

1 Details of the model

We denote the frequencies on the island of haplotypes A_1B_1 , A_1B_2 , A_2B_1 , and A_2B_2 by x_1 , x_2 , x_3 , and x_4 , respectively. The haplotype frequencies are related to the allele frequencies (p , q) and the linkage disequilibrium (D) as follows (e.g. Bürger 2000). The frequencies of A_1 and B_1 on the island can be expressed as $p = x_1 + x_2$ and $q = x_1 + x_3$. Accordingly, the frequencies of A_2 and B_2 are $1 - p = x_3 + x_4$ and $1 - q = x_2 + x_4$. Moreover, $x_1 = pq + D$, $x_2 = p(1 - q) - D$, $x_3 = (1 - p)q - D$, and $x_4 = (1 - p)(1 - q) + D$, and the linkage disequilibrium can be expressed in terms of the haplotype frequencies as $D = x_1x_4 - x_2x_3$. Thereby, we must recall the constraints $x_i \geq 0$ ($i = 1, \dots, 4$) and $\sum_{i=1}^4 x_i = 1$, which are equivalent to $0 \leq p, q \leq 1$ and

$$-\min\{pq, (1-p)(1-q)\} \leq D \leq \min\{p(1-q), (1-p)q\}. \quad (26)$$

The matrix of relative fitnesses on the island is

$$\mathbf{W} = \begin{matrix} & B_1B_1 & B_1B_2 & B_2B_2 \\ \begin{matrix} A_1A_1 \\ A_1A_2 \\ A_2A_2 \end{matrix} & \begin{pmatrix} w_{11} & w_{12} & w_{22} \\ w_{13} & w_{14} = w_{23} & w_{24} \\ w_{33} & w_{34} & w_{44} \end{pmatrix} \end{matrix}, \quad (27)$$

where w_{ij} is the relative fitness of the genotype composed of haplotypes i and j ($i, j \in \{1, 2, 3, 4\}$). For additive fitnesses, we use Eq. (1) in the main text. The marginal fitness of haplotype i on the island is defined as $w_{i\bullet} = \sum_{j=1}^4 w_{ij}x_j$ and the mean fitness of the island population as $\bar{w} = \sum_{i,j} w_{ij}x_ix_j = \sum_{i=1}^4 w_{i\bullet}x_i$.

Straightforward extension of two-locus models without migration (cf. Lewontin and Kojima 1960 or Bürger 2000, chap. 2) yields the recursion equations for the haplotype frequencies,

$$x'_1 = (1 - m)(x_1w_{1\bullet} - rw_{14}D)/\bar{w}, \quad (28a)$$

$$x'_2 = (1 - m)(x_2w_{2\bullet} + rw_{14}D)/\bar{w}, \quad (28b)$$

$$x'_3 = (1 - m)(x_3w_{3\bullet} + rw_{14}D)/\bar{w} + mq_c, \quad (28c)$$

$$x'_4 = (1 - m)(x_4w_{4\bullet} - rw_{14}D)/\bar{w} + m(1 - q_c), \quad (28d)$$

where r is the recombination rate, m the migration rate, and q_c the frequency of B_1 on the continent. For a monomorphic continent, $q_c = 0$. For this case, a continuous-time version of Eq. (28) has been fully described (Bürger and Akerman 2011).

2 Approximating the dynamics for rare A_1

Because A_1 arises as a novel mutation in our scenario (see main text), the haplotype frequencies x_1 and x_2 are initially small. We therefore ignore terms of order x_ix_j ($i, j \in \{1, 2\}$) and higher in Eq. (28). Moreover, we assume that, upon invasion of A_1 , the frequency of B_1 stays constant at the one-locus migration–selection equilibrium ($q = \hat{q}_B$). In principle, q approaches an internal equilibrium \hat{q}_+ , but the change is small compared to the change in p (Bürger and Akerman 2011). We then have $x_3 = q - x_1 \approx \hat{q}_B$ and $x_4 = 1 - q - x_2 \approx 1 - \hat{q}_B$ for x_1 and x_2 small. As a consequence, the dynamics in Eq. (28) reduces to a system with only two equations in x_1 and x_2 ,

$$x'_1 = (1 - m)[w_1x_1 + rw_{14}x_2\hat{q}_B - rw_{14}x_1(1 - \hat{q}_B)]/\bar{w}, \quad (29a)$$

$$x'_2 = (1 - m)[w_2x_2 - rw_{14}x_2\hat{q}_B + rw_{14}x_1(1 - \hat{q}_B)]/\bar{w}, \quad (29b)$$

where w_1 and w_2 are the marginal fitnesses of the A_1B_1 and A_1B_2 haplotypes, respectively. These are given by

$$w_1 = w_{13}\hat{q}_B + w_{14}(1 - \hat{q}_B), \quad (30a)$$

$$w_2 = w_{24}(1 - \hat{q}_B) + w_{14}\hat{q}_B. \quad (30b)$$

Moreover, \bar{w} is the mean fitness of the resident population on the island, which is assumed to be monomorphic at locus A:

$$\bar{w} = \hat{q}_B^2 w_{33} + 2\hat{q}_B(1 - \hat{q}_B)w_{34} + (1 - \hat{q}_B)^2 w_{44}. \quad (31)$$

This holds approximately if A_1 is rare on the island. Equation (29) can be written more compactly in matrix form as $\mathbf{x}' = \mathbf{x}\mathbf{L}$, where $\mathbf{x} = (x_1, x_2)$ is a row vector, and

$$\mathbf{L} = \begin{pmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \end{pmatrix}, \quad (32)$$

with

$$\lambda_{11} = (1 - m) [w_1 - r(1 - \hat{q}_B)w_{14}] / \bar{w}, \quad (33a)$$

$$\lambda_{12} = (1 - m)r(1 - \hat{q}_B)w_{14} / \bar{w}, \quad (33b)$$

$$\lambda_{21} = (1 - m)r\hat{q}_B w_{14} / \bar{w}, \quad (33c)$$

$$\lambda_{22} = (1 - m) [w_2 - r\hat{q}_B w_{14}] / \bar{w}. \quad (33d)$$

Setting $m = 0$, we recover the dynamics derived by Ewens (1967) for a panmictic population and a focal mutation occurring in linkage to a background locus at which overdominant selection maintains B_1 at frequency \hat{q}_B . We note that Eqs. (29) to (33) are valid for both a monomorphic and a polymorphic continent. The difference comes in only via \hat{q}_B , which is derived in the following section. Matrix \mathbf{L} will be encountered again as the *mean matrix* of the two-type branching process used to study the invasion probability of A_1 (see also the following section).

Note the difference between w_i and $w_{i\cdot}$: the former refers to the resident population under the assumption of the branching process (this section), whereas the latter applies to the island population in the general two-locus model (previous section). The same distinction holds for \bar{w} and \tilde{w} .

3 Marginal one-locus migration–selection model

We denote the marginal one-locus migration–selection equilibrium by $E_B = (p = 0, q = \hat{q}_B, D = 0)$. This equilibrium is assumed to be realised on the island before occurrence of the A_1 mutation. The equilibrium frequency \hat{q}_B of allele B_1 plays an important role. It determines the division of the resident island population into two genetic backgrounds and provides the weights for computing the average invasion probability of A_1 given the haplotype-specific invasion probabilities (see sections 2 and 4). Analysis of the one-locus dynamics (File S2) shows that \hat{q}_B is obtained by solving

$$q'_B = (1 - m) \frac{\tilde{w}_1}{\tilde{w}} q_B + m q_c = q_B \quad (34)$$

for q_B , where $\tilde{w}_1 = w_{33}q_B + w_{34}(1 - q_B)$ is the marginal relative fitness of the B_1 allele and

$$\tilde{w} = q_B^2 w_{33} + 2q_B(1 - q_B)w_{34} + (1 - q_B)^2 w_{44} \quad (35)$$

the mean fitness in the island population. From Eq. (34), one obtains

$$\hat{q}_B = \frac{w_{34}(1 - m) - \tilde{w} + \sqrt{4(1 - m)m q_c (w_{34} - w_{33})\tilde{w} + [\tilde{w} - (1 - m)]^2}}{2(1 - m)(w_{34} - w_{33})}, \quad (36)$$

which simplifies to $\hat{q}_B = [w_{34}(1 - m) - \tilde{w}] / [(1 - m)(w_{34} - w_{33})]$ for a monomorphic continent ($q_c = 0$). The equilibrium E_B is asymptotically stable if the migration rate is smaller than a critical value,

$$m < \frac{w_{34} - \tilde{w}}{w_{34}}. \quad (37)$$

We note that \tilde{w} is a (non-linear) function of q_B , and hence of m . Therefore, Eq. (36) is only an implicit solution and condition (37) not immediately informative. However, for additive fitnesses (see Eq. 1 of the main text) and a monomorphic continent ($q_c = 0$) we find the explicit solution given in Eq. (2). This is an admissible polymorphic equilibrium (i.e. $0 < \hat{q}_B < 1$), if the migration rate is below a critical value,

$$m < \frac{b}{1 - a} =: m_B. \quad (38)$$

Because $a < 1$ was assumed, m_B is always positive. Straightforward calculations show that Eq. (38) is also the condition for asymptotic stability of E_B within its marginal one-locus system. That is, under the marginal one-locus dynamics, E_B is stable whenever it is admissible (see File S2, or Nagylaki 1992, chap. 6.1).

When the mutation A_1 occurs, there is a transition from one- to two-locus dynamics. It is therefore crucial to study the stability of E_B also under the full two-locus dynamics. We find that E_B is not hyperbolic if $m = m^*$ or if $m = m_B > m^*$, with m^* given in Eq. (10). In the first case, E_B changes stability from unstable to asymptotically stable as m increases above m^* ; in the second case, E_B leaves the state space as m increases beyond m_B . We do not have a complete stability and bifurcation analysis of E_B . However, some numerical and analytical results suggest that the qualitative behaviour is the same as in the continuous-time model (Bürger and Akerman 2011). Then, the following holds. If E_B exists and is asymptotically stable under the one-locus dynamics, (i.e. $m < \min(b, m_B)$), but unstable under the two-locus dynamics (i.e. $m < m^*$), then a fully-polymorphic internal equilibrium E_+ ($0 < \hat{p}_+, \hat{q}_+ < 1$ and $\hat{D}_+ > 0$) exists and is asymptotically stable. Therefore, if $m < m^*$, a novel mutation A_1 can invade via E_B . Presumably, the internal equilibrium E_+ is reached. Comprehensive numerical computations under the discrete-time dynamics corroborate this conjecture (see File S2 and Figure S1).

With a polymorphic continent ($0 < q_c < 1$) and additive fitnesses, the frequency of B_1 at the marginal one-locus migration–selection polymorphism (E_B) is

$$\hat{q}_B = \frac{b - (1 - a)m + 2bm q_c + \sqrt{R}}{2b(1 + m)}, \quad (39)$$

where

$$R = 4b(1 - a - b)m(1 + m)q_c + [b - (1 - a)m + 2bm q_c]^2 \geq 0. \quad (40)$$

In contrast to the case of a monomorphic continent, where E_B exists only if $m < m_B$, with a polymorphic continent, both alleles B_1 and B_2 are introduced by migration and hence E_B always exists and is always asymptotically stable under the one-locus dynamics if $0 < q_c < 1$ and $0 < m < 1$.

