The Effect of Linkage on Establishment and Survival of Locally Beneficial Mutations

Simon Aeschbacher*,†,1 and Reinhard Bürger*
*Department of Mathematics, University of Vienna, 1090 Vienna, Austria and †Department of Evolution and Ecology, University of California, Davis, California 95616

ABSTRACT We study invasion and survival of weakly beneficial mutations arising in linkage to an established migration–selection polymorphism. Our focus is on a continent–island model of migration, with selection at two biallelic loci for adaptation to the island environment. Combining branching and diffusion processes, we provide the theoretical basis for understanding the evolution of islands of divergence, the genetic architecture of locally adaptive traits, and the importance of so-called "divergence hitchhiking" relative to other mechanisms, such as "genomic hitchhiking", chromosomal inversions, or translocations. We derive approximations to the invasion probability and the extinction time of a *de novo* mutation. Interestingly, the invasion probability is maximized at a nonzero recombination rate if the focal mutation is sufficiently beneficial. If a proportion of migrants carries a beneficial background allele, the mutation is less likely to become established. Linked selection may increase the survival time by several orders of magnitude. By altering the timescale of stochastic loss, it can therefore affect the dynamics at the focal site to an extent that is of evolutionary importance, especially in small populations. We derive an effective migration rate experienced by the weakly beneficial mutation, which accounts for the reduction in gene flow imposed by linked selection. Using the concept of the effective migration rate, we also quantify the long-term effects on neutral variation embedded in a genome with arbitrarily many sites under selection. Patterns of neutral diversity change qualitatively and quantitatively as the position of the neutral locus is moved along the chromosome. This will be useful for population-genomic inference. Our results strengthen the emerging view that physically linked selection is biologically relevant if linkage is tight or if selection at the background locus is strong.

ADAPTATION to local environments may generate a selective response at several loci, either because the fitness-related traits are polygenic or because multiple traits are under selection. However, populations adapting to spatially variable environments often experience gene flow that counteracts adaptive divergence. The dynamics of polygenic adaptation is affected by physical linkage among selected genes, and hence by recombination (Barton 1995). Recombination allows contending beneficial mutations to form optimal haplotypes, but it also breaks up existing beneficial associations (Fisher 1930; Muller 1932; Hill and Robertson 1966; Lenormand and Otto 2000). On top of that, finite population size causes random fluctuations of allele frequencies that may lead to fixation or loss. Migration and selection create statistical associations even among physically unlinked loci.

The availability of genome-wide marker and DNA-sequence data has spurred both empirical and theoretical work on the interaction of selection, gene flow, recombination, and genetic drift. Here, we study the stochastic fate of a locally beneficial mutation that arises in linkage to an established migration–selection polymorphism. We also investigate the long-term effect on linked neutral variation of adaptive divergence with gene flow.

Empirical insight on local adaptation with gene flow emerges from studies of genome-wide patterns of genetic differentiation between populations or species. Of particular interest are studies that have either related such patterns to function and fitness (*e.g.*, Nadeau *et al.* 2012, 2013) or detected significant deviations from neutral expectations (*e.g.*, Karlsen *et al.* 2013), thus implying that some of this divergence is adaptive. One main observation is that in some organisms putatively adaptive differentiation (*e.g.*, measured by elevated F_{ST}) is clustered at

doi: 10.1534/genetics.114.163477
Manuscript received December 10, 2013; accepted for publication February 27, 2014; published Early Online March 7, 2014.
Available freely online through the author-supported open access option.
Supporting information is available online at http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.114.163477/-/DC1.
[1]Corresponding author: Department of Evolution and Ecology, University of California, Davis Room 2320 Storer Hall, 1 Shields Ave., Davis, CA 95616.
E-mail: saeschbacher@mac.com

certain positions in the genome (Nosil and Feder 2012 and references therein). This has led to the metaphor of genomic islands of divergence or speciation (Turner *et al.* 2005). Other studies did not identify such islands, however (see Strasburg *et al.* 2012, for a review of plant studies).

These findings have stimulated theoretical interest in mechanistic explanations for the presence or absence of genomic islands. Polygenic local adaptation depends crucially on the genetic architecture of the selected traits, but, in the long run, local adaptation may also lead to the evolution of this architecture. Here, we define genetic architecture as the number of, and physical distances between, loci contributing to local adaptation, and the distribution of selection coefficients of established mutations.

Using simulations, Yeaman and Whitlock (2011) have shown that mutations contributing to adaptive divergence in a quantitative trait may physically aggregate in the presence of gene flow. In addition, these authors reported cases where the distribution of mutational effects changed from many divergent loci with mutations of small effect to few loci with mutations of large effect. Such clustered architectures reduce the likelihood of recombination breaking up locally beneficial haplotypes and incorporating maladaptive immigrant alleles. This provides a potential explanation for genomic islands of divergence. However, it is difficult to explain the variability in the size of empirically observed islands of divergence, especially the existence of very long ones. Complementary mechanisms have been proposed, such as the accumulation of adaptive mutations in regions of strongly reduced recombination (*e.g.*, at chromosomal inversions; Guerrero *et al.* 2012; McGaugh and Noor 2012) or the assembly of adaptive mutations by large-scale chromosomal rearrangements (*e.g.*, transpositions of loci under selection; Yeaman 2013).

It is well established that spatially divergent selection can cause a reduction in the effective migration rate (Charlesworth *et al.* 1997; Kobayashi *et al.* 2008; Feder and Nosil 2010). This is because migrants tend to carry combinations of alleles that are maladapted, such that selection against a locally deleterious allele at one locus also eliminates incoming alleles at other loci. The effective migration rate can be reduced either by physical linkage to a gene under selection or by statistical associations among physically unlinked loci. Depending on whether physical or statistical linkage is involved, the process of linkage-mediated differentiation with gene flow has, by some authors, been called "divergence hitchhiking" or "genomic hitchhiking," respectively (Nosil and Feder 2012; Feder *et al.* 2012; Via 2012). These two processes are not mutually exclusive, and, recently, interest in assessing their relative importance in view of explaining observed patterns of divergence has been growing. If not by inversions or translocations, detectable islands of divergence are expected as a consequence of so-called divergence hitchhiking, but not of genomic hitchhiking. This is because physical linkage reduces the effective migration rate only locally (*i.e.*, in the neighborhood of selected sites), whereas statistical linkage may reduce it across the whole genome. Yet, if many loci are under selection, it is unlikely that all of them are physically unlinked (Barton 1983), and so the two sources of linkage disequilibrium may be confounded.

A number of recent studies have focused on the invasion probability of neutral or locally beneficial *de novo* mutations in the presence of divergently selected loci in the background (Feder and Nosil 2010; Yeaman and Otto 2011; Feder *et al.* 2012; Flaxman *et al.* 2013; Yeaman 2013). They showed that linkage elevates invasion probabilities only over very short map distances, implying that physical linkage provides an insufficient explanation for both the abundance and size of islands of divergence. Such conclusions hinge on assumptions about the distribution of effects of beneficial mutations, the distribution of recombination rates along the genome, and the actual level of gene flow. These studies were based on time-consuming simulations (Feder and Nosil 2010; Feder *et al.* 2012; Flaxman *et al.* 2013; Yeaman 2013) or heuristic *ad hoc* aproximations (Yeaman and Otto 2011; Yeaman 2013) that provide limited understanding. Although crucial, invasion probabilities on their own might not suffice to gauge the importance of physical linkage in creating observed patterns of divergence. In finite populations, the time to extinction of adaptive mutations is also relevant. It codetermines the potential of synergistic interactions among segregating adaptive alleles.

Here, we fill a gap in existing theory to understand the role of physical linkage in creating observed patterns of divergence with gene flow. First, we provide numerical and analytical approximations to the invasion probability of locally beneficial mutations arising in linkage to an existing migration–selection polymorphism. This sheds light on the ambiguous role of recombination and allows for an approximation to the distribution of fitness effects of successfully invading mutations. Second, we obtain a diffusion approximation to the proportion of time the beneficial mutation segregates in various frequency ranges (the sojourn-time density) and the expected time to its extinction (the mean absorption time). From these, we derive an invasion-effective migration rate experienced by the focal mutation. Third, we extend existing approximations of the effective migration rate at a neutral site linked to two migration–selection polymorphisms (Bürger and Akerman 2011) to an arbitrary number of such polymorphisms. These formulae are used to predict the long-term footprint of polygenic local adaptation on linked neutral variation. We extend some of our analysis to the case of standing, rather than *de novo*, adaptive variation at the background locus.

Methods

Model

We consider a discrete-time version of a model with migration and selection at two biallelic loci (Bürger and Akerman 2011). Individuals are monoecious diploids and reproduce sexually. Soft selection occurs at the diploid stage and then a proportion m ($0 < m < 1$) of the island population is replaced by immigrants from the continent (Haldane 1930). Migration is followed by

gametogenesis, recombination with probability r ($0 \le r \le 0.5$), and random union of gametes including population regulation. Generations do not overlap.

We denote the two loci by A and B and their alleles by A_1 and A_2, and B_1 and B_2, respectively. Locus A is taken as the focal locus and locus B as background locus. The four haplotypes 1, 2, 3, and 4 are A_1B_1, A_1B_2, A_2B_1, and A_2B_2. On the island, the frequencies of A_1 and B_1 are p and q, and the linkage disequilibrium is denoted by D (see Supporting Information, File S1, sect.1, for details).

Biological scenario

We assume that the population on the continent is fixed for alleles A_2 and B_2. The island population is of size N and initially fixed for A_2 at locus A. At locus B, the locally beneficial allele B_1 has arisen some time ago and is segregating at migration–selection balance. Then, a weakly beneficial mutation occurs at locus A, resulting in a single copy of A_1 on the island. Its fate is jointly determined by direct selection on locus A, linkage to the selected locus B, migration, and random genetic drift. If A_1 occurs on the beneficial background (B_1), the fittest haplotype is formed and invasion is likely unless recombination transfers A_1 to the deleterious background (B_2). If A_1 initially occurs on the B_2 background, a suboptimal haplotype is formed (A_1B_2; Equation 1 below) and A_1 is doomed to extinction unless it recombines onto the B_1 background early on. These two scenarios occur proportionally to the marginal equilibrium frequency $\hat{q}_B$ of B_1. Overall, recombination is therefore expected to play an ambiguous role.

Two aspects of genetic drift are of interest: random fluctuations when A_1 is initially rare and random sampling of alleles between successive generations. In the first part of the article, we focus exclusively on the random fluctuations when A_1 is rare, assuming that N is so large that the dynamics is almost deterministic after an initial stochastic phase. In the second part, we allow for small to moderate population size N on the island. The long-term invasion properties of A_1 are expected to differ in the two cases (Ewens 2004, pp. 167–171). With N sufficiently large and parameter combinations for which a fully polymorphic internal equilibrium exists under deterministic dynamics, the fate of A_1 is decided very early on. If it survives the initial phase of stochastic loss, it will reach the (quasi-) deterministic equilibrium frequency and stay in the population for a very long time (Petry 1983). This is what we call invasion, or establishment. Extinction will finally occur, because migration introduces A_2, but not A_1. Yet, extinction occurs on a timescale much longer than is of interest for this article. For small or moderate N, however, genetic drift will cause extinction of A_1 on a much shorter timescale, even for moderately strong selection. In this case, stochasticity must be taken into account throughout, and interest shifts to the expected time A_1 spends in a certain range of allele frequencies (sojourn time) and the expected time to extinction (absorption time).

As an extension of this basic scenario, we allow the background locus to be polymorphic on the continent. Allele B_1 is assumed to segregate at a constant frequency q_c. This reflects, for instance, a polymorphism maintained at drift–mutation or mutation–selection balance. It could also apply to the case where the continent is a metapopulation or receives migrants from other populations. A proportion q_c of haplotypes carried by immigrants to the focal island will then be A_2B_1, and a proportion $1 - q_c$ will be A_2B_2.

Fitness and evolutionary dynamics

We define the relative fitness of a genotype as its expected relative contribution to the gamete pool from which the next generation of zygotes is formed. We use w_{ij} for the relative fitness of the genotype composed of haplotypes i and j ($i, j \in \{1, 2, 3, 4\}$). Ignoring parental and position effects in heterozygotes, we distinguish nine genotypes. We then have $w_{ij} = w_{ji}$ for all $i \neq j$ and $w_{23} = w_{14}$.

The extent to which analytical results can be obtained for general fitnesses is limited (Ewens 1967; Karlin and McGregor 1968). Unless otherwise stated, we therefore assume absence of dominance and epistasis, *i.e.*, allelic effects combine additively within and between loci. The matrix of relative genotype fitnesses w_{ij} (Equation 27 in File S1) may then be written as

$$\begin{array}{c} \\ A_1A_1 \\ A_1A_2 \\ A_2A_2 \end{array} \begin{array}{c} \begin{array}{ccc} B_1B_1 & B_1B_2 & B_2B_2 \end{array} \\ \begin{pmatrix} 1+a+b & 1+a & 1+a-b \\ 1+b & 1 & 1-b \\ 1-a+b & 1-a & 1-a-b \end{pmatrix} \end{array}, \tag{1}$$

where a and b are the selective advantages on the island of alleles A_1 and B_1 relative to A_2 and B_2, respectively. To enforce positive fitnesses, we require that $0 < a, b < 1$, and $a + b < 1$. We assume that selection in favor of A_1 is weaker than selection in favor of B_1 ($a < b$). Otherwise, A_1 could be maintained in a sufficiently large island population independently of B_1, whenever B_1 is not swamped by gene flow (Haldane 1930). As our focus is on the effect of linkage on establishment of A_1, this case is not of interest.

The deterministic dynamics of the haplotype frequencies are given by the recursion equations in File S1, Equation 28 (see also File S2). A crucial property of these dynamics is the following. Whenever a marginal one-locus migration–selection equilibrium E_B exists such that the background locus B is polymorphic and locus A is fixed for allele A_2, this equilibrium is asymptotically stable. After occurrence of A_1, E_B may become unstable, in which case a fully polymorphic (internal) equilibrium emerges and is asymptotically stable, independently of whether the continent is monomorphic ($q_c = 0$) or polymorphic ($0 < q_c < 1$) at the background locus. Therefore, in the deterministic model, invasion of A_1 via E_B is always followed by an asymptotic approach toward an internal equilibrium (see File S1, sect. 3 and 6).

Casting our model into a stochastic framework is difficult in general. By focusing on the initial phase after occurrence of A_1, the four-dimensional system in Equation 28 can be simplified to a two-dimensional system (Equation 29 in File S1).

This allows for a branching-process approach as described in the following.

Two-type branching process

As shown in File S1, sect. 2, for rare A_1, we need to follow only the frequencies of haplotypes A_1B_1 and A_1B_2. This corresponds to A_1 initially occurring on the B_1 or B_2 background, respectively, and holds as long as A_1 is present in heterozygotes only. Moreover, it is assumed that allele B_1 is maintained constant at the marginal one-locus migration–selection equilibrium E_B of the dynamics in Equation 28. At this equilibrium, the frequency of B_1 is

$$\hat{q}_B = \frac{b - m(1-a)}{b(1+m)} \tag{2}$$

for a monomorphic continent (see File S1, sect. 3, for details, and Equation 39 for a polymorphic continent).

To model the initial stochastic phase after occurrence of A_1 for large N, we employed a two-type branching process in discrete time (Harris 1963). We refer to haplotypes A_1B_1 and A_1B_2 as types 1 and 2, respectively. They are assumed to propagate independently and contribute offspring to the next generation according to type-specific distributions. We assume that the number of j-type offspring produced by an i-type parent is Poisson-distributed with parameter λ_{ij} ($i \in \{1, 2\}$). Because of independent offspring distributions, the probability-generating function (pgf) for the number of offspring of any type produced by an i-type parent is $f_i(s_1, s_2) = \prod_{j=1}^{2} f_{ij}(s_j)$, where $f_{ij}(s_j) = e^{-\lambda_{ij}(1-s_j)}$ for i, $j \in \{1, 2\}$ (File S1, sect. 4). The λ_{ij} depend on fitness, migration, and recombination and are derived from the deterministic model (Equation 33 in File S1). The matrix $\mathbf{L} = (\lambda_{ij})$, $i, j \in \{1, 2\}$, is called the mean matrix. Allele A_1 has a strictly positive invasion probability if $\nu > 1$, where ν is the leading eigenvalue of $\mathbf{L}$. The branching process is called supercritical in this case.

We denote the probability of invasion of A_1 conditional on initial occurrence on background B_1 (B_2) by π_1 (π_2), and the corresponding probability of extinction by Q_1 (Q_2). The latter are found as the smallest positive solution of

$$f_1(s_1, s_2) = s_1 \tag{3a}$$

$$f_2(s_1, s_2) = s_2 \tag{3b}$$

such that $s_i < 1$ ($i \in \{1, 2\}$). Then, $\pi_1 = 1 - Q_1$ and $\pi_2 = 1 - Q_2$ (Haccou *et al.* 2005). The overall invasion probability of A_1 is given as the weighted average of the two conditional probabilities,

$$\bar{\pi} = \hat{q}_B \pi_1 + (1 - \hat{q}_B)\pi_2 \tag{4}$$

(*cf.* Ewens 1967, 1968; Kojima and Schaffer 1967). File S1, sect. 4, gives further details and explicit expressions for additive fitnesses.

Diffusion approximation

The branching process described above models the initial phase of stochastic loss and applies as long as the focal mutant A_1 is rare. To study long-term survival of A_1, we employ a diffusion approximation. We start from a continuous-time version of the deterministic dynamics in Equation 28, assuming additive fitnesses as in Equation 1. For our purpose, it is convenient to express the dynamics in terms of the allele frequencies (p, q) and the linkage disequilibrium (D), as given in Equation 87 in File S1. Changing to the diffusion scale, we measure time in units of $2N_e$ generations, where N_e is the effective population size.

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 $\mu = 2N_e m$. As it is difficult to obtain analytical results for the general two-locus diffusion problem (Ethier and Nagylaki 1980, 1988, 1989; Ewens 2004), we assume that recombination is much stronger than selection and migration. Then, linkage disequilibrium decays on a faster timescale, whereas allele frequencies evolve on a slower one under quasi-linkage equilibrium (QLE) (Kimura 1965; Nagylaki *et al.* 1999; Kirkpatrick *et al.* 2002). In addition, we assume that the frequency of the beneficial background allele B_1 is not affected by establishment of A_1 and stays constant at $q = \hat{\tilde{q}}_B$. Here, $\hat{\tilde{q}}_B$ is the frequency of B_1 at the one-locus migration–selection equilibrium when time is continuous, $\tilde{E}_B$ (Equations 88 and 89 in File S1). As further shown in File S1, sect. 6, these assumptions lead to a one-dimensional diffusion process. The expected change in p per unit time is

$$M(p) = \alpha p(1-p) - \mu p + \frac{\mu(\beta - \mu)}{\beta - \mu - \alpha(1-2p) + \rho}\, p \tag{5}$$

if the continent is monomorphic. The first term is due to direct selection on the focal locus, the second reflects migration, and the third represents the interaction of all forces.

For a polymorphic continent, $M(p)$ is given by Equation 116 in File S1, and the interaction term includes the continental frequency q_c of B_1. In both cases, assuming random genetic drift according to the Wright–Fisher model, the expected squared change in p per unit time is $V(p) = p(1 - p)$ (Ewens 2004). We call $M(p)$ the *infinitesimal mean* and $V(p)$ the *infinitesimal variance* (Karlin and Taylor 1981, p. 159).

Let the initial frequency of A_1 be p_0. We introduce the sojourn-time density (STD) $t(p; p_0)$ such that the integral $\int_{p_1}^{p_2} t(p; p_0)dp$ approximates the expected time A_1 segregates at a frequency between p_1 and p_2 before extinction, conditional on p_0. Following Ewens (2004, Equations 4.38 and 4.39), we define

$$t_{QLE}(p; p_0) = \begin{cases} t_{1,QLE}(p; p_0) & \text{if} \quad 0 \le p \le p_0 \\ t_{2,QLE}(p; p_0) & \text{if} \quad p_0 \le p \le 1 \end{cases}, \tag{6}$$

with subscript QLE for the assumption of quasi-linkage equilibrium. The densities $t_{i,QLE}(p; p_0)$ are

$$t_{1,QLE}(p; p_0) = \frac{2}{V(p)\psi(p)} \int_0^p \psi(y)dy, \tag{7a}$$

$$t_{2,\mathrm{QLE}}(p;p_0) = \frac{2}{V(p)\psi(p)} \int_0^{p_0} \psi(y)dy, \tag{7b}$$

where $\psi(p) = \exp\left[-2\int_0^p (M(z)/V(z))dz\right]$. Integration over p yields the expected time to extinction,

$$\bar{t}_{\mathrm{QLE}} = \int_0^{p_0} t_{1,\mathrm{QLE}}(p;p_0)dp + \int_{p_0}^1 t_{2,\mathrm{QLE}}(p;p_0)dp, \tag{8}$$

or the *mean absorption time*, in units of $2N_e$ generations. A detailed exposition is given in File S1, sect. 7. See File S10 for Mathematica Notebooks.

Simulations

We conducted two types of simulation, one for the branching-process regime and another for a finite island population with Wright–Fisher random drift. In the branching-process regime, we simulated the absolute frequency of the two types of interest (A_1B_1 and A_1B_2) over time. Each run was initiated with a single individuum and its type determined according to Equation 2. Every generation, each individual produced a Poisson-distributed number of offspring of either type (see above). We performed $n = 10^6$ runs. Each run was terminated if either the mutant population went extinct (no invasion), reached a size of $500/(2a)$ (invasion), or survived for more than 5×10^4 generations (invasion). We estimated the invasion probability from the proportion $\hat{\pi}$ of runs that resulted in invasion, and its standard error as $\sqrt{\hat{\pi}(1-\hat{\pi})/n}$.

In the Wright–Fisher-type simulations, each generation was initiated by zygotes built from gametes of the previous generation. Viability selection, migration, and gamete production including recombination (meiosis) were implemented according to the deterministic recursions for the haplotype frequencies in Equation 28. Genetic drift was simulated through the formation of N_e (rather, the nearest integer) zygotes for the next generation by random union of pairs of gametes. Gametes were sampled with replacement from the gamete pool in which haplotypes were represented according to the deterministic recursions. Replicates were terminated if either allele A_1 went extinct or a maximum of 10^9 generations was reached. Unless otherwise stated, for each parameter combination we performed 1000 runs, each with 1000 replicates. Replicates within a given run provided one estimate of the mean absorption time, and runs provided a distribution of these estimates. Java source code and JAR files are available in File S11.

Results

Establishment in a large island population

We first describe the invasion properties of the beneficial mutation A_1, which arises in linkage to a migration–selection polymorphism at the background locus B. Because we assume that the island population is large, random genetic drift is ignored after A_1 has overcome the initial phase during which stochastic loss is likely. Numerical and analytical results were obtained from the two-type branching process and confirmed by simulations (see *Methods*). We turn to the case of small to moderate population size further below. (See lines 511, 515, and 516.)

Conditions for the invasion of A_1: Mutation A_1 has a strictly positive invasion probability whenever

$$rw_{14}\left[\frac{\overline{w}}{1-m} - \hat{q}_B w_1 - \left(1-\hat{q}_B\right)w_2\right] < -\left(\frac{\overline{w}}{1-m} - w_1\right)\left(\frac{\overline{w}}{1-m} - w_2\right) \tag{9}$$

(File S1, sect. 4, and File S3). Here, w_i is the marginal fitness of type i and $\overline{w}$ the mean fitness of the resident population (see Equations 30 and 31 in File S1). Setting $m = 0$, we recover the invasion condition obtained by Ewens (1967) for a panmictic population in which allele B_1 is maintained at frequency $\hat{q}_B$ by overdominant selection. All remaining results in this subsection assume additive fitnesses as in Equation 1.

For a monomorphic continent ($q_c = 0$), it follows from Equation 9 that A_1 can invade only if $m < m^*$, where

$$m^* = \frac{a(b-a+r)}{(a-r)(a-b)+r(1-a)}. \tag{10}$$

In terms of the recombination rate, A_1 can invade only if $r < r^*$, where

$$r^* = \begin{cases} \dfrac{1}{2} & \text{if } m \le \dfrac{a}{1-2a+b} \\ \dfrac{a(a-b)(1+m)}{a(1+2m)-(1+b)m} & \text{otherwise} \end{cases} \tag{11}$$

(see File S1, sects. 3 and 4, File S2, and Figure S1 for details).

For a polymorphic continent ($0 < q_c < 1$), A_1 has a strictly positive invasion probability whenever r and q_c are below the critical values r^* and q_c^* derived in File S1, sect. 3 (*cf.* File S4, Figure S2). In this case, we could not determine the critical migration rate m^* explicitly. For an analysis in continuous time, see File S1, sect. 6, and Figure S7, Figure S8, and Figure S9.

