

Establishment of locally adapted mutations – Supplemental Material

Appendix A: Derivation of equations (4) and (5) in the main text

Establishment probabilities in single-type branching processes

We first review some basic results for single type branching processes before we move on to multi-type branching processes. In particular, we consider the evolution of a single panmictic population of mutant copies and model it with a Galton-Watson branching process. In every generation $n = 0, 1, 2, \dots$, each mutant copy produces a random number of offspring independently of other individuals, and then dies. Let $X(n) \in \mathbb{N}$ denote the number of mutant copies present at time n and $\zeta_j^{(n)} \in \mathbb{N}$ the random number of offspring of mutant copy j in generation n . We assume that the random variables $\zeta_j^{(n)}$ are independent and identically distributed. The evolution of $X(n)$ then follows

$$X(n+1) = \sum_{j=1}^{X(n)} \zeta_j^{(n)}. \quad (\text{S1})$$

We assume that there is a single copy of the mutation at time $n = 0$ such that $X(0) = 1$. Let $q = \lim_{n \rightarrow \infty} P(X(n) = 0)$ be the probability of eventual extinction of the lineage. We define the establishment probability as $p = 1 - q$, that is, the probability that the lineage will survive indefinitely. A useful tool to calculate establishment and extinction probabilities in branching processes are probability generating functions, defined as

$$f(x) = \mathbb{E}[x^\zeta] = \sum_{k=0}^{\infty} \mathbb{P}(\zeta = k) x^k = \sum_{k=0}^{\infty} p_k x^k, \quad 0 \leq x \leq 1, \quad (\text{S2})$$

where $p_i = \mathbb{P}(\zeta = i)$ for $i \in \mathbb{N}$. It has been shown that (Haccou *et al.* 2005) q is the smallest positive root of the equation

$$f(q) = q. \quad (\text{S3})$$

Establishment probabilities in multi-type branching processes

One can readily extend this framework to include variation in the offspring distributions among individuals using multi-type branching processes. We will focus on a process with two types of individuals, labeled 1 and 2. In our model for the establishment of locally adapted mutations, these types will correspond to the deme in which an individual resides. We will first present the theory necessary to calculate establishment probabilities and then apply it to our model for the evolution of locally adapted mutations.

Let the random variables $\zeta_k^{(i,j)} \in \mathbb{N}$ denote the the number of offspring of type i of the k -th individual of

type j . The total number of individuals of type i , that is, the number of copies of the mutant allele in deme i , at time n is then given by $X_i(n)$ and the evolution of the X_i 's is given by

$$X_i(n+1) = \sum_{k=1}^{X_1(n)} \zeta_k^{(i,1)} + \sum_{l=1}^{X_2(n)} \zeta_l^{(i,2)}, \quad (\text{S4})$$

and the total number of mutants across both demes is $X_{\text{tot}}(n) = X_1(n) + X_2(n)$. The probability generation function of the offspring distribution is defined as $\mathbf{f}(x_1, x_2) = (f_1(x_1, x_2), f_2(x_1, x_2))$ where

$$f_h(x_1, x_2) = f_h(\mathbf{x}) = \mathbb{E}(x_1^{\zeta_1^{(1,h)}} x_2^{\zeta_2^{(2,h)}}), \quad h \in \{1, 2\}. \quad (\text{S5})$$

The extinction probabilities are defined as $q^{(i)} = \lim_{n \rightarrow \infty} P(X_{\text{tot}}(n) = 0)$ with the initial conditions $X_i(0) = 1$ and $X_j(0) = 0$ ($i, j \in \{1, 2\}, i \neq j$). In other words, $q^{(i)}$ is the probability that a lineage starting with one individual of type i will go extinct. Using $\mathbf{q} = (q^{(1)}, q^{(2)})$, the extinction probabilities are given by the smallest positive roots of equation

