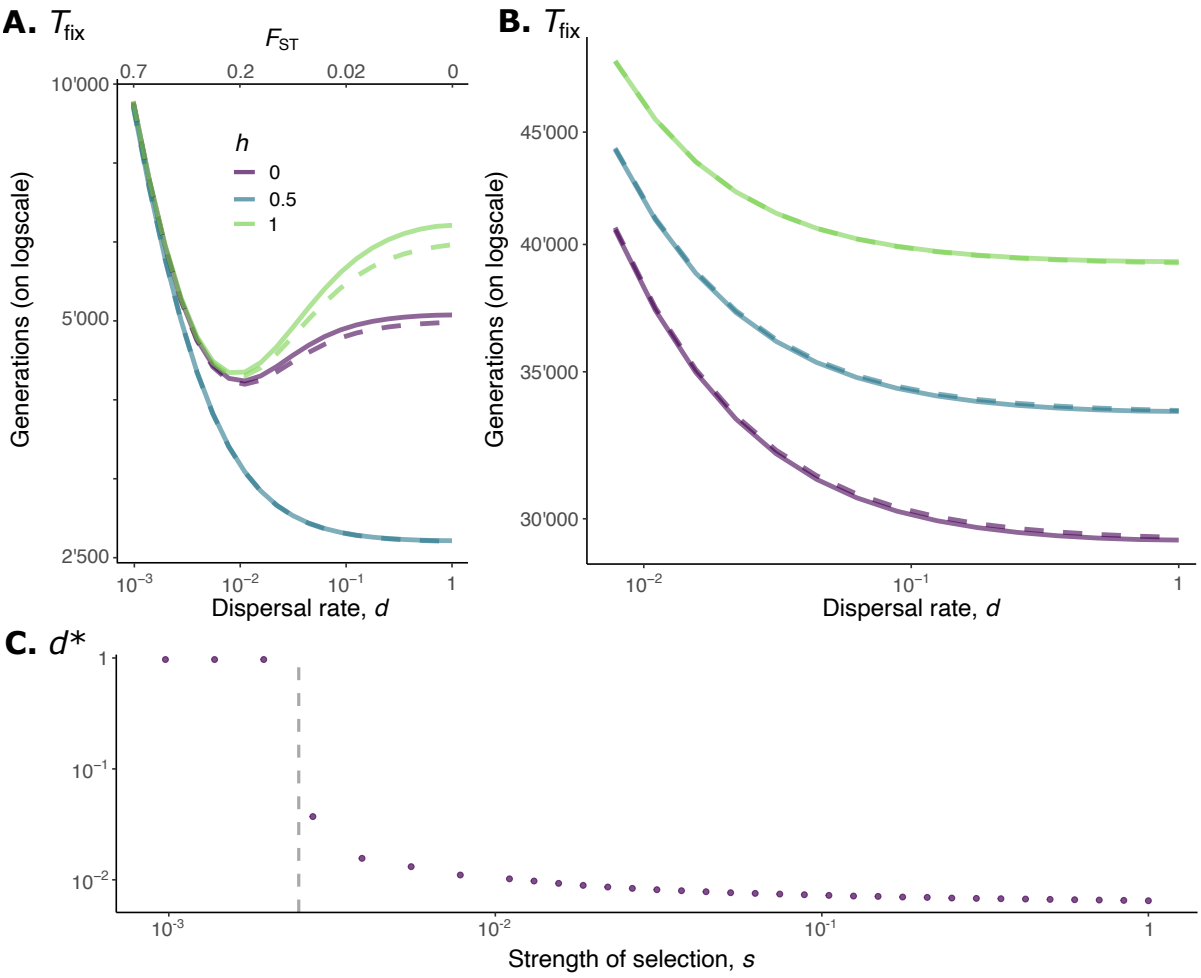
SM1 Selection threshold for faster fixation under dispersal limitation