Invasion probability: We obtained exact conditional invasion probabilities, π_1 and π_2, of A_1 by numerical solution of the pair of transcendental equations in Equation 3. From these, we calculated the average invasion probability $\bar{\pi}$ according to Equation 4, with $\hat{q}_B$ as in Equation 2 (Figure 1 and Figure S3 for a monomorphic continent). Haldane (1927) approximated the invasion probability without migration and linked selection by $2a$, *i.e.*, twice the selective advantage of A_1 in a heterozygote. With linked selection, the map distance over which $\bar{\pi}$ is above, say, 10% of $2a$ can be large despite gene flow (Figure 1, A and B).

Analytical approximations were obtained by assuming that the branching process is slightly supercritical, *i.e.*, that the leading eigenvalue of the mean matrix **L** is of the form

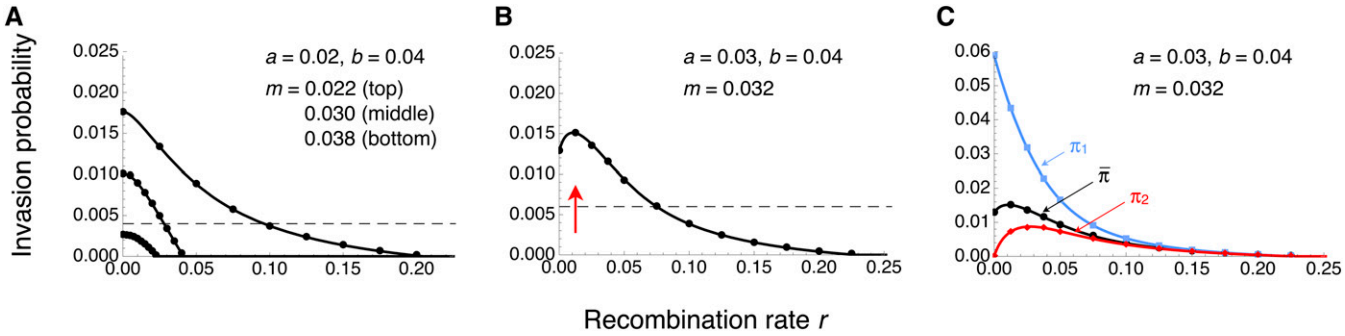

Figure 1 Invasion probability of A_1 as a function of the recombination rate for a monomorphic continent. (A and B) Weighted average invasion probabilitiy $\bar{\pi}$ across the two genetic backgrounds B_1 and B_2 (Equations 2 and 4). For comparison, horizontal dashed lines give 10% of Haldane's (1927) approximation $2a$, valid for $m = 0$ and $r = 0.5$. (B) The optimal recombination rate r_{opt}, defined as the recombination rate at which $\bar{\pi}$ is maximized (red arrow), is nonzero. (C) Same as in B, but in addition to the weighted average, the invasion probabilities of A_1 conditional on initial occurrence on the B_1 or B_2 background are shown in blue or red, respectively. Note the difference in the scale of the vertical axis between B and C. In A–C, curves show exact numerical solutions to the branching process. Dots represent the point estimates across 10^6 simulations under the branching-process assumptions (see *Methods*). Error bars span twice the standard error on each side of the point estimates, but are too short to be visible.

$\nu = 1 + \xi$, with $\xi > 0$ small. We denote these approximations by $\pi_1(\xi)$ and $\pi_2(\xi)$. The expressions are long (File S5) and not shown here. For weak evolutionary forces (a, b, m, $r \ll 1$), $\pi_1(\xi)$ and $\pi_2(\xi)$ can be approximated by

$$\tilde{\pi}_1(\xi) = \max\left[0, \frac{a\left(b + r + \sqrt{R_2}\right) - 2mr}{\sqrt{R_2}}\right], \tag{12a}$$

$$\tilde{\pi}_2(\xi) = \max\left[0, \frac{b^2 - 2mr + b\left(r - \sqrt{R_2}\right) - a\left(b - r - \sqrt{R_2}\right)}{\sqrt{R_2}}\right], \tag{12b}$$

where $R_2 = b^2 + 2br - 4mr + r^2$ and $\xi \approx \frac{1}{2}\left(2a - b - r + \sqrt{R_2}\right)$. The approximate average invasion probability $\bar{\tilde{\pi}}(\xi)$ is obtained according to Equation 4, with $\hat{q}_B$ as in Equation 2. Formally, these approximations are justified if $\xi \ll 1$ (File S1, sect. 4). Figure 2 suggests that the assumption of weak evolutionary forces is more crucial than ξ small and that if it is fulfilled, the approximations are very good (compare Figure 2, A–D).

For a polymorphic continent, exact and approximate invasion probabilities are derived in File S3 and File S5 (see also sect. 4 in File S1). The most important, and perhaps surprising, effect is that the average invasion probability decreases with increasing continental frequency q_c of the beneficial background allele B_1 (Figure S4). As a consequence, invasion requires tighter linkage if $q_c > 0$. This is because the resident island population has a higher mean fitness when a proportion $q_c > 0$ of immigrating haplotypes carry the B_1 allele, which makes it harder for A_1 to become established. Competition against fitter residents therefore compromises the increased probability of recombining onto a beneficial background (B_1) when A_1 initially occurs on the deleterious background (B_2). However, a closer look suggests that if A_1 is sufficiently beneficial and recombination sufficiently weak ($r \ll a$), there are cases where the critical migration rate below which A_1 can invade is maximized at an intermediate q_c (Figure S5, right column). In other words, for certain combinations of m and r, the average invasion probability as a function of q_c is maximized at an intermediate (nonzero) value of q_c (Figure S6).

For every combination of selection coefficients (a, b) and recombination rate (r), the mean invasion probability decreases as a function of the migration rate m. This holds for a monomorphic and a polymorphic continent (Figure S3 and Figure S5, respectively). In both cases, migrants carry only allele A_2 and, averaged across genetic backgrounds, higher levels of migration make it harder for A_1 to invade (*cf.* Bürger and Akerman 2011).

Optimal recombination rate: Deterministic analysis showed that A_1 can invade if and only if recombination is sufficiently weak; without epistasis, large r is always detrimental to establishment of A_1 (Bürger and Akerman 2011; File S1, sect. 3). In this respect, stochastic theory is in line with deterministic predictions. However, considering the average invasion probability $\bar{\pi}$ as a function of r, we could distinguish two qualitatively different regimes. In the first one, $\bar{\pi}(r)$ decreases monotonically with increasing r (Figure 1A). In the second one, $\bar{\pi}(r)$ is maximized at an intermediate recombination rate r_{opt} (Figure 1B). A similar dichotomy was previously found for a panmictic population in which the background locus is maintained polymorphic by heterozygote superiority (Ewens 1967) and has recently been reported in the context of migration and selection in simulation studies (Feder and Nosil 2010; Feder *et al.* 2012). As shown in File S1, sect. 5, $r_{opt} > 0$ holds in our model whenever

$$w_1 - w_2 > \overline{w}\,\frac{\pi_1^\circ}{(1-m)\left(1-\pi_1^\circ\right)}, \tag{13}$$

where w_1 (w_2) is the marginal fitness of type 1 (2) and $\overline{w}$ the mean fitness of the resident population (defined in Equations 30 and 31 in File S1). Here, π_1° is the invasion probability of A_1

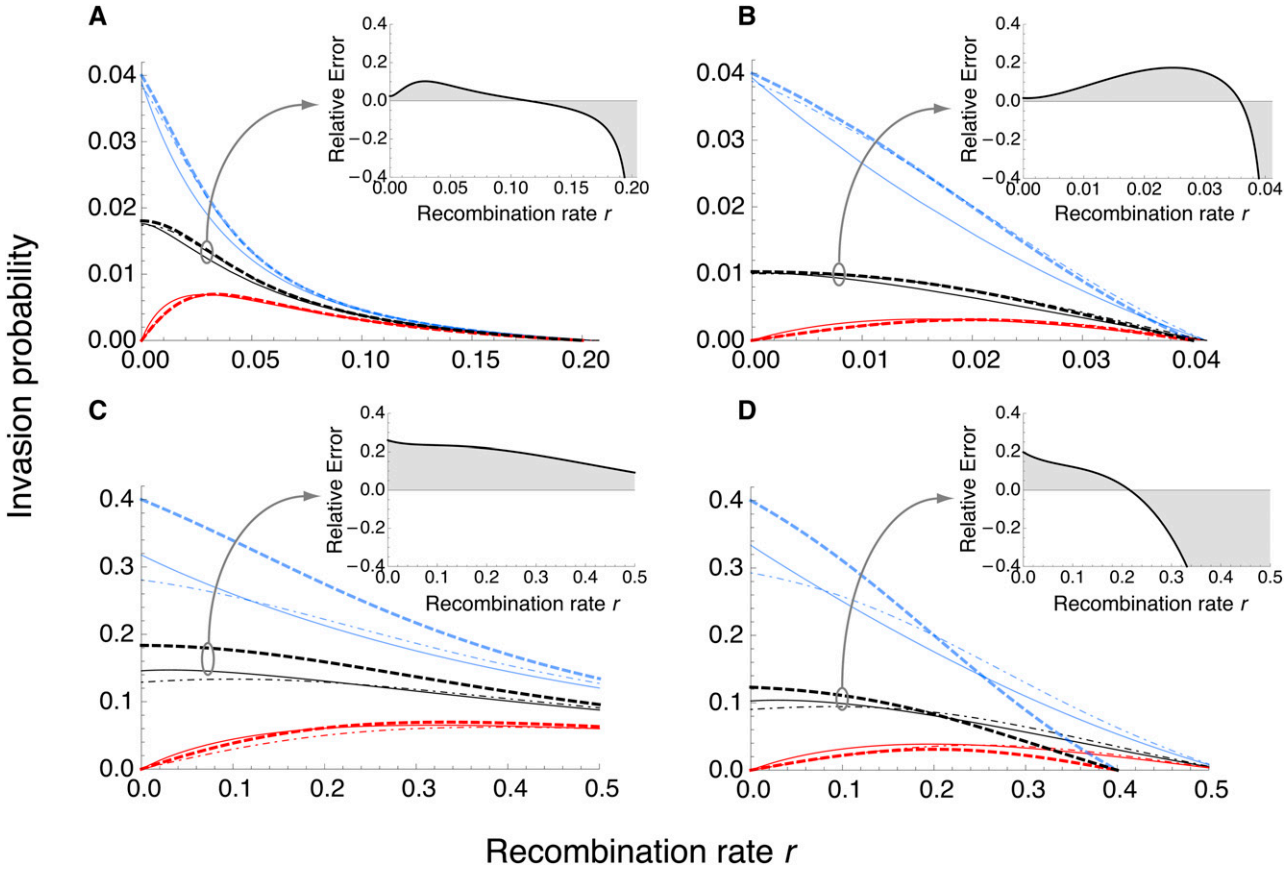

Figure 2 Approximation to the invasion probability of A_1 for a monomorphic continent. Invasion probabilities are shown for A_1 initially occurring on the beneficial background B_1 (blue), on the deleterious background B_2 (red), and as a weighted average across backgrounds (black). Analytical approximations assuming a slightly supercritical branching process (dot-dashed curves) and, in addition, weak evolutionary forces (Equation 12; thick dashed curves) are compared to the exact numerical branching-process solution (solid curves). Inset figures show the error of the analytical approximation $\tilde{\bar{\pi}}(\zeta)$ (thick dashed black curve) relative to $\bar{\pi}$ (solid black curve), $\tilde{\bar{\pi}}(\zeta)/\bar{\pi} - 1$. (A) $a = 0.02$, $b = 0.04$, $m = 0.022$. (B) $a = 0.02$, $b = 0.04$, $m = 0.03$. (C) $a = 0.2$, $b = 0.4$, $m = 0.22$. (D) $a = 0.2$, $b = 0.4$, $m = 0.3$. As expected, the analytical approximations are very good for weak evolutionary forces (top row), but less so for strong forces (bottom row).

conditional on background B_1 and complete linkage ($r = 0$). Setting $m = 0$, we recover Equation 36 of Ewens (1967) for a panmictic population with overdominance at the background locus.

Inequality (13) is very general. In particular, it also holds with epistasis or dominance. However, explicit conclusions require calculation of π_1°, $\overline{w}$, and w_i, which themselves depend on $\hat{q}_B$ and hence on m (*cf.* Equation 2). For mathematical convenience, we resorted to the assumption of additive fitnesses (Equation 1). For a monomorphic continent, $\pi_1^\circ \approx 2a(1+m)/(1+b)$ to first order in a. Moreover, we found that

$$a > a^* \tag{14}$$

is a necessary condition for $r_{\text{opt}} > 0$, where

$$a^* = \frac{1}{2}\left\{1 + b(2+m) - \sqrt{1 + 2b(1+m) + b^2[2 + m(4+m)]}\right\}$$

(File S1, sect. 5). Thus, A_1 must be sufficiently beneficial for $r_{\text{opt}} > 0$ to hold. Figure 3 shows the division of the parameter space where A_1 can invade into two areas where $r_{\text{opt}} = 0$ or $r_{\text{opt}} > 0$ holds.

The two regimes $r_{\text{opt}} = 0$ and $r_{\text{opt}} > 0$ arise from the ambiguous role of recombination. On the one hand, when A_1 initially occurs on the deleterious background (B_2), some recombination is needed to transfer A_1 onto the beneficial background (B_1) and rescue it from extinction. This is reminiscent of Hill and Robertson's (1966) result that recombination improves the efficacy of selection in favor of alleles that are partially linked to other selected sites (Barton 2010). On the other hand, when A_1 initially occurs on the beneficial background, recombination is always deleterious, as it breaks up the fittest haplotype on the island (A_1B_1). This interpretation is confirmed by considering π_1 and π_2 separately as functions of r (Figure 1C). Whereas $\pi_1(r)$ always decreases monotonically with increasing r, $\pi_2(r)$ is always 0 at $r = 0$ (File S1, sect. 5) and then increases to a maximum at an intermediate recombination rate (compare blue to red curve in Figure 1C). As r increases further, $\pi_1(r)$ and $\pi_2(r)$ both approach 0. We recall from Equation 4 that the average invasion probability $\bar{\pi}$ is given by $\hat{q}_B\pi_1 + (1-\hat{q}_B)\pi_2$. Depending on $\hat{q}_B$, either π_1 or π_2 makes a stronger conbribution to $\bar{\pi}$, which then leads to either $r_{\text{opt}} > 0$ or $r_{\text{opt}} = 0$.

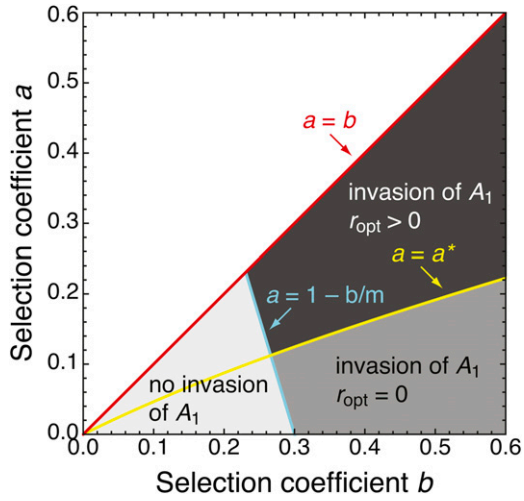

Figure 3 Optimal recombination rate and regions of invasion. The dark shaded area indicates where the optimal recombination rate r is positive ($r_{opt} > 0$; *cf.* Figure 1B). The medium shaded area shows the parameter range for which $r_{opt} = 0$ (cf. Figure 1A). Together, these two areas indicate where A_1 can invade via the marginal one-locus migration–selection equilibrium E_B if r is sufficiently small. The light shaded area shows where E_B does not exist and A_1 cannot invade via E_B. The area above $a = b$ is not of interest, as we focus on mutations that are weakly beneficial compared to selection at the background locus ($a < b$). The critical selection coefficient a^* is given in Equation 14 and the migration rate is $m = 0.3$ (other values of m yield qualitatively similar diagrams). The continent is monomorphic ($q_c = 0$).

A more intuitive interpretation of Equation 14 is as follows. If A_1 conveys a weak advantage on the island ($a < a^*$), it will almost immediately go extinct when it initially arises on background B_2. Recombination has essentially no opportunity of rescuing A_1, even if r is large. Therefore, π_2 contributes little to $\bar{\pi}$. If A_1 is sufficiently beneficial on the island ($a > a^*$), however, it will survive for some time even when arising on the deleterious background. Recombination now has time to rescue A_1 if r is sufficiently different from 0 (but not too large). In this case, π_2 makes an important contribution to $\bar{\pi}$ and leads to $r_{opt} > 0$. For a polymorphic continent, $r_{opt} > 0$ may also hold (File S1, sect. 5). However, in such cases, r_{opt} approaches zero quickly with increasing q_c (File S6 and Figure S4).

Distribution of fitness effects of successful mutations: Using Equation 12 we can address the distribution of fitness effects (DFE) of successfully invading mutations. This distribution depends on the distribution of selection coefficients a of novel mutations (Kimura 1979), which in general is unknown (Orr 1998). In our scenario, the island population is at the marginal one-locus migration–selection equilibrium E_B before the mutation A_1 arises. Unless linkage is very tight, the selection coefficient a must be above a threshold for A_1 to effectively withstand gene flow (this threshold is implicitly defined by Equation 10). Therefore, we assumed that a is drawn from the tail of the underlying distribution, which we took to be exponential (Gillespie 1983, 1984; Orr 2002, 2003; Barrett *et al.* 2006; Eyre-Walker and Keightley 2007) (for alternatives, see Cowperthwaite *et al.* 2005; Barrett *et al.* 2006; Martin and Lenormand 2008). We further assumed that selection is directional with a constant fitness gradient (Equation 1). We restricted the analysis to the case of a monomorphic continent. As expected, linkage to a migration–selection polymorphism shifts the DFE of successfully invading mutations toward smaller effect sizes (Figure 4). Comparison to simulated histograms in Figure 4 suggests that the approximation based on Equation 12 is very accurate.

Survival in a finite island population

We now turn to island populations of small to moderate size N. In this case, genetic drift is strong enough to cause extinction on a relevant timescale even after successful initial establishment. Our focus is on the sojourn-time density and the mean absorption time of the locally beneficial mutation A_1 (see *Methods*). We also derive an approximation to the effective migration rate experienced by A_1.

Sojourn-time density: A general expression for the STD was given in Equation 7. Here, we describe some properties of the exact numerical solution and then discuss analytical approximations (see also File S7). Because A_1 is a *de novo* mutation, it has an initial frequency of $p_0 = 1/(2N)$. For simplicity, we assumed that the effective population size on the island is equal to the actual population size, *i.e.*, $N_e = N$ (this assumption is relaxed later). As $p_0 = 1/(2N)$ is very close to zero in most applications, we used $t_{2,QLE}(p; p_0)$ as a proxy for $t_{QLE}(p; p_0)$ (*cf.* Equation 6).

The STD always has a peak at $p = 0$, because most mutations go extinct after a very short time (Figure 5). However, for parameter combinations favorable to invasion of A_1 (migration weak relative to selection, or selection strong relative to genetic drift), the STD has a second mode at an intermediate allele frequency p. Then, allele A_1 may spend a long time segregating in the island population before extinction. The second mode is usually close to—but slightly greater than—the corresponding deterministic equilibrium frequency (solid black curves in Figure 5, C–F for a monomorphic continent). The peak at this mode becomes shallower as the continental frequency q_c of B_1 increases (Figure S10).

The effect on the STD of linkage is best seen from a comparison to the one-locus model (OLM), for which the STD is given by Ewens (2004) as

$$t_{OLM}(p;p_0) = \begin{cases} t_{1,OLM}(p;p_0) = 2e^{2p\alpha}(1-p)^{2\mu-1} & \text{if } 0 \le p \le p_0, \\ t_{2,OLM}(p;p_0) = 2p_0 e^{2p\alpha} p^{-1}(1-p)^{2\mu-1} & \text{if } p_0 \le p \le 1. \end{cases} \quad (15)$$

If invasion of A_1 is unlikely without linkage, but selection at the background locus is strong, even loose linkage has a large effect and causes a pronounced second mode in the STD (compare orange to black curves in Figure 5C). In cases where A_1 can be established without linkage, the STD of the one-locus model also shows a second mode at an intermediate allele frequency p. Yet, linkage to a background polymorphism leads to a much higher peak, provided that selection at the background locus is

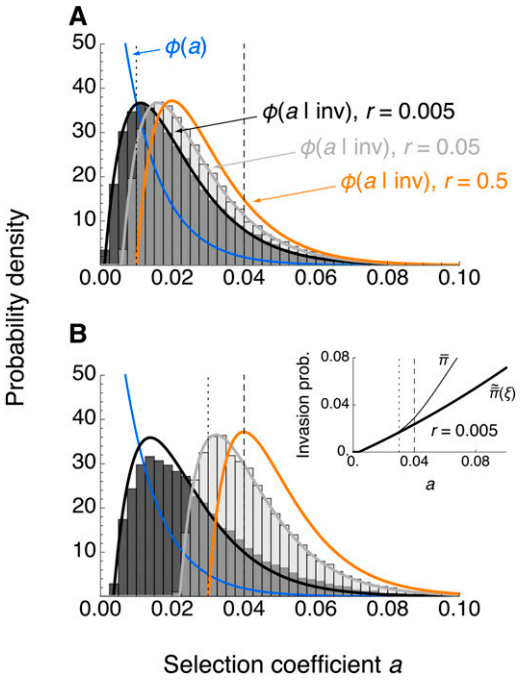

Figure 4 DFE of successfully invading mutations for a monomorphic continent. The DFE of successfully invading mutations was obtained as $\phi(a|\text{inv}) = \phi(\text{inv}|a)\phi(a)/\int_0^\infty \phi(\text{inv}|a)\phi(a)da$, where $\phi(\text{inv}|a) = \bar{\pi}(\xi) = \hat{q}_B\bar{\pi}_1(\xi) + (1-\hat{q}_B)\bar{\pi}_2(\xi)$, with $\hat{q}_B$ and $\bar{\pi}_i(\xi)$ as in Equations 2 and 12, respectively. The mutational input distribution was assumed to be exponential, $\phi(a) = \lambda e^{-a\lambda}$ (blue). Vertical lines denote $a = m$ (dotted) and $a = b$ (dashed). Histograms were obtained from simulations under the branching-process assumptions (intermediate shading indicates where histograms overlap). Each represents 2.5×10^4 realizations in which A_1 successfully invaded (see *Methods*). As a reference, the one-locus model (no linkage) is shown in orange. (A) Relatively weak migration: $b = 0.04$, $m = 0.01$. (B) Migration three times stronger: $b = 0.04$, $m = 0.03$. In A and B, $\lambda = 100$ and $\phi(a|\text{inv})$ is shown for a recombination rate of $r = 0.005$ (black) and $r = 0.05$ (gray). The inset in B shows why the fit is worse for $r = 0.005$: in this case, $\bar{\pi}(\xi)$ underestimates the exact invasion probability $\bar{\pi}$ (Equation 4) for large a.

strong and the recombination rate not too high (Figure 5, D–F). Specifically, comparison of Figure 5E with Figure 5F suggests that the effect of linkage becomes weak if the ratio of the (scaled) recombination rate to the (scaled) selection coefficient at the background locus, ρ/β, becomes much larger than $\sim$10. In other words, for a given selective advantage b of the beneficial background allele, a weakly beneficial mutation will profit from linkage if it occurs within $\sim b \times 10^3$ map units (centimorgans) from the background locus. This assumes that one map unit corresponds to $r = 0.01$.

An analytical approximation of the STD can be obtained under two simplifying assumptions. The first is that the initial frequency p_0 of A_1 is small (p_0 on the order of $1/(2N_e) \ll 1$). The second concerns the infinitesimal mean $M(p)$ of the change in the frequency of A_1: assuming that recombination is much stronger than selection and migration, we may approximate Equation 5 by

$$M_{\rho\gg 0}(p) = \alpha p(1-p) - \mu p + \frac{\mu(\beta-\mu)}{\rho}p \qquad (16)$$

for a monomorphic continent. The STDs in Equation 7 can then be approximated by

$$\tilde{t}_{1,\text{QLE},\rho\gg 0}(p;p_0) = 2e^{2p\alpha}(1-p)^{2\mu(\mu-\beta+\rho)/\rho-1}, \qquad (17a)$$

$$\tilde{t}_{2,\text{QLE},\rho\gg 0}(p;p_0) = 2p_0e^{2p\alpha}p^{-1}(1-p)^{2\mu(\mu-\beta+\rho)/\rho-1}. \qquad (17b)$$

Here, we use $\sim$ to denote the assumption of p_0 small, and a subscript $\rho \gg 0$ for the assumption of $\rho \gg \max(\alpha, \beta, \mu)$. For a polymorphic continent, expressions analogous to Equations 16 and 17 are given in Equations 117 and 119 in File S1.