$$\mathbf{f}(\mathbf{q}) = \mathbf{q}. \quad (\text{S6})$$

To obtain our main results, we use a theorem by [Haccou *et al.* \(2005\)](#) (pp. 127–128) for slightly supercritical branching processes, that is, branching processes where the establishment probability is small but positive. This theorem allows to calculate the probability of survival (*i.e.* $\mathbf{p} = (p^{(1)}, p^{(2)})$ where $p^{(h)} = 1 - q^{(h)}$) as a function of the leading eigenvalue of the mean reproduction matrix M , defined as

$$M = (M_{ij}), \quad M_{ij} = \frac{\partial f_i}{\partial x_j}(1, 1), \quad i, j \in \{1, 2\}. \quad (\text{S7})$$

The mean reproduction matrix is assumed to be positively regular (meaning that an integer $n \geq 1$ exists such that all entries of M^n are strictly positive). To calculate the extinction probabilities, we first need to calculate the maximal eigenvalue of M , ρ , and the left and right eigenvectors $\mathbf{u} = (u_1, u_2)$ and $\mathbf{v} = (v_1, v_2)$, such that

$$u_h > 0, \quad v_h > 0, \quad h \in \{1, 2\}, \quad (\text{S8})$$

$$\sum_h u_h v_h = \sum_h u_h = 1.$$

The leading eigenvalue ρ determines supercriticality. If $\rho > 1$, then the process is supercritical, and it is sub-critical if $\rho < 1$. We assume slight supercriticality and write the eigenvalue and the eigenvectors as

$\rho = 1 + \epsilon \tilde{c}$, where ϵ is a parameter of the model used to determine offspring distributions and \tilde{c} is a constant that depends on parameters of the model other than ϵ , such that

$$\rho(\epsilon) \rightarrow 1, \quad \text{when } \epsilon \rightarrow 0. \quad (\text{S9})$$

Theorem 5.6 of [Haccou *et al.* \(2005\)](#) states that the probability of indefinite survival of a lineage that starts with a single copy of a mutant of type h is then given by

$$p^{(h)} = \frac{2[\rho - 1]}{B} v_h + \mathcal{O}(\epsilon), \quad (\text{S10})$$

where the dependence of ρ , B and v_h on ϵ is omitted for simplicity, and B is given by

$$B = \sum_h u_h \text{Var} \left[\sum_j v_j \xi_{hj} \right] + \rho(\rho - 1) \sum_j u_j v_j^2. \quad (\text{S11})$$

Application to our model

We now apply the Theorem to our model. The types in the branching process correspond to the two demes in our model and the evolutionary forces of drift, selection and migration determine the offspring distributions of the two types. The strength and direction of selection varies across demes, and s_i is the selection coefficient of a mutant copy in deme i . As in the main text, we assume that $s_1 > 0 > s_2$. To derive the offspring distributions of the two types of individuals, we assume a Wright-Fisher model for reproduction, selection and migration. In a Wright-Fisher model mutants are sampled for the next generation with probability

$$\pi_i^{(r)} = \frac{k_i(1 + s_i)}{k_i(1 + s_i) + N_i - k_i}, \quad (\text{S12})$$

where N_i is the current number of wild-type individuals in deme i , and k_i the number of mutant individuals. In a branching process, we assume that mutant copies evolve independently of each other. Hence we can set $k_i = 1$ for the rest of the derivation and follow the fate of a single mutant copy. We assume that reproduction follows migration (but note that changing the order of the life-cycle does not affect our final conclusions). An individual carrying the mutant copy can either remain in the same deme or migrate to the other deme. We denote the probability that an individual migrates from deme i to deme j by m_{ij} . The probability of switching

from type i to type j due to migration is denoted $\pi_{i \rightarrow j}^{(m)}$, and is then given by

$$\pi_{i \rightarrow j}^{(m)} = \begin{cases} 1 - m_{12} & \text{if } i = j = 1, \\ m_{12} & \text{if } i = 1, j = 2, \\ m_{21} & \text{if } i = 2, j = 1, \\ 1 - m_{21} & \text{if } i = j = 2. \end{cases} \quad (\text{S13})$$