A comprehensive analysis of the stability of E_B involves solving a complicated cubic equation, which results in expressions that are not informative. We could not accomplish a complete analytical treatment, but a combination of analytical, numerical and graphical approaches suggests the following. Upon occurrence of A_1 at locus A, E_B may either become unstable, in which case A_1 can invade and a fully-polymorphic internal equilibrium E_+ is reached, or E_B may stay asymptotically stable, in which case A_1 cannot invade. The transition between these two scenarios occurs at a critical recombination rate

$$r^* = \begin{cases} \frac{1}{2} & \text{if } m \leq m_{r^*}, \\ \tilde{r}^*(m) & \text{otherwise,} \end{cases} \quad (41)$$

where $\tilde{r}^*(m)$ is a complicated function of m that we do not present here (but see Eq. 3 in File S2, and Eq. 92 in section 6), and m_{r^*} is the migration rate at which $\tilde{r}^*(m)$ has a pole. Then, for a given combination of values for a , b , m and q_c , A_1 can invade if and only if $r < r^*$ (Figure S2). A similar argument holds for a critical continental frequency q_c^* of B_1 , such that for a given combination of values for a , b , m and r , A_1 can invade if and only if $q_c < q_c^*$ (see File S2 for details). We were not able to find an explicit expression for a critical migration rate m^* with an interpretation analogous to that of r^* or q_c^* . However, m^* is implicitly defined by r^* or q_c^* and can be computed numerically.

As a final remark, we note that for weak evolutionary forces, Eqs. (2), (38) and (10) can be approximated by the corresponding equations derived by Bürger and Akerman (2011) for the continuous-time model with a monomorphic continent. Specifically, scaling a , b , m and r by ϵ and expanding Eqs. (2), (38) and (10) into a Taylor series around $\epsilon = 0$ yields

$$\hat{q}_B \approx 1 - \frac{m}{b}, \quad (42)$$

$$m_B \approx b, \quad (43)$$

and

$$m^* \approx a \left(1 + \frac{b - a}{r} \right) \quad (44)$$

to first order of ϵ and after rescaling. Equations (42) and (44) correspond to Eqs. (3.9) and (3.11) in Bürger and Akerman (2011).

4 Branching-process approximation to the invasion probability

For a proper stochastic treatment, the evolution of haplotype frequencies has to be modelled by a Markov process. In the context of invasion of novel mutations, particularly useful approximations can be obtained using branching processes (Fisher 1922) and diffusion processes (Kimura 1962). Both approaches deal with the probabilistic effect due to the initially small absolute number of copies of the mutant allele. The effect of finite population size is only accounted for by the diffusion approximation, however. In the first part of the main paper, we are concerned only with initial rareness of the mutation.

We employ a two-type branching process (Harris 1963; Ewens 1968, 1967) to study the dynamics of the two haplotypes of interest, A_1B_1 (type 1) and A_1B_2 (type 2) after occurrence of mutation A_1 (see section 2 above). Let λ_{ij} be the mean number of j -type offspring produced by an i -type parent each generation, and x_i the proportion of type i in the island population. Then the expected proportion of types A_1B_1 and A_1B_2 in the next generation is

$$\mathbb{E}[x'_1] = \lambda_{11}x_1 + \lambda_{21}x_2, \quad (45a)$$

$$\mathbb{E}[x'_2] = \lambda_{12}x_1 + \lambda_{22}x_2, \quad (45b)$$

or, in matrix form

$$\mathbb{E}[\mathbf{x}'] = \mathbf{x}\mathbf{L}, \quad (46)$$

where $\mathbf{x} = (x_1, x_2)$, and $\mathbf{L} = (\lambda_{ij})$, $i, j \in \{1, 2\}$, is called the *mean matrix* (cf. Eq. 32 in section 2). The leading eigenvalue ν of \mathbf{L} determines whether the branching process is supercritical ($\nu > 1$) and A_1 has a strictly positive invasion probability, or subcritical ($\nu < 1$), in which case A_1 goes extinct with probability 1. Expressions for the λ_{ij} were given in Eq. (33).

The leading eigenvalue of \mathbf{L} is

$$\nu = \frac{1-m}{2\bar{w}} \left[w_1 + w_2 - rw_{14} + \sqrt{(w_1 - w_2)^2 + 2rw_{14}(2\hat{q}_B - 1)(w_1 - w_2) + r^2w_{14}^2} \right], \quad (47)$$

where w_1 and w_2 are the marginal fitnesses of type 1 and type 2 defined in Eq. (30), and \bar{w} is the mean fitness of the resident population on the island as defined in Eq. (31) (section 2). After some algebra (see File S3), the condition for invasion of A_1 , $\nu > 1$, is found to be equivalent to Eq. (9) in the main text. Equations (47) and (9) hold for both a monomorphic and a polymorphic continent.

Let ζ_{ij} be the random number of j -type offspring produced by a single i -type parent. We assume that ζ_{i1} and ζ_{i2} are independent and Poisson-distributed with mean λ_{i1} and λ_{i2} , respectively ($i \in \{1, 2\}$). Then, the probability-generating function (pgf) of ζ_{ij} is

$$f_{ij}(s_j) = \mathbb{E}[s_j^{\zeta_{ij}}] = \sum_{k=0}^{\infty} p_k s_j^k = e^{-\lambda_{ij}(1-s_j)}, \quad i, j \in \{1, 2\}, \quad (48)$$

where $p_k = \mathbb{P}[\zeta_{ij} = k]$ is the probability that an i -type parent has k offspring of type j . The first two equalities follow from the definition of the pgf (e.g. Harris 1963), and the third from the properties of the Poisson distribution. Because of independent offspring distributions for each type, the pgf for the number of offspring (of any type) produced by an i -type parent is given by

$$f_i(s_1, s_2) = \prod_{j=1}^2 f_{ij}(s_j). \quad (49)$$

Inserting Eq. (48) into Eq. (49), we obtain

$$f_1(s_1, s_2) = e^{-\lambda_{11}(1-s_1)} \cdot e^{-\lambda_{12}(1-s_2)}, \quad (50a)$$

$$f_2(s_1, s_2) = e^{-\lambda_{21}(1-s_1)} \cdot e^{-\lambda_{22}(1-s_2)}. \quad (50b)$$

We use Q_i for the extinction probability of allele A_1 conditional on initial occurrence on background B_i , and $\pi_i = 1 - Q_i$ for the respective probability of invasion. The extinction probabilities Q_i are found as the smallest positive solution to Eq. (3) in the main text. The average invasion probability $\bar{\pi}$ is found as the weighted average of π_1 and π_2 (see Eq. 4 in the main text). As the problem stated in Eq. (3) amounts to solving a system of transcendental equations, an explicit solution cannot be found in general. Numerical solutions can be obtained, however (see File S3).

We proceed by assuming additive fitnesses as defined in Eq. (1) of the main text. The entries λ_{ij} of the mean matrix \mathbf{L} in Eq. (32) are then given by

$$\lambda_{11} = E + Fr, \quad (51a)$$

$$\lambda_{12} = -Fr, \quad (51b)$$

$$\lambda_{21} = Hr, \quad (51c)$$

$$\lambda_{22} = J - Hr, \quad (51d)$$

where

$$E = \frac{1 + b + am}{1 - a + b}, \quad (52a)$$

$$F = -\frac{m}{b}, \quad (52b)$$

$$H = \frac{b - (1 - a)m}{b(1 - a + b)}, \quad (52c)$$

$$J = \frac{1 + m(a - b)}{1 - a + b}. \quad (52d)$$

Assuming weak evolutionary forces, i.e. replacing a, b, m and r by $\alpha\epsilon, \beta\epsilon, \mu\epsilon$ and $\rho\epsilon$, respectively, and expanding into a Taylor series around $\epsilon = 0$, the terms in Eq. (51) are approximated to first order in ϵ by

$$\begin{aligned} \lambda_{11} &\approx 1 + a - \frac{m}{b}r, & \lambda_{12} &\approx \frac{m}{b}r, \\ \lambda_{21} &\approx \left(1 - \frac{m}{b}\right)r, & \lambda_{22} &\approx 1 + a - b - \left(1 - \frac{m}{b}\right)r, \end{aligned}$$

after resubstituting $\alpha \rightarrow a/\epsilon, \beta \rightarrow b/\epsilon, \mu \rightarrow m/\epsilon$ and $\rho \rightarrow r/\epsilon$.

With additive fitnesses and a monomorphic continent, the dominant eigenvalue of \mathbf{L} is

$$\nu = \frac{2 + b - r + m(2a - b - r) + \sqrt{R_1}}{2(1 - a + b)}, \quad (53)$$

where

$$R_1 = (1 + m) \{b^2(1 + m) + 2b(1 - m)r + r[r - m(4 - 4a - r)]\}. \quad (54)$$

The branching process is supercritical ($\nu > 1$) if $m < m^*$ or, alternatively, if $r < r^*$, with m^* and r^* the critical migration and recombination rates defined in Eqs. (10) and (11) of the main text, respectively (see File S3 for details). Assuming weak evolutionary forces, ν simplifies to

$$\nu \approx 1 + \frac{1}{2} \left(2a - b - r + \sqrt{R_2}\right),$$

where

$$R_2 = b^2 + 2br - 4mr + r^2. \quad (55)$$

Then, m^* is approximated by Eq. (44) and

$$r^* \approx \tilde{r}^* = \begin{cases} \infty & \text{if } m \leq a, \\ \frac{a(b-a)}{m-a} & \text{otherwise} \end{cases} \quad (56)$$

(see File S3). Note that the critical migration and recombination rates for invasion of A_1 obtained under the deterministic model (section 3) and the corresponding two-type branching process are identical. In File S4 we show that this agreement is generically expected.

To obtain the extinction probabilities of A_1 given initial occurrence on background B_1 or B_2 , we plug Eq. (51) into (50) and solve

$$f_1(s_1, s_2) = e^{(E+Fr)s_1 - Frs_2 - E} = s_1 \quad (57a)$$

$$f_2(s_1, s_2) = e^{Hrs_1 + (J-Hr)s_2 - J} = s_2 \quad (57b)$$

for s_1 and s_2 . The smallest solutions between 0 and 1 are the extinction probabilities $Q_1 = 1 - \pi_1$ and $Q_2 = 1 - \pi_2$ (cf. Eq. 3 in the main text). An explicit solution is not available and we need to use numerical methods to obtain exact results (File S3).