Better approximations than those in Equations 17 and 119 are obtained by making only one of the two assumptions above. We denote by $\tilde{t}_{1,\text{QLE}}(p;p_0)$ and $\tilde{t}_{2,\text{QLE}}(p;p_0)$ the approximations of the STD in Equation 7 based on the assumption $p_0 \ll 1$ (Equations 108 and 109 in File S1). Alternatively, the approximations obtained from the assumption $\rho \gg \max(\alpha, \beta, \mu)$ in $M(p)$ are called $t_{1,\text{QLE},\rho\gg 0}(p; p_0)$ and $t_{2,\text{QLE},\rho\gg 0}(p; p_0)$ (Equation 113).

In the following, we compare the different approximations to each other and to stochastic simulations. Conditional on $p_0 = 1/(2N)$, the approximation $\tilde{t}_{2,\text{QLE}}(p;p_0)$ (Equations 109) is indeed very close to the exact numerical value $t_{2,\text{QLE}}(p; p_0)$ from Equation 7b. This holds across a wide range of parameter values, as seen from comparing solid to dashed curves in Figure 5 (monomorphic continent) and Figure S10 (polymorphic continent). The accuracy of the approximation $\tilde{t}_{2,\text{QLE},\rho\gg 0}(p;p_0)$ from Equation 17b is rather sensitive to violation of the assumption $\rho \gg \max(\alpha, \beta, \mu)$, however (dotted curves deviate from other black curves in Figure 5, B and C). The same applies to a polymorphic continent, but the deviation becomes smaller as q_c increases from zero (Figure S10A).

Comparison of the diffusion approximation $\tilde{t}_{2,\text{QLE}}(p;p_0)$ to sojourn-time distributions obtained from stochastic simulations shows a very good agreement, except at the boundary $p = 0$. There, the continuous solution of the diffusion approximation is known to provide a suboptimal fit to the discrete distribution (Figure S11 and Figure S12).

Based on the analytical approximations above, we may summarize the effect of weak linkage relative to the one-locus model as follows. For a monomorphic continent, the ratio of $\tilde{t}_{2,\text{QLE},\rho\gg 0}(p;p_0)$ to $t_{2,\text{OLM}}(p; p_0)$ is $\tilde{R} = (1-p)^{-\gamma}$, where $\gamma = 2\mu(\beta-\mu)/\rho$. The exponent γ is a quadratic function of μ and linear in β. For weak migration, $\tilde{R} \approx 1 - 2(\beta\mu/\rho)\ln(1-p)$, suggesting the following rule of thumb. For the focal allele to spend at least the $\tilde{R}$-fold amount of time at frequency p compared to the case without linkage, we require

$$\frac{\beta\mu}{\rho} > \frac{\tilde{R}-1}{-2\ln(1-p)}. \qquad (18)$$

For example, allele A_1 will spend at least twice as much time at frequency $P = 0.5$ (0.8) if $\beta\mu \gtrsim 0.72\rho$ (0.31ρ). Because we

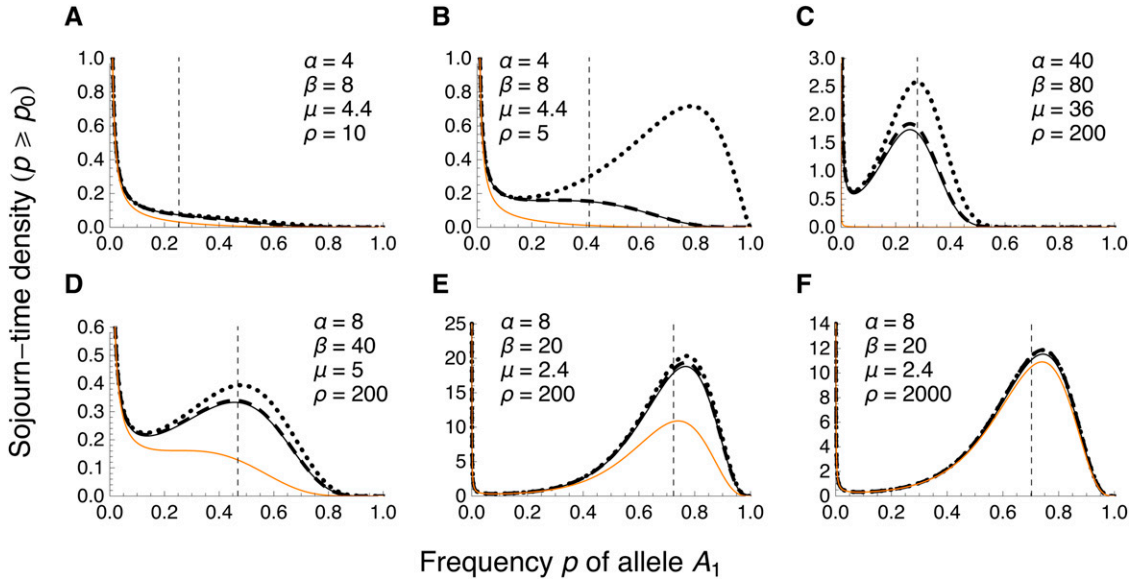

Figure 5 Diffusion approximation to the sojourn-time density of A_1 under quasi-linkage equilibrium for a monomorphic continent. Comparison of the STD $t_{2,\mathrm{QLE}}(p; p_0)$ (thin black; Equation 7b) to the approximation valid for small p_0, $\tilde{t}_{2,\mathrm{QLE}}(p; p_0)$ (dashed black; Equation 109 in File S1), and the one based on the additional assumption of $\rho \gg \max(\alpha, \beta, \mu)$, $\tilde{t}_{2,\mathrm{QLE},\rho\gg 0}(p; p_0)$ (dotted; Equation 17b). The STD for the one-locus model, $\tilde{t}_{2,\mathrm{OLM}}(p; p_0)$, is shown in orange as a reference. Vertical lines give the deterministic frequency $\hat{p}_+$ of A_1 at the fully polymorphic equilibrium (computed in File S7). (A) Weak evolutionary forces relative to genetic drift. (B) As in A, but with half the scaled recombination rate ρ. The assumption of $\rho \gg \max(\alpha, \beta, \mu)$ is violated and hence $\tilde{t}_{2,\mathrm{QLE},\rho\gg 0}(p; p_0)$ is a poor approximation of $t_{2,\mathrm{QLE}}(p; p_0)$. (C) Strong evolutionary forces relative to genetic drift. The STD has a pronounced mode different from $p = 0$, but $\tilde{t}_{2,\mathrm{QLE},\rho\gg 0}(p; p_0)$ overestimates $t_{2,\mathrm{QLE}}(p; p_0)$ considerably. (D) Strong assymmetry in selection coefficients and moderate migration. As in C, the STD has a pronounced mode different from $p = 0$, but $\tilde{t}_{2,\mathrm{QLE},\rho\gg 0}(p; p_0)$ now approximates $t_{2,\mathrm{QLE}}(p; p_0)$ better. (E) Recombination 10 times stronger than selection at locus B. (F) As in E, but with recombination 100 times stronger than selection at locus B. In A–F, $p_0 = 0.005$, which corresponds to an island population of size $N = 100$ and a single initial copy of A_1.

assumed weak migration and QLE, we conducted numerical explorations to check when this rule is conservative, meaning that it does not predict a larger effect of linked selection than is observed in simulations. We found that, first, genetic drift must not dominate, *i.e.*, $1 < \alpha, \beta, \mu, \rho$ holds. Second, migration, selection at the background locus and recombination should roughly satisfy $\mu < \beta/4 < 0.1\rho$. This condition applies only to the validity of Equation 18, which is based on $\tilde{t}_{2,\mathrm{QLE},\rho\gg 0}(p; p_0)$ in Equation 17b. It does not apply to $\tilde{t}_{2,\mathrm{QLE}}(p; p_0)$, which fits simulations very well if ρ is as low as 1.25β (Figure S11D). For related observations in different models, see Slatkin (1975) and Barton (1983).

Mean absorption time: The mean absorption time is obtained by numerical integration of the STD as outlined in *Methods*. Comparison to stochastic simulations shows that the diffusion approximation $\bar{t}_{\mathrm{QLE}}$ from Equation 8 is fairly accurate: the absolute relative error is <15%, provided that the QLE assumption is not violated and migration is not too weak (Figure 6).

Given the approximations to the STD derived above, various degrees of approximation are available for the mean absorption time, too. Their computation is less prone to numerical issues than that of the exact expressions. Extensive numerical computations showed that if $p_0 = 1/(2N)$ and $N_e = N$, the approximations based on the assumption of p_0 small ($\tilde{\bar{t}}_{\mathrm{QLE}}$ and $\tilde{\bar{t}}_{\mathrm{QLE},\rho\gg 0}$ as given in Equations 110 and 115) provide an excellent fit to their more exact counterparts ($\bar{t}_{\mathrm{QLE}}$ and $\bar{t}_{\mathrm{QLE},\rho\gg 0}$ in Equations 8 and 114, respectively). See also Table S2 and S4. Across a wide range of parameter values, the absolute relative error never exceeds 1.8% (Figure S13, A and C). In contrast, the approximation based on the assumption of $\rho \gg 0$, $\bar{t}_{\mathrm{QLE},\rho\gg 0}$, is very sensitive to violations of this assumption. For large effective population sizes and weak migration, the relative error becomes very high if recombination is not strong enough (Figure S13B; Table S3).

The effect of linkage is again demonstrated by a comparison to the one-locus model. If selection is strong relative to recombination, the mean absorption time with linkage, $\tilde{\bar{t}}_{\mathrm{QLE}}$, is increased by several orders of magnitude compared to the one-locus case, $\tilde{\bar{t}}_{\mathrm{OLM}}$ (Figure 7, A and D; Table S5). The effect is reduced, but still notable, when the recombination rate becomes substantially higher than 10 times the strength of selection in favor of the beneficial background allele, *i.e.*, $\rho/\beta \gg 10$ (Figure 7, B and E). Importantly, large ratios of $\tilde{\bar{t}}_{\mathrm{QLE}}/\tilde{\bar{t}}_{\mathrm{OLM}}$ are not an artifact of $\tilde{\bar{t}}_{\mathrm{OLM}}$ being very small, as Figure 7, C and F confirm. Moreover, $\tilde{\bar{t}}_{\mathrm{QLE}}/\tilde{\bar{t}}_{\mathrm{OLM}}$ is maximized at intermediate migration rates: for very weak migration, A_1 has a fair chance of surviving for a long time even without linkage ($\tilde{\bar{t}}_{\mathrm{OLM}}$ is large); for very strong migration, $\tilde{\bar{t}}_{\mathrm{OLM}}$ and $\tilde{\bar{t}}_{\mathrm{QLE}}$ both tend to zero and $\tilde{\bar{t}}_{\mathrm{QLE}}/\tilde{\bar{t}}_{\mathrm{OLM}}$ approaches unity.

As expected from deterministic theory (Bürger and Akerman 2011; see also File S1, sect. 3) and invasion probabilities calculated above, the mean absorption time decreases as a function of the migration rate m (Figure S14). A noteworthy interaction

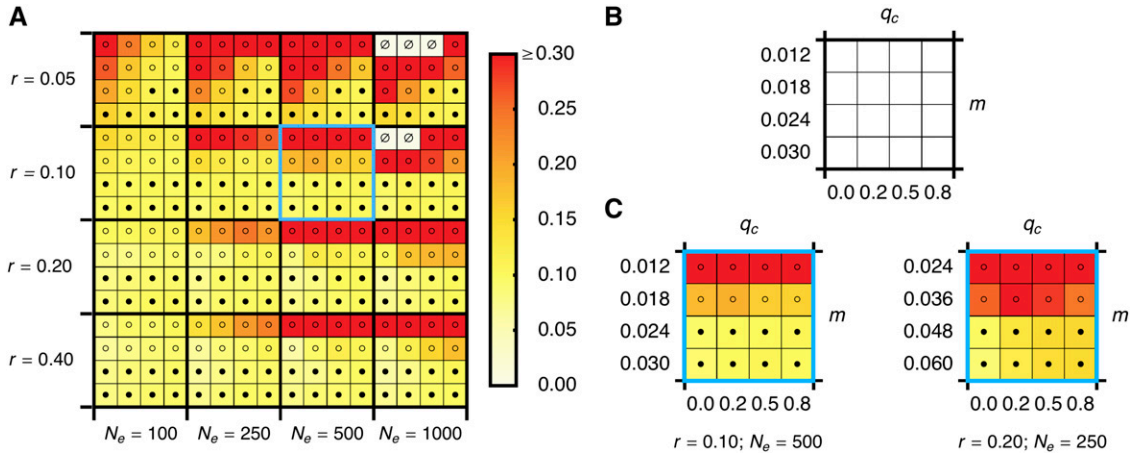

Figure 6 Relative error of the diffusion approximation to the mean absorption time of A_1. (A) The error of $\bar{t}_{QLE}$ from Equation 8 relative to simulations for various parameter combinations. Squares bounded by thick lines delimit combinations of values of the recombination rate r and the effective population size N_e. Within each of them, values of the migration rate m and the continental frequency q_c of B_1 are as shown in B. No negative relative errors were observed. For better resolution, we truncated values >0.30 (the maximum was 3.396 for $N_e = 1000$, $r = 0.05$, $m = 0.018$, $q_c = 0.0$). Open (solid) circles indicate that the marginal one-locus equilibrium $\tilde{E}_B$ is unstable (stable) and A_1 can (not) be established under deterministic dynamics. Parameter combinations for which simulations were too time consuming are indicated by ∅. Selection coefficients are $a = 0.02$ and $b = 0.04$. (C) The left plot corresponds to the square in A that is framed in blue. The right plot shows the fit of the diffusion approximation to simulations conducted with unscaled parameters twice as large and N_e half as large, as on the left side. Scaled parameters are equal on both sides. As expected, the diffusion approximation is worse on the right side. Simulations were as described in *Methods*. See Table S1 for numerical values.

exists between m and the effective population size N_e. For small m, the mean absorption time increases with N_e, whereas for large m, it decreases with N_e. Interestingly, the transition occurs at a value of m lower than the respective critical migration rate below which A_1 can invade in the deterministic model (Figure S14). Hence, there exists a small range of intermediate values of m for which deterministic theory suggests that A_1 will invade, but the stochastic model suggests that survival of A_1 lasts longer in island populations of small rather than large effective size. Similar, but inverted, relations hold for the dependence of the mean absorption time on the selective advantage a of allele A_1 and N_e (Figure S15).

For the parameter ranges we explored, the mean absorption time decreases with increasing continental frequency q_c of B_1. As for the invasion probabilities, competition against a fitter resident population has a negative effect on maintenance of the focal mutation A_1. For a given recombination rate, the effect depends on the relative strength of migration and selection, though: increasing q_c from 0 to 0.8 decreases the mean absorption time by a considerable amount only if m is low or a is large enough; otherwise, genetic drift dominates (Figure S16). This effect is more pronounced for weak than for strong recombination (Figure S17).

So far, we assumed that the initial frequency of A_1 is small, *i.e.*, $p_0 = 1/(2N)$, and that $N_e = N$. In many applications, $N_e < N$ holds and hence $1/(2N) < 1/(2N_e)$. Approximations based on the assumption of p_0 being small, *i.e.*, on the order of $1/(2N_e)$ or smaller, then cause no problem. However, $N_e > N$ may hold in certain models, *e.g.*, with spatial structure (Whitlock and Barton 1997), and $p_0 = 1/(2N)$ may be much greater than $1/(2N_e)$. We therefore investigated the effect of violating the assumption of $p_0 \leq 1/(2N_e)$. For this purpose, we fixed the initial frequency at $p_0 = 0.005$ (*e.g.*, a single copy of A_1 in a population of actual size $N = 100$) and then assessed the relative error of our approximations for various $N_e \geq 100$. As expected, the approximate mean absorption times based on the assumption of p_0 small ($\tilde{\bar{t}}_{QLE}$ and $\tilde{\bar{t}}_{QLE,\rho\gg 0}$) deviate further from their exact conterparts ($\bar{t}_{QLE}$ and $\bar{t}_{QLE,\rho\gg 0}$, respectively) as N_e increases from 100 to 10^4 (Figure S13, D and F). See also Table S6 and S7. For strong migration, the relative error tends to be negative, while it is positive for weak migration (blue *vs.* red boxes in Figure S13, D and F). The assumption of $\rho \gg \max(\alpha, \beta, \mu)$ in $M(p)$ does not lead to any further increase of the relative error, though (Figure S13E; Table S7). Moreover, violation of $p_0 \leq 1/(2N_e)$ has almost no effect on the ratio of the two-locus to the one-locus absorption time, $\bar{t}_{QLE}/\bar{t}_{OLM}$ (compare Table S9 to Table S5).

Invasion-effective migration rate: Comparison of the sojourn-time densities given in Equations 15 and 17 suggests that if μ in the one-locus model is replaced by $\mu_e = \mu(\mu - \beta + \rho)/\rho$, one obtains the STD for the two-locus model. Hence, μ_e denotes the scaled migration rate in a one-locus model such that allele A_1 has the same sojourn properties as it would have if it arose in linkage (decaying at rate ρ) to a background polymorphism maintained by selection against migration at rate μ. In other words, if the assumptions stated above hold, we may use single-locus migration–selection theory, with μ replaced by μ_e, to describe two-locus dynamics. Transforming from the diffusion to the natural scale, we therefore define an invasion-effective migration rate as

$$m_e = m\frac{m + r - b}{r}, \tag{19}$$

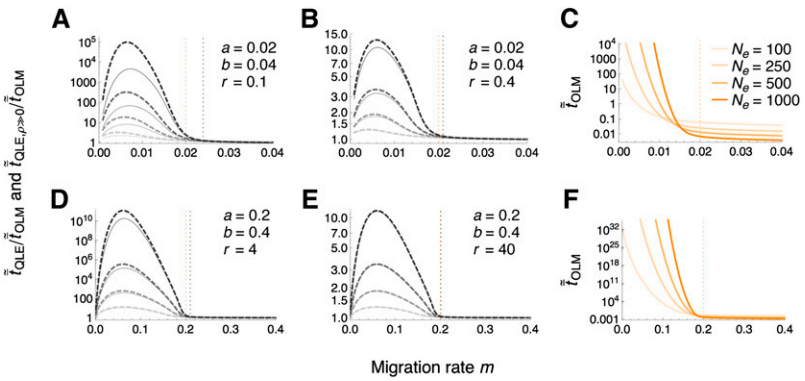

Figure 7 Mean absorption time of A_1 under quasi-linkage equilibrium relative to the one-locus model (OLM). In panels (A), (B), (D) and (E), thin solid curves show the ratio $\tilde{\bar{t}}_{QLE}/\bar{\tilde{t}}_{OLM}$ and thick dashed curves $\tilde{\bar{t}}_{QLE,\rho\gg0}/\bar{\tilde{t}}_{OLM}$, as a function of the migration rate m. The effective population size N_e increases from light to dark gray, taking values of 100, 250, 500, and 1000. Vertical lines denote the migration rate below which A_1 can invade in the deterministic one-locus (orange) and two-locus (black) model. (A) Recombination is too weak for the assumption $\rho \gg \max(\alpha, \beta, \mu)$ to hold. (B) As in (A), but with recombination four times stronger. (D) Evolutionary forces – other than drift – are ten times stronger than in (B). (E) As in (D), but with recombination ten times stronger. Panels (C) and (F) show the mean absorption time (in multiples of $2N_e$) under the one-locus model for the respective row. For m close to 0, numerical procedures are unstable and we truncated the curves. As m converges to 0, $\tilde{\bar{t}}_{QLE}/\bar{\tilde{t}}_{OLM}$ and $\tilde{\bar{t}}_{QLE,\rho\gg0}/\bar{\tilde{t}}_{OLM}$ are expected to approach unity, however.

which, for small m, is approximately

$$\tilde{m}_e = m\left(1-\frac{b}{r}\right) \tag{20}$$

(Figure S18A). Note that m_e and $\tilde{m}_e$ are nonnegative only if $r \geq b - m$ and $r \geq b$, respectively. As we assumed quasi-linkage equilibrium in the derivation, these conditions do not impose any further restriction.

Petry (1983) previously derived an effective migration rate for a neutral site linked to a selected site. In our notation, it is given by

$$m_e^{(P)} = m\left(1+\frac{b}{r}\right)^{-1} = m\frac{r}{b+r} \tag{21}$$

(see Bengtsson 1985 and Barton and Bengtsson 1986 for an extension of the concept). Petry (1983) obtained this approximation by comparing the moments of the stationary allele-frequency distribution for the two-locus model to those for the one-locus model. He assumed that selection and recombination are strong relative to migration and genetic drift. To first order in r^{-1}, *i.e.*, for loose linkage, Petry's $m_e^{(P)}$ is equal to our $\tilde{m}_e$ in Equation 20. As we derived $\tilde{m}_e$ under the assumption of QLE, convergence of $m_e^{(P)}$ to $\tilde{m}_e$ is reassuring. Effective gene flow decreases with the strength of background selection b, but increases with the recombination rate r (Figure S18, B and C).

Long-term effect on linked neutral variation

Selection maintaining genetic differences across space impedes the homogenizing effect of gene flow at closely linked sites (Bengtsson 1985; Barton and Bengtsson 1986). This has consequences for the analysis of sequence or marker data, as patterns of neutral diversity may reveal the action of recent or past selection at nearby sites (Maynard Smith and Haigh 1974; Kaplan *et al.* 1989; Takahata 1990; Barton 1998). We investigated the impact of a two-locus polymorphism contributing to local adaptation on long-term patterns of linked genetic variation. For this purpose, we included a neutral locus C with alleles C_1 and C_2. Allele C_1 segregates on the continent at a constant frequency n_c ($0 \leq n_c < 1$), for example at drift–mutation equilibrium. This may require that the continental population is very large, such that extinction or fixation of C_1 occurs over sufficiently long periods of time compared to the events of interest on the island. The neutral locus is on the same chromosome as A and B, to the left (C–A–B), in the middle (A–C–B), or to the right (A–B–C) of the two selected loci (without loss of generality, A is to the left of B). We denote the recombination rate between locus X and Y by r_{XY}, where $r_{XY} = r_{YX}$, and assume that the recombination rate is additive. For example, if the configuration is A–C–B, we set $r_{AB} = r_{AC} + r_{CB}$.

Unless linkage to one of the selected loci is complete, under deterministic dynamics, allele C_1 will reach the equilibrium frequency $\hat{n} = n_c$ on the island, independently of its initial frequency on the island. Recombination affects only the rate of approach to this equilibrium, not its value. We focus on the case where the continent is monomorphic at locus B ($q_c = 0$). Selection for local adaptation acts on loci A and B, and migration–selection equilibrium will be reached at each of them (File S1, sect. 6). Gene flow from the continent will be effectively reduced in their neighborhood on the chromosome. Although the expected frequency of C_1 remains n_c throughout, drift will cause variation around this mean to an extent that depends on the position of C on the chromosome. It may take a long time for this drift–migration equilibrium to be established, but the resulting signal should be informative for inference.

To investigate the effect of selection at two linked loci, we employed the concept of an effective migration rate according to Bengtsson (1985), Barton and Bengtsson (1986), and Kobayashi *et al.* (2008). As derived in File S1, sect. 8, and File S8, for continuous time and weak migration, the effective migration rates for the three configurations are

$$m_e^{CAB} = m\frac{r_{CA}(a+r_{CB})}{(a+r_{CA})(a+b+r_{CB})}, \tag{22a}$$

$$m_e^{ACB} = m\frac{r_{AC}r_{CB}}{(a+r_{AC})(b+r_{CB})}, \tag{22b}$$

$$m_e^{ABC} = m\frac{r_{BC}(b + r_{AC})}{(b + r_{BC})(a + b + r_{AC})}. \tag{22c}$$

We note that m_e^{ACB} has been previously derived (Bürger and Akerman 2011, Equation 4.30). From Equation 22, we define the effective migration rate experienced at a neutral site as

$$m_e^{(n)} = \begin{cases} m_e^{CAB} & \text{if C}-\text{A}-\text{B holds,} \\ m_e^{ACB} & \text{if A}-\text{C}-\text{B holds,} \\ m_e^{ABC} & \text{if A}-\text{B}-\text{C holds.} \end{cases} \tag{23}$$

Equation 23 subsumes the effect on locus C of selection at loci A and B. It can be generalized to an arbitrary number of selected loci. Let A_i ($i = 1,\ldots, I$) and B_j ($j = 1,\ldots, J$) be the ith and jth locus to the left and right of the neutral locus, respectively. We find that the effective migration rate at the neutral locus is

$$m_e^{(n)} = m\left[\prod_{i=1}^{I}\left(1 + \frac{a_i}{\sum_{k=1}^{i-1} a_k + r_{A_i}}\right)^{-1}\right] \times \left[\prod_{j=1}^{J}\left(1 + \frac{b_j}{\sum_{k=1}^{j-1} b_k + r_{B_j}}\right)^{-1}\right], \tag{24}$$

where a_i (b_j) is the selection coefficient at locus A_i (B_j), and r_{A_i} (r_{B_j}) the recombination rate between the neutral locus and A_i (B_j). Each of the terms in the round brackets in Equation 24 is reminiscent of Petry's (1983) effective migration rate for a neutral linked site (Equation 21). For weak linkage, these terms are also similar to the invasion-effective migration rate experienced by a weakly beneficial mutation (Equation 20). This suggests that the effective migration rate experienced by a linked neutral site is approximately the same as that experienced by a linked weakly beneficial mutation, which corroborates the usefulness of Equation 24. In the following, we study different long-term properties of the one-locus drift-migration model by substituting effective for actual migration rates.