We assume a Wright-Fisher model of reproduction and selection and denote by O_{ij} the offspring of type j of an individual of type i . With probability m_{ij} ($i \neq j$) the offspring of a mutant of type i is

$$\begin{aligned} O_{ii} &= 0, \\ O_{ij} &= \mathcal{B}(N_j, \pi_j^{(r)}), \end{aligned} \quad (\text{S14})$$

and with probability $1 - m_{ij}$ the offspring of a mutant of type i is

$$\begin{aligned} O_{ii} &= \mathcal{B}(N_i, \pi_i^{(r)}), \\ O_{ij} &= 0. \end{aligned} \quad (\text{S15})$$

In general we can thus write the expected number of offspring of type j from a parent of type i as

$$E [O_{ij}] = N_j \frac{1 + s_j}{s_j + N_j} \cdot \pi_{i \rightarrow j}^{(m)} + 0 \cdot (1 - \pi_{i \rightarrow j}^{(m)}), \quad (\text{S16})$$

where $i, j \in \{1, 2\}$. For $N_j \rightarrow \infty$, this yields

$$E [O_{ij}] \approx (1 + s_j) \pi_{i \rightarrow j}^{(m)}. \quad (\text{S17})$$

Thus, during the crucial phase where mutants are rare and hence can get lost from the population due to random fluctuations, the offspring distributions are given by a binomial distribution with the average per individual offspring as summarized in table 1. Furthermore, for $N_i \rightarrow \infty$ ($\forall i$) the binomial distribution

Type of the parent	Average number of offspring of type...	
	Type 1	Type 2
Type 1	$(1 + s_1)(1 - m_{12})$	$(1 + s_2)m_{12}$
Type 2	$(1 + s_1)m_{21}$	$(1 + s_2)(1 - m_{21})$

Table 1 The average of offspring corresponding to the different types.

of offspring numbers converges to a Poisson distribution with the same expected offspring (S17) and we hence assume Poisson distributed offspring in the branching process. The probability generating functions of Poisson-distributed random variables are given by

$$f_h(x_1, x_2) = e^{\sum_{i=1}^2 M_{hi}(x_i-1)} . \quad (\text{S18})$$

where M_{hi} is the mean number of offspring of type i from a parent of type h (also the elements of the reproduction mean matrix M). Hence, the probability generating functions in our model are

$$\begin{aligned} f_1(x_1, x_2) &= e^{(1+s_1)(1-m_{12})(x_1-1)+(1+s_2)m_{12}(x_2-1)} , \\ f_2(x_1, x_2) &= e^{(1+s_1)m_{21}(x_1-1)+(1+s_2)(1-m_{21})(x_2-1)} . \end{aligned} \quad (\text{S19})$$

The reproduction mean matrix is then

$$M = \begin{pmatrix} (1 + s_1)(1 - m_{12}) & (1 + s_2)m_{12} \\ (1 + s_1)m_{21} & (1 + s_2)(1 - m_{21}) \end{pmatrix} , \quad (\text{S20})$$

and the maximal eigenvalue of this matrix is

$$\begin{aligned} \rho = \frac{1}{2} \left(\sqrt{4(m_{12} + m_{21} - 1)(1 + s_1)(1 + s_2) + (m_{12}(1 + s_1) + m_{21}(1 + s_2) - 2 - s_1 - s_2)^2} + \right. \\ \left. 2 - m_{12}(1 + s_1) - m_{21}(1 + s_2) + s_1 + s_2 \right) . \end{aligned} \quad (\text{S21})$$

We normalize the left and right eigenvectors corresponding to ρ and normalize them according to (S8). The full forms of \mathbf{u} and \mathbf{v} are neither very readable nor informative, and hence we do not show them here. All calculations can be readily checked using a computer algebra software such as *Mathematica*. Following

Aeschbacher and Bürger (2014), B can be written as

$$B = \sum_{i=1}^2 u_i \sum_{j=1}^2 v_j M_{ij} + \rho(\rho - 1) \sum_{j=1}^2 u_j v_j^2, \quad (\text{S22})$$

where, M_{ij} are the elements of the reproduction mean matrix M (see (S20)).