We now turn to the case of a polymorphic continent ($0 < q_c < 1$), still assuming additive fitnesses. Then,

$$\lambda_{11} = \tilde{E} + \tilde{F}r, \quad (58a)$$

$$\lambda_{12} = \tilde{G}r, \quad (58b)$$

$$\lambda_{21} = \tilde{H}r, \quad (58c)$$

$$\lambda_{22} = \tilde{J} + \tilde{I}r, \quad (58d)$$

with

$$\begin{aligned} \tilde{E} &= \frac{(1-m)(2+b+m+am+2bmq_c+\sqrt{R})}{2[1-a-bm(1-2q_c)+\sqrt{R}]}, \\ \tilde{F} &= -\frac{(1-m)[b+(1-a)m+2bm(1-q_c)-\sqrt{R}]}{2b[1-a-bm(1-2q_c)+\sqrt{R}]}, \\ \tilde{G} &= \frac{b+m[1-a-2b(1-q_c)]-\sqrt{R}}{2b(1-a-b)}, \\ \tilde{H} &= \frac{b-(1-a)m-2bmq_c+\sqrt{R}}{2b(1-a+b)}, \\ \tilde{I} &= -\frac{(1-m)[b-(1-a)m+2bmq_c+\sqrt{R}]}{2b[1-a-bm(1-2q_c)+\sqrt{R}]}, \\ \tilde{J} &= \frac{(1-m)[2+m+am-b(1+2m(1-q_c))+\sqrt{R}]}{2[1-a-bm(1-2q_c)+\sqrt{R}]}. \end{aligned}$$

Here, R is as defined in Eq. (40). Assuming weak evolutionary forces, i.e. scaling a , b , m and r by ϵ and expanding into a Taylor series around $\epsilon = 0$, Eq. (58) is approximated to first order in ϵ by

$$\begin{aligned} \lambda_{11} &\approx \frac{1}{2}(2+2a+b-m-\sqrt{R_3}) - \frac{b+m-\sqrt{R_3}}{2b}r, & \lambda_{12} &\approx \frac{b+m-\sqrt{R_3}}{2b}r, \\ \lambda_{21} &\approx \frac{b-m+\sqrt{R_3}}{2b}r, & \lambda_{22} &\approx \frac{1}{2}(2+2a-b-m-\sqrt{R_3}) - \frac{b-m+\sqrt{R_3}}{2b}r, \end{aligned}$$

where

$$R_3 = (b-m)^2 + 4bmq_c > 0. \quad (59)$$

Note that the continental frequency q_c of B_1 enters these equations only via $4bmq_c$ in the radicand R_3 . For a polymorphic continent, the eigenvalues of \mathbf{L} are complicated expressions, which we do not show here (but see File S3). The leading eigenvalue can be identified, though. For weak evolutionary forces, and to first order in ϵ , it is approximately

$$\nu \approx 1 + \frac{1}{2} \left[2a - m - r - \sqrt{R_3} + \sqrt{b^2 - r(2m - r - 2\sqrt{R_3})} \right] \quad (60)$$

(see File S5). Finally, the system of transcendental equations to be solved in order to obtain the extinction probabilities of A_1 becomes

$$f_1(s_1, s_2) = e^{-(\tilde{E}+\tilde{F}r)(1-s_1)-\tilde{G}r(1-s_2)} = s_1 \quad (61a)$$

$$f_2(s_1, s_2) = e^{-\tilde{H}r(1-s_1)-(\tilde{J}+\tilde{I}r)(1-s_2)} = s_2 \quad (61b)$$

(cf. Eq. 3 of main text).

To obtain analytical approximations to the invasion probability of A_1 , we follow Haccou (2005, pp. 127–128) and assume that the branching process is slightly supercritical (see also Eshel 1984; Hoppe 1992; Athreya 1992, 1993). This means that the leading eigenvalue of the mean matrix \mathbf{L} is of the form

$$\nu = \nu(\xi) = 1 + \xi, \quad (62)$$

where ξ is small and positive. To make explicit the dependence on ξ , we write $Q_i = Q_i(\xi)$ and $\pi_i = \pi_i(\xi)$ for the extinction and invasion probabilities, respectively ($i \in 1, 2$). Using the Ansatz in Eq. (62), Haccou et al. state in their Theorem 5.6 that, as $\xi \rightarrow 0$, $q_i(\xi)$ converges to 1 and

$$\pi_i(\xi) = 1 - q_i(\xi) = \frac{2[\nu(\xi) - 1]}{\mathbf{B}(\xi)} v_i(\xi) + o(\xi). \quad (63)$$

Here, $v_k = i$ is the i th entry of the right eigenvector $\mathbf{v} = (v_1, v_2)^\top$ pertaining to the leading eigenvalue ν of the mean matrix \mathbf{L} . The matrix $\mathbf{B}(\xi)$ is defined as

$$\mathbf{B}(\xi) = \sum_{i=1}^2 u_i \sum_{j=1}^2 v_j \lambda_{ij} + \nu(\xi) [1 - \nu(\xi)] \sum_{j=1}^2 u_j v_j^2, \quad (64)$$

where u_i is the i th entry of the normalised left eigenvector $\mathbf{u} = (u_1, u_2)$ associated with the leading eigenvalue ν of \mathbf{L} . By normalised we mean that $\sum_k^2 u_k = 1$. For Eq. (63) to hold, \mathbf{u} and \mathbf{v} must in addition fulfill $\sum_{k=1}^2 u_k v_k = 1$.

For additive fitnesses (Eq. 1) and a monomorphic continent ($q_c = 0$), we combine Eqs. (53) and (62) to identify ξ as

$$\xi = \frac{2a(1+m) - b - r - m(b+r) + \sqrt{R_1}}{2(1-a+b)}, \quad (65)$$

where R_1 is defined in Eq. (54). Therefore, the assumption of a slightly supercritical branching process will hold for all parameter combinations that result in a small positive ξ in Eq. (65). For weak evolutionary forces, Eq. (65) is approximated by the simpler expression below Eq. (12) in the main text. After some algebra using *Mathematica* (File S5), we obtain the appropriately normalised left and right eigenvectors of \mathbf{L} as

$$\mathbf{u} = \left(\frac{b(1+m) - (1+m)r + \sqrt{R_1}}{2b(1+m)}, \frac{2(1-a+b)mr}{b[b+r+m(b+r) + \sqrt{R_1}]} \right)^\top \quad (66)$$

and

$$\mathbf{v} = \left(\frac{b^2(1+m) - 2(1-a)mr + b(r - mr + \sqrt{R_1})}{(b+r)^2 + m[(b-r)^2 - 4(1-a)r] + (b-r)\sqrt{R_1}}, \frac{2[b - (1-a)m]r}{(b+r)^2 + m[(b-r)^2 - 4(1-a)r] + (b-r)\sqrt{R_1}} \right), \quad (67)$$

respectively. Combining Eqs. (51), (53), (66), (67), and (64), we find analytical expressions for the conditional invasion probabilities $\pi_1(\xi)$ and $\pi_2(\xi)$ under a slightly supercritical branching process. The weighted average invasion probability $\bar{\pi}(\xi)$ is obtained according to Eq. (4) with \hat{q}_B given in Eq. (2). The resulting expressions are long and not very informative (see File S5 for details and Figure 2 for a graphical comparison to numerical solutions). However, if we assume weak evolutionary forces, we obtain the analytical approximations $\tilde{\pi}_1(\xi)$ and $\tilde{\pi}_2(\xi)$ given in Eq. (12) of the main text. The corresponding average invasion probability $\bar{\tilde{\pi}}(\xi)$ is obtained by insertion of Eqs. (12) and (2) into Eq. (4) (see main text).

For a polymorphic continent ($0 < q_c < 1$), the procedure is analogous to the one outlined above. Intermediate and final expressions are more complicated as those obtained for the monomorphic continent, though. We therefore refer to File S5 for details and to Figures S5 and S6 for a graphical comparison to numerical solutions. The approximations $\tilde{\pi}_1(\xi)$, $\tilde{\pi}_1(\xi)$ and $\bar{\tilde{\pi}}(\xi)$ given in Eqs. (7)–(9) in File S5 for weak evolutionary forces and $0 < q_c < 1$ are accurate if ξ is small, where

$$\xi \approx \frac{1}{2} \left[m + r\sqrt{R_3} - \sqrt{b^2 - r(2m - r - 2\sqrt{R_3})} \right]$$

and R_3 is defined in Eq. (59). Then, the branching process is slightly supercritical (cf. Eq. 62). In practice, the approximations derived for a polymorphic continent are useful for efficient plotting, but otherwise not very intuitive. Letting $q_c \rightarrow 0$ and assuming $m < m_B$ (cf. Eq. 38 in section 3), we recover the respective analytical expressions for the case of a monomorphic continent.

5 Condition for a non-zero optimal recombination rate

Observation of the mean invasion probability $\bar{\pi}$ of allele A_1 as a function of the recombination rate r suggests that $\bar{\pi}(r)$ may have a maximum at a non-zero recombination rate ($r_{\text{opt}} > 0$) in some cases, whereas it is maximised at $r_{\text{opt}} = 0$ in other cases (Figures 1A and 1B). To distinguish between these two regimes, we note that $r_{\text{opt}} > 0$ holds whenever the derivative of $\bar{\pi}(r)$ with respect to r , evaluated at $r = 0$, is positive. This is because $\bar{\pi}(r)$ will always decay for sufficiently large r . We denote the derivative of interest by

$$\bar{\pi}'(0) := \frac{d}{dr} [\hat{q}_B \pi_1(r) + (1 - \hat{q}_B) \pi_2(r)] \Big|_{r=0} = \hat{q}_B \frac{d\pi_1(r)}{dr} \Big|_{r=0} + (1 - \hat{q}_B) \frac{d\pi_2(r)}{dr} \Big|_{r=0}, \quad (68)$$

where π_1 and π_2 are the invasion probabilities of A_1 conditional on initial occurrence on the B_1 and B_2 background, respectively, and \hat{q}_B is the equilibrium frequency of B_1 before invasion of A_1 . In the following, we obtain $\bar{\pi}'(0)$ via implicit differentiation. We will first derive a general, implicit condition for $\bar{\pi}'(0) > 0$, and then proceed by assuming additive fitnesses to obtain explicit conditions. We will do so first for a monomorphic ($q_c = 0$) and then for a polymorphic ($0 < q_c < 1$) continent.

We start from Eq. (3) of the main text with probability generating functions $f_i(s_1, s_2)$ ($i \in \{1, 2\}$) as defined in Eq. (50) in section 4. Recall that the extinction probabilities $Q_i = 1 - \pi_i$ are the smallest positive solutions to Eq. (3). Assuming that these solutions have been identified, we know that the invasion probabilities π_i satisfy

$$1 - \pi_1 = e^{-\lambda_{11}\pi_1} \cdot e^{-\lambda_{12}\pi_2}$$

$$1 - \pi_2 = e^{-\lambda_{21}\pi_1} \cdot e^{-\lambda_{22}\pi_2}.$$

Taking the logarithm on both sides and making the dependence of both π_i and λ_{ij} on r explicit, we have

$$\ln [1 - \pi_1(r)] = -\lambda_{11}(r)\pi_1(r) - \lambda_{12}(r)\pi_2(r) \quad (69a)$$

$$\ln [1 - \pi_2(r)] = -\lambda_{21}(r)\pi_1(r) - \lambda_{22}(r)\pi_2(r). \quad (69b)$$

Applying the formulae for the $\lambda_{ij}(r)$ given in Eq. (33), Eq. (69) becomes

$$\ln [1 - \pi_1(r)] = -\frac{1-m}{\bar{w}} \left\{ [w_1 - (1 - \hat{q}_B)r w_{14}] \pi_1(r) + (1 - \hat{q}_B)r w_{14} \pi_2(r) \right\} \quad (70a)$$

$$\ln [1 - \pi_2(r)] = -\frac{1-m}{\bar{w}} \left\{ \hat{q}_B r w_{14} \pi_1(r) + (w_2 - \hat{q}_B r w_{14}) \pi_2(r) \right\}. \quad (70b)$$

Differentiating both sides with respect to r , and setting $r = 0$ yields

$$\frac{\pi_1'(0)}{1 - \pi_1^\circ} = (1-m) \frac{w_1 \pi_1'(0) - (1 - \hat{q}_B) w_{14} (\pi_1^\circ - \pi_2^\circ)}{\bar{w}} \quad (71a)$$

$$\frac{\pi_2'(0)}{1 - \pi_2^\circ} = (1-m) \frac{w_2 \pi_2'(0) + \hat{q}_B w_{14} (\pi_1^\circ - \pi_2^\circ)}{\bar{w}}, \quad (71b)$$

where $\pi_i'(0) = \frac{d\pi_i(r)}{dr} \Big|_{r=0}$ for $i \in \{1, 2\}$. Moreover, $\pi_1^\circ = \pi_1(0)$ and $\pi_2^\circ = \pi_2(0)$ are the conditional invasion probabilities of A_1 if it initially occurs on background B_1 and B_2 , respectively, and if there is no recombination ($r = 0$). Solving the system in Eq. (71) for $\pi_1'(0)$ and $\pi_2'(0)$, and plugging the solutions into Eq. (68), we find after some algebra

$$\bar{\pi}'(0) = (1-m) \hat{q}_B (1 - \hat{q}_B) (\pi_2^\circ - \pi_1^\circ) \frac{w_{14}}{\bar{w}} \left(\frac{1 - \pi_1^\circ}{1 - (1-m)(1 - \pi_1^\circ) w_1 / \bar{w}} - \frac{1 - \pi_2^\circ}{1 - (1-m)(1 - \pi_2^\circ) w_2 / \bar{w}} \right). \quad (72)$$