Mean absorption time: Suppose that C_1 is absent from the continent ($n_c = 0$), but present on the island as a *de novo* mutation. Although any such mutant allele is doomed to extinction, recurrent mutation may lead to a permanent influx and, at mutation–migration equilibrium, to a certain level of neutral differentiation between the continent and the island. Here, we ignore recurrent mutation and focus on the fate of a mutant population descending from a single copy of C_1. We ask how long it will survive on the island, given that a migration–selection polymorphism is maintained at equilibrium at both selected loci in the background (A, B). Standard diffusion theory predicts that the mean absorption (extinction) time of C_1 is approximately $\tilde{\bar{t}}_{neut} = N_e^{-1}\int_{1/(2N)}^{1} n^{-1}(1-n)^{2\mu-1}\,dn$ (Ewens 2004, pp. 171–175). We replace the scaled actual migration rate μ by $\mu_e^{(n)} = 2N_e m_e^{(n)}$, with $m_e^{(n)}$ from Equation 23. This assumes that the initial frequency of C_1 on the island is $n_0 = 1/(2N)$ and that $N_e = N$. For moderately strong migration ($\mu_e^{(n)} \approx 1$), $\tilde{\bar{t}}_{neut}$ is of order $\log(2N_e)$, meaning that C_1 will on average remain in the island population for a short time. However, if locus C is tightly linked to one of the selected loci, or if configuration A–C–B applies and A and B are sufficiently close, the mean absorption time of C_1 is strongly elevated (Figure S19).

Stationary distribution of allele frequencies: In contrast to above, assume that C_1 is maintained at a constant frequency $n_c \in (0, 1)$ on the continent. Migrants may therefore carry both alleles, and genetic drift and migration will lead to a stationary distribution of allele frequencies given by

$$\phi(n) = \frac{\Gamma(2\mu)}{\Gamma(2\mu n_c)\Gamma(2\mu[1-n_c])}\, n^{2\mu n_c - 1}(1-n)^{2\mu(1-n_c)-1},$$

where $\Gamma(x)$ is the Gamma function (Wright 1940, pp. 239–241). As above, we replace μ by $\mu_e^{(n)}$ to account for the effect of linked selection. The mean of the distribution $\phi(n)$ is n_c, independently of μ_e, whereas the stationary variance is $\mathrm{var}(n) = n_c(1-n_c)/(1+2\mu_e^{(n)})$ (Wright 1940). The expected heterozygosity is $H = 4\mu_e^{(n)} n_c(1-n_c)/(1+2\mu_e^{(n)})$, and the divergence from the continental population is $F_{ST} = \mathrm{var}(n)/[n_c(1-n_c)] = 1/(1+2\mu_e^{(n)})$ (see File S9 for details). Depending on the position of the neutral locus, $\phi(n)$ may change considerably in shape, for example, from L- to U- to bell-shaped (Figure 8). The pattern of $\phi(n)$, H and F_{ST} along the chromosome reveals the positions of the selected loci, and their rate of change per base pair contains information about the strength of selection if the actual migration rate is known.

Rate of coalescence: As a third application, we study the rate of coalescence for a sample of size two taken from the neutral locus C, assuming that migration–selection equilibrium has been reached a long time ago at the selected loci A and B. We restrict the analysis to the case of strong migration compared to genetic drift, for which results by Nagylaki (1980) (forward in time) and Notohara (1993) (backward in time) apply (see Wakeley 2009, for a detailed review). The strong-migration limit follows from a separation of timescales: going back in time, migration spreads the lineages on a faster timescale, whereas genetic drift causes lineages to coalesce on a slower one.

For a moment, let us assume that there are two demes of size N_1 and N_2, and denote the total number of diploids by $\overline{N} = N_1 + N_2$. We define the relative deme size $c_i = N_i/\overline{N}$ and let the backward migration rates m_1 and m_2 denote the fractions of individuals in deme 1 and 2 in the current generation that were in deme 2 and 1 in the previous generation, respectively. The strong-migration limit then requires that $N_i m_i = \overline{N} c_i m_i$ is large (Wakeley 2009). Importantly, the relative

deme sizes c_i are constant in the limit of $\overline{N} \to \infty$. Under these assumptions, it can be shown that the rate of coalescence for a sample of two is independent of whether the two lineages were sampled from the same or different demes. The rate of coalescence is given by

$$G = \frac{m_2^2}{(m_1 + m_2)^2} \frac{1}{c_1} + \frac{m_1^2}{(m_1 + m_2)^2} \frac{1}{c_2} \qquad (25)$$

(Wakeley 2009, p. 193). The coalescent-effective population size is defined as the actual total population size times the inverse of the rate of coalescence, $N_e^{(coal)} = \overline{N}/G$ (Sjödin *et al.* 2005).

In our context, we substitute $m_e^{(n)}$ from Equation 23 for m_1 in G. To be consistent with the assumption of continent–island migration—under which we studied the migration–selection dynamics at A and B—we require $N_2 \gg N_1$ and $m_2 \ll m_1$. This way, the assumptions of $N_1 m_1$ and $N_2 m_2$ being large can still be fulfilled. However, note that $m_2 \ll m_1$ does not automatically imply $m_2 \ll m_e^{(n)}$; depending on the strength of selection and recombination, $m_e^{(n)}$ may become very small. Hence, in applying the theory outlined here, one should bear in mind that the approximation may be misleading if $m_e^{(n)}$ is small (for instance, if locus C is tightly linked to either A or B). The neutral coalescent rate G is strongly increased in the neighborhood of selected sites; accordingly, $N_e^{(coal)}$ is increased (Figure S20). Reassuringly, this pattern parallels those for linked neutral diversity and divergence in Figure 8.

Discussion

We have provided a comprehensive analysis of the fate of a locally beneficial mutation that arises in linkage to a previously established migration–selection polymorphism. In particular, we obtained explicit approximations to the invasion probability. These reveal the functional dependence on the key parameters and substitute for time-consuming simulations. Further, we found accurate approximations to the mean extinction time, showing that a unilateral focus on invasion probabilities yields an incomplete understanding of the effects of migration and linkage. Finally, we derived the effective migration rate experienced by a neutral or weakly beneficial mutation that is linked to arbitrarily many migration–selection polymorphisms. This opens up a genome-wide perspective of local adaptation and establishes a link to inferential frameworks.

Insight from stochastic modeling

Previous theoretical studies accounting for genetic drift in the context of polygenic local adaptation with gene flow were mainly simulation based (Yeaman and Whitlock 2011; Feder *et al.* 2012; Flaxman *et al.* 2013) or did not model recombination explicitly (Lande 1984, 1985; Barton 1987; Rouhani and Barton 1987; Barton and Rouhani 1991; but see Barton and Bengtsson 1986). Here, we used stochastic processes to model genetic drift and to derive explicit expressions that provide an alternative to simulations. We distinguished between the stochastic effects due to initial rareness of a *de novo* mutation on the one hand and the long-term effect of finite population size on the other.

For a two-locus model with a steady influx of maladapted genes, we found an implicit condition for invasion of a single locally beneficial mutation linked to the background locus (Equation 9). This condition is valid for arbitrary fitnesses, *i.e.*, any regime of dominance or epistasis. It also represents an extension to the case of a panmictic population in which the background polymorphism is maintained by overdominance, rather than migration–selection balance (Ewens 1967). Assuming additive fitnesses, we derived simple explicit conditions for invasion in terms of a critical migration or recombination rate (Equations 10 or 11, respectively). Whereas these results align with deterministic theory (Bürger and Akerman 2011), additional quantitative and qualitative insight emerged from studying invasion probabilities and extinction times. Specifically, invasion probabilities derived from a two-type branching process (Equations 3 and 12) capture the ambiguous role of recombination breaking up optimal haplotypes on the one hand and creating them on the other. Diffusion approximations to the sojourn and mean absorption time shed light on the long-term effect of finite population size. A comparison between the dependence of invasion probabilities and extinction times on migration and recombination rate revealed important differences (discussed further below). Deterministic theory fails to represent such aspects, and simulations provide only limited understanding of functional relationships.

Recently, Yeaman (2013) derived an *ad hoc* approximation of the invasion probability, using the so-called "splicing approach" (Yeaman and Otto 2011). There, the leading eigenvalue of the appropriate Jacobian (Bürger and Akerman 2011) is taken as a proxy for the selection coefficient and inserted into Kimura's (1962) formula for the one-locus invasion probability in a panmictic population. Yeaman's (2013) method provides a fairly accurate approximation to the invasion probability if A_1 initially occurs on the beneficial background B_1 (at least for tight linkage). However, it does not describe the invasion probability of an average mutation (Figure S22) and hence does not predict the existence of a nonzero optimal recombination rate. As a consequence, Yeaman's (2013) conclusion that physically linked selection alone is of limited importance for the evolution of clustered architectures is likely conservative, because it is based on an approximation that inflates the effect of linked selection.

Nonzero optimal recombination rate

We have shown that the average invasion probability of a linked beneficial mutation can be maximized at a nonzero recombination rate ($r_{opt} > 0$). Equation 13 provides a general condition for when this occurs. With additive fitnesses, the local advantage of the focal mutation must be above a critical value (Equation 14). Otherwise, the invasion probability is maximized at $r_{opt} = 0$.

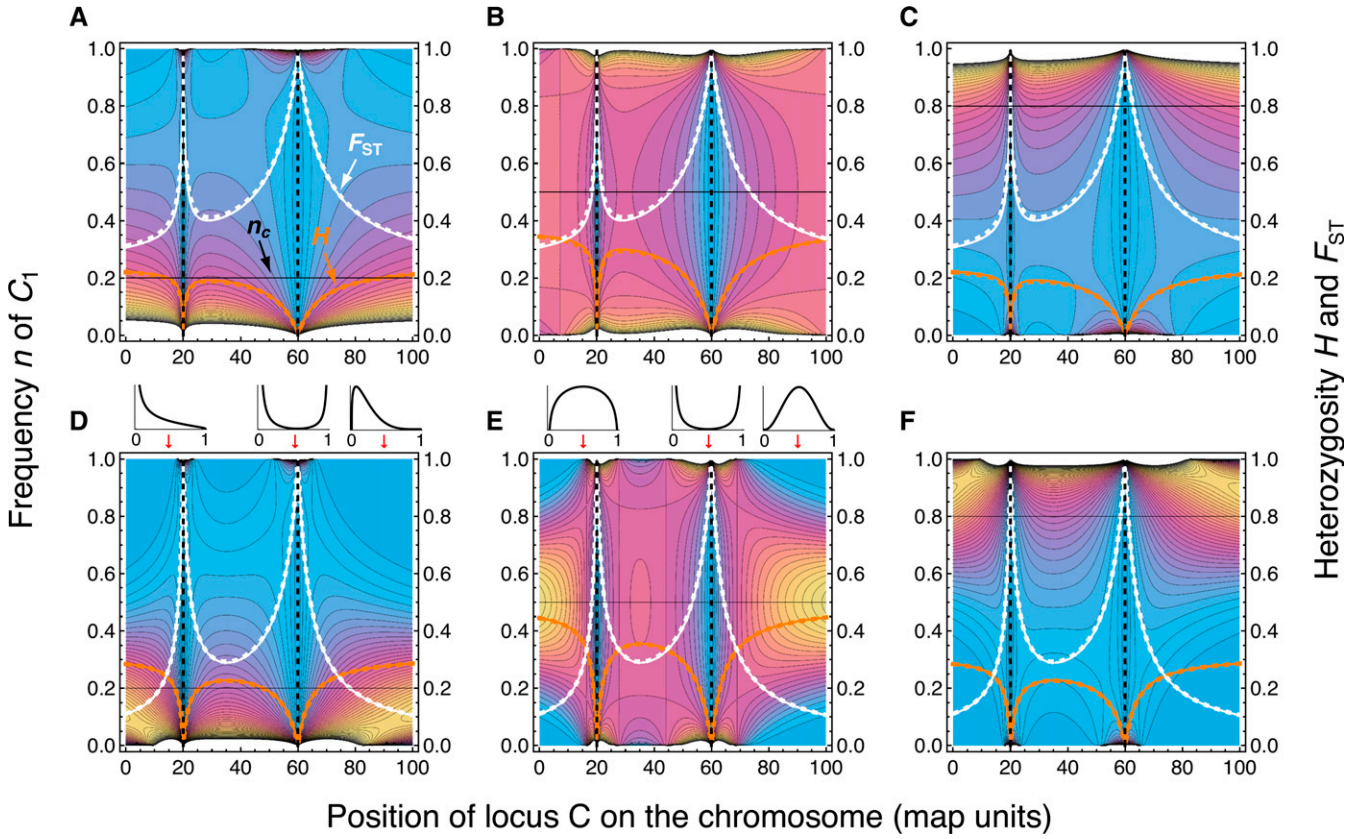

Figure 8 The effect of linked selection on neutral diversity and population divergence. Shown are top views of the stationary allele frequency distribution on the island for a neutral biallelic locus C linked to two selected sites at (locus A) 20 and (locus B) 60 map units from the left end of the chromosome. Density increases from light blue to yellow (high peaks were truncated for better resolution). Orange and white curves show the expected diversity (heterozygosity H) and population divergence (F_{ST}) as a function of the position of the neutral site. Solid curves use exact, numerically computed values of the effective migration rate and dashed curves use the approximations given in Equation 23. One map unit (cM) corresponds to $r = 0.01$ and the effective size of the island population is $N_e = 100$. The continental frequency n_c of allele C_1 is indicated by a horizontal black line and, from left to right, equal to 0.2, 0.5 and 0.8. (A–C) Relatively strong drift and weak migration compared to selection: $\alpha = 4$, $\beta = 80$, $\mu = 2$. (D–F) Relatively weak drift and migration on the same order of magnitude as selection at locus A: $\alpha = 40$, $\beta = 800$, $\mu = 48$. Note that H is sensitive to n_c, whereas F_{ST} is not. On top of D and E, allele frequency distributions that result from taking vertical slices at positions indicated by red arrows are shown (15, 59 and 90 cM).

Existence of a nonzero optimal recombination rate in the absence of epistasis and dominance is noteworthy. For a panmictic population in which the polymorphism at the background locus is maintained by overdominance, Ewens (1967) has shown that the optimal recombination rate may be nonzero, but this requires epistasis. In the context of migration, the existence of $r_{opt} > 0$ has been noted and discussed in a simulation study (Feder *et al.* 2012, Figure 5), but no analytical approximation or explanation that captures this feature has been available. In principle, $r_{opt} > 0$ suggests that the genetic architecture of polygenic adaptation may evolve such as to optimize the recombination rates between loci harboring adaptive mutations. Testing this prediction requires modifier-of-recombination theory (*e.g.*, Otto and Barton 1997; Martin *et al.* 2006; Roze and Barton 2006; Kermany and Lessard 2012). While we expect evolution toward the optimal recombination rate in a deterministic model (Lenormand and Otto 2000), it is important to determine if and under which conditions this occurs in a stochastic model and what the consequences for polygenic adaptation are.

For instance, in a model with two demes and a quantitative trait for which the fitness optima are different in the two demes, Yeaman and Whitlock (2011) have shown that mutations contributing to adaptive divergence in the presence of gene flow may cluster with respect to their position on the chromosome. Moreover, architectures with many weakly adaptive mutations tended to become replaced by architectures with fewer mutations of larger effect. Although our migration model is different, existence of a nonzero optimal recombination rate suggests that there might be a limit to the degree of clustering of locally adaptive mutations. It is worth recalling that our result of $r_{opt} > 0$ applies to the *average* invading mutation (black curves in Figure 1 and Figure 2). For any *particular* mutation that arises on the beneficial background, $r = 0$ is (almost) always optimal (blue curves in Figure 1 and Figure 2; see Figure S3D for an exception).

Long-term dynamics of adaptive divergence

Finite population size on the island eventually leads to extinction of a locally beneficial mutation even after successful

initial establishment. This is accounted for neither by deterministic nor branching-process theory. Employing a diffusion approximation, we have shown that linkage of the focal mutation to a migration–selection polymorphism can greatly increase the time to extinction and thus alter the long-term evolutionary dynamics. In such cases, the timescale of extinction may become similar to that on which mutations occur. This affects the rate at which an equilibrium between evolutionary forces is reached. We provided a rule of thumb for when the time spent by the focal allele at a certain frequency exceeds a given multiple of the respective time without linkage. Essentially, the product of the background selection coefficient times the migration rate must be larger than a multiple of the recombination rate (Equation 18).

The effect of linked selection can also be expressed in terms of an invasion-effective migration rate (Equations 19 and 20). Both our rule of thumb and the formula for the effective migration rate provide a means of quantifying the importance of linkage to selected genes in the context of local adaptation. In practice, however, their application requires accurate estimates of the recombination map, the selective advantage of the beneficial background allele, and the actual migration rate.

A nontrivial effect of gene flow

Our stochastic modeling allows for a more differentiated understanding of the role of gene flow in opposing adaptive divergence. Whereas deterministic theory specifies a critical migration rate beyond which a focal mutation of a given advantage cannot be established (Bürger and Akerman 2011; see also Figure S7 and Figure S8), the potential of invasion is far from uniform if migration is below this critical value (Figure S3, Figure S5, and Figure S14). For instance, we may define the relative advantage of linkage to a migration–selection polymorphism as the ratio of the quantity of interest *with* a given degree of linkage to that *without* linkage.

A comparison of the two quantities of interest in our case—invasion probability and mean extinction time—with respect to migration is instructive (Figure 7 and Figure S21). Starting from zero migration, the relative advantage of linkage in terms of the invasion probability initially increases with the migration rate very slowly, but then much faster as the migration rate approaches the critical value beyond which an unlinked focal mutation cannot invade (Figure S21A). Beyond this critical value, the relative advantage is infinite until migration is so high that even a fully linked mutation cannot be established. In contrast, we have shown that the relative advantage of linkage in terms of the mean extinction time is maximized at an intermediate migration rate (Figure 7).

In conclusion, for very weak migration, the benefit of being linked to a background polymorphism is almost negligible. For intermediate migration rates, the potential of invasion is elevated by linked selection; this is mainly due to a substantially increased mean extinction time of those still rather few mutations that successfully survive the initial phase of stochastic loss. This argument is based on the increase of the mean extinction time *relative* to unlinked selection. Because, for large populations, *absolute* extinction times become very large as the migration rate decreases (Figure 7, C and F), the biological relevance of this comparison may be confined to cases in which the mean extinction time of an unlinked mutation is not extremely high. For migration rates close to the critical migration rate, however, any relative advantage of linkage seems to arise via an increased invasion probability, not via an increased mean extinction time. This is because, in this case, the latter is close to that for no linkage (compare Figure S21A to Figure 7, A and B). A final statement about the relative importance of invasion probability *vs.* mean extinction time is not appropriate at this point. This would require extensive numerical work, along with a derivation of a diffusion approximation to the mean extinction time for tight linkage. However, for small populations, our results show that linked selection can increase the mean extinction time to an extent that is biologically relevant, while, at the same time, not affecting the invasion probability much. This suggests that invasion probabilities may not be a sufficient measure for the importance of physical linkage in adaptive divergence.

Standing variation at the background locus

We have extended some of our analyses to the case where the background locus is polymorphic on the continent and immigrants may therefore carry both the locally beneficial or deleterious allele. This represents a compromise between the extremes of adaptation from standing *vs. de novo* genetic variation. We have shown that the presence of the beneficial background allele on the continent, and hence among immigrants, leads to a lower invasion probability and a shorter extinction time for the focal *de novo* mutation. This effect is due to increased competition against a fitter resident population. While this result is of interest as such, it should not be abused to gauge the relative importance of standing *vs. de novo* variation in the context of local adaptation. For this purpose, invasion probabilities and extinction times of single mutations do matter, but are not sufficient metrics on their own. Factors such as the mutation rate, the mutational target size, and the distribution of selection coefficients must be taken into account (Hermisson and Pennings 2005).

Footprint of polygenic local adaptation

A number of previous studies have quantified the effect of divergent selection or genetic conflicts on linked neutral variation in discrete (Bengtsson 1985; Charlesworth *et al.* 1997) and continuous space (Barton 1979; Petry 1983). They all concluded that a single locus under selection leads to a pronounced reduction in effective gene flow only if selection is strong or if linkage to the neutral site is tight. Whereas Bengtsson (1985) found that additional, physically unlinked, loci under selection had no substantial effect on neutral differentiation, Feder and Nosil (2010) recently suggested that such loci may have an appreciable effect as long as they are not too numerous. When these authors added a large number of unlinked loci under selection, this resulted in a genome-wide reduction of the effective migration rate, such that the baseline level of neutral divergence was elevated and

any effect of linkage to a single selected locus unlikely to be detected. However, for large numbers of selected loci, it is no longer justified to assume that all of them are physically unlinked. This was noted much earlier by Barton and Bengtsson (1986), who therefore considered a linear genome with an arbitrary number of selected loci linked to a focal neutral site. They showed that a large number of linked selected loci is needed to cause a strong reduction in effective migration rate. In such cases, the majority of other genes must be linked to some locus under selection.

The concept of an effective migration rate has played a key role in most of the studies mentioned above (see Barton and Bengtsson 1986, and Charlesworth *et al.* 1997 for a more comprehensive review). However, for models with more than one linked locus under selection, previous studies relied on numerical solutions or simulations to compute the effective migration rate. Recently, Bürger and Akerman (2011) derived an analytical approximation for a neutral site that is flanked by two selected loci. We have generalized their result to alternative genetic architectures and an arbitrary number of selected loci (Equation 24). From this, we predicted the long-term footprint of polygenic local adaptation in terms of the distribution of allele frequencies, population divergence, and coalescent rate at the neutral site. When considered as a function of the position of the neutral site on the chromosome, these quantities reveal patterns that can hopefully be used for inference about the selective process (Figure 8 and Figure S20).

We have considered only the case where migration–selection equilibrium has been reached at the selected loci. It would be interesting, although more demanding, to study the transient phase during which locally beneficial mutations (such as A_1 in our case) rise in frequency from $p_0 = 1/(2N)$ to the (pseudo-) equilibrium frequency. We expect this to create a temporary footprint similar to that of a partial sweep (Pennings and Hermisson 2006a,b; Pritchard *et al.* 2010; Ralph and Coop 2010). Theoretical progress hinges on a description of the trajectory of the linked sweeping alleles, accounting in particular for the stochastic "lag phase" at the beginning. It will then be of interest to study recurrent local sweeps and extend previous theory for panmictic populations (Coop and Ralph 2012; Lessard and Kermany 2012) to include population structure, migration, and spatially heterogeneous selection. The hitchhiking effect of a beneficial mutation in a subdivided population has been described in previous studies (*e.g.*, Slatkin and Wiehe 1998; Kim and Maruki 2011), but these did not account for additional linked loci under selection.

One limitation to our prediction of the coalescence rate at linked neutral sites is the assumption of strong migration relative to genetic drift (Nagylaki 1980; Notohara 1993). As the effective migration rate decays to zero if the neutral site is very closely linked to a selected site (Figure S18C), this assumption will be violated. Therefore, our predictions should be interpreted carefully when linkage is tight. Moreover, and even though this seems widely accepted, we are not aware of a rigorous proof showing that an effective migration rate can sufficiently well describe the effect of local selection on linked neutral genealogies (*cf*. Barton and Etheridge 2004).

Another limitation is that our prediction of linked neutral diversity and divergence (Figure 8) holds only for drift–migration equilibrium. For closely linked neutral sites, which experience very low rates of effective migration, it may take a long time for this equilibrium to be reached. By that time, other evolutionary processes such as background selection and mutation will have interfered with the dynamics at the focal site.