Now we have all the ingredients necessary to calculate $p^{(h)}$ according to (S10). To proceed further, however, we first identify conditions for slight supercriticality and then perform a weak selection approximation to obtain our main results.

We notice that for $s_1 = 0$ and $s_2 = 0$, the maximal eigenvalue is $\rho = 1$. Furthermore, for $s_2 < 0$ and $s_1 = 0$, we get that $\rho < 1$, and for $s_2 = 0$ and $s_1 > 0$, we get $\rho > 1$. We will thus consider a weak-selection approximation and rescale all model parameters by s_1 :

$$\chi_{ij} = \frac{m_{ij}}{s_1}, \quad \zeta = \frac{s_2}{s_1}, \quad \chi_{ij}, \zeta < \infty. \quad (\text{S23})$$

The Taylor expansion of ρ with respect to s_1 is then

$$\rho = 1 + \frac{1}{2} \left(1 - \chi_{12} - \chi_{21} + \zeta + \sqrt{(\chi_{12} + \chi_{21})^2 + (1 - \zeta)^2 - 2(\chi_{12} - \chi_{21})(1 - \zeta)} \right) s_1 + \mathcal{O}(s_1^2). \quad (\text{S24})$$

and we write $\rho = 1 + s_1 \tilde{c} + \mathcal{O}(s_1^2)$. For small s_1 , \tilde{c} needs to be finite and positive. We find that $\tilde{c} > 0$ if $\chi_{21} > \zeta(1 - \chi_{12})$. Also, for constant χ_{ij} and ζ , $\tilde{c} < \infty$. As a matter of fact, when $\chi_{21} \rightarrow \infty$, $\tilde{c} = 1$, and the assumption of slight supercriticality holds. This corresponds to a scenario in which migration from the deme in which the mutant has a disadvantage acts as a source. If $\chi_{12} \rightarrow \infty$, $\tilde{c} = \zeta < 0$, and the branching process will be sub-critical in this case. Then, the deme where the mutation is detrimental acts a sink. Finally, when $\zeta \rightarrow -\infty$, $\tilde{c} = 1 - \chi_{12}$, and the process is slightly supercritical as long as $\chi_{12} < 1$, that is, as long as migration from deme 1 to deme 2 is of the same order of magnitude of s_1 .

When looking at the symmetric case where $m_{12} = m_{21} = m/2 = \chi s_1$, we can calculate that $\tilde{c} = 1/2(1 + \zeta - 2\chi + \sqrt{1 - 2\zeta + \zeta^2 + 4\chi^2})$. In general, we find that $\tilde{c} > 0$ if $\chi > \zeta/(1 + \zeta)$. When looking at a mutation strongly deleterious in deme 2, when $\zeta \rightarrow -\infty$, $\tilde{c} = 1 - \chi$, meaning that if $\chi < 1$ the process is supercritical (hence for $m/2 < s_1$). When migration is very strong, $\chi \rightarrow \infty$, $\tilde{c} = 1 + \zeta$, meaning that the process remains supercritical for $\zeta > -1$ (this means that $|s_2| < s_1$). We next calculate equation (S10) using the rescaled variables defined in (S23), and perform a Taylor expansion in s_1 , ignoring second- and higher-order terms. The

full calculations are technically not very difficult but lengthy and cumbersome and are hence not shown here. After changing back to the original variables according to (S23), and defining $\mu_{ij} = 2m_{ij}/\lambda$, $\sigma = (s_1 - s_2)/\lambda$, with $\lambda = \sqrt{(m_{12} + m_{21})^2 + (s_1 - s_2)^2 - 2(m_{12} - m_{21})(s_1 - s_2)}$ we can write the establishment probabilities as

$$\begin{aligned} p^{(1)} &= s_1 \left[1 - \left(\frac{\mu_{12} - \mu_{21}}{2} \right) + \sigma \right] + s_2 \mu_{12} , \\ p^{(2)} &= s_1 \mu_{21} + s_2 \left[1 + \left(\frac{\mu_{12} - \mu_{21}}{2} \right) - \sigma \right] . \end{aligned} \quad (\text{S25})$$