SM2 Fixation under strong selection

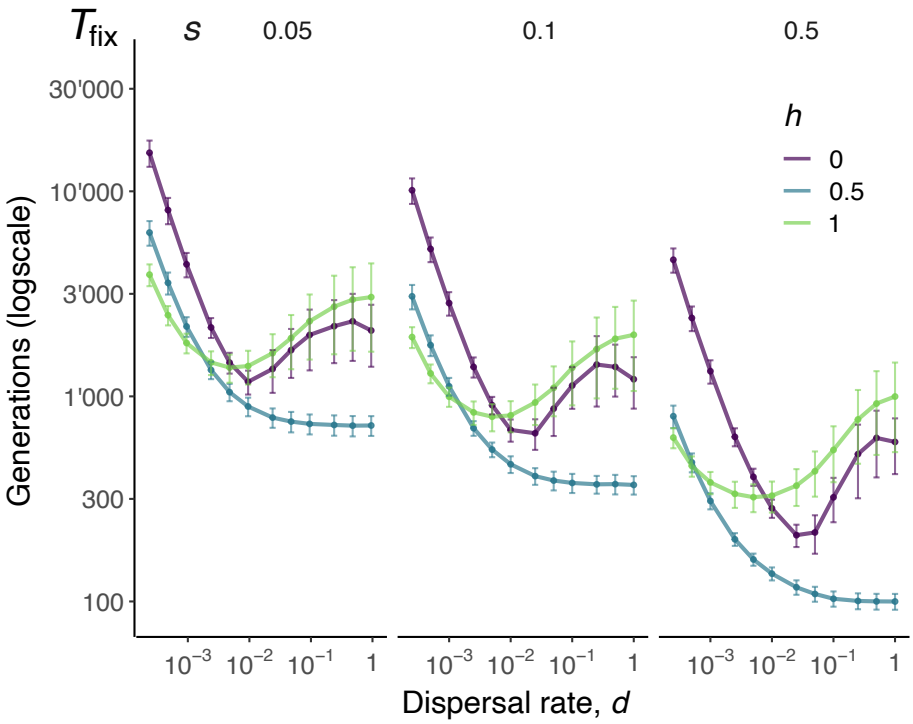
SM3 Effects of dispersal limitation on the signature of sweeps

SM4 Summary statistics of fixation events

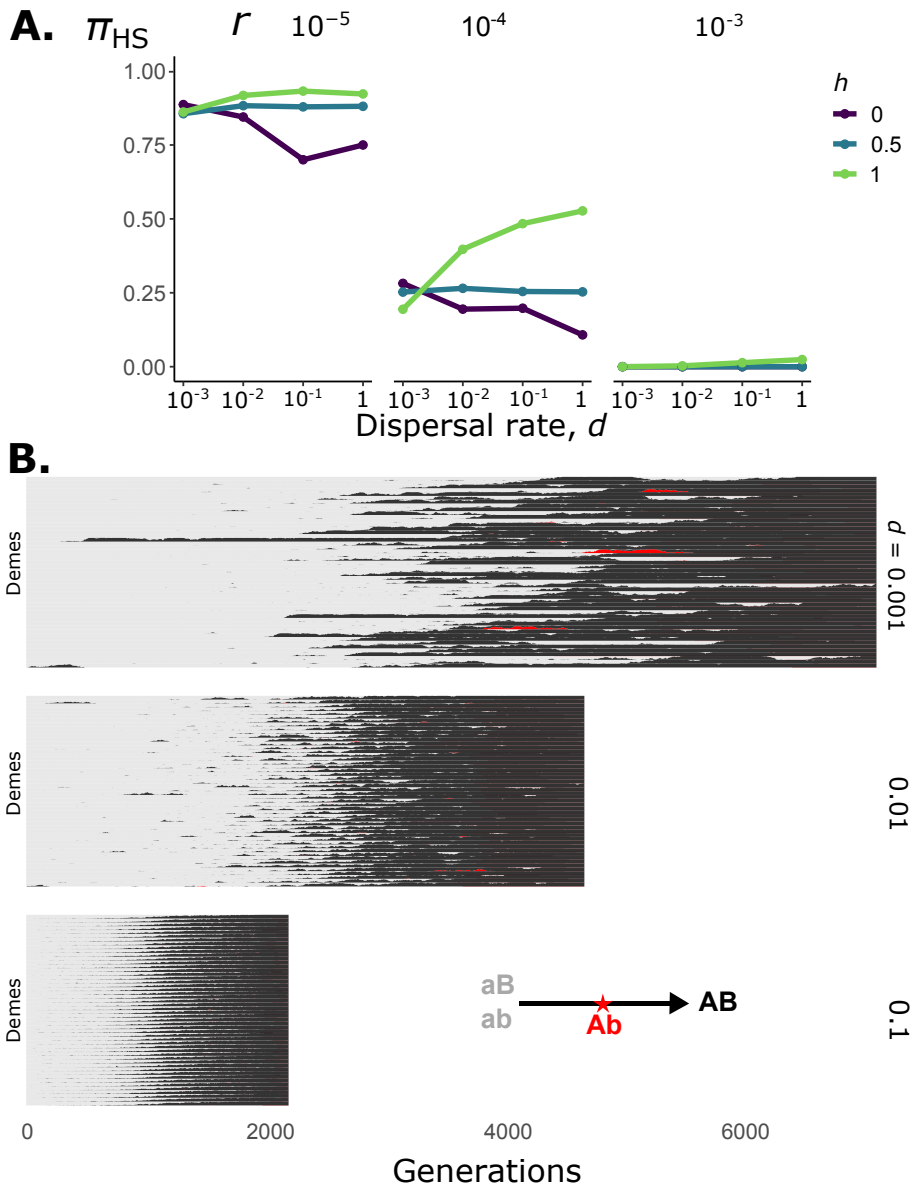
[2] Correspondence: `charles.mullon@unil.ch`