Further limitations and future extensions

We assumed no dominance and no epistasis. Both are known to affect the rate of adaptation and the maintenance of genetic variation (*e.g.*, Charlesworth *et al.* 1987; Bank *et al.* 2012). Empirical results on dominance effects of beneficial mutations are ambiguous (Vicoso and Charlesworth 2006). Some studies showed no evidence for a deviation from additivity, whereas others suggested weak recessivity (reviewed in Orr 2010 and Presgraves 2008). Empirical evidence for epistasis comes from studies reporting genetic incompatibilities between hybridizing populations (Lowry *et al.* 2008; Presgraves 2010). In the classical Dobzhansky–Muller model (Bateson 1909; Dobzhansky 1936; Muller 1942), such incompatibilities may become expressed during secondary contact after allopatry, even if divergence is neutral. With gene flow, genetic incompatibilities can be maintained only if the involved alles are locally beneficial (Bank *et al.* 2012). Bank *et al.* (2012) derived respective conditions using deterministic theory. An extension to a stochastic model focusing on invasion probabilities and extinction times would be desirable.

Our model assumed one-way migration. While this is an important limiting case and applies to a number of natural systems (*e.g.*, King and Lawson 1995), an extension to two-way migration is of interest, because natural populations or incipient species often exchange migrants mutually (*e.g.*, Janssen and Mundy 2013; Nadeau *et al.* 2013). Such theory will allow for a direct comparison to recent simulation studies (Feder and Nosil 2010; Yeaman and Whitlock 2011; Feder *et al.* 2012; Flaxman *et al.* 2013; Yeaman 2013). It will also have a bearing on the evolution of suppressed recombination in sex chromosomes (*e.g.*, Rice 1984, 1987; Fry 2010; Jordan and Charlesworth 2012; Charlesworth 2013). Deterministic theory suggests that linkage becomes less crucial for the maintenance of locally beneficial alleles the more symmetric gene flow is (Akerman and Bürger 2014).

When describing the distribution of fitness effects of successful beneficial mutations, we considered only a single mutation. Future studies should investigate a complete adaptive walk, allowing for mutations at multiple loci to interact via dominance, epistasis, and linkage. Moreover, it would then seem justified to relax the assumption of a constant fitness gradient, especially in the proximity of an optimum, and to account for the fact that the input DFE is not necessarily exponential (Martin and Lenormand 2008).

In our derivations of sojourn and mean absorption times, we assumed QLE. As expected, the approximations break down if recombination is weak (*e.g.*, Figure 6). For tight linkage, when linked selection is most beneficial, an alternative diffusion process needs to be developed. However, to determine how weak physical linkage may be such that an invading mutation still has an advantage, an approximation that is accurate for moderate and loose linkage is required. Therefore, the assumption of QLE does not restrict the scope of our results that address the limits to the importance of linked selection.

Conclusion

This study advances our understanding of the effects of physical linkage and maladaptive gene flow on local adaptation. We derived explicit approximations to the invasion probability and extinction time of benefical *de novo* mutations that arise in linkage to an established migration–selection polymorphsim. In addition, we obtained an analytical formula for the effective migration rate experienced by a neutral or weakly beneficial site that is linked to an arbitrary number of selected loci. These approximations provide an efficient alternative to simulations (*e.g.*, Feder and Nosil 2010; Feder *et al.* 2012). Our results strengthen the emerging view that physically linked selection (and hence so-called divergence hitchhiking) is biologically relevant only if linkage is tight or if selection at the background locus is strong (Petry 1983; Barton and Bengtsson 1986; Feder *et al.* 2012; Flaxman *et al.* 2013). When these conditions are met, however, the effect of linkage can be substantial. A definite statement about the importance of "divergence hitchhiking" *vs.* "genome hitchhiking" and complementary processes (*cf.* Yeaman 2013) seems premature, though; it will require further empirical and theoretical work. We suggest that future theoretical studies (i) obtain analogous approximations for bi- rather than unidirectional gene flow, (ii) account for epistasis and dominance, (iii) incorporate the distribution of fitness effects of beneficial mutations, and (iv) employ a stochastic modifier-of-recombination model to assess the importance of nonzero optimal recombination rates. Extensions of this kind will further enhance our understanding of polygenic local adaptation and its genetic footprint.

Acknowledgments

We thank Ada Akerman for sharing an unpublished manuscript and *Mathematica* code, Josef Hofbauer for help with finding the derivative of the invasion probability as a function of the recombination rate, the Biomathematics Group at the University of Vienna for stimulating discussions, and Samuel Flaxman and Samuel Yeaman for useful comments on the manuscript. S.A. and R.B. acknowledge financial support by the Austrian Science Fund (FWF), projects P21305 and P25188. The computational results presented have been achieved in parts using the Vienna Scientific Cluster (VSC).

Literature cited

Akerman, A., and R. Bürger, 2014 The consequences of gene flow for local adaptation and differentiation: a two-locus two-deme model. J. Math. Biol. 68: 1135–1198.

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.

Barrett, R. D. H., L. K. M'Gonigle, and S. P. Otto, 2006 The distribution of beneficial mutant effects under strong selection. Genetics 174: 2071–2079.

Barton, N., 1979 Gene flow past a cline. Heredity 43: 333–339.

Barton, N., and B. O. Bengtsson, 1986 The barrier to genetic exchange between hybridising populations. Heredity 57: 357–376.

Barton, N. H., 1983 Multilocus clines. Evolution 37: 454–471.

Barton, N. H., 1987 The probability of establishment of an advantageous mutant in a subdivided population. Genet. Res. 50: 35–40.

Barton, N. H., 1995 Linkage and the limits to natural selection. Genetics 140: 821–841.

Barton, N. H., 1998 The effect of hitchhiking on neutral genealogies. Genet. Res. 72: 123–133.

Barton, N. H., 2010 Genetic linkage and natural selection. Philos. Trans. R. Soc. Lond. B Biol. Sci. 365: 2559–2569.

Barton, N. H., and A. M. Etheridge, 2004 The effect of selection on genealogies. Genetics 166: 1115–1131.

Barton, N. H., and S. Rouhani, 1991 The probability of fixation of a new karyotype in a continuous population. Evolution 45: 499–517.

Bateson, W., 1909 Heredity and variation in modern lights, pp. 85–101 in *Darwin and Modern Science*, edited by A. C. Seward. Cambridge University Press, Cambridge, UK.

Bengtsson, B. O., 1985 The flow of genes through a genetic barrier, pp. 31–42 in *Evolution: Essays in Honour of John Maynard Smith*, Vol. 1. Chap. 3, edited by P. J. Greenwood, P. Harvey, and M. Slatkin. Cambridge University Press, New York.

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.

Charlesworth, B., J. A. Coyne, and N. H. Barton, 1987 The relative rates of evolution of sex chromosomes and autosomes. Am. Nat. 130: 113–146.

Charlesworth, B., M. Nordborg, and D. Charlesworth, 1997 The effects of local selection, balanced polymorphism and background selection on equilibrium patterns of genetic diversity in subdivided populations. Genet. Res. 70: 155–174.

Charlesworth, D., 2013 Plant sex chromosome evolution. J. Exp. Bot. 64: 405–420.

Coop, G., and P. Ralph, 2012 Patterns of neutral diversity under general models of selective sweeps. Genetics 192: 205–224.

Cowperthwaite, M. C., J. J. Bull, and L. A. Meyers, 2005 Distributions of beneficial fitness effects in RNA. Genetics 170: 1449–1457.

Dobzhansky, T., 1936 Studies on hybrid sterility. II. Localization of sterility factors in *Drosophila pseudoobscura* hybrids. Genetics 21: 113–135.

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.

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*, Ed. 2. Springer-Verlag, New York.
Eyre-Walker, A., and P. D. Keightley, 2007 The distribution of fitness effects of new mutations. Nat. Rev. Genet. 8: 610–618.
Feder, J. L., and P. Nosil, 2010 The efficacy of divergence hitchhiking in generating genomic islands during ecological speciation. Evolution 64: 1729–1747.
Feder, J. L., R. Gejji, S. Yeaman, and P. Nosil, 2012 Establishment of new mutations under divergence and genome hitchhiking. Philos. Trans. R. Soc. Lond. B Biol. Sci. 367: 461–474.
Fisher, R. A., 1930 *The Genetical Theory of Natural Selection*. Oxford University Press, New York.
Flaxman, S. M., J. L. Feder, and P. Nosil, 2013 Genetic hitchhiking and the dynamic buildup of genomic divergence during speciation with gene flow. Evolution 67: 2577–2591.
Fry, J. D., 2010 The genomic location of sexually antagonistic variation: some cautionary comments. Evolution 64: 1510–1516.
Gillespie, J. H., 1983 A simple stochastic gene substitution model. Theor. Popul. Biol. 23: 202–215.
Gillespie, J. H., 1984 Molecular evolution over the mutational landscape. Evolution 38: 1116–1129.
Guerrero, R. F., F. Rousset, and M. Kirkpatrick, 2012 Coalescent patterns for chromosomal inversions in divergent populations. Philos. Trans. R. Soc. Lond. B Biol. Sci. 367: 430–438.
Haccou, P., J. Peter, and V. A. Vatutin, 2005 *Branching Processes: Variation, Growth, and Extinction of Populations, Vol. 5, Cambridge Studies in Adaptive Dynamics*. Cambridge University Press, New York.
Haldane, J. B. S., 1927 A mathematical theory of natural and artificial selection. v. Selection and mutation. Math. Proc. Camb. Philos. Soc. 23: 838–844.
Haldane, J. B. S., 1930 A mathematical theory of natural and artificial selection. VI. Isolation. Math. Proc. Camb. Philos. Soc. 26: 220–230.
Harris, T. E., 1963 *The Theory of Branching Processes, Vol. 119, Die Grundlehren der Mathematischen Wissenschaften*, Ed. 1. Springer-Verlag, Berlin.
Hermisson, J., and P. S. Pennings, 2005 Soft sweeps: molecular population genetics of adaptation from standing genetic variation. Genetics 169: 2335–2352.
Hill, W. G., and A. Robertson, 1966 The effect of linkage on limits to artificial selection. Genet. Res. 8: 269–294.
Janssen, K., and N. I. Mundy, 2013 Molecular population genetics of the melanic plumage polymorphism in Arctic skuas (*Stercorarius parasiticus*): evidence for divergent selection on plumage colour. Mol. Ecol. 22: 4634–4643.
Jordan, C. Y., and D. Charlesworth, 2012 The potential for sexually antagonistic polymorphism in different genome regions. Evolution 66: 505–516.
Kaplan, N. L., R. R. Hudson, and C. H. Langley, 1989 The "hitchhiking effect" revisited. Genetics 123: 887–899.
Karlin, S., and J. McGregor, 1968 Rates and probabilities of fixation for two locus random mating finite populations without selection. Genetics 58: 141–159.
Karlin, S., and H. M. Taylor, 1981 *A Second Course in Stochastic Processes*, Vol. 2. Academic Press, San Diego, CA.
Karlsen, B. O., K. Klingan, Å. Emblem, T. E. Jørgensen, A. Jueterbock *et al.*, 2013 Genomic divergence between the migratory and stationary ecotypes of Atlantic cod. Mol. Ecol. 22: 5098–5111.
Kermany, A. R., and S. Lessard, 2012 Effect of epistasis and linkage on fixation probability in three-locus models: an ancestral recombination-selection graph approach. Theor. Popul. Biol. 82: 131–145.
Kim, Y., and T. Maruki, 2011 Hitchhiking effect of a beneficial mutation spreading in a subdivided population. Genetics 189: 213–226.
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.
Kimura, M., 1979 Model of effectively neutral mutations in which selective constraint is incorporated. Proc. Natl. Acad. Sci. USA 76: 3440–3444.
King, R. B., and R. Lawson, 1995 Color-pattern variation in Lake Erie water snakes: the role of gene flow. Evolution 49: 885–896.
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.
Kojima, K., and H. E. Schaffer, 1967 Survival process of linked mutant genes. Evolution 21: 518–531.
Lande, R., 1984 The expected fixation rate of chromosomal inversions. Evolution 38: 743–752.
Lande, R., 1985 The fixation of chromosomal rearrangements in a subdivided population with local extinction and colonization. Heredity 54: 323–332.
Lenormand, T., and S. P. Otto, 2000 The evolution of recombination in a heterogeneous environment. Genetics 156: 423–438.
Lessard, S., and A. R. Kermany, 2012 Fixation probability in a two-locus model by the ancestral recombination–selection graph. Genetics 190: 691–707.
Lowry, D. B., J. L. Modliszewski, K. M. Wright, C. A. Wu, and J. H. Willis, 2008 The strength and genetic basis of reproductive isolating barriers in flowering plants. Philos. Trans. R. Soc. Lond. B Biol. Sci. 363: 3009–3021.
Martin, G., and T. Lenormand, 2008 The distribution of beneficial and fixed mutation fitness effects close to an optimum. Genetics 179: 907–916.
Martin, G., S. P. Otto, and T. Lenormand, 2006 Selection for recombination in structured populations. Genetics 172: 593–609.
Maynard Smith, J., and J. Haigh, 1974 Hitch-hiking effect of a favorable gene. Genet. Res. 23: 23–35.
McGaugh, S. E., and M. A. F. Noor, 2012 Genomic impacts of chromosomal inversions in parapatric drosophila species. Philos. Trans. R. Soc. Lond. B Biol. Sci. 367: 422–429.
Muller, H. J., 1932 Some genetic aspects of sex. Am. Nat. 66: 118–138.
Muller, H. J., 1942 Isolating mechanisms, evolution and temperature. Biol. Symp. 6: 71–125.
Nadeau, N. J., A. Whibley, R. T. Jones, J. W. Davey, K. K. Dasmahapatra *et al.*, 2012 Genomic islands of divergence in hybridizing *Heliconius* butterflies identified by large-scale targeted sequencing. Philos. Trans. R. Soc. Lond. B Biol. Sci. 367: 343–353.
Nadeau, N. J., S. H. Martin, K. M. Kozak, C. Salazar, K. K. Dasmahapatra *et al.*, 2013 Genome-wide patterns of divergence and gene flow across a butterfly radiation. Mol. Ecol. 22: 814–826.
Nagylaki, T., 1980 The strong-migration limit in geographically structured populations. J. Math. Biol. 9: 101–114.
Nagylaki, T., J. Hofbauer, and P. Brunovský, 1999 Convergence of multilocus systems under weak epistasis or weak selection. J. Math. Biol. 38: 103–133.
Nosil, P., and J. L. Feder, 2012 Genomic divergence during speciation: causes and consequences. Philos. Trans. R. Soc. Lond. B Biol. Sci. 367: 332–342.
Notohara, M., 1993 The strong-migration limit for the genealogical process in geographically structured populations. J. Math. Biol. 31: 115–122.
Orr, H. A., 1998 The population genetics of adaptation: the distribution of factors fixed during adaptive evolution. Evolution 52: 935–949.

Orr, H. A., 2002 The population genetics of adaptation: the adaptation of DNA sequences. Evolution 56: 1317–1330.
Orr, H. A., 2003 The distribution of fitness effects among beneficial mutations. Genetics 163: 1519–1526.
Orr, H. A., 2010 The population genetics of beneficial mutations. Philos. Trans. R. Soc. Lond. B Biol. Sci. 365: 1195–1201.
Otto, S. P., and N. H. Barton, 1997 The evolution of recombination: removing the limits to natural selection. Genetics 147: 879–906.
Pennings, P. S., and J. Hermisson, 2006a Soft sweeps II: molecular population genetics of adaptation from recurrent mutation or migration. Mol. Biol. Evol. 23: 1076–1084.
Pennings, P. S., and J. Hermisson, 2006b Soft sweeps III: the signature of positive selection from recurrent mutation. PLoS Genet. 2: e186.
Petry, D., 1983 The effect on neutral gene flow of selection at a linked locus. Theor. Popul. Biol. 23: 300–313.
Presgraves, D. C., 2008 Sex chromosomes and speciation in *Drosophila*. Trends Genet. 24: 336–343.
Presgraves, D. C., 2010 The molecular evolutionary basis of species formation. Nat. Rev. Genet. 11: 175–180.
Pritchard, J. K., J. K. Pickrell, and G. Coop, 2010 The genetics of human adaptation: hard sweeps, soft sweeps, and polygenic adaptation. Curr. Biol. 20: R208–R215.
Ralph, P., and G. Coop, 2010 Parallel adaptation: One or many waves of advance of an advantageous allele? Genetics 186: 647–668.
Rice, W. R., 1984 Sex chromosomes and the evolution of sexual dimorphism. Evolution 38: 735–742.
Rice, W. R., 1987 The accumulation of sexually antagonistic genes as a selective agent promoting the evolution of reduced recombination between primitive sex chromosomes. Evolution 41: 911–914.
Rouhani, S., and N. Barton, 1987 Speciation and the "shifting balance" in a continuous population. Theor. Popul. Biol. 31: 465–492.
Roze, D., and N. H. Barton, 2006 The Hill–Robertson effect and the evolution of recombination. Genetics 173: 1793–1811.
Sjödin, P., I. Kaj, S. Krone, M. Lascoux, and M. Nordborg, 2005 On the meaning and existence of an effective population size. Genetics 169: 1061–1070.
Slatkin, M., 1975 Gene flow and selection in a two-locus system. Genetics 81: 787–802.
Slatkin, M., and T. Wiehe, 1998 Genetic hitch-hiking in a subdivided population. Genet. Res. 71: 155–160.
Strasburg, J. L., N. A. Sherman, K. M. Wright, L. C. Moyle, J. H. Willis *et al.*, 2012 What can patterns of differentiation across plant genomes tell us about adaptation and speciation? Philos. Trans. R. Soc. Lond. B Biol. Sci. 367: 364–373.
Takahata, N., 1990 A simple genealogical structure of strongly balanced allelic lines and transspecies evolution of polymorphism. Proc. Natl. Acad. Sci. USA 87: 2419–2423.
Turner, T. L., M. W. Hahn, and S. V. Nuzhdin, 2005 Genomic islands of speciation in *Anopheles gambiae*. PLoS Biol. 3: e285.
Via, S., 2012 Divergence hitchhiking and the spread of genomic isolation during ecological speciation-with-gene-flow. Philos. Trans. R. Soc. Lond. B Biol. Sci. 367: 451–460.
Vicoso, B., and B. Charlesworth, 2006 Evolution on the X chromosome: unusual patterns and processes. Nat. Rev. Genet. 7: 645–653.
Wakeley, J., 2009 *Coalescent Theory: An Introduction*, Roberts & Company, Greenwood Village, CO.
Whitlock, M. C., and N. H. Barton, 1997 The effective size of a subdivided population. Genetics 146: 427–441.
Wright, S., 1940 Breeding structure of populations in relation to speciation. Am. Nat. 74: 232–248.
Yeaman, S., 2013 Genomic rearrangements and the evolution of clusters of locally adaptive loci. Proc. Natl. Acad. Sci. USA 110: 1743–1751.
Yeaman, S., and S. P. Otto, 2011 Establishment and maintenance of adaptive genetic divergence under migration, selection, and drift. Evolution 65: 2123–2129.
Yeaman, S., and M. C. Whitlock, 2011 The genetic architecture of adaptation under migration–selection balance. Evolution 65: 1897–1911.

Communicating editor: L. M. Wahl

GENETICS

Supporting Information
http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.114.163477/-/DC1

The Effect of Linkage on Establishment and Survival of Locally Beneficial Mutations

Simon Aeschbacher and Reinhard Bürger

DOI: 10.1534/genetics.114.163477

Table S1 **Simulated sojourn and absorption times and comparison to diffusion approximation. (Electronically only) Available for download at http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.114.163477/-/DC1**

Table S2 **The effect of assuming p_0 small in the diffusion approximation of the mean absorption time.**

		$N_e = 100$				$N_e = 10^3$				$N_e = 10^4$			
r	m	$q_c = 0$	$q_c = 0.2$	$q_c = 0.5$	$q_c = 0.8$	$q_c = 0$	$q_c = 0.2$	$q_c = 0.5$	$q_c = 0.8$	$q_c = 0$	$q_c = 0.2$	$q_c = 0.5$	$q_c = 0.8$
0.05	0.006	0.017	0.016	0.015	0.014	0.017	0.017	0.016	0.015	0.017	0.017	0.016	0.015
0.05	0.012	0.014	0.012	0.010	0.009	0.014	0.012	0.011	0.009	0.014	0.012	0.011	0.009
0.05	0.018	0.009	0.007	0.005	0.003	0.010	0.008	0.005	0.003	0.010	0.008	0.005	0.003
0.05	0.024	0.004	0.002	0.000	−0.002	0.004	0.002	0.000	−0.002	0.004	0.002	0.000	−0.002
0.10	0.006	0.016	0.015	0.015	0.014	0.016	0.015	0.015	0.014	0.016	0.015	0.015	0.014
0.10	0.012	0.011	0.010	0.009	0.008	0.011	0.010	0.009	0.009	0.011	0.010	0.009	0.009
0.10	0.018	0.005	0.005	0.004	0.002	0.006	0.005	0.004	0.003	0.006	0.005	0.004	0.003
0.10	0.024	0.000	−0.001	−0.002	−0.003	0.000	−0.001	−0.002	−0.003	0.000	−0.001	−0.002	−0.003
0.20	0.006	0.015	0.015	0.014	0.014	0.015	0.015	0.015	0.014	0.015	0.015	0.015	0.014
0.20	0.012	0.009	0.009	0.008	0.008	0.010	0.009	0.009	0.008	0.010	0.009	0.009	0.008
0.20	0.018	0.004	0.003	0.003	0.002	0.004	0.003	0.003	0.002	0.004	0.004	0.003	0.002
0.20	0.024	−0.002	−0.002	−0.003	−0.003	−0.002	−0.002	−0.003	−0.003	−0.002	−0.002	−0.003	−0.003

The relative error $\tilde{\bar{t}}_{\mathrm{QLE}}/\bar{t}_{\mathrm{QLE}} - 1$ is tablulated. The initial frequency of the focal mutant A_1 is $p_0 = 1/(2N)$ (we assumed $N_e = N$). Other parameters are $a = 0.02$ and $b = 0.04$. For a graphical representation, see Figure S13A.

Table S3 **The effect of assuming ρ large in $M(p)$ when deriving the diffusion approximation to the mean absorption time.**

		$_e = 100$				$_e = 10^3$				$_e = 10^4$			
		$_c = 0$	$_c = 0\ 2$	$_c = 0\ 5$	$_c = 0\ 8$	$_c = 0$	$_c = 0\ 2$	$_c = 0\ 5$	$_c = 0\ 8$	$_c = 0$	$_c = 0\ 2$	$_c = 0\ 5$	$_c = 0\ 8$
0.05	0.006	3.818	1.995	0.770	0.211	6.558×10^{5}	1.892×10^{4}	242.595	6.217	1.397×10^{57}	1.503×10^{42}	4.312×10^{23}	3.395×10^{8}
0.05	0.012	1.391	0.901	0.382	0.106	1.288×10^{4}	2.086×10^{3}	77.967	3.411	8.300×10^{40}	1.700×10^{33}	1.534×10^{19}	4.275×10^{6}
0.05	0.018	0.280	0.299	0.167	0.054	27.006	41.590	5.361	0.371	6.349×10^{14}	1.583×10^{17}	1.140×10^{10}	448.323
0.05	0.024	0.025	0.118	0.092	0.035	−0.014	0.620	0.235	0.056	−0.427	1.367×10^{4}	1.808	0.075
0.10	0.006	0.404	0.301	0.169	0.060	22.759	11.577	3.772	0.843	3.891×10^{13}	7.582×10^{10}	5.379×10^{6}	435.536
0.10	0.012	0.149	0.129	0.080	0.030	4.922	4.152	1.995	0.539	5.884×10^{7}	1.595×10^{7}	7.357×10^{4}	86.049
0.10	0.018	0.033	0.048	0.037	0.015	0.372	0.565	0.292	0.073	48.506	487.778	88.493	3.323
0.10	0.024	0.003	0.022	0.022	0.010	−0.008	0.038	0.037	0.015	−0.029	0.077	0.053	0.018
0.20	0.006	0.083	0.067	0.042	0.017	1.181	0.915	0.528	0.191	2.241×10^{3}	623.130	66.327	4.677
0.20	0.012	0.027	0.027	0.019	0.008	0.439	0.442	0.313	0.125	39.447	41.132	15.525	2.415
0.20	0.018	0.006	0.010	0.009	0.004	0.029	0.056	0.044	0.017	0.643	1.783	1.417	0.405
0.20	0.024	0.000	0.005	0.006	0.003	−0.001	0.006	0.008	0.004	−0.003	0.008	0.010	0.005

The relative error $\bar{}_{\mathrm{QLE},\rho\gg0}/\bar{}_{\mathrm{QLE}} - 1$ is tabulated. The initial frequency of the focal mutant $_1$ is $_0 = 1/(2\ \ )$ (we assumed $_e =\ \ $). Other parameters are $= 0\ 02$ and $= 0\ 04$. For a graphical representation, see Figure S13B.