Setting $r = 0$ in Eq. (70) and rearranging, we obtain

$$(1-m) \frac{w_i}{\bar{w}} = -\ln(1 - \pi_i^\circ) / \pi_i^\circ \quad i \in \{1, 2\}. \quad (73)$$

Insertion of Eq. (73) into Eq. (72) yields

$$\bar{\pi}'(0) = (1-m) \hat{q}_B (1 - \hat{q}_B) (\pi_2^\circ - \pi_1^\circ) \frac{w_{14}}{\bar{w}} \left(\frac{1 - \pi_1^\circ}{1 + \ln(1 - \pi_1^\circ) / \pi_1^\circ} - \frac{1 - \pi_2^\circ}{1 + \ln(1 - \pi_2^\circ) / \pi_2^\circ} \right). \quad (74)$$

At this point, a closer inspection of Eq. (73) is worthwhile. Straightforward rearrangement leads to

$$1 - \pi_i^\circ = \exp\left[-(1-m)\frac{w_i}{\bar{w}}\pi_i^\circ\right] \quad i \in \{1, 2\}, \quad (75)$$

which has a solution π_i° in $(0, 1]$ if and only if $(1-m)w_i/\bar{w} > 0$. Otherwise, the only solution is $\pi_i^\circ = 0$. In our setting, we always assumed that when A_1 occurs on the deleterious background (B_2), it will form a suboptimal haplotype (A_1B_2 less fit on the island than A_1B_1) and go extinct in the absence recombination. This assumption translates into $w_2 < \bar{w}$. As $0 < m < 1$, we immediately note that for $i = 2$, the only possible solution of Eq. (75) is $\pi_2^\circ = 0$. Therefore, whenever $w_1 > \bar{w}/(1-m)$ holds, the derivative of interest in Eq. (72) simplifies to

$$\bar{\pi}'(0) = (1-m)\hat{q}_B(1-\hat{q}_B)\pi_1^\circ\frac{w_{14}}{\bar{w}}\left(\frac{\bar{w}}{\bar{w}-(1-m)w_2} - \frac{1-\pi_1^\circ}{1-(1-m)(1-\pi_1^\circ)w_1/\bar{w}}\right). \quad (76)$$

After some algebra (File S6), we find that $\bar{\pi}'(0) > 0$, and hence $r_{\text{opt}} > 0$, is equivalent to Eq. (13) in the main text. Again, if we set $m = 0$ in the derivation above, we obtain expressions previously derived by Ewens for a panmictic population in which the background locus is maintained polymorphic by heterozygote superiority (Ewens 1967).

To obtain more explicit conditions, we assume additive fitnesses (Eq. 1). We start directly from Eq. (57), replacing s_i by the smallest solution Q_i between 0 and 1. Taking the logarithm on both sides and making the dependence of Q_i on r explicit, we find

$$\ln Q_1(r) = (E + Fr)Q_1(r) - FrQ_2(r) - E \quad (77a)$$

$$\ln Q_2(r) = HrQ_1(r) + (J - Hr)Q_2(r) - J, \quad (77b)$$

where E, F, J and H are independent of r and as defined in Eq. (52). Differentiating Eq. (77) on both sides, setting $r = 0$ and rearranging, we obtain

$$\frac{Q_1'(0)}{Q_1^\circ} = F(Q_1^\circ - Q_2^\circ) + EQ_1'(0)$$

$$\frac{Q_2'(0)}{Q_2^\circ} = H(Q_1^\circ - Q_2^\circ) + JQ_2'(0),$$

with $Q_i'(0) = \frac{dQ_i(r)}{dr}\big|_{r=0}$. Here, we used $Q_i^\circ = Q_i(0)$ for the extinction probability of A_1 conditional on initial occurrence on background B_i ($i \in \{1, 2\}$). Solving for $Q_1'(0)$ and $Q_2'(0)$ yields

$$Q_1'(0) = \frac{FQ_1^\circ(Q_1^\circ - Q_2^\circ)}{1 - EQ_1^\circ} \quad (78a)$$

$$Q_2'(0) = \frac{HQ_2^\circ(Q_1^\circ - Q_2^\circ)}{1 - JQ_2^\circ}. \quad (78b)$$

To obtain an explicit solution, we aim at approximating the Q_i° in the following. Going back to Eq. (57) again, but setting $r = 0$ directly, we find

$$Q_i^\circ = e^{-Z_i(1-Q_i^\circ)} \quad i \in \{1, 2\}, \quad (79)$$

where

$$Z_1 := E = \frac{1+b+am}{1-a+b}, \quad (80a)$$

$$Z_2 := J = \frac{1+m(a-b)}{1-a+b}. \quad (80b)$$

Importantly, the equations for Q_1° and Q_2° in (79) are now decoupled. Moreover, we note that Eq. (79) has a solution Q_i° in $[0, 1)$ if and only if $Z_i > 1$; if $Z_i \leq 1$, the solution is $Q_i^\circ = 1$. In other words, in the case of complete linkage ($r = 0$), type i has a non-zero invasion probability if and only if $Z_i > 1$ (recall that $\pi_i^\circ = 1 - Q_i^\circ$). Closer inspection of Eq. (80) shows that, given our assumptions of $a < b$ and $0 < m < 1$, $Z_1 > 1$ and $Z_2 < 1$ hold always. Hence, we have $\pi_2^\circ = 1 - Q_2^\circ = 0$, and we are left with finding an approximate solution of Eq. (79) for $i = 1$. For this purpose, we focus on the case where invasion is

just possible, i.e. π_1° is close to 0 and hence Q_1° close to 1. This is equivalent to Z_1 being close to, but larger than, 1. We therefore use the Ansatz

$$Z_1 = 1 + \epsilon \quad (81)$$

with $\epsilon > 0$ small. We then have $Q_1^\circ = e^{-(1+\epsilon)(1-Q_1^\circ)}$. Noting that $Q_1^\circ(\epsilon)$ must be close to 1 for ϵ small, we expand the right-hand side into a Taylor series around $Q_1^\circ = 1$, which results in

$$Q_1^\circ = 1 - (1 - Q_1^\circ)(1 + \epsilon) + \frac{1}{2}(1 - Q_1^\circ)^2(1 + \epsilon)^2 + \mathcal{O}(Q_1^\circ)^3 \quad (82)$$

Neglecting terms beyond $\mathcal{O}(Q_1^\circ)^2$ and solving for Q_1° , we obtain $Q_1^\circ = (1 + \epsilon^2)/(1 + \epsilon)^2$ (excluding the trivial solution $Q_1^\circ = 1$). To first order in ϵ , this is approximated by

$$Q_1^\circ = 1 - \pi_1^\circ \approx 1 - 2\epsilon. \quad (83)$$

We identify ϵ by inserting Eq. (80a) into Eq. (81) and solving for ϵ . To first order in a , this yields $\epsilon \approx a(1 + m)/(1 + b)$ and hence, from Eq. (83), we find

$$Q_1^\circ = 1 - 2 \frac{a(1 + m)}{(1 + b)} + \mathcal{O}(a)^2. \quad (84)$$

Note that if we set $m = 0$ (no migration) and $b = 0$ (no background selection), we recover Haldane's (1927) well-known approximation $\pi \approx 2a$.

Comparison of Eqs. (83) and (84) suggests that the invasion probability π_1° increases with the migration rate m . This may seem counterintuitive. However, with complete linkage ($r = 0$), the cases of A_1 occurring on background B_1 or B_2 can be considered separately. If A_1 occurs on background B_1 , it forms haplotype A_1B_1 . From then on it competes against the resident population consisting of haplotypes A_2B_1 and A_2B_2 at frequencies \hat{q}_B and $1 - \hat{q}_B$, respectively. Because, initially, A_1B_1 types do not interfere nor contribute to the resident population, what matters is the ratio of the marginal fitness w_1 of A_1B_1 to the mean fitness \bar{w} of the resident population. This follows directly from Eq. (73). Equations (30a) and (31) in section 2 show that both w_1 and \bar{w} depend on \hat{q}_B . For additive fitnesses, \hat{q}_B is given by Eq. (2) in the main text; it depends on m . Therefore, to understand the apparently paradoxical increase of π_1° on m , we must compare the dependence on m of w_1 and \bar{w} . We have $w_1 = (1 + b + am)/(1 + m)$ and $\bar{w} = (1 - m)(1 - a + b)/(1 + m)$. Both decrease with m , but \bar{w} does so faster. The ratio $w_1/\bar{w} = (1 + b + am)/[(1 - a + b)(1 - m)]$ increases quickly with m (File S6). This explains why π_1° increases with m . It also explains why π_1 increases with small m in Figure S3D for very weak recombination. If recombination is too strong, the effect vanishes (Figures S3E and S3F).

Finally, plugging Eq. (52) from section 4 and Eq. (84) into Eq. (78), we obtain the explicit approximations

$$Q_1'(0) \approx \frac{2m(1 - a + b)[1 + b - 2a(1 + m)]}{b(1 + b)(1 + b + 2am)}, \quad (85a)$$

$$Q_2'(0) \approx \frac{2a[b - (1 - a)m]}{b(1 + b)(a - b)}, \quad (85b)$$

valid for a small relative to m and b . Noting that $\bar{\pi}'(0) = -[\hat{q}_B Q_1'(0) + (1 - \hat{q}_B) Q_2'(0)]$ and using \hat{q}_B from Eq. (2) of the main text for additive fitnesses and a monomorphic continent, we find the approximate derivative of the mean invasion probability $\bar{\pi}$ at $r = 0$ as

$$\bar{\pi}'(0) \approx \frac{2m(1 - a + b)[b - (1 - a)m]\{2a^2 + b + b^2 - 2a[1 + b(2 + m)]\}}{b^2(1 + b)(a - b)(1 + m)(1 + b + 2am)}. \quad (86)$$

After some algebra, one can show that $\bar{\pi}'(0) > 0$, and hence $r_{\text{opt}} > 0$, if $a > 1 - b/m$ and $a > a^*$, with a^* defined in Eq. (14) of the main text. Combination of Eq. (14) with our assumption $a < b$ and the condition for existence of the marginal one-locus equilibrium E_B ($a > 1 - b/m$, from Eq. 38 in section 3) yields a sufficient condition for $r_{\text{opt}} > 0$ (Figure 3). For further details, we refer to File S6.

For the case of a polymorphic continent ($q_c > 0$), we were not able to derive informative analytical conditions for $r_{\text{opt}} > 0$. Analytical and numerical computations in File S6 suggest that if we start with a monomorphic continent ($q_c = 0$) in a constellation where $r_{\text{opt}} > 0$ holds, and then increase q_c , the maximum in $\bar{\pi}(r)$ shifts to 0 ($r_{\text{opt}} \rightarrow 0$). There must be a critical value of q_c at which the shift from $r_{\text{opt}} > 0$ to $r_{\text{opt}} = 0$ occurs, but we could not determine it analytically.