In the main text we further defined $\Delta\mu = (\mu_{12} - \mu_{21})/2$ for brevity. Since equations (S25) can have negative values when the branching process is sub-critical, we set the establishment probabilities to zero when this is the case. Hence, we redefine $p^{(i)} = \max(p^{(i)}, 0)$. Equations (4)–(5) from the main text are readily obtained from (S25) with $m_{12} = m_{21} = m/2$.

Maximum of $p^{(i)}$

For the following computations, we write m_{12} and m_{21} as a function of their ratio, $\gamma = m_{21}/m_{12}$. Hence we define $m_{12} = \tilde{m}$ and $m_{21} = \tilde{m}\gamma$. Then, the maximum of $p^{(2)}$ with respect to \tilde{m} is at

$$m_{\max}^{(2)} = \frac{s_1(s_2 - s_1)}{s_1(\gamma - 1) + 2s_2} . \quad (\text{S26})$$

Meanwhile, $p^{(1)}$ does not have any valid maximum for the range $s_1 > 0 > s_2$. On one hand, the derivative is

$$\frac{\partial p^{(1)}}{\partial \tilde{m}} = 0 \quad \Longrightarrow \quad m_{\max}^{(1)} = \frac{s_2(s_1 - s_2)}{2s_1\gamma + s_2(1 - \gamma)} . \quad (\text{S27})$$

We can easily calculate when (S27) yields a positive result:

$$m_{\max}^{(1)} > 0 \quad \Longrightarrow \quad \gamma < \frac{s_2}{s_2 - 2s_1} . \quad (\text{S28})$$

On the other hand, $p^{(1)}(\tilde{m} = m_{\max}^{(1)}) > 0$ if

$$\gamma > \frac{s_2}{s_2 - 2s_1} . \quad (\text{S29})$$

But condition (S29) for γ does not overlap with (S28), which means that $p^{(1)}$ has no valid positive maximum. Furthermore, the limit for no migration is of course $p^{(1)} = 2s_1$, while for strong migration we will have $p^{(1)} = 2(\gamma s_1 + s_2)/(1 + \gamma)$. In the scenario where $s_2 < 0 < s_1$, the limit of strong migration will always be smaller of the limit for weak migration, which means that $p^{(1)}$ is monotonically decreasing.

Appendix B: Correction for density regulation

Here we present a version of the symmetric migration model in which we modify the offspring distribution in the branching process to account for the effects of deme-independent density regulation in finite populations (soft selection, sensu Wallace (1975)). Consider a model with finite population size and let κ_1 and κ_2 denote the carrying capacities of deme 1 and 2, respectively. The larger deme then acts as a source, that is, it sends out more migrants than it receives. We assume that density regulation acts after migration and brings each deme back to its carrying capacity instantaneously. The number of individuals in deme i after migration but before density regulation are given by

$$N'_1 = \kappa_1(1 - m) + \kappa_2 m, \quad \text{and} \quad N'_2 = \kappa_1 m + \kappa_2(1 - m). \quad (\text{S30})$$

Density regulation will then change the number of individuals in each deme by a factor

$$\delta_i = \frac{\kappa_i}{\kappa_i(1 - m) + \kappa_j m}; \quad i, j \in \{1, 2\}, \quad i \neq j, \quad (\text{S31})$$