Table S4 **The effect of assuming p_0 small in the diffusion approximation to the mean absorption time, given the assumption of ρ large in $M(p)$.**

		$_e = 100$				$_e = 10^3$				$_e = 10^4$			
		$_c = 0$	$_c = 0\ 2$	$_c = 0\ 5$	$_c = 0\ 8$	$_c = 0$	$_c = 0\ 2$	$_c = 0\ 5$	$_c = 0\ 8$	$_c = 0$	$_c = 0\ 2$	$_c = 0\ 5$	$_c = 0\ 8$
0.05	0.006	0.018	0.017	0.016	0.015	0.018	0.017	0.016	0.015	0.018	0.017	0.016	0.015
0.05	0.012	0.015	0.013	0.011	0.009	0.015	0.014	0.012	0.009	0.015	0.014	0.012	0.009
0.05	0.018	0.010	0.009	0.006	0.004	0.010	0.009	0.007	0.004	0.010	0.009	0.007	0.004
0.05	0.024	0.003	0.003	0.001	−0.002	0.004	0.004	0.001	−0.002	0.004	0.004	0.001	−0.002
0.10	0.006	0.016	0.016	0.015	0.014	0.016	0.016	0.015	0.014	0.016	0.016	0.015	0.014
0.10	0.012	0.011	0.010	0.009	0.008	0.011	0.011	0.010	0.009	0.011	0.011	0.010	0.009
0.10	0.018	0.006	0.005	0.004	0.003	0.006	0.005	0.004	0.003	0.006	0.005	0.004	0.003
0.10	0.024	0.000	0.000	−0.001	−0.003	0.000	0.000	−0.001	−0.003	0.000	0.000	−0.001	−0.003
0.20	0.006	0.015	0.015	0.014	0.014	0.015	0.015	0.015	0.014	0.015	0.015	0.015	0.014
0.20	0.012	0.009	0.009	0.009	0.008	0.010	0.009	0.009	0.008	0.010	0.009	0.009	0.008
0.20	0.018	0.004	0.003	0.003	0.002	0.004	0.004	0.003	0.002	0.004	0.004	0.003	0.002
0.20	0.024	−0.002	−0.002	−0.002	−0.003	−0.002	−0.002	−0.002	−0.003	−0.002	−0.002	−0.002	−0.003

The relative error $\tilde{}_{\mathrm{QLE},\rho\gg0}/\bar{}_{\mathrm{QLE},\rho\gg0} - 1$ is tabulated. The initial frequency of the focal mutant $_1$ is $_0 = 1/(2\ \ )$ (we assumed $_e =\ \ $). Other parameters are $= 0\ 02$ and $= 0\ 04$. For a graphical representation, see Figure S13C.

Table S5 The mean absorption time under the QLE approximation relative to the one without linkage.

		$N_e = 100$				$N_e = 10^3$				$N_e = 10^4$			
r	m	$q_c = 0$	$q_c = 0.2$	$q_c = 0.5$	$q_c = 0.8$	$q_c = 0$	$q_c = 0.2$	$q_c = 0.5$	$q_c = 0.8$	$q_c = 0$	$q_c = 0.2$	$q_c = 0.5$	$q_c = 0.8$
0.05	0.006	3.887	2.738	1.763	1.228	1.377×10^{6}	4.280×10^{4}	468.965	9.837	1.037×10^{181}	1.053×10^{166}	3.122×10^{146}	5.560×10^{129}
0.05	0.012	2.563	1.897	1.386	1.116	3.898×10^{10}	7.261×10^{8}	8.047×10^{6}	2.895×10^{5}	3.186×10^{94}	1.105×10^{77}	1.511×10^{57}	2.154×10^{42}
0.05	0.018	1.679	1.407	1.187	1.059	2.092×10^{4}	1.212×10^{3}	118.033	46.914	5.463×10^{38}	3.218×10^{25}	1.172×10^{13}	1.268×10^{6}
0.05	0.024	1.335	1.224	1.111	1.037	246.649	150.044	111.839	97.857	NA	NA	NA	NA
0.10	0.006	2.183	1.815	1.418	1.140	4.032×10^{3}	608.988	45.294	4.288	6.045×10^{155}	4.043×10^{147}	2.273×10^{136}	1.374×10^{126}
0.10	0.012	1.580	1.403	1.209	1.071	5.262×10^{7}	9.268×10^{6}	9.804×10^{5}	1.497×10^{5}	2.960×10^{65}	6.234×10^{57}	6.338×10^{47}	2.248×10^{39}
0.10	0.018	1.256	1.189	1.104	1.037	224.875	118.051	60.712	41.759	7.771×10^{16}	1.157×10^{13}	2.039×10^{8}	1.085×10^{5}
0.10	0.024	1.131	1.109	1.064	1.024	116.181	111.037	102.175	95.712	NA	NA	NA	NA
0.20	0.006	1.519	1.386	1.218	1.079	91.941	34.912	8.699	2.325	2.582×10^{139}	1.660×10^{135}	1.588×10^{129}	3.019×10^{123}
0.20	0.012	1.247	1.188	1.108	1.040	1.550×10^{6}	7.361×10^{5}	2.525×10^{5}	9.391×10^{4}	6.923×10^{49}	3.274×10^{46}	5.148×10^{41}	1.717×10^{37}
0.20	0.018	1.112	1.091	1.055	1.021	63.566	55.483	45.480	38.849	4.712×10^{8}	3.624×10^{7}	6.508×10^{5}	2.380×10^{4}
0.20	0.024	1.059	1.054	1.035	1.014	101.157	100.356	97.307	94.249	NA	NA	NA	NA

Tabulated is the ratio $\bar{t}_{\mathrm{QLE}}/\bar{t}_{\mathrm{OLM}}$. The initial frequency of the focal mutant A_1 is $p_0 = 1/(2N)$ (we assumed $N_e = N$). Other parameters are $a = 0.02$ and $b = 0.04$. NA denotes cases where $\bar{t}_{\mathrm{OLM}}$ is numerically indistinguishable from 0 and hence the ratio $\bar{t}_{\mathrm{QLE}}/\bar{t}_{\mathrm{OLM}}$ not defined.

Table S6 **The error of $\tilde{\bar{t}}_{\mathrm{QLE}}$ relative to $\bar{t}_{\mathrm{QLE}}$ as in Table S2, but for $p_0 = 0.005$ fixed instead of $p_0 = 1/(2N)$.**

		e = 100				e = 10^3				e = 10^4			
		c = 0	c = 0 2	c = 0 5	c = 0 8	c = 0	c = 0 2	c = 0 5	c = 0 8	c = 0	c = 0 2	c = 0 5	c = 0 8
0.05	0.006	0.017	0.016	0.015	0.014	0.182	0.174	0.163	0.153	2.551	2.418	2.239	2.079
0.05	0.012	0.014	0.012	0.010	0.009	0.144	0.128	0.108	0.092	1.943	1.690	1.389	1.144
0.05	0.018	0.009	0.007	0.005	0.003	0.099	0.077	0.051	0.030	1.249	0.934	0.599	0.347
0.05	0.024	0.004	0.002	0.000	−0.002	0.042	0.020	−0.004	−0.023	0.492	0.229	−0.039	−0.175
0.10	0.006	0.016	0.015	0.015	0.014	0.166	0.162	0.156	0.150	2.297	2.228	2.131	2.039
0.10	0.012	0.011	0.010	0.009	0.008	0.115	0.108	0.097	0.088	1.490	1.379	1.227	1.088
0.10	0.018	0.005	0.005	0.004	0.002	0.059	0.050	0.037	0.025	0.698	0.590	0.434	0.295
0.10	0.024	0.000	−0.001	−0.002	−0.003	0.000	−0.005	−0.016	−0.027	−0.003	−0.047	−0.131	−0.199
0.20	0.006	0.015	0.015	0.014	0.014	0.157	0.155	0.152	0.148	2.148	2.113	2.062	2.012
0.20	0.012	0.009	0.009	0.008	0.008	0.099	0.096	0.090	0.085	1.250	1.200	1.124	1.050
0.20	0.018	0.004	0.003	0.003	0.002	0.038	0.035	0.029	0.022	0.445	0.404	0.331	0.257
0.20	0.024	−0.002	−0.002	−0.003	−0.003	−0.018	−0.019	−0.024	−0.029	−0.141	−0.149	−0.181	−0.215

The relative error is computed as $\tilde{\bar{t}}_{\mathrm{QLE}}/\bar{t}_{\mathrm{QLE}} - 1$. It quantifies the effect of assuming p_0 small in the derivation of the diffusion approximation of the mean absorption time. Other parameters are $h = 0.02$ and $s = 0.04$. For a graphical representation, see Figure S13D.

Table S7 **The error of $\tilde{\bar{t}}_{\mathrm{QLE},\rho\gg 0}$ relative to $\bar{t}_{\mathrm{QLE}}$ as in Table S3, but for $p_0 = 0.005$ fixed instead of $p_0 = 1/(2N)$.**

		$N_e = 100$				$N_e = 10^3$				$N_e = 10^4$			
r	m	$q_c = 0$	$q_c = 0.2$	$q_c = 0.5$	$q_c = 0.8$	$q_c = 0$	$q_c = 0.2$	$q_c = 0.5$	$q_c = 0.8$	$q_c = 0$	$q_c = 0.2$	$q_c = 0.5$	$q_c = 0.8$
0.05	0.006	3.818	1.995	0.770	0.211	6.508×10^{5}	1.879×10^{4}	241.353	6.201	1.336×10^{57}	1.441×10^{42}	4.175×10^{23}	3.344×10^{8}
0.05	0.012	1.391	0.901	0.382	0.106	1.278×10^{4}	2.064×10^{3}	77.230	3.392	7.871×10^{40}	1.579×10^{33}	1.430×10^{19}	4.120×10^{6}
0.05	0.018	0.280	0.299	0.167	0.054	26.941	41.405	5.751	0.457	6.210×10^{14}	1.430×10^{17}	1.010×10^{10}	419.102
0.05	0.024	0.025	0.118	0.092	0.035	−0.011	0.830	0.347	0.086	−0.393	1.274×10^{4}	4.193	0.192
0.10	0.006	0.404	0.301	0.169	0.060	22.709	11.552	3.765	0.842	3.840×10^{13}	7.486×10^{10}	5.327×10^{6}	433.553
0.10	0.012	0.149	0.129	0.080	0.030	4.909	4.137	1.987	0.537	5.790×10^{7}	1.559×10^{7}	7.196×10^{4}	85.052
0.10	0.018	0.033	0.048	0.037	0.015	0.388	0.611	0.342	0.092	48.191	472.445	85.111	3.239
0.10	0.024	0.003	0.022	0.022	0.010	−0.010	0.056	0.056	0.023	−0.058	0.174	0.131	0.046
0.20	0.006	0.083	0.067	0.042	0.017	1.179	0.914	0.528	0.191	2.233×10^{3}	620.925	66.145	4.670
0.20	0.012	0.027	0.027	0.019	0.008	0.438	0.441	0.312	0.125	39.267	40.862	15.423	2.404
0.20	0.018	0.006	0.010	0.009	0.004	0.034	0.067	0.055	0.022	0.640	1.758	1.391	0.401
0.20	0.024	0.000	0.005	0.006	0.003	−0.002	0.010	0.012	0.006	−0.006	0.020	0.024	0.012

The relative error is computed as $\tilde{\bar{t}}_{\mathrm{QLE},\rho\gg 0}/\bar{t}_{\mathrm{QLE}} - 1$. It quantifies the effect of assuming ρ very large in $M(p)$ when deriving the diffusion approximation of the mean absorption time. Other parameters are $a = 0.02$ and $b = 0.04$. For a graphical representation, see Figure S13E.

Table S8 **The error of $\tilde{\tilde{t}}_{\mathrm{QLE},\rho\gg 0}$ relative to $\tilde{\tilde{t}}_{\mathrm{QLE},\rho\gg 0}$ as in Table S4, but for $p_0 = 0.005$ fixed instead of $p_0 = 1/(2N)$.**

		$_e = 100$				$_e = 10^3$				$_e = 10^4$			
		$_c = 0$	$_c = 0\,2$	$_c = 0\,5$	$_c = 0\,8$	$_c = 0$	$_c = 0\,2$	$_c = 0\,5$	$_c = 0\,8$	$_c = 0$	$_c = 0\,2$	$_c = 0\,5$	$_c = 0\,8$
0.05	0.006	0.018	0.017	0.016	0.015	0.192	0.183	0.169	0.156	2.716	2.569	2.347	2.126
0.05	0.012	0.015	0.013	0.011	0.009	0.154	0.141	0.120	0.097	2.107	1.900	1.566	1.226
0.05	0.018	0.010	0.009	0.006	0.004	0.102	0.092	0.068	0.038	1.298	1.145	0.806	0.443
0.05	0.024	0.003	0.003	0.001	−0.002	0.035	0.035	0.013	−0.015	0.410	0.410	0.141	−0.120
0.10	0.006	0.016	0.016	0.015	0.014	0.169	0.164	0.158	0.151	2.342	2.271	2.162	2.053
0.10	0.012	0.011	0.010	0.009	0.008	0.118	0.111	0.101	0.090	1.531	1.435	1.277	1.113
0.10	0.018	0.006	0.005	0.004	0.003	0.060	0.055	0.042	0.028	0.709	0.642	0.491	0.323
0.10	0.024	0.000	0.000	−0.001	−0.003	−0.002	−0.002	−0.011	−0.024	−0.016	−0.016	−0.096	−0.183
0.20	0.006	0.015	0.015	0.014	0.014	0.158	0.155	0.152	0.149	2.160	2.124	2.071	2.016
0.20	0.012	0.009	0.009	0.009	0.008	0.100	0.097	0.091	0.086	1.260	1.214	1.138	1.057
0.20	0.018	0.004	0.003	0.003	0.002	0.038	0.036	0.030	0.023	0.447	0.416	0.346	0.266
0.20	0.024	−0.002	−0.002	−0.002	−0.003	−0.018	−0.018	−0.023	−0.029	−0.143	−0.143	−0.173	−0.211

The relative error is computed as $\tilde{\tilde{}}_{\mathrm{QLE},\rho\gg 0}/\tilde{\tilde{}}_{\mathrm{QLE},\rho\gg 0} - 1$. It quantifies the effect of assuming $_0$ small, given the assumption of ρ large in ( ) when deriving the diffusion approximation of the mean absorption time. Other parameters are $= 0\ 02$ and $= 0\ 04$. For a graphical representation, see Figure S13F.

Table S9 **The ratio $\bar{t}_{\mathrm{QLE}}/\bar{t}_{\mathrm{OLM}}$ as in Table S5, but for $p_0 = 0.005$ instead of $p_0 = 1/(2N)$.**

		$N_e = 100$				$N_e = 10^3$				$N_e = 10^4$			
r	m	$q_c = 0$	$q_c = 0.2$	$q_c = 0.5$	$q_c = 0.8$	$q_c = 0$	$q_c = 0.2$	$q_c = 0.5$	$q_c = 0.8$	$q_c = 0$	$q_c = 0.2$	$q_c = 0.5$	$q_c = 0.8$
0.05	0.006	3.887	2.738	1.763	1.228	1.340×10^{6}	4.191×10^{4}	463.139	9.790	2.964×10^{180}	3.124×10^{165}	9.765×10^{145}	1.828×10^{129}
0.05	0.012	2.563	1.897	1.386	1.116	3.451×10^{10}	6.511×10^{8}	7.330×10^{6}	2.672×10^{5}	1.095×10^{94}	4.148×10^{76}	6.376×10^{56}	1.011×10^{42}
0.05	0.018	1.679	1.407	1.187	1.059	2.634×10^{4}	1.542×10^{3}	141.905	49.620	2.614×10^{38}	1.786×10^{25}	7.849×10^{12}	1.006×10^{6}
0.05	0.024	1.335	1.224	1.111	1.037	351.786	190.096	126.558	103.788	NA	NA	NA	NA
0.10	0.006	2.183	1.815	1.418	1.140	3.971×10^{3}	601.729	44.963	4.276	1.858×10^{155}	1.269×10^{147}	7.351×10^{135}	4.574×10^{125}
0.10	0.012	1.580	1.403	1.209	1.071	4.768×10^{7}	8.447×10^{6}	9.010×10^{5}	1.386×10^{5}	1.199×10^{65}	2.641×10^{57}	2.866×10^{47}	1.083×10^{39}
0.10	0.018	1.256	1.189	1.104	1.037	280.052	142.014	67.583	42.921	4.907×10^{16}	7.793×10^{12}	1.521×10^{8}	8.927×10^{4}
0.10	0.024	1.131	1.109	1.064	1.024	133.739	125.264	110.775	100.350	NA	NA	NA	NA
0.20	0.006	1.519	1.386	1.218	1.079	91.196	34.687	8.664	2.321	8.305×10^{138}	5.399×10^{134}	5.247×10^{128}	1.014×10^{123}
0.20	0.012	1.247	1.188	1.108	1.040	1.423×10^{6}	6.775×10^{5}	2.333×10^{5}	8.711×10^{4}	3.099×10^{49}	1.498×10^{46}	2.439×10^{41}	8.426×10^{36}
0.20	0.018	1.112	1.091	1.055	1.021	71.310	60.791	47.766	39.139	3.490×10^{8}	2.761×10^{7}	5.223×10^{5}	1.997×10^{4}
0.20	0.024	1.059	1.054	1.035	1.014	109.131	107.834	102.913	98.013	NA	NA	NA	NA

Here, $\bar{t}_{\mathrm{QLE}}$ is the mean absorption time assuming quasi-linkage equilibrium (QLE), and $\bar{t}_{\mathrm{OLM}}$ the one for the one-locus model (no linkage). Parameters are $a = 0.02$ and $b = 0.04$. NA denotes cases where $\bar{t}_{\mathrm{OLM}}$ is numerically indistinguishable from 0 and hence the ratio $\bar{t}_{\mathrm{QLE}}/\bar{t}_{\mathrm{OLM}}$ not defined.

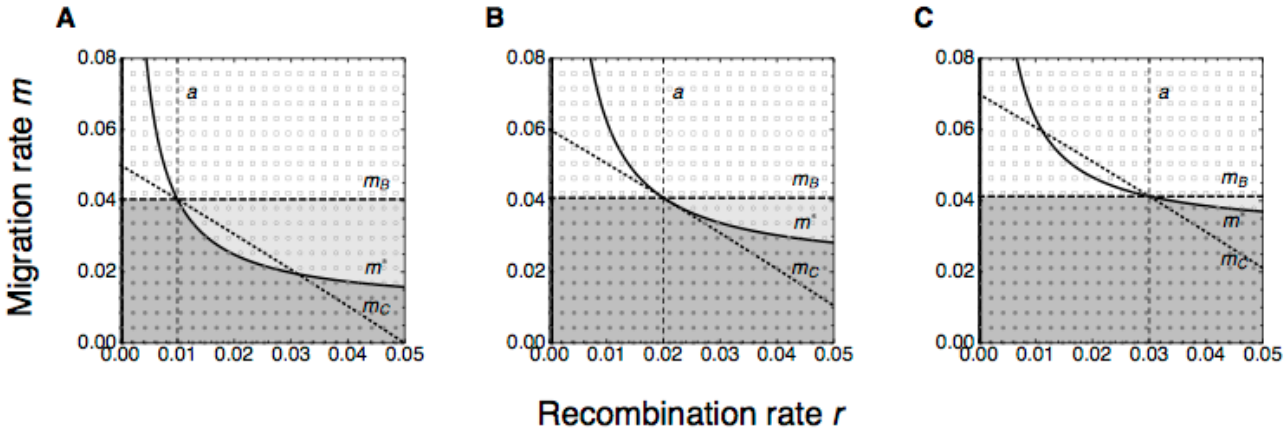

Figure S1 Critical migration rates and invasion of A_1 for a monomorphic continent. Dark grey: invasion of A_1 via the unstable marginal equilibrium E_{B}; light grey: no invasion of A_1, stable marginal equilibrium E_{B}; white: no invasion, fixation of continental haplotype A_2B_2 and convergence to the monomorphic equilibrium E_C, at which the island population is fixed for the continental haplotype A_2B_2. Numerical iterations of invasion dynamics where performed at coordinates indicated by grey symbols (File S2). Different symbols show which equilibrium is reached: • E_+; ∘ E_{B}; ▫ E_C. Initial values for iterations were $\{p_0, q_0, D_0\} = \{0, \hat{q}_{\mathrm{B}}, 0\}$, where $\hat{q}_{\mathrm{B}}$ is the frequency of B_1 at E_{B}. Iterations were stopped when successive changes in each coordinate became smaller than the numerical machine precision. The thick, almost-vertical line close to $r = 0$ is for the critical migration rate m^*. This curve crosses the r axis at $r = a(b-a)/(1-2a+b)$, which is denoted by a vertical dashed line that can hardly be seen. The second vertical dashed line corresponds to $r = a$. (A) $a = 0.01, b = 0.04$. (B) $a = 0.02, b = 0.04$. (C) $a = 0.03, b = 0.04$.

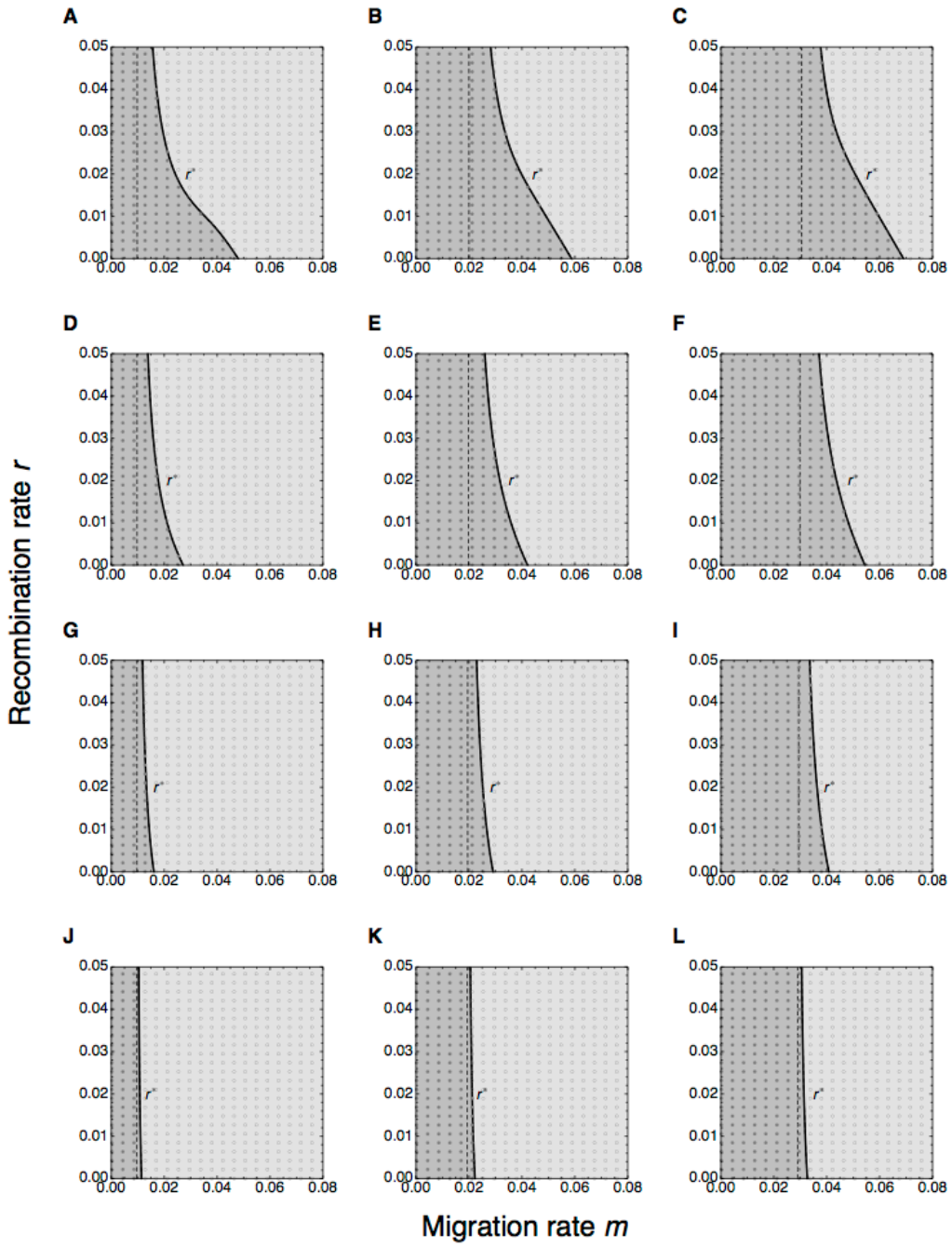

Figure S2 Critical recombination rate and invasion of A_1 for a polymorphic continent. Dark grey: invasion of A_1 via the unstable marginal equilibrium E_{B}; light grey: no invasion of A_1, stable marginal equilibrium E_{B}. Numerical iterations of invasion dynamics where performed at coordinates indicated by grey symbols (File S2). Different symbols show which equilibrium is reached: • E_+; ∘ E_{B}. Initial values for iterations were $\{p_0, q_0, D_0\} = \{0, \hat{q}_{\mathrm{B}}, 0\}$, where $\hat{q}_{\mathrm{B}}$ is the frequency of B_1 at E_{B}. The vertical dashed line indicates the pole of the function $r^*(m)$ from Eq. (41). In the left column (A, D, G, and J), the selection coefficients are $a = 0.01$, $b = 0.04$ ($a < b/2$), in the middle column (B, E, H, and K) they are $a = 0.02$, $b = 0.04$ ($a = b/2$), and in the right column (C, F, I, and L) they are $a = 0.03$, $b = 0.04$ ($a > b/2$). From top to bottom, the continental frequency of B_1 increases and takes values of $q_c = 0.01$ in (A)–(C), $q_c = 0.2$ in (D)–(F), $q_c = 0.5$ in (G)–(I), and $q_c = 0.8$ in (J)–(L).