6 Analysis of the deterministic model in continuous time

For the diffusion approximation in the following section we will need a continuous-time version of our model as a starting point. Here, we derive this model from the discrete-time version. We will analyse some properties of interest in the context of invasion and survival of a weakly beneficial mutation arising in linkage to a migration–selection polymorphism. The continuous-time version with a monomorphic continent ($q_c = 0$) has been completely analysed by Bürger and Akerman (2011). Therefore, we only summarise some of their results and focus on the extension to a polymorphic continent ($0 < q_c < 1$). We use a tilde (\sim) to distinguish continuous-time expressions from their analogous terms in discrete time. For ease of typing, though, this distinction is not made in all *Mathematica* Notebooks provided in the Supporting Information.

We start from the recursion equations for the haplotype frequencies given in Eq. (28) of this text, with relative fitnesses w_{ij} according to Eq. (1). As we will assume quasi-linkage equilibrium (QLE) in the following section, it is more convenient to express the dynamics in terms of allele frequencies (p, q) and linkage disequilibrium (D), rather than haplotype frequencies. This is achieved by recalling the relationships between D, p, q , and the x_i ($i = 1, \dots, 4$) given in section 1. The resulting difference equations are complicated and only shown in File S7. We obtain the differential equations by assuming that the changes due to selection, migration and recombination are small during a short time interval Δt . Scaling a, b, m and r by Δt and taking the limit $\lim_{\Delta t \rightarrow 0} \frac{\Delta x}{\Delta t}$ for $x \in \{p, q, D\}$ results in

$$\dot{p} = \frac{dp}{dt} = ap(1-p) - mp + bD, \quad (87a)$$

$$\dot{q} = \frac{dq}{dt} = bq(1-q) - m(q - q_c) + aD, \quad (87b)$$

$$\dot{D} = \frac{dD}{dt} = [a(1-2p) + b(1-2q)]D + m[p(q - q_c) - D] - rD. \quad (87c)$$

For a monomorphic continent ($q_c = 0$), one finds the marginal one-locus migration–selection equilibrium \tilde{E}_B for locus B by setting $p = D = 0$ and solving $\dot{q} = 0$ for q , which yields

$$\hat{q}_B = 1 - \frac{m}{b} \quad (88)$$

as the solution of interest (cf. Eq. 42). Bürger and Akerman (2011) have shown that this equilibrium is asymptotically stable in its one-locus dynamics whenever it exists, i.e. when $m < b = \tilde{m}_B$. Moreover, it is asymptotically stable under the two-locus dynamics if and only if $\tilde{m}^* < m < b$, where $\tilde{m}^* = a \left(1 + \frac{b-a}{r}\right)$ (cf. Eq. 44 in section 3, and Eq. 3.13 in Bürger and Akerman 2011). Note that Bürger and Akerman used m_B for what we call \tilde{m}^* . Invasion of A_1 via \tilde{E}_B requires $m < \min(b, \tilde{m}^*)$. After invasion, the system reaches an asymptotically stable, fully-polymorphic equilibrium \tilde{E}_+ . There may exist a second fully-polymorphic equilibrium \tilde{E}_- , but this is never stable and does not exist when \tilde{E}_B is unstable. It is therefore of limited interest to us. Bürger and Akerman give the coordinates of these equilibria in their Eq. (3.15).

For a polymorphic continent ($0 < q_c < 1$), we find the frequency \hat{q}_B of B_1 at the marginal one-locus migration–selection equilibrium \tilde{E}_B as

$$\hat{q}_B = \frac{b - m + \sqrt{R_3}}{2b}, \quad (89)$$

with $R_3 = (b - m)^2 + 4bm q_c > 0$ as previously encountered in Eq. (59) in section 4. Equilibrium \tilde{E}_B always exists and is always asymptotically stable under its one-locus dynamics (File S7). To know when a weakly beneficial mutation at locus A can invade, we investigate the stability properties of \tilde{E}_B under the two-locus dynamics. The Jacobian matrix evaluated at $\tilde{E}_B = (p = 0, q = \hat{q}_B, D = 0)$ is

$$\mathbf{J}_{\tilde{E}_B} = \begin{pmatrix} a - m & 0 & b \\ 0 & -\sqrt{R_3} & a \\ m(b - m - 2bq_c + \sqrt{R_3})/(2b) & 0 & a - r - \sqrt{R_3} \end{pmatrix} \quad (90)$$

and its leading eigenvalue is

$$\tilde{\nu} = \frac{1}{2} \left[2a - m - r - \sqrt{R_3} + \sqrt{b^2 - r(2m - r - 2\sqrt{R_3})} \right] \quad (91)$$

(cf. Eq. 60). Equilibrium \tilde{E}_B is unstable if and only if $\tilde{\nu} > 0$. To obtain explicit conditions, we determine values of r and q_c at which \tilde{E}_B is not hyperbolic (i.e. $\tilde{\nu} = 0$) and may therefore enter or leave the state space, or change its stability. Equilibrium \tilde{E}_B is not hyperbolic if the recombination rate is equal to

$$\tilde{r}^{**} = \frac{2a^2 - 2a(m + \sqrt{R_3}) + m[m - b(1 - 2q_c) + \sqrt{R_3}]}{2(a - m)} \quad (92)$$

(File S7). As a function of m , \tilde{r}^{**} has a pole at $m = a$, and $\tilde{r}^{**} = 0$ if $m = a(a + b)/(a + bq_c)$. This holds for $a < b$, which is one of our general assumptions. We conclude that \tilde{E}_B is unstable and A_1 can invade whenever $r < \tilde{r}_B$, where

$$\tilde{r}_B = \begin{cases} \infty & \text{if } 0 \leq m \leq a, \\ \tilde{r}^{**} & \text{if } m > a. \end{cases} \quad (93)$$

Figure S7 shows the division of the (m, r) -parameter space into areas where \tilde{E}_B is asymptotically stable (blue) and unstable (orange), respectively.

By solving $\tilde{\nu} = 0$ for q_c , we obtain two critical continental frequencies of B_1 at which \tilde{E}_B is not hyperbolic. These are given by

$$\tilde{q}_{c\pm}^{**} = \frac{1}{2} + \frac{(a - m)(a + r)}{bm} \pm \frac{(2a - m)\sqrt{R_4}}{2bm}, \quad (94)$$

where $R_4 = 4r(a - m) + b^2$. We first investigate the properties of \tilde{q}_{\pm}^{**} as a function of the migration rate m . A combination of algebra and graphical exploration given in File S7 suggests that the following cases must be distinguished:

Case 1 $2a \leq b$ and $(r \leq a$ or $b - a \leq r)$. Then \tilde{E}_B is unstable if $q_c < \tilde{q}_{c,B}$, with $\tilde{q}_{c,B}$ defined as

$$\tilde{q}_{c,B} = \begin{cases} \infty & \text{if } m < a, \\ \tilde{q}_{c+}^{**} & \text{if } a \leq m < a + b - r, \\ 0 & \text{if } a + b - r \leq m. \end{cases} \quad (95)$$

Case 2 $(2a < b$ and $a < r < b - a)$ or $(2a > b$ and $b - a < r < a)$. Then \tilde{E}_B is unstable if $q_c < \tilde{q}_{c,B}$, with $\tilde{q}_{c,B}$ defined as

$$\tilde{q}_{c,B} = \begin{cases} \infty & \text{if } m < a, \\ \tilde{q}_{c+}^{**} & \text{if } a \leq m < a(b - a + r)/r, \\ 0 & \text{if } a(b - a + r)/r \leq m. \end{cases} \quad (96)$$

Case 3 $2a > b$ and $2r > b$ and $a \leq r \Leftrightarrow 2a > b$ and $a \leq r$. We distinguish four subcases:

- 3a** $m < a$. Then \tilde{E}_B is always unstable.
- 3b** $a \leq m \leq a(b - a + r)/r$. Then \tilde{E}_B is unstable if $q_c < \tilde{q}_{c+}^{**}$.
- 3c** $a(b - a + r)/r < m < a + b^2/(4r)$. Then \tilde{E}_B is unstable if $\tilde{q}_{c-}^{**} < q_c < \tilde{q}_{c+}^{**}$.
- 3d** $a + b^2/(4r) \leq m$. Then \tilde{E}_B is asymptotically stable.

Case 4 $2a > b$ and $2r > b$ and $a > r \Leftrightarrow 2r > b$ and $a > r$. We distinguish four subcases:

- 4a** $m < a$. Then \tilde{E}_B is always unstable.
- 4b** $a \leq m \leq a + b - r$. Then \tilde{E}_B is unstable if $q_c < \tilde{q}_{c+}^{**}$.
- 4c** $a + b - r < m < a + b^2/(4r)$. Then \tilde{E}_B is unstable if $\tilde{q}_{c-}^{**} < q_c < \tilde{q}_{c+}^{**}$.
- 4d** $a + b^2/(4r) \leq m$. Then \tilde{E}_B is asymptotically stable.

Figure S8 shows the partition of the (m, q_c) -parameter space into areas where \tilde{E}_B is asymptotically stable (blue) and unstable (orange), respectively. There are parameter combinations such that \tilde{E}_B is asymptotically stable for very low and for high values of q_c , but unstable for intermediate q_c (Figures S8B and S8C). This effect is weak and constrained to a small proportion of the parameter space (q_c small).

Alternatively, we assess the properties of \tilde{q}_{\pm}^{**} as a function of the recombination rate r . Graphical exploration (File S7) suggests the following, provided that $a < \min(m, b)$ holds. If recombination is weak, i.e. $r < a(b - a)/(m - a) = \tilde{r}^*$,

then \tilde{E}_B is unstable if $q_c < \tilde{q}_{c+}^{**}$. If recombination is intermediate, i.e. $\tilde{r}^* < r < b^2/[4(m-a)]$, then \tilde{E}_B is unstable if $\tilde{q}_{c-}^{**} < q_c < \tilde{q}_{c+}^{**}$. Last, if recombination is strong, i.e. $r \geq b^2/[4(m-a)]$, then \tilde{E}_B is asymptotically stable. Note that \tilde{r}^* was previously encountered in Eq. (56) in the context of the branching process. Figure S9 shows the division of the (r, q_c) -parameter space into areas where \tilde{E}_B is asymptotically stable (blue) and unstable (orange), respectively. As just shown, there are parameter combinations such that \tilde{E}_B is asymptotically stable for very low and for high values of q_c , but unstable for intermediate q_c (Figures S9A–S9C).

In principle, analogous conditions for asymptotic stability of \tilde{E}_B under the two-locus dynamics could be obtained in terms of a critical migration rate m^{**} at which \tilde{E}_B is not hyperbolic ($\tilde{\nu} = 0$). However, we were not able to derive informative explicit conditions (see File S7 for a graphical exploration).

So far, we have described the conditions for instability of the marginal one-locus migration–selection equilibrium \tilde{E}_B under the two-locus dynamics, both for a monomorphic ($q_c = 0$) and a polymorphic ($0 < q_c < 1$) continent. In both cases, there is no other stable equilibrium on the boundary for $0 < m < 1$. As mentioned above, for the case of a monomorphic continent, the coordinates of the fully-polymorphic equilibria can be found (Bürger and Akerman 2011) and asymptotic stability proved (Bank *et al.* 2012). For a polymorphic continent, simple explicit expressions are not available, but we could show analytically that at most three candidates for a fully-polymorphic equilibrium exist. Numerical and graphical explorations suggest that if \tilde{E}_B is unstable, at most one of these candidates is an admissible equilibrium, and it is asymptotically stable (see File S7 for details). Figures S7–S9 therefore directly tell us when A_1 can be established if introduced near \tilde{E}_B (orange areas).