We now introduce this in the branching process framework by modifying the absolute fitness of individuals in deme i (see table 1) to $w_i = (1 + s_i)\delta_i$. Then, the absolute fitness of a mutant in deme i will be

$$1 + s_{i,\text{eff}} = (1 + s_1)\delta_i. \quad (\text{S32})$$

The generating functions are

$$\begin{aligned} f_1(x_1, x_2) &= e^{(1+s_{1,\text{eff}})(1-\frac{m}{2})(x_1-1)+(1+s_{2,\text{eff}})\frac{m}{2}(x_2-1)}, \\ f_2(x_1, x_2) &= e^{(1+s_{1,\text{eff}})\frac{m}{2}(x_1-1)+(1+s_2)(1-\frac{m}{2})(x_2-1)}, \end{aligned} \quad (\text{S33})$$

and the mean reproduction matrix is

$$M = \begin{pmatrix} (1 + s_{1,\text{eff}})(1 - \frac{m}{2}) & (1 + s_{2,\text{eff}})\frac{m}{2} \\ (1 + s_{1,\text{eff}})\frac{m}{2} & (1 + s_2)(1 - \frac{m}{2}) \end{pmatrix}, \quad (\text{S34})$$

From this, following the same procedure as Appendix A, we find the linearized solution

$$\begin{aligned}
p_{\text{dens}}^{(1)} &= \frac{1}{2}m \left[\left(1 - \frac{\kappa_2}{\kappa_1}\right) - \frac{1}{2}\mu(\kappa_1 - \kappa_2) \left(\kappa_1 + \frac{\kappa_2^2}{\kappa_1}\right) \right] \\
&\quad + s_1 \left[1 + \sigma\kappa_1\kappa_2 + \frac{1}{2}\mu(\kappa_1 - \kappa_2)(\kappa_1 + 2\kappa_2) \right], \\
&\quad \quad \quad + s_2 \left[\mu\kappa_2 \frac{\kappa_1 + \kappa_2}{2} \right] \\
p_{\text{dens}}^{(2)} &= \frac{1}{2}m \left[\left(1 - \frac{\kappa_1}{\kappa_2}\right) + \frac{1}{2}\mu(\kappa_1 - \kappa_2) \left(\frac{\kappa_1^2}{\kappa_2} + \kappa_2\right) \right] \\
&\quad + s_1 \left[\mu\kappa_1 \frac{\kappa_1 + \kappa_2}{2} \right] \\
&\quad + s_2 \left[1 - \sigma\kappa_1\kappa_2 - \frac{1}{2}\mu(\kappa_1 - \kappa_2)(2\kappa_1 + \kappa_2) \right],
\end{aligned} \tag{S35}$$

where we defined $\sigma = (s_1 - s_2)/\lambda_{\text{dens}}$ and $\mu = m/\lambda_{\text{dens}}$, with

$$\lambda_{\text{dens}} = \sqrt{(s_1 - s_2)^2\kappa_1^2\kappa_2^2 + 2m(s_1 - s_2)\kappa_1\kappa_2(\kappa_1 - \kappa_2)\frac{\kappa_1 + \kappa_2}{2} + \frac{1}{4}m^2(\kappa_1^2 + \kappa_2^2)^2}.$$

Appendix C: Comparisons with simulations and previous results

Gavrilets and Gibson (2002) calculated the probability of fixation of an allele in a two-deme model with migration using a diffusion approximation. Comparing their numerical result to our analytical form, we see good agreement between the two models. Gavrilets and Gibson (2002) calculated fixation probabilities, while