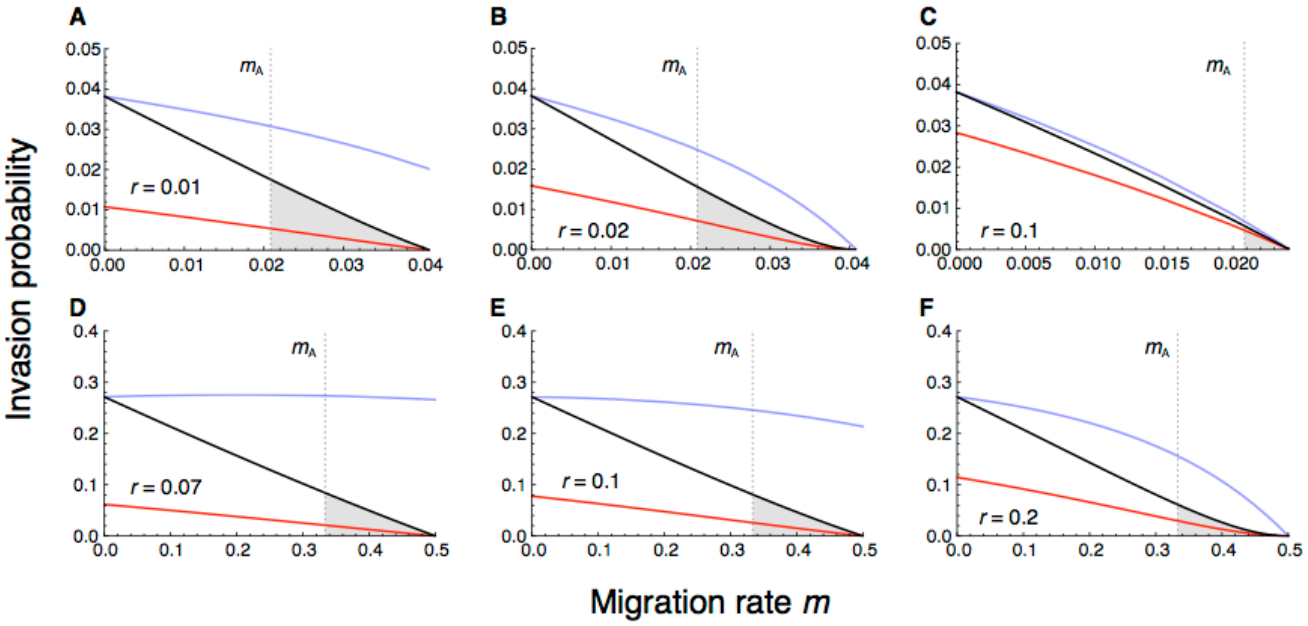

Figure S3 Invasion probability of A_1 as a function of the migration rate for a monomorphic continent. Shown are numerical solutions to the branching process, conditional on initial occurrence of A_1 on background B_1 (blue), on background B_2 (red), and when averaged across backgrounds with weights determined by the equilibrium frequency $\hat{q}_B$ of B_1 (black). The vertical dashed line shows $m_A = a/(1-b)$, the critical migration rate below which A_1 can invade without linkage to the background locus. The shaded area thus indicates where A_1 has a non-zero average invasion probability exclusively due to linkage to locus B. (A)–(C) Weak selection: $a = 0.02$, $b = 0.04$ and r as given in the panels. (D)–(F) Strong selection: $a = 0.2$, $b = 0.4$ and r as given in the panels. In this case, if linkage is tight (r small), the invasion probability conditional on the beneficial background increases with m as long as m is sufficiently small, and only starts decreasing if m is much larger (blue curve in panel D). This is because migration reduces the fitness of the resident population (consisting of A_2B_1 and A_2B_2) more strongly than it reduces the marginal fitness of type A_1B_1, which is favourable to type A_1B_1. As migration becomes stronger, though, the reduction in marginal fitness of A_1B_1 becomes dominant. The parameter combination in (D) was arbitrarily chosen to illustrate this effect (for a detailed explanation, see section 5 of File S1). For $r < 0.07$, the maximum of the blue curve is shifted further to the right.

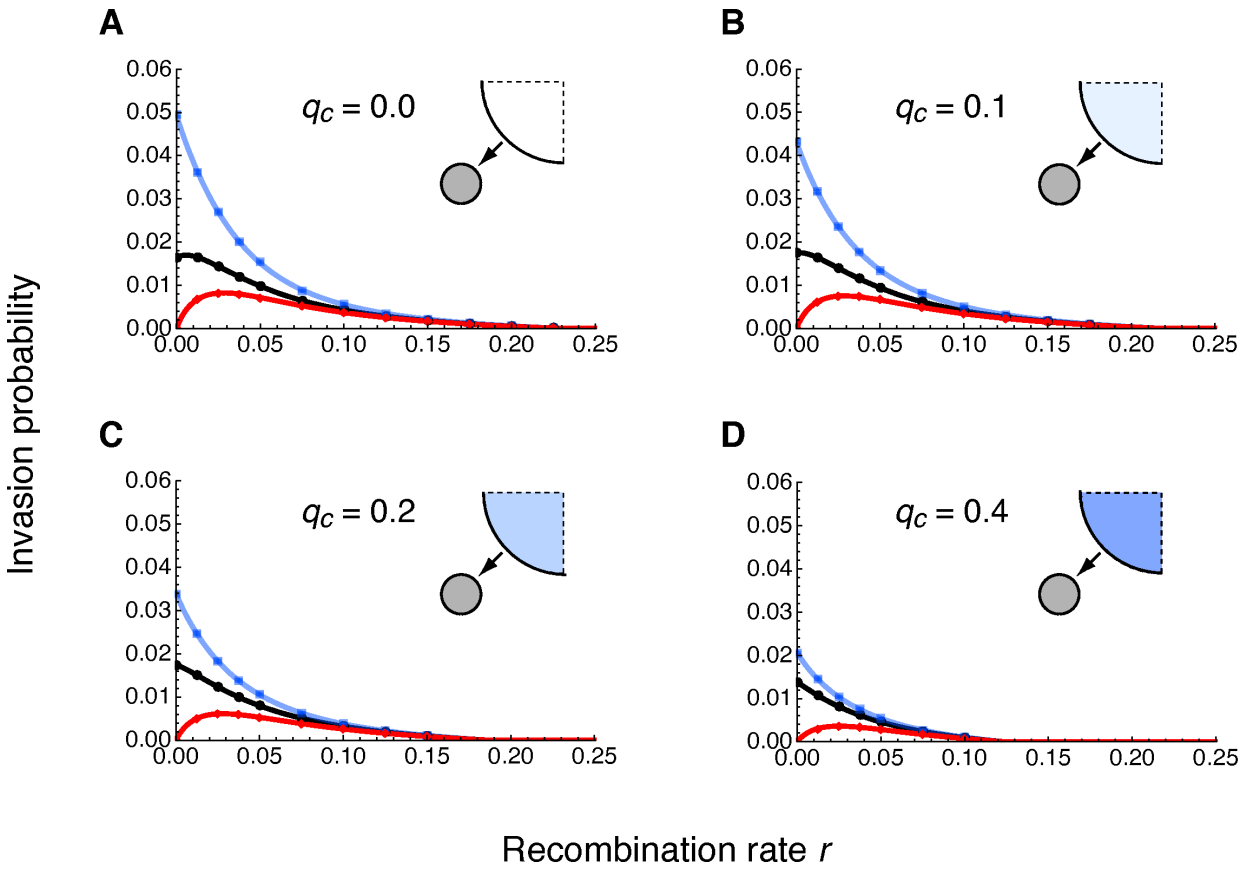

Figure S4 Invasion probability of A_1 as a function of the recombination rate and the continental frequency of B_1. Panels are for different values of the continental frequency q_c of the beneficial background allele (B_1). Curves show numerical solutions to the branching process (Eq. 61), conditional on initial occurrence of A_1 on background B_1 (blue), on background B_2 (red), and when averaged across backgrounds with weights determined by the equilibrium frequency $\hat{q}_B$ of B_1 (black). Dots represent the point estimates across 10^6 simulations under the branching-process assumptions (see Methods). Error bars span twice the standard error on each side of the point estimates, but are too short to be visible. Parameters other than q_c are the same in all panels: $a = 0.03$, $b = 0.04$ and $m = 0.032$.

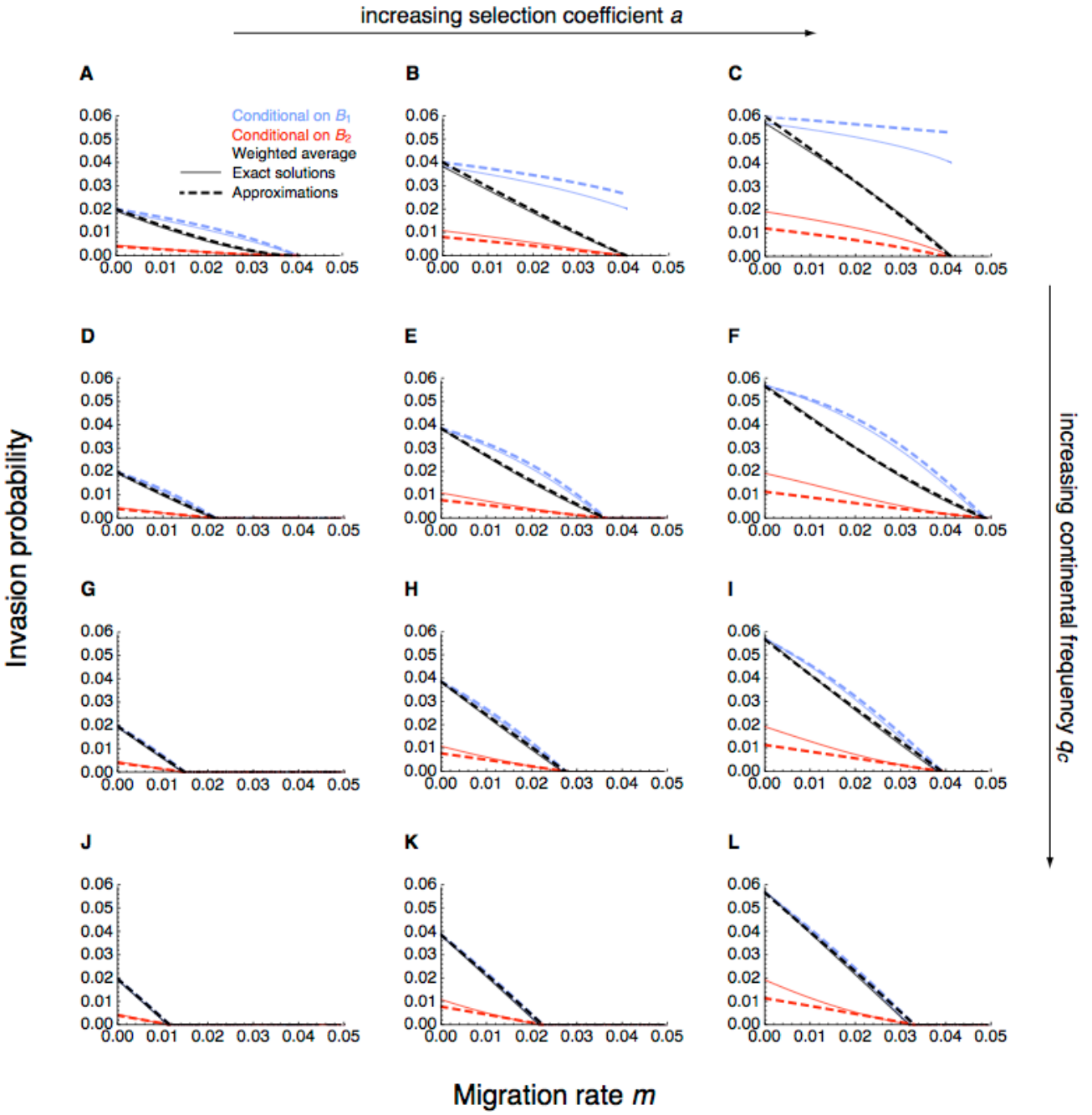

Figure S5 Invasion probability of A_1 as a function of the migration rate. Panels are for varying selective advantage a and continental frequency q_c of the beneficial background allele B_1. Invasion probabilities are shown conditional on initial occurrence of A_1 on background B_1 (blue), on background B_2 (red), and as a weighted average across the two backgrounds (black). Solid curves show exact numerical solutions to the branching process, whereas thick dashed curves show the analytical approximations valid for weak evolutionary forces and a slightly supercritical branching process (see section 4 of File S1, and Eqs. 7–9 in File S5). In all panels, $b = 0.04$ and $r = 0.01$. The selective advantage a of A_1 increases from left to right, taking values of $a = 0.01$ in panels (A), (D), (G), (J), $a = 0.02$ in panels (B), (E), (H), (K), and $a = 0.03$ in panels (C), (F), (I) and (L). The continental frequency q_c of B_1 increases from top to bottom, taking values of $q_c = 0$ in panels (A)–(C), $q_c = 0.2$ in panels (D)–(F), $q_c = 0.5$ in panels (G)–(I), and $q_c = 0.8$ in panels (J)–(L).

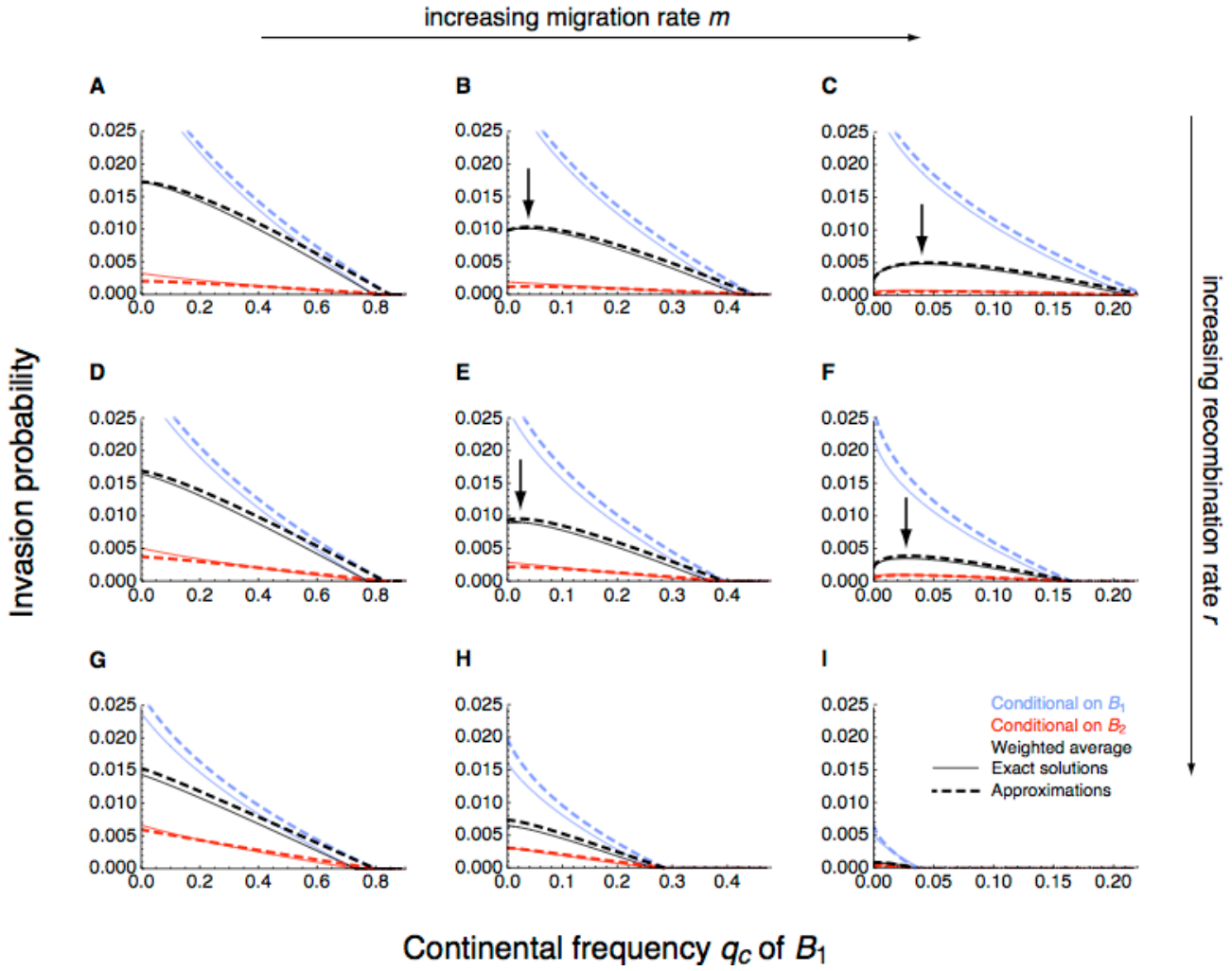

Figure S6 Invasion probability of A_1 as a function of the continental frequency of B_1. Panels are for varying migration and recombination rates. Invasion probabilities are shown conditional on initial occurrence of A_1 on background B_1 (blue), on background B_2 (red), and as a weighted average across the two backgrounds (black). Solid curves show exact numerical solutions to the branching process, whereas thick dashed curves show the analytical approximations valid for weak evolutionary forces and a slightly supercritical branching process (see section 3 of File S1, and Eqs. 7–9 in File S5). In all panels, $a = 0.02$ and $b = 0.04$. The migration rate m increases from left to right, taking values of $m = 0.022$ in panels (A), (D), (G), $m = 0.03$ in panels (B), (E), (H), and $m = 0.038$ in panels (C), (F), and (I). The recombination rate increases from top to bottom, taking values of $r = 0.005$ in panels (A)–(C), $r = 0.01$ in panels (D)–(F), and $r = 0.02$ in panels (G)–(I). Arrows indicate where the optimal q_c is non-zero.

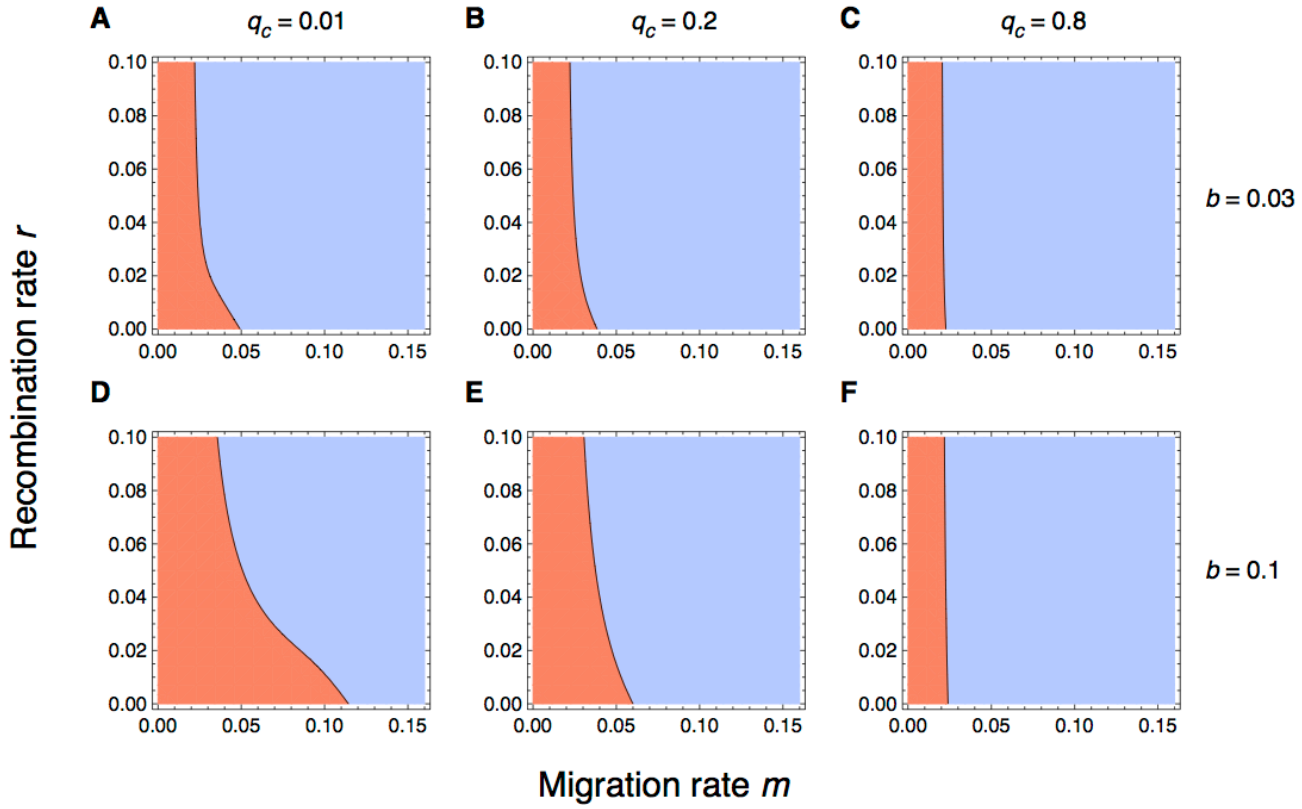

Figure S7 Asymptotic stability of the marginal one-locus migration–selection equilibrium $\tilde{E}_{\mathrm{B}}$ in continuous time I. Light blue areas indicate where $\tilde{E}_{\mathrm{B}}$ is asymptotically stable and A_1 cannot invade ($\tilde{\nu} < 0$; $\tilde{\nu}$ as in Eq. 91, File S1). Orange areas indicate where $\tilde{E}_{\mathrm{B}}$ is unstable and A_1 may potentially invade ($\tilde{\nu} > 0$). The black curve represents the critical recombination rate $\tilde{r}_B$ given in Eq. (93), as a function of the migration rate. The selection coefficient a in favour of A_1 is 0.02 throughout, the selection doefficient b in favour of B_1 is 0.03 in the first row (A–C) and 0.1 in the second (D–F). In each row, the continental frequency of B_1 increases from left to right, taking values of $q_c = 0.01$ in (A) and (D), $q_c = 0.2$ in (B) and (E), and $q_c = 0.8$ in (C) and (F).