In the following section, we will derive a diffusion approximation of sojourn and absorption times under the assumption of quasi-linkage equilibrium (QLE), i.e. for $r \gg \max(m, b)$. Therefore, we briefly discuss the properties under the QLE assumption of the fully-polymorphic, asymptotically stable, equilibria mentioned in the previous paragraphs. For a monomorphic continent, \tilde{E}_+ is approximated to first order in $1/r$ by

$$\hat{p}_+ = \frac{bm + ar - m(m+r)}{ar} = 1 - \frac{m}{a} + \frac{m}{r} \frac{(b-m)}{a}, \quad (97a)$$

$$\hat{q}_+ = \frac{am + br - m(m+r)}{br} = 1 - \frac{m}{b} + \frac{m}{r} \frac{(a-m)}{b}, \quad (97b)$$

$$\hat{D}_+ = \frac{(a-m)(b-m)m}{abr} = \frac{m}{r} \left(1 - \frac{m}{a}\right) \left(1 - \frac{m}{b}\right), \quad (97c)$$

(cf. Eq. 4.3 in Bürger and Akerman 2011). As $r \rightarrow \infty$, Eq. (97) converges to the case of no linkage, where $\hat{p}_+ = 1 - m/a$, $\hat{q}_+ = 1 - m/b$, and $\hat{D}_+ = 0$. Turning to the case of a polymorphic continent, we recall from above that there is at most one admissible fully-polymorphic equilibrium. To first order in $1/r$, its coordinates are

$$\hat{p}_+ = \frac{2ar + m(b - 2bq_c - m - 2r + \sqrt{R_3})}{2ar} = 1 - \frac{bmq_c}{ar} + \frac{m(b-m)}{2ar} + \frac{m\sqrt{R_3}}{2ar} - \frac{m}{a}, \quad (98a)$$

$$\hat{q}_+ = \frac{1}{2} - \frac{am(2bq_c - b + m - \sqrt{R_3})}{2br\sqrt{R_3}} + \frac{m(m+r)(m - \sqrt{R_3})}{2br\sqrt{R_3}} + \frac{b}{2\sqrt{R_3}} + \frac{m(2q_c - 1)(m + 2r)}{2r\sqrt{R_3}}, \quad (98b)$$

$$\hat{D}_+ = \frac{m(a-m)[b(1-2q_c) - m + \sqrt{R_3}]}{2abr}. \quad (98c)$$

Setting $q_c = 0$ and recalling that $m < \tilde{m}^* = a(1 + \frac{b-a}{r})$ must hold for invasion in this case (section 3), it is easy to verify that Eq. (98) coincides with Eq. (97). This is why we call the equilibrium in Eq. (98) \tilde{E}_{+QLE} . Graphical exploration in File S7 confirms that \tilde{E}_{+QLE} is asymptotically stable whenever it exists under the QLE regime.

Finally, we ask when \tilde{E}_{+QLE} exists in the admissible state space. We note that \hat{p}_{+QLE} is a strictly decreasing function of the recombination rate r , independently of the migration rate m . In contrast, \hat{q}_{+QLE} is a strictly decreasing function of r if and only if $m \leq a$, which is of limited interest, because A_1 can then be established in any case. We denote by $r_{\hat{p}_{+QLE}}^0$ and $r_{\hat{p}_{+QLE}}^1$ the recombination rates at which \hat{p}_{+QLE} equals 0 and 1, respectively. Analogously, we use $r_{\hat{q}_{+QLE}}^0$ and $r_{\hat{q}_{+QLE}}^1$ for the recombination rates at which \hat{q}_{+QLE} equals 0 and 1, respectively. These critical recombination rates are found to

be

$$r_{\hat{p}_{+QLE}^0} = m \frac{m - b(1 - 2q_c) - \sqrt{R_3}}{2(a - m)}, \quad (99a)$$

$$r_{\hat{p}_{+QLE}^1} = \frac{1}{2} \left(b - m - 2bq_c + \sqrt{R_3} \right), \quad (99b)$$

and

$$r_{\hat{q}_{+QLE}^0} = (m - a) \frac{b + m - \sqrt{R_3}}{2\sqrt{R_3}}, \quad (100a)$$

$$r_{\hat{q}_{+QLE}^1} = (a - m) \frac{b - m + \sqrt{R_3}}{2\sqrt{R_3}}. \quad (100b)$$

As shown in File S7, if $m < a$, \tilde{E}_{+QLE} exists in the admissible state space if and only if $r > \max\left(r_{\hat{p}_{+QLE}^1}, r_{\hat{q}_{+QLE}^1}\right)$. If $m \geq a$, \tilde{E}_{+QLE} exists in the admissible state space if and only if $\max\left(r_{\hat{p}_{+QLE}^1}, r_{\hat{q}_{+QLE}^1}\right) < r < r_{\hat{p}_{+QLE}^0}$. At a first glance, it may seem surprising to obtain an upper limit on r . However, as is easily verified, $r_{\hat{p}_{+QLE}^0}$ is also the critical value at which \tilde{E}_{+QLE} coincides with the QLE approximation of \tilde{E}_B , which becomes asymptotically stable. Thus, with looser linkage, allele A_1 is lost.

7 Diffusion approximation to sojourn and absorption times assuming quasi-linkage equilibrium

Although some two-locus diffusion theory has been developed (Ewens 2004; Ethier and Nagylaki 1989, 1988, 1980), explicit calculation of quantities of interest, such as absorption probabilities or times, seems difficult. Substantial progress can be made, though, by assuming that recombination is much stronger compared to selection (and migration). Then, linkage disequilibrium decays on a fast time scale, whereas allele frequencies evolve on a slow time scale under quasi-linkage equilibrium (QLE) (Kimura 1965; Nagylaki *et al.* 1999; Kirkpatrick *et al.* 2002). Here, we employ the QLE assumption to approximate the expected amount of time the focal allele A_1 spends in a certain range of allele frequencies (the sojourn times), as well as the expected time to extinction (the mean absorption time). We do so in detail for a monomorphic continent ($q_c = 0$) first. For a polymorphic continent ($0 < q_c < 1$), we will only give a brief outline and refer to File S7 for details. Throughout, we closely follow Ewens (2004) in our application of diffusion theory.

We start from the continuous-time dynamics of the allele frequencies (p, q) and the linkage disequilibrium (D) in Eq. (87), setting $q_c = 0$ for a monomorphic continent. Given that recombination is strong compared to selection and migration, D will be close to an equilibrium, so that $\dot{D} = dD/dt \approx 0$ may be assumed. Moreover, we assume that the frequency of the beneficial background allele B_1 is not affected by establishment of A_1 . Specifically, $q = \hat{q}_B$ constant, where $\hat{q}_B = 1 - m/b$ is the frequency of B_1 at the one-locus migration–selection equilibrium in continuous time (Eq. 88). Equation (87) is therefore approximated by

$$\dot{p} = \frac{dp}{dt} = ap(1 - p) - mp + bD, \quad (101a)$$

$$\dot{q} = \frac{dq}{dt} = 0, \quad (101b)$$

$$\dot{D} = \frac{dD}{dt} = [a(1 - 2p) + b(1 - 2q)]D + m(pq - D) - rD = 0. \quad (101c)$$

Solving Eq. (101c) for D , plugging the solution into Eq. (101a) and setting $q = \hat{q}_B$, we obtain a single differential equation in p :

$$\dot{p} = ap(1 - p) - mp + \frac{m(b - m)}{b - m - a(1 - 2p) + r} p. \quad (102)$$

In the limit of $r \rightarrow \infty$, we recover the one-locus migration–selection dynamics for the continent–island model, $\dot{p} = ap(1 - p) - mp$.

We now consider the diffusion process obtained from the Wright–Fisher model (Fisher 1930; Wright 1931). More precisely, we measure time in units of $2N_e$ generations, where N_e is the effective population size, and use T for time on the

diffusion scale. Further, we introduce the scaled selection coefficients $\alpha = 2N_e a$ and $\beta = 2N_e b$, the scaled recombination rate $\rho = 2N_e r$, and the scaled migration rate $\mu = 2N_e m$. Equation (102) yields the infinitesimal mean

$$M(p) = \alpha p(1-p) - \mu p + \frac{\mu(\beta - \mu)}{\beta - \mu - \alpha(1-2p) + \rho} p$$

(cf. Eq. 5 in the main text). It expresses the mean change in p per unit of time on the diffusion scale. The infinitesimal variance is

$$V(p) = p(1-p) \quad (103)$$

(Karlin and Taylor 1981, p. 159).

Later, we will need the ratio of $M(p)$ to $V(p)$, which is

$$\frac{M(p)}{V(p)} = \alpha - \frac{\mu}{1-p} \left(1 - \frac{\beta - \mu}{\beta - \alpha(1-2p) - \mu + \rho} \right). \quad (104)$$

We define the function $\psi(p)$ according to Eq. (4.16) in Ewens (2004) as

$$\psi(p) := \exp \left[-2 \int_0^p \frac{M(z)}{V(z)} dz \right]. \quad (105)$$

Inserting Eq. (104), we find,

$$\psi(p) = e^{-2\alpha p} (1-p)^{-\frac{2\mu(\alpha+\rho)}{\alpha+\beta-\mu+\rho}} (\beta - \alpha - \mu + \rho)^{\frac{2\mu(\beta-\mu)}{\alpha+\beta-\mu+\rho}} [\beta - (1-2p)\alpha - \mu + \rho]^{\frac{2\mu(\mu-\beta)}{\alpha+\beta-\mu+\rho}}. \quad (106)$$

The derivation assumes that $(\alpha - \beta + \mu - \rho)/(\alpha p) < 0$ holds. Recalling from section 3 that, for instability of the marginal one-locus equilibrium \tilde{E}_B , it is required that $m < \tilde{m}^* = a(1 + \frac{b-a}{r})$ and that then $a < \min(b, r)$, one can show that $(\alpha - \beta + \mu - \rho)/(\alpha p) < 0$ holds indeed (see File S7).

We now turn to the sojourn times as defined in Ewens (2004, pp. 141–144). We denote the initial frequency of the focal mutation A_1 by p_0 and introduce the function $t(p; p_0)$ to describe the sojourn-time density (STD). The interpretation of $t(p; p_0)$ is the following. The integral

$$\int_{p_1}^{p_2} t(p; p_0) dp$$

approximates the mean time in units of $2N_e$ generations allele A_1 spends at a frequency in the interval (p_1, p_2) , conditional on the initial frequency p_0 . According to Eqs. (4.38) and (4.39) in Ewens (2004), we define

$$t(p; p_0) = \begin{cases} t_1(p; p_0) & \text{if } 0 \leq p \leq p_0, \\ t_2(p; p_0) & \text{if } p_0 \leq p \leq 1. \end{cases} \quad (107)$$

To make the assumption of quasi-linkage equilibrium explicit, we will add the subscript QLE to relevant quantities from now on. The densities $t_{i,\text{QLE}}(p; p_0)$ are given by Eq. (7) in the main text, with $\psi(y)$ as in Eq. (105). The integral $\int_0^x \psi(y) dy$ cannot be found explicitly. However, because Eq. (7a) takes the form $t_{1,\text{QLE}}(p; p_0) = 2\psi(y)^{-1}(1-p)^{-1}p^{-1} \int_0^p \psi(y) dy$ and $p^{-1} \int_0^p \psi(y) dy \rightarrow 1$ as $p \rightarrow 0$ (File S7), we approximate $t_{1,\text{QLE}}(p; p_0)$ by

$$\tilde{t}_{1,\text{QLE}}(p; p_0) = \frac{2p}{V(p)\psi(p)} \quad (108)$$

whenever p is small. Recall from Eq. (107) that $t_1(p; p_0)$ is needed only if $0 \leq p \leq p_0$. We are in general interested in a de-novo mutation, i.e. $p_0 = 1/(2N)$, with population size N at least about 100. Hence, $p \leq p_0$ automatically implies that p is small whenever $t_{1,\text{QLE}}(p; p_0)$ is employed. The approximation in Eq. (108) is therefore valid for our purpose.