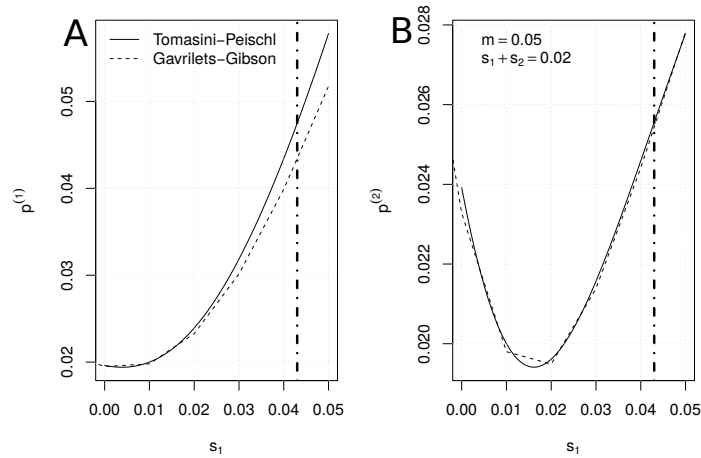


Figure S1 Comparison between the model described in Gavrilets and Gibson (2002) and the one presented in this work (equations (4) and (5)). The fixation probability for the model presented in Gavrilets and Gibson (2002) is calculated for 6 points. The thick dotted line shows the limit for which a polymorphic equilibrium is maintained in that model, corresponding to a case with establishment but not fixation. (A) Comparison with $p^{(1)}$. (B) Comparison with $p^{(2)}$. For both figures, $m = 0.05$, $s_1 + s_2 = 0.02$.

our formula deals with establishment. While for some parameters establishment is equivalent to fixation, this is not true in general ; an allele may become permanently established in a polymorphic equilibrium without ever reaching fixation. The black dotted vertical line in figure S1 represents the deterministic limit for the existence of a polymorphic equilibrium. The fact that our solution also takes into account established polymorphisms explains why the two models do not match on the right of the dotted line.

In a similar model to the one that we study in the present work, Yeaman and Otto (2011) used the asymptotic rate of increase in frequency of a rare allele to approximate the selection coefficient of the mutation itself. They then used Kimura's formula to calculate the fixation probability of such mutation. Figure S2 shows the comparison between the two approaches.

Setup of simulations (see figure 5 in main text)

In the simulations, we start with a population of $N = \kappa_1 + \kappa_2$ individuals (κ_i is the carrying capacity of deme i), and we inject one mutant copy in either deme 1 or deme 2. Both the wildtype and the mutant individuals

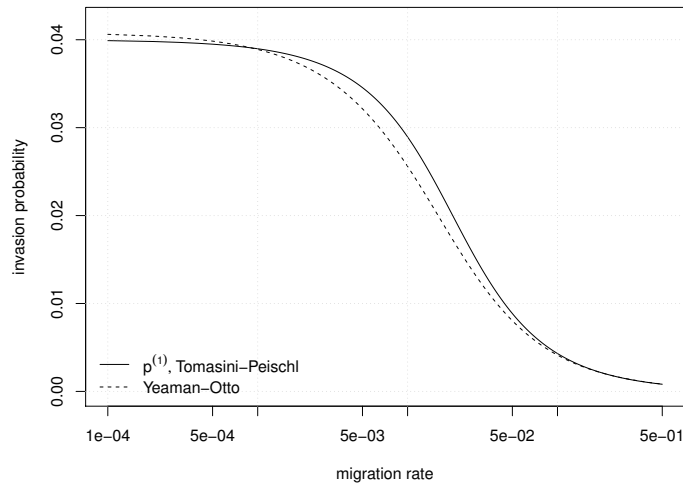


Figure S2 Comparison between the model described in [Yeaman and Otto \(2011\)](#) and our own. For this to work, we assume co-dominance in Yeaman and Otto’s model, while we take a very large population size. $s_1 = -s_2 = 0.02$.

follow a logistic growth model. The number of wildtype individuals in the next generation N^* is a Poisson distributed variable with mean

$$N^* = N \cdot R = N \cdot \frac{R_0}{1 + (R_0 - 1) \frac{N}{K}} \quad (\text{S36})$$