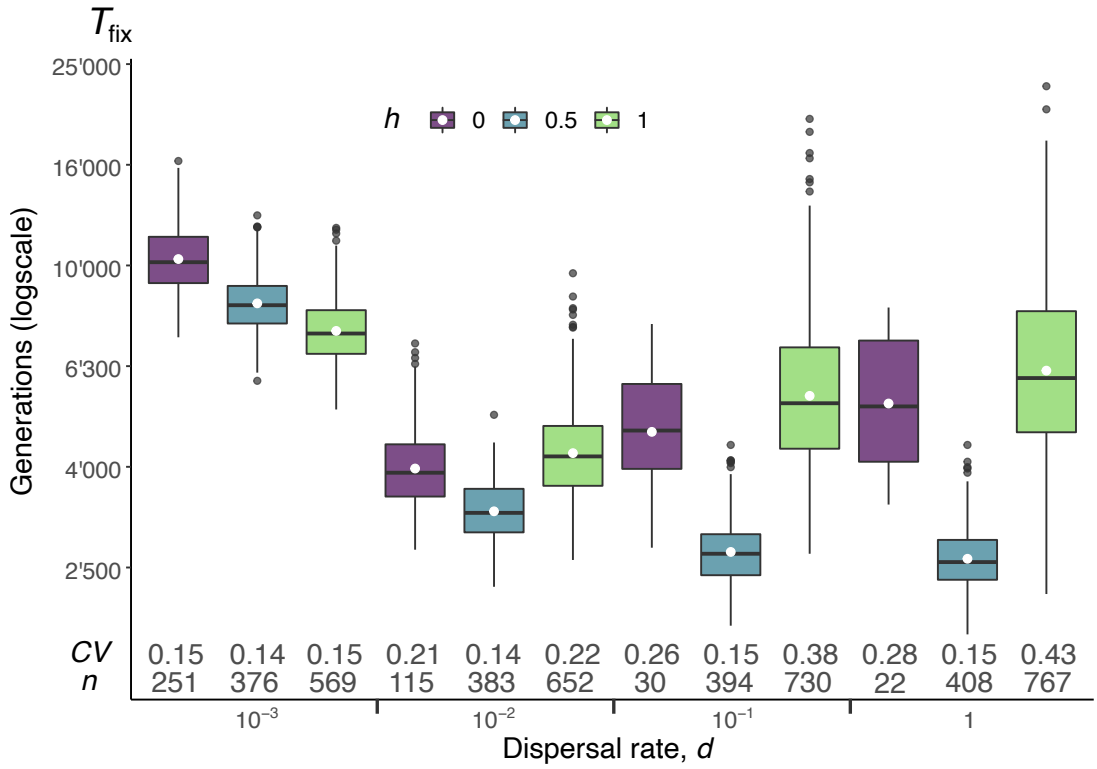
Supplementary Material, Figure SM1: Selection threshold for faster fixation under dispersal limitation. **A.** Expected time to fixation $T_{\text{fix}}(p_0)$ (on a log scale) of recessive ($h = 0$), additive ($h = 0.5$) and dominant ($h = 1$) alleles (in different colors, see legend) arising as single copies $p_0 = 1/(2N_{\text{tot}})$, with solid lines for general deme size and dispersal (i.e. using second column of Table 1) and dashed lines for large deme - weak dispersal limit (i.e. using third column of Table 1). Parameters: same as Fig. 2, in particular $s = 0.01$. **B.** Same as A, but with very weak selection, $s = 2 \cdot 10^{-4}$ (so that $N_{\text{tot}}s$ is as in Whitlock 2003). This shows that when selection is very weak, fixation is always slower under limited dispersal, irrespective of dominance. **C.** Dispersal rate d^* that minimizes the time to fixation $T_{\text{fix}}(p_0)$ of a recessive beneficial *de novo* mutation, according to the strength of selection s. This was calculated based on numerical integration of eq. (5) for $s < 0.01$. To speed up calculation, we used semi deterministic approximation eq. (7) for $s > 0.01$. This graph shows that when s is small, $d^* = 1$, i.e. fixation is fastest under panmixia and limited dispersal increases the time to fixation. But provided approximately $s > 0.0025$ (vertical dashed line), $d^* < 1$, indicating that limited dispersal reduces the time to fixation. Other parameters: same as Fig. 2.