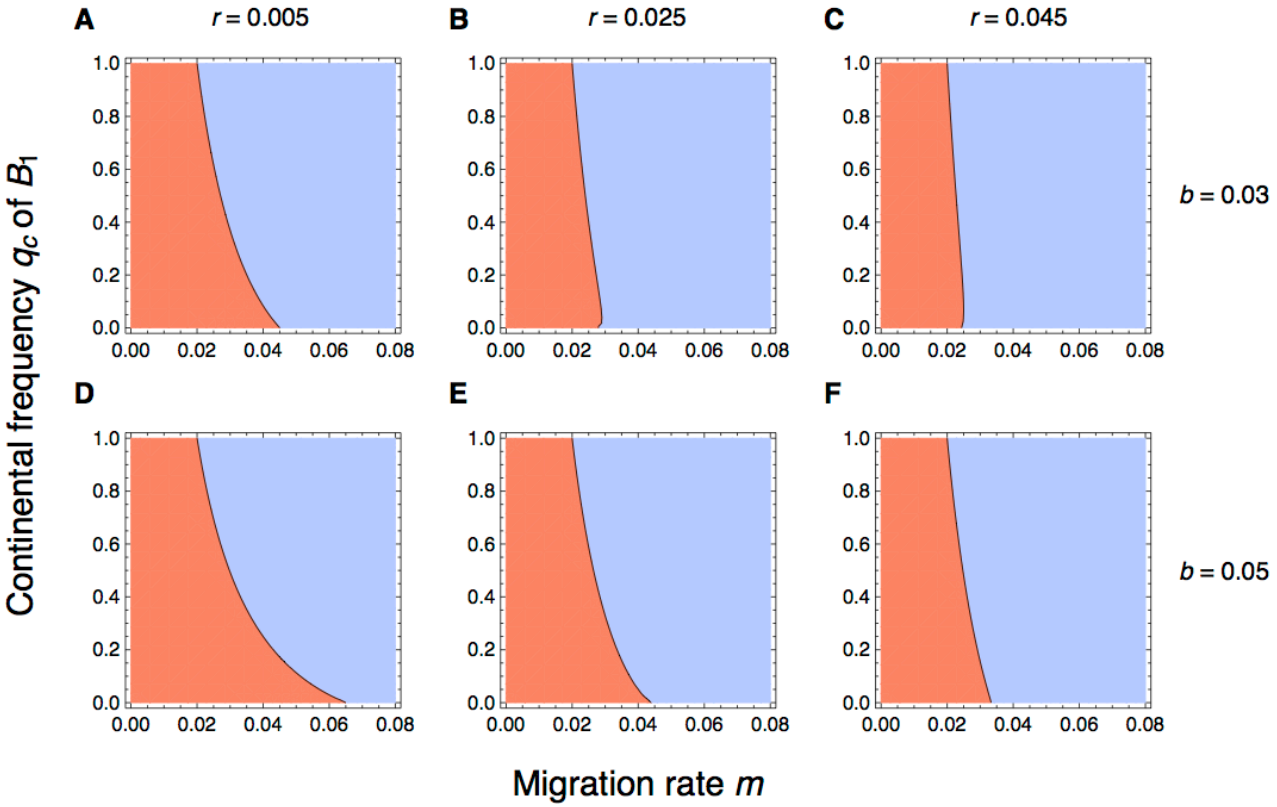

Figure S8 Asymptotic stability of the marginal one-locus migration–selection equilibrium $\tilde{E}_{\mathrm{B}}$ in continuous time II. Light blue areas indicate where $\tilde{E}_{\mathrm{B}}$ is asymptotically stable and A_1 cannot invade ($\tilde{\nu} < 0$; $\tilde{\nu}$ as in Eq. 91, File S1). Orange areas indicate where $\tilde{E}_{\mathrm{B}}$ is unstable and A_1 may potentially invade ($\tilde{\nu} > 0$). The black curve corresponds to a combination of $\tilde{q}_{c-}^{**}$ and $\tilde{q}_{c+}^{**}$ as described in section 6 of File S1, as a function of the migration rate. The selection coefficient a in favour of A_1 is 0.02 throughout, the selection doefficient b in favour of B_1 is 0.03 in the first row (A–C) and 0.05 in the second (D–F). In each row, the recombination rate increases from left to right, taking values of $r = 0.005$ in (A) and (D), $r = 0.025$ in (B) and (E), and $r = 0.045$ in (C) and (F).

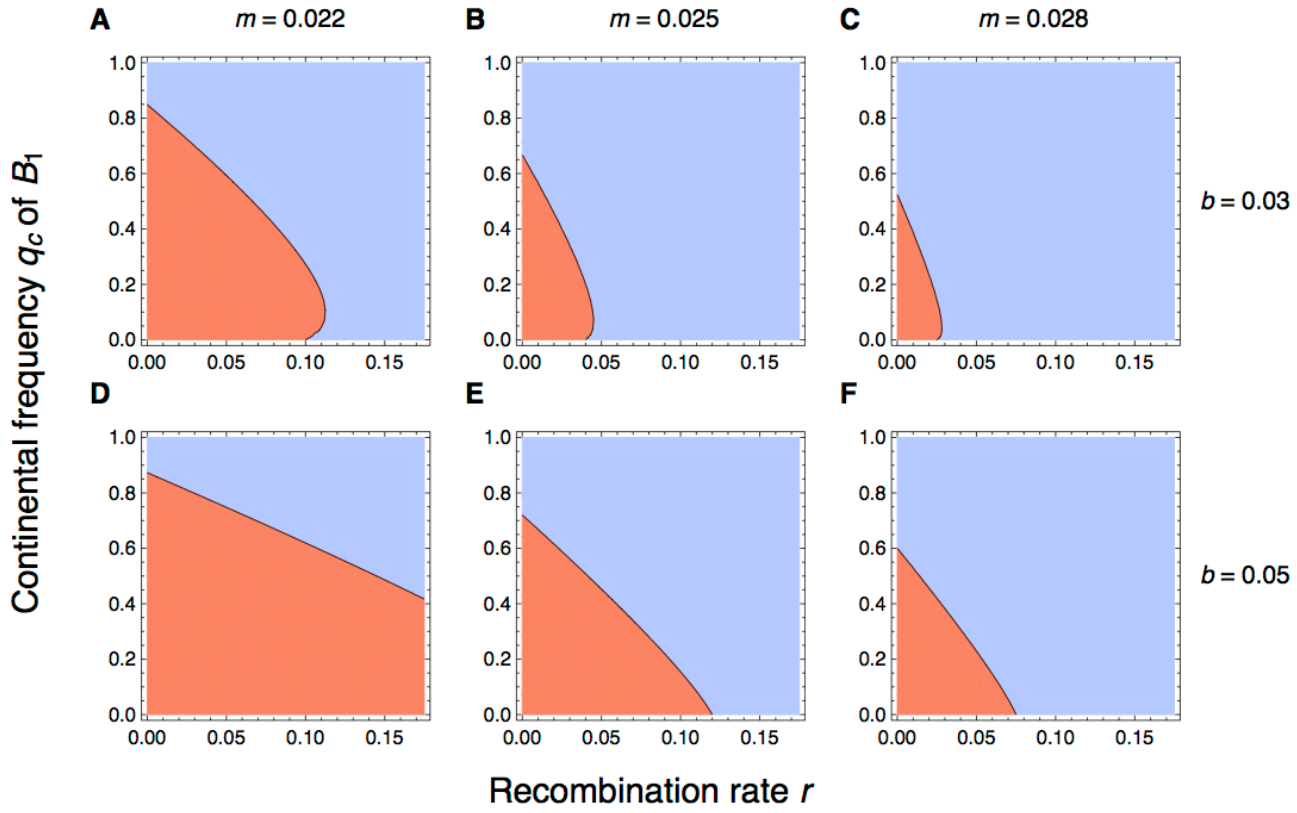

Figure S9 Asymptotic stability of the marginal one-locus migration–selection equilibrium $\tilde{E}_{\mathrm{B}}$ in continuous time III. Light blue areas indicate where $\tilde{E}_{\mathrm{B}}$ is asymptotically stable and A_1 cannot invade ($\tilde{\nu} < 0$; $\tilde{\nu}$ as in Eq. 91, File S1). Orange areas indicate where $\tilde{E}_{\mathrm{B}}$ is unstable and A_1 may potentially invade ($\tilde{\nu} > 0$). The black curve corresponds to a combination of $\tilde{q}_{c-}^{**}$ and $\tilde{q}_{c+}^{**}$ as described in section 6 of File S1, as a function of the recombination rate. The selection coefficient a in favour of A_1 is 0.02 throughout, the selection doefficient b in favour of B_1 is 0.03 in the first row (A–C) and 0.05 in the second (D–F). In each row, the migration rate increases from left to right, taking values of $m = 0.022$ in (A) and (D), $m = 0.025$ in (B) and (E), and $m = 0.028$ in (C) and (F).

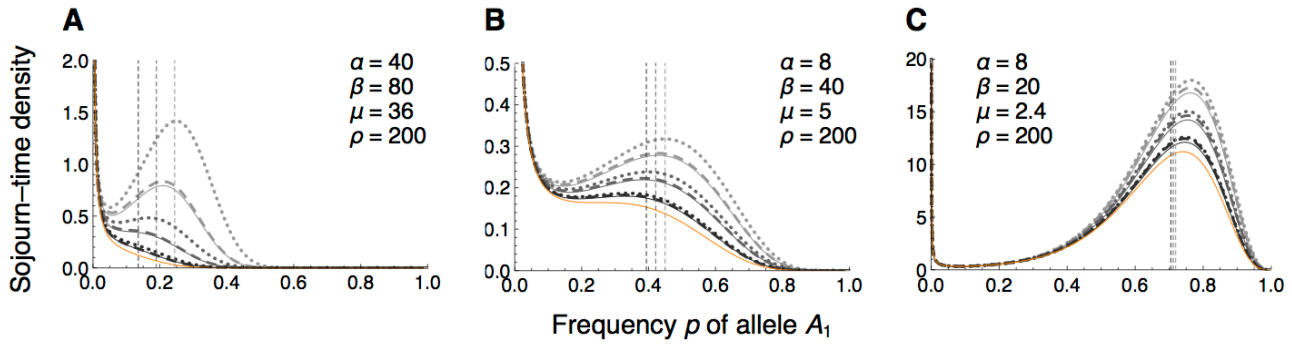

Figure S10 Diffusion approximation to the sojourn-time density of A_1 under quasi-linkage equilibrium for a polymorphic continent. Comparison of the sojourn-time density (STD) $t_{2,\mathrm{QLE}}(p;p_0)$ (thin curves, Eq. 7b) to the approximation valid for small p_0, $\tilde{t}_{2,\mathrm{QLE}}(p;p_0)$ (dashed curves, analogous to Eq. 109 in File S1) and the one based on the additional assumption of $\rho \gg 0$, $\tilde{t}_{2,\mathrm{QLE},\rho\gg 0}(p;p_0)$ (dotted curves, Eq. 119b) assuming a polymorphic continent. The continental frequency q_c of B_1 increases from light to dark grey, taking values of 0.2, 0.5, and 0.8. The STD for the one-locus model, $\tilde{t}_{2,\mathrm{OLM}}(p;p_0)$, is shown in orange as a reference. Vertical lines give the deterministic frequency $\hat{p}_+$ of A_1 at the respective fully-polymorphic equilibrium (computed in File S7). (A) Strong evolutionary forces relative to genetic drift. (B) Strong asymmetry in selection coefficients, and moderate migration. (C) Recombination ten times stronger than selection at locus B. In all panels, $p_0 = 0.005$, which corresponds to an island population of size $N = 100$ and a single initial copy of A_1. Panels (A), (B) and (C) correspond to Figures 5C, 5D and 5E for a monomorphic continent ($q_c = 0$), respectively.

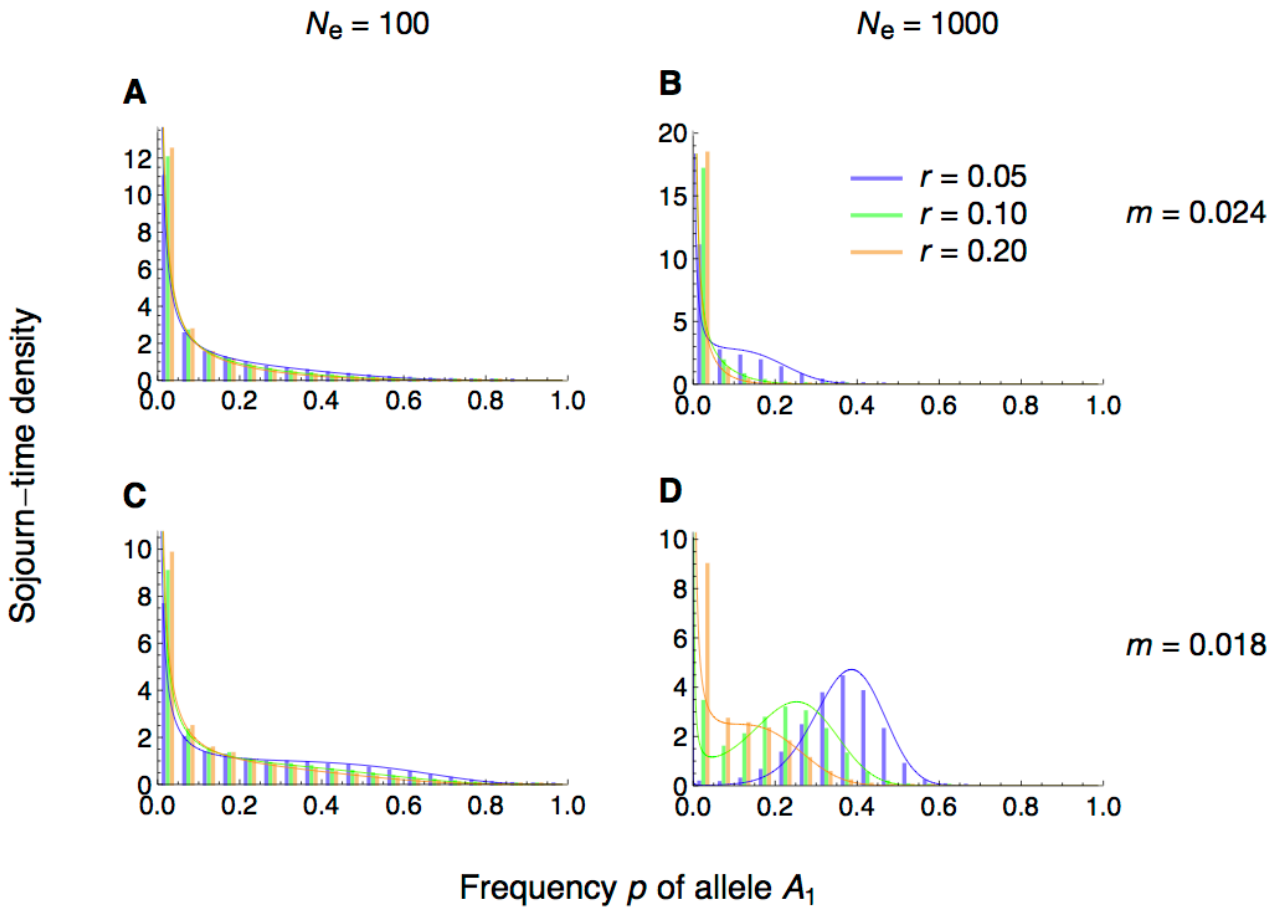

Figure S11 Comparison of analytical and simulated sojourn-time densites of A_1 for a monomorphic continent. Results are shown for various recombination rates r. Histograms were obtained from 10^6 simulations (see Methods) and curves give the diffusion approximation $\tilde{t}_{2,\mathrm{QLE}}(p;p_0)$ from Eq. (109). Throughout, $a = 0.02$, $b = 0.04$ and $p_0 = 1/(2N)$ (we assumed $N_e = N$). In the first row, migration is relatively strong compared to selection in favour of A_1 ($m = 0.024 > a$), in the second row it is relatively weak ($m = 0.018 < a$). In the left column, the effective population size is small ($N_e = 100$) and drift dominates, whereas in the right column, $N_e = 1000$ and deterministic forces become more important.

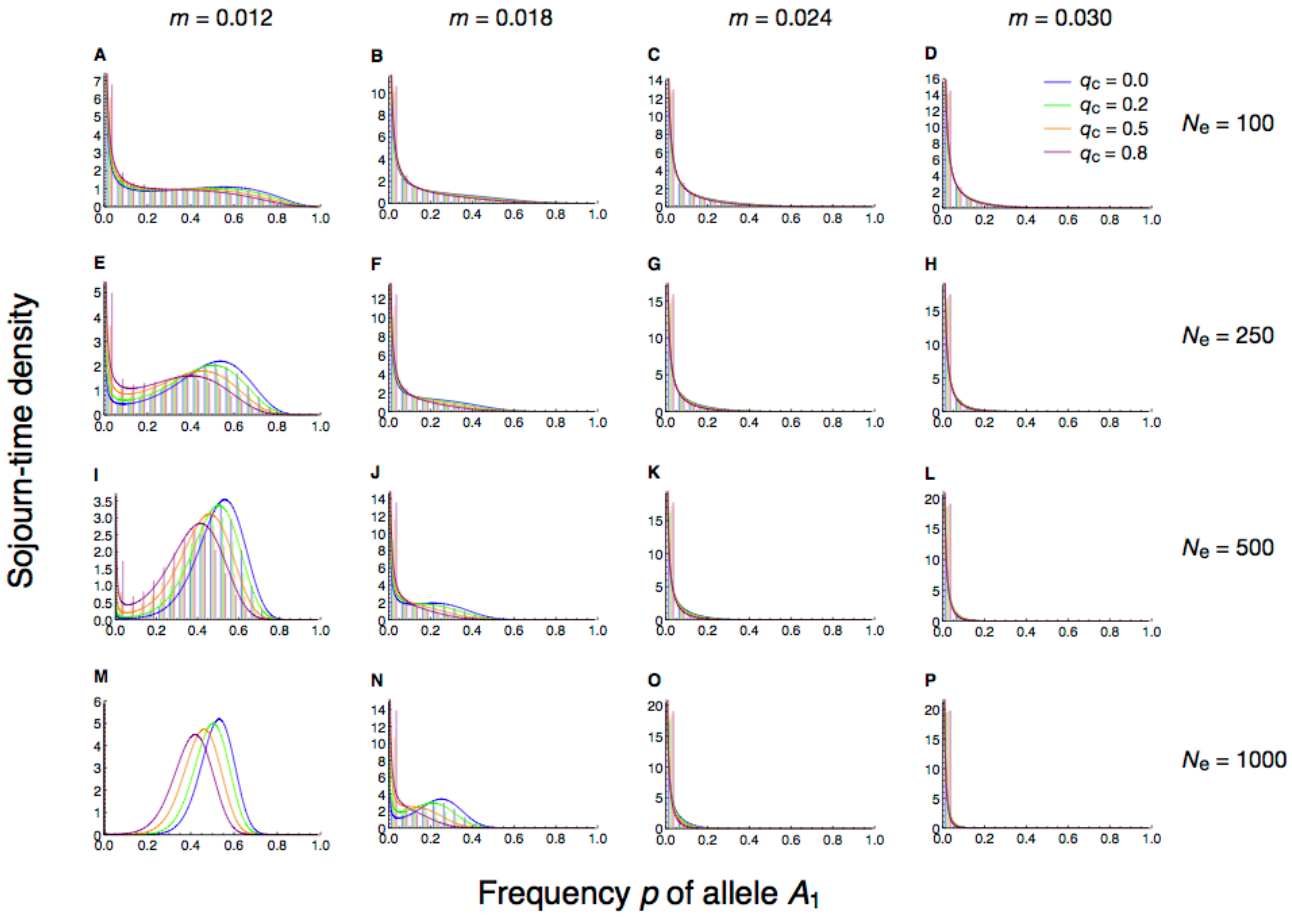

Figure S12 Comparison of analytical and simulated sojourn-time densites of A_1 for a polymorphic continent. Results are shown for various migration rates m and continental frequencies q_c of B_1. Histograms were obtained from 10^6 simulations (see Methods) and curves give the diffusion approximation under the assumption of quasi-linkage equilibrium, $\tilde{t}_{2,\mathrm{QLE}}(p;p_0)$, from Eq. (109). Throughout, $a = 0.02$, $b = 0.04$, $r = 0.1$ and $p_0 = 1/(2N)$ (we assumed $N_e = N$). From the top to the bottom row, the effective population size N_e increases and therefore genetic drift becomes less important. From the left to the right column, the migration rate m increases, making it more difficult for A_1 to survive. No simulations were completed for the parameter combination in panel (M), as they were too time-consuming.

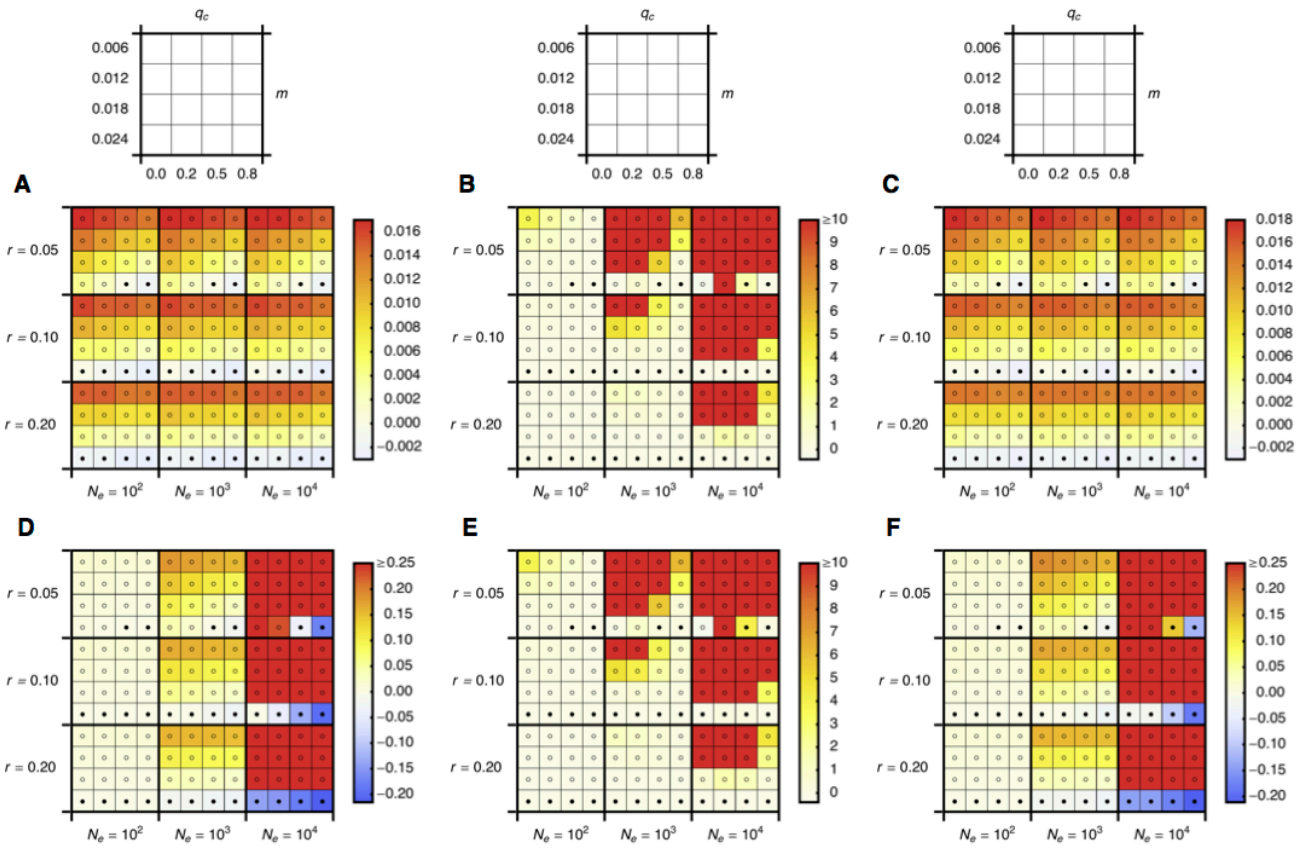

Figure S13 Comparison of various diffusion approximations of the mean absorption time of A_1. (A) The error of $\tilde{\bar{t}}_{\mathrm{QLE}}$ (Eq. 110 in File S1) relative to $\bar{t}_{\mathrm{QLE}}$ (Eq. 8) for various parameter combinations and an initial frequency of A_1 equal to $p_0 = 1/(2N)$ (we assumed $N_e = N$). Squares bounded by thick lines delimit combinations of values of the recombination rate r and the effective population size N_e. Within each of them, squares bounded by thin lines correspond to combinations of values of the migration rate m and the continental frequency q_c of A_1, as shown in the small panels on top. The colour code assigns deeper blue to more negative, and deeper red to more positive values. Empty (filled) circles indicate that the marginal one-locus equilibrium $\tilde{E}_{\mathrm{B}}$ is unstable (stable) and A_1 can (not) be established under deterministic dynamics. Selection coefficients are $a = 0.02$ and $b = 0.04$. (B) The error of $\bar{t}_{\mathrm{QLE},\rho\gg 0}$ (Eq. 114 in File S1) relative to $\bar{t}_{\mathrm{QLE}}$ for $p_0 = 1/(2N)$. Other details as for panel (A). (C) The error of $\tilde{\bar{t}}_{\mathrm{QLE},\rho\gg 0}$ (Eq. 115 in File S1) relative to $\bar{t}_{\mathrm{QLE},\rho\gg 0}$ for $p_0 = 1/(2N)$. Other details as for panel (A). (D) As in panel (A), but for an initial frequency of A_1 equal to $p_0 = 0.005$, independently of N. (E) As in panel (B), but for $p_0 = 0.005$ fixed. (F) As in panel (C), but for $p_0 = 0.005$ fixed. Simulations were as described in Methods. Numerical values for errors represented in panels (A) to (C) and (D) to (F) are shown in Tables S2 to S4 and S6 to S8, respectively.