Similarly, we may multiply $t_{2,\text{QLE}}(p; p_0)$ by p_0 and $1/p_0$ and write

$$t_{2,\text{QLE}}(p; p_0) = 2p_0 \psi(y)^{-1} (1-p)^{-1} p^{-1} p_0^{-1} \int_0^{p_0} \psi(y) dy.$$

Again, $p_0^{-1} \int_0^{p_0} \psi(y) dy \rightarrow 1$ as $p_0 \rightarrow 0$ (File S7). We therefore approximate $t_{2,\text{QLE}}(p; p_0)$ by

$$\tilde{t}_{2,\text{QLE}}(p; p_0) = \frac{2p_0}{V(p)\psi(p)} \quad (109)$$

whenever p_0 is small. In the following, we use a tilde (\sim) to denote the assumption of small p_0 .

The expected time to extinction of allele A_1 in our model is identical to the mean absorption time, because extinction is the only absorbing state. For arbitrary initial frequency p_0 , the approximate mean absorption time under the QLE approximation is obtained from the sojourn-time densities as shown in Eq. (8) of the main text. Assuming small p_0 , this simplifies to

$$\tilde{t}_{\text{QLE}} = \int_0^{p_0} \tilde{t}_{1,\text{QLE}}(p; p_0) dp + \int_{p_0}^1 \tilde{t}_{2,\text{QLE}}(p; p_0) dp. \quad (110)$$

In both cases, the integrals must be computed numerically. As a further approximation for very small p_0 , one may omit the first integral on the right-hand side of Eq. (110), as its contribution becomes negligible when $p_0 \rightarrow 0$.

The predictions for the sojourn-time densities (STDs) and the mean absorption time derived above are accurate if the QLE assumption holds (Figures 7, S11 and S12). However, the analytical expressions for the STDs in Eqs. (108) and (109) are not very informative once we plug in explicit formulae for $V(p)$ and $\psi(p)$ (see File S7). In the following, we will gain more insight by making an additional assumption.

We assume that recombination is much stronger than selection and migration, and expand $M(p)$ from Eq. (5) as a function of ρ^{-1} to first order into a Taylor series. This yields

$$M(p) \approx M_{\rho \gg 0}(p) = \alpha p(1-p) - \mu p + \frac{\mu(\beta - \mu)}{\rho} p$$

and hence Eq. (16) in the main text. The infinitesimal variance $V(p)$ from Eq. (103) remains unchanged, but the ratio of $M(p)$ to $V(p)$ simplifies to

$$\frac{M_{\rho \gg 0}(p)}{V(p)} = \alpha - \frac{\mu}{1-\rho} \left(1 - \frac{\beta - \mu}{\rho}\right). \quad (111)$$

Insertion into Eq. (105), integration and some algebra yields

$$\psi_{\rho \gg 0}(p) = e^{-2\alpha p} (1-p)^{-\frac{2\mu(\mu-\beta+\rho)}{\rho}}. \quad (112)$$

The sojourn-time density (STD) is then given by

$$t_{1,\text{QLE},\rho \gg 0}(p; p_0) = \frac{2}{V(p)\psi_{\rho \gg 0}(p)} \int_0^p \psi_{\rho \gg 0}(y) dy, \quad (113a)$$

$$t_{2,\text{QLE},\rho \gg 0}(p; p_0) = \frac{2}{V(p)\psi_{\rho \gg 0}(p)} \int_0^{p_0} \psi_{\rho \gg 0}(y) dy. \quad (113b)$$

As before, $x^{-1} \int_0^x \psi_{\rho \gg 0}(p) dp \rightarrow 1$ as $x \rightarrow 0$. Arguments analogous to those leading to Eqs. (108) and (109) show that, for a small initial frequency p_0 , the STD is approximated by

$$\begin{aligned} \tilde{t}_{1,\text{QLE},\rho \gg 0}(p; p_0) &= \frac{2p}{V(p)\psi_{\rho \gg 0}(p)} = 2e^{2p\alpha} (1-p)^{\frac{2\mu(\mu-\beta+\rho)}{\rho}-1}, \\ \tilde{t}_{2,\text{QLE},\rho \gg 0}(p; p_0) &= \frac{2p_0}{V(p)\psi_{\rho \gg 0}(p)} = 2p_0 e^{2p\alpha} p^{-1} (1-p)^{\frac{2\mu(\mu-\beta+\rho)}{\rho}-1} \end{aligned}$$

(cf. Eq. 17 of the main text). For details, we refer to File S7. The mean absorption time is again obtained as

$$\bar{t}_{\text{QLE},\rho \gg 0} = \int_0^{p_0} t_{1,\text{QLE},\rho \gg 0}(p; p_0) dp + \int_{p_0}^1 t_{2,\text{QLE},\rho \gg 0}(p; p_0) dp \quad (114)$$

using the STD in Eq. (113) for arbitrary initial frequency p_0 , or as

$$\tilde{t}_{\text{QLE},\rho \gg 0} = \int_0^{p_0} \tilde{t}_{1,\text{QLE},\rho \gg 0}(p; p_0) dp + \int_{p_0}^1 \tilde{t}_{2,\text{QLE},\rho \gg 0}(p; p_0) dp \quad (115)$$

using the STD in Eq. (17) for small p_0 . Figure 5 compares the various approximations to the STD derived under the QLE assumption for a monomorphic continent (q_c). It also includes a comparison to the STD for a one-locus model (OLM), which is specified by

$$\begin{aligned} \tilde{t}_{1,\text{OLM}}(p; p_0) &= 2e^{2p\alpha} (1-p)^{2\mu-1} && \text{if } 0 \leq p \leq p_0, \\ \tilde{t}_{2,\text{OLM}}(p; p_0) &= 2p_0 e^{2p\alpha} p^{-1} (1-p)^{2\mu-1} && \text{if } p_0 \leq p \leq 1 \end{aligned}$$

for small p_0 (cf. Eq. 15 in the main text).

A comparison of the STD given in Eq. (17) for two loci with large ρ and small p_0 to the corresponding one-locus STD in Eq. (15) is interesting. The difference is that μ in the one-locus model is replaced by $\mu(\mu - \beta + \rho)/\rho$ to obtain the formulae for the two-locus model. Hence, for strong recombination, we may define an effective scaled migration rate

$$\mu_e = \mu \frac{\mu + \rho - \beta}{\rho} = \mu - \frac{\beta\mu}{\rho} + \frac{\mu^2}{\rho} \approx \mu \left(1 - \frac{\beta}{\rho}\right),$$

where the approximation holds for $\mu \ll \min(\beta, \rho)$. The interpretation is that μ_e denotes the scaled migration rate in a one-locus migration–selection model for which allele A_1 has the same sojourn-time properties as if it arose in a two-locus model with scaled migration rate μ and linkage to a previously established polymorphism that decays at a scaled recombination rate ρ . Transforming back from the diffusion to the natural scale, we obtain the invasion-effective migration rates m_e and \tilde{m}_e given in Eqs. (19) and (20) of the main text, respectively (see also Figure S18A).

We now turn to the case of a polymorphic continent ($0 < q_c < 1$). Derivations are analogous to those shown above for the monomorphic continent, but more cumbersome. We therefore give only a rough summary here and refer to File S7 for details.

The mean change in p per unit of time on the diffusion scale and under the assumption of quasi-linkage equilibrium (QLE) is

$$M(p) := \frac{dp}{dT} = \alpha p(1-p) - \mu p - \frac{\mu(\beta - \mu - 2\beta q_c + \sqrt{R_5})}{2[\alpha(1-2p) - \rho - \sqrt{R_5}]} p, \quad (116)$$

where $R_5 = (\beta - \mu)^2 + 4\beta\mu q_c > 0$.

Equation (116) can be used to numerically compute the sojourn-time densities (STDs) and the mean absorption time analogous to Eqs. (7) and (8) (see File S7). To obtain informative analytical results for the STDs, however, it is necessary to assume that recombination is strong compared to selection and migration, i.e. $\rho \gg \min(b, m)$. Then, the infinitesimal mean is approximated by

$$M(p) \approx M_{\rho \gg 0}(p) = \alpha p(1-p) - \mu p + \frac{\mu(\beta - \mu - 2\beta q_c + \sqrt{R_5})}{2\rho} p \quad (117)$$

The infinitesimal variance is the same as for a monomorphic continent, $V(p) = p(1-p)$. Inserting $M_{\rho \gg 0}(p)$ from Eq. (117) and $V(p)$ into the definition of $\psi(p)$ in Eq. (105), we obtain

$$\psi_{\rho \gg 0}(p) = e^{-2\alpha p} (1-p)^{\frac{\mu(\beta - \mu - 2\beta q_c + \sqrt{R_5})}{\rho}}. \quad (118)$$

The STDs $t_{1, \text{QLE}, \rho \gg 0}(p; p_0)$ and $t_{2, \text{QLE}, \rho \gg 0}(p; p_0)$ are found by insertion of $\psi_{\rho \gg 0}(p)$ from Eq. (118) into Eq. (113). Exploiting the fact that $x^{-1} \int_0^x \psi_{\rho \gg 0}(p) dp$ converges to 1 as x approaches 0, the STDs can be approximated by

$$\tilde{t}_{1, \text{QLE}, \rho \gg 0}(p; p_0) = 2e^{2p\alpha} (1-p)^{\frac{\mu(\mu - \beta + 2\beta q_c + 2\rho - \sqrt{R_5})}{\rho} - 1}, \quad (119a)$$

$$\tilde{t}_{2, \text{QLE}, \rho \gg 0}(p; p_0) = 2p_0 e^{2p\alpha} p^{-1} (1-p)^{\frac{\mu(\mu - \beta + 2\beta q_c + 2\rho - \sqrt{R_5})}{\rho} - 1} \quad (119b)$$

This approximation is valid if the initial frequency p_0 is small and ρ is large. The mean absorption time for arbitrary p_0 is found according to Eq. (114). For small p_0 , it is given by Eq. (115), with $\tilde{t}_{i, \text{QLE}, \rho \gg 0}(p; p_0)$ from Eq. (119).

8 Effective migration rate at a neutral site linked to two migration–selection polymorphisms

We derive the effective migration rate experienced by a neutral locus (C) linked to two loci (A and B) that are maintained polymorphic at migration–selection balance. Locus C has two alleles C_1 and C_2 , which are assumed to segregate at constant frequencies n_c and $1 - n_c$ on the continent. The frequency of C_1 on the island at time t is denoted by $n(t)$. Loci A and B are as above, with alleles A_1 and B_1 segregating at frequencies p and q on the island, respectively. Without loss of generality, we assume that A is located to the left of B on the chromosome. We denote by r_{XY} the recombination rate between loci X and Y , where $r_{XY} = r_{YX}$. Because we consider a continuous-time model here, we may assume that the recombination rate increases additively with distance on the chromosome. For simplicity, we restrict the analysis to the case of a monomorphic continent, i.e. alleles A_2 and B_2 are fixed on the continent.