Supplementary Material, Figure SM2: Fixation time under strong selection. Average time to fixation (on a log scale) of recessive ($h = 0$), additive ($h = 0.5$), and dominant ($h = 1$) alleles arising as single copies $p_0 = 1/(2N_{\text{tot}})$, for various fecundity advantage s (columns) from simulations (fixation events among at least 40'000 replicates for each set of parameters, error bars show standard deviation, Appendix B for details). Parameters: $N_{\text{d}} = 200$, $N_{\text{ind}} = 100$. As expected, stronger selection decreases time to fixation, but for each s investigated here, mild dispersal limitation speeds up the fixation of non-additive alleles (as in our baseline model).



Supplementary Material, Figure SM3: Effects of limited dispersal on the probability of hard vs. soft sweeps. **A.** Probability π_{HS} of a hard sweep, i.e. that polymorphism at a neutral locus is lost with fixation of a beneficial allele at a linked locus (Appendix D for details on simulations). Each column corresponds to a different rate of recombination between the selected and the neutral loci, see legend for level of dominance of beneficial allele. These graphs show that limited dispersal does not influence π_{HS} under additive fitness effects ($h = 1/2$), but increases (decreases) π_{HS} for recessive (dominant) effects. Other parameters: same as Fig. 2. **B.** Evolutionary dynamics of hard sweeps in populations under additive effects ($h = 0.5$) for three different dispersal rates ($d = 0.001, 0.01, 0.1$). For each dispersal rate, each line corresponds to one deme (50 shown out of 200 for illustration), showing haplotype frequencies within that deme through time. The population is initially composed of maladapted haplotypes (aB and ab in gray) and eventually fixes haplotype AB (in black) leading to a hard sweep. In red is shown the frequency of the alternative Ab haplotype that is created through recombination. These plots show that although this Ab haplotype emerges multiple times under low dispersal rates ($d = 0.001$ and $d = 0.01$), it is eventually lost due to genetic drift within demes. When dispersal rate is high ($d = 0.1$), there is not enough time for the Ab haplotype to be generated by recombination. Parameters: $r = 10^{-4}$, other parameters: same as A.



Supplementary Material, Figure SM4: Summary statistics of fixation events. Boxplot of the distribution of fixation times (given fixation) of a recessive ($h = 0$), additive ($h = 0.5$), and dominant ($h = 1$) allele for $d = 1, 10^{-1}, 10^{-2}$ and 10^{-3} (columns) for fixations among 40'000 replicates for each set of parameters. Boxes are from first quantile (Q1) to the third quantile (Q3). The median value (Q2) is shown with black line, and a white dot shows the mean. Outliers are shown as black dots. The coefficient of variation (CV) and the number of fixation events (n) are given for each set of parameters at the bottom of each boxplot. Parameters: same as Fig. 2. Overall, these distributions are close to being symmetric around the mean.