(also see [Beverton and Holt \(1957\)](#)). Similarly, the number of mutants in the next generation is a Poisson distributed variable with mean

$$N_{\text{mut}}^* = N_{\text{mut}} \cdot R \cdot (1 + s_i), \quad (\text{S37})$$

where s_i is the selective coefficient in deme i . We let the system evolve until the number of mutants reaches a predefined limit or until there are no mutants left in the population. The threshold for which we stop has to be of the order of $1/s_1$ in order to escape drift. We assume that if this threshold is reached in the deme where the mutation is advantaged, establishment has occurred. In our simulations we set the limit to 200 individual which is 4 times larger than $1/s_1$ in all considered cases, but we noticed low variability with respect to this parameter. If after 20’000 generations the threshold is not reached, the replicate is counted as not having reached establishment. We performed 50’000 replicates for each parameter combination.

Strongly deleterious mutation in one deme

Our approximation holds remarkably well also in cases where the mutation is strongly deleterious in one of the two demes (see figure [S3](#)).

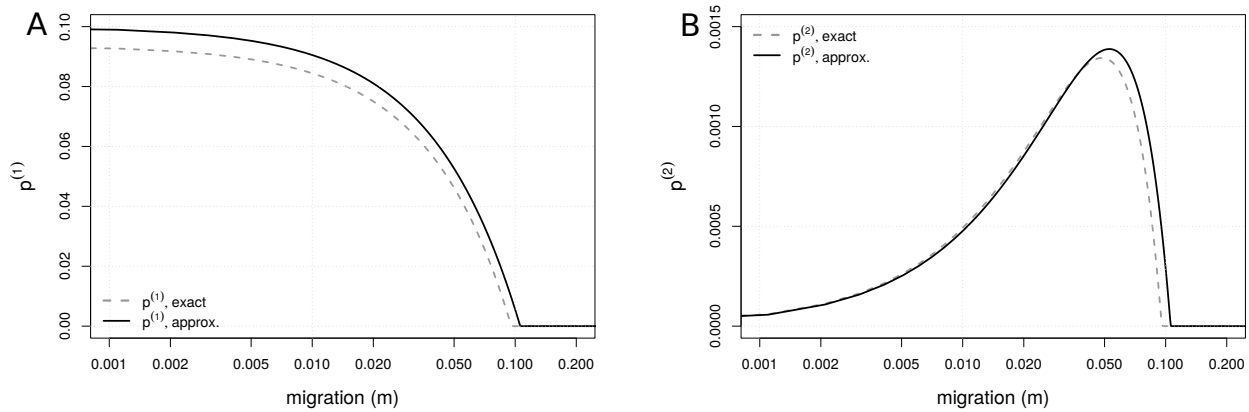


Figure S3 Comparison between simulations and analytical prediction for the establishment probabilities of a mutation strongly deleterious in deme 2. For both figures, $s_1 = 0.05$ and $s_2 = -0.9$. The exact solution is calculated numerically through 10'000 iterations of the probability generating functions (2) and (3) (see equation 1 in main text). (A) Probability of establishment as a function of the migration rate for a mutant born in deme 1. (B) Probability of establishment as a function of the migration rate for a mutant born in deme 2.

Literature Cited

- Aeschbacher, S. and R. Bürger, 2014 The effect of linkage on establishment and survival of locally beneficial mutations. *Genetics* **197**: 317–336.
- Beverton, R. J. H. and S. J. Holt, 1957 *On the dynamics of exploited fish populations*, volume 19. Ministry of Agriculture, Fisheries and Food.
- Gavrilets, S. and N. Gibson, 2002 Fixation probabilities in a spatially heterogeneous environment. *Population Ecology* **44**: 51–58.
- Haccou, P., P. Jagers, and V. A. Vatutin, 2005 *Branching processes: variation, growth, and extinction of populations*. Number 5, Cambridge university press.
- Wallace, B., 1975 Hard and soft selection revisited. *Evolution* **29**: 465–473.
- Yeaman, S. and S. P. Otto, 2011 Establishment and maintenance of adaptive genetic divergence under migration, selection, and drift. *Evolution* **65**: 2123–2129.