Following Bürger and Akerman (2011), we define the effective migration rate as the asymptotic rate of convergence of $n(t)$ to the fully-polymorphic three-locus equilibrium. This rate of convergence is defined by the leading eigenvalue λ_N of the Jacobian of the system that describes the evolution of the frequency of C_1 and the linkage disequilibria associated with locus C. Specifically, we define the effective migration rate as $m_e = -\lambda_N$ (cf. Kobayashi *et al.* 2008).

We start by assuming that the neutral locus is located between the two selected ones (configuration A–C–B). We denote by $D_{AB} = D$, D_{AC} and D_{CB} the linkage disequilibria between the indicated loci, and by $D_{ACB} = y_1 - pqn - pD_{CB} - qD_{AC} - nD_{AB}$ the three-way linkage disequilibrium, where y_1 is the frequency of gamete $A_1C_1B_1$. The changes due to selection, migration and recombination in p , q , and D_{AB} are given by Eq. (87) of this text, with r replaced by r_{AB} . The frequency of C_1 evolves according to

$$\dot{n} = m(n_c - n) + aD_{AC} + bD_{CB} \quad (120)$$

and the differential equations for the linkage disequilibria associated with locus C are

$$\dot{D}_{AC} = a(1 - 2p)D_{AC} + bD_{ACB} - mD_{AC} - mp(n_c - n) - r_{AC}D_{AC}, \quad (121a)$$

$$\dot{D}_{CB} = aD_{ACB} + b(1 - 2q)D_{CB} - mD_{CB} - mq(n_c - n) - r_{CB}D_{CB}, \quad (121b)$$

$$\begin{aligned} \dot{D}_{ACB} = & [a(1 - 2p) + b(1 - 2q)] D_{ACB} - 2(aD_{AC} + bD_{CB})D_{AB} + m(pD_{CB} + qD_{AC} - D_{ACB}) \\ & + m(pq - D_{AB})(n_c - n) - r_{AB}D_{ACB} \end{aligned} \quad (121c)$$

(we use \dot{x} for the differential of x with respect to time, dx/dt). We refer to File S8 for the derivation. Recall that $r_{AB} = r_{AC} + r_{CB}$. This system has an asymptotically stable equilibrium such that the selected loci are at the equilibrium \tilde{E}_+ (Eq. 3.15 in Bürger and Akerman 2011), and $n = n_c$ and $D_{AC} = D_{CB} = D_{ACB} = 0$ hold. The Jacobian at this equilibrium has the block structure

$$\mathbf{J} = \begin{pmatrix} \mathbf{J}_S & 0 \\ 0 & \mathbf{J}_N \end{pmatrix},$$

where \mathbf{J}_S is the Jacobian approximating convergence of (p, q, D_{AB}) to \tilde{E}_+ , and \mathbf{J}_N is the Jacobian approximating convergence of $(n, D_{AC}, D_{CB}, D_{ACB})$ to $(n_c, 0, 0, 0)$. In the limit of weak migration, i.e. $m \ll (a, b, r)$, the latter is given by

$$\mathbf{J}_N^{\text{ACB}} = \begin{pmatrix} -m & a & b & 0 \\ m & -a - r_{AC} + \frac{m(a-b+r_{AB})}{a+b+r_{AB}} & 0 & b \\ m & 0 & -b - r_{CB} + \frac{m(b-a+r_{AB})}{a+b+r_{AB}} & a \\ -m & \frac{m(b-a+r_{AB})}{a+b+r_{AB}} & \frac{m(a-b+r_{AB})}{a+b+r_{AB}} & -a - b - r_{AB} + \frac{m(a+b+3r_{AB})}{a+b+r_{AB}} \end{pmatrix}. \quad (122)$$

As shown previously (Bürger and Akerman 2011), to first order in m , the leading eigenvalue of $\mathbf{J}_N^{\text{ACB}}$ is given by

$$\lambda_N^{\text{ACB}} = m \frac{r_{AC}r_{CB}}{(a + r_{AC})(b + r_{CB})}, \quad (123)$$

and hence the approximation of the effective migration rate in Eq. (22b) in the main text is obtained (see File S8 for details). We note that Eqs. (120), (121) and (122) correct errors in Eqs. (4.25), (4.26) and (4.28) of Bürger and Akerman (2011), respectively. The main results by Bürger and Akerman (2011) were not affected, though.

If the neutral locus is located to the right of the two selected ones (configuration A–B–C), Eqs. (120) and (121) remain the same (recall that $r_{XY} = r_{YX}$ and in this case $r_{AC} = r_{AB} + r_{BC}$). In Eq. (87c), r must be replaced by r_{AC} . Then, the Jacobian $\mathbf{J}_N^{\text{ABC}}$ approximating convergence of $(n, D_{AC}, D_{BC} = D_{CB}, D_{ABC} = D_{ACB})$ to $(n_c, 0, 0, 0)$ in the limit of weak migration is equal to $\mathbf{J}_N^{\text{ACB}}$ with the last entry of the last row replaced by $-a - b - r_{AC} + \frac{m(a+b+3r_{AB})}{a+b+r_{AB}}$. To first order in m , the leading eigenvalue of $\mathbf{J}_N^{\text{ABC}}$ is

$$\lambda_N^{\text{ABC}} = m \frac{r_{BC}(b + r_{AC})}{(b + r_{BC})(a + b + r_{AC})}, \quad (124)$$

and hence Eq. (22c) in the main text. Details are given in File S8.

Last, the leading eigenvalue for configuration C–A–B follows directly by symmetry,

$$\lambda_N^{\text{CAB}} = m \frac{r_{CA}(a + r_{CB})}{(a + r_{CA})(a + b + r_{CB})}, \quad (125)$$

and hence Eq. (22a) in the main text.

Recall that the Jacobian matrices $\mathbf{J}_N^{\text{ACB}}$ and $\mathbf{J}_N^{\text{ABC}}$ hold under the assumption of weak migration. In File S8, we derive analogous matrices under the assumption of weak recombination, i.e. $r \ll (a, b, m)$. These are too complicated to be shown here, but importantly, to first order in m , their leading eigenvalues are identical to Eqs. (123) and (124), respectively. By symmetry, this also applies to the configuration C–A–B. Therefore, the approximate effective migration rates in Eq. (22) are valid also for tight linkage between the neutral locus and the selected loci.

To test the robustness of our results against violation of the assumption of weak migration, we numerically computed exact effective migration rates. In most cases, the deviation is very small; compare dashed to solid curves in Figures 8 and S19, and dots to curves in Figure S20.

Literature Cited

- Athreya, K. B., 1992 Rates of decay for the survival probability of a mutant gene. *J. Math. Biol.* 30: 577--581, 10.1007/BF00948892.
- Athreya, K. B., 1993 Rates of decay for the survival probability of a mutant gene II The multitype case. *J. Math. Biol.* 32: 45--53, 10.1007/BF00160373.
- Bank, C., R. Bürger, and J. Hermisson, 2012 The limits to parapatric speciation: Dobzhansky--Muller incompatibilities in a continent--island model. *Genetics* 191: 845--863.
- Bengtsson, B. O., 1985 The flow of genes through a genetic barrier. In *Evolution -- Essays in honour of John Maynard Smith*, edited by P. J. Greenwood, P. Harvey, and M. Slatkin, volume 1, chapter 3, pp. 31--42, Cambridge University Press, New York, NY.
- Bürger, R., 2000 *The Mathematical Theory of Selection, Recombination, and Mutation*. Wiley Series in Mathematical and Computational Biology, John Wiley & Sons Ltd, Chichester, UK.
- Bürger, R. and A. Akerman, 2011 The effects of linkage and gene flow on local adaptation: A two-locus continent--island model. *Theor. Popul. Biol.* 80: 272--288.
- Eshel, I., 1984 On the survival probability of a slightly advantageous mutant gene in a multitype population: A multidimensional branching process model. *J. Math. Biol.* 19: 201--209, 10.1007/BF00277746.
- Ethier, S. N. and T. Nagylaki, 1980 Diffusion approximations of Markov-chains with 2 time scales with applications to population-genetics. *Adv. Appl. Probab.* 12: 14--49.
- Ethier, S. N. and T. Nagylaki, 1988 Diffusion approximations of Markov-chains with 2 time scales with applications to population-genetics 2. *Adv. Appl. Probab.* 20: 525--545.
- Ethier, S. N. and T. Nagylaki, 1989 Diffusion approximations of the two-locus Wright--Fisher model. *J. Math. Biol.* 27: 17--28, 10.1007/BF00276078.
- Ewens, W. J., 1967 The probability of fixation of a mutant: The two-locus case. *Evolution* 21: 532--540.
- Ewens, W. J., 1968 Some applications of multiple-type branching processes in population genetics. *J. R. Stat. Soc. Ser. B Stat. Methodol.* 30: 164--175.
- Ewens, W. J., 2004 *Mathematical Population Genetics*. Springer-Verlag, New York, NY, second edition.
- Fisher, R. A., 1922 On the dominance ratio. *Proc. R. Soc. Edinb.* 42: 321--341.
- Fisher, R. A., 1930 *The Genetical Theory of Natural Selection*. Oxford University Press Inc, New York, NY.
- Haccou, P., J. Peter, and V. A. Vatutin, 2005 *Branching Processes. Variation, Growth, and Extinction of Populations*, volume 5 of *Cambridge Studies in Adaptive Dynamics*. Cambridge University Press, New York, NY.
- Haldane, J. B. S., 1927 A mathematical theory of natural and artificial selection, part v: Selection and mutation. *Math. Proc. Camb. Phil. Soc.* 23: 838--844.
- Harris, T. E., 1963 *The Theory of Branching Processes*, volume 119 of *Die Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin, DE, first edition.
- Hoppe, F. M., 1992 The survival probability of a mutant in a multidimensional population. *J. Math. Biol.* 30: 567--575, 10.1007/BF00948891.
- Karlin, S. and H. M. Taylor, 1981 *A second course in stochastic processes*, volume 2. Academic Press, San Diego, CA.
- Kimura, M., 1962 On the probability of fixation of mutant genes in a population. *Genetics* 47: 713--719.
- Kimura, M., 1965 Attainment of quasi linkage equilibrium when gene frequencies are changing by natural selection. *Genetics* 52: 875--890.

- Kirkpatrick, M., T. Johnson, and N. Barton, 2002 General models of multilocus evolution. *Genetics* 161: 1727--1750.
- Kobayashi, Y., P. Hammerstein, and A. Telschow, 2008 The neutral effective migration rate in a mainland--island context. *Theor. Popul. Biol.* 74: 84--92.
- Lewontin, R. C. and K. Kojima, 1960 The evolutionary dynamics of complex polymorphisms. *Evolution* 14: 458--472.
- Nagylaki, T., 1992 *Introduction to Theoretical Population Genetics*, volume 21 of *Biomathematics*. Springer-Verlag, Berlin, DE.
- Nagylaki, T., J. Hofbauer, and P. Brunovský, 1999 Convergence of multilocus systems under weak epistasis or weak selection. *J. Math. Biol.* 38: 103--133.
- Petry, D., 1983 The effect on neutral gene flow of selection at a linked locus. *Theor. Popul. Biol.* 23: 300--313.
- Wakeley, J., 2009 *Coalescent theory -- An introduction*. Roberts & Company Publishers, Greenwood Village, CO.
- Wright, S., 1931 Evolution in Mendelian populations. *Genetics* 16: 97--159.
- Yeaman, S., 2013 Genomic rearrangements and the evolution of clusters of locally adaptive loci. *Proc. Natl. Acad. Sci. U.S.A.* 110: 1743--1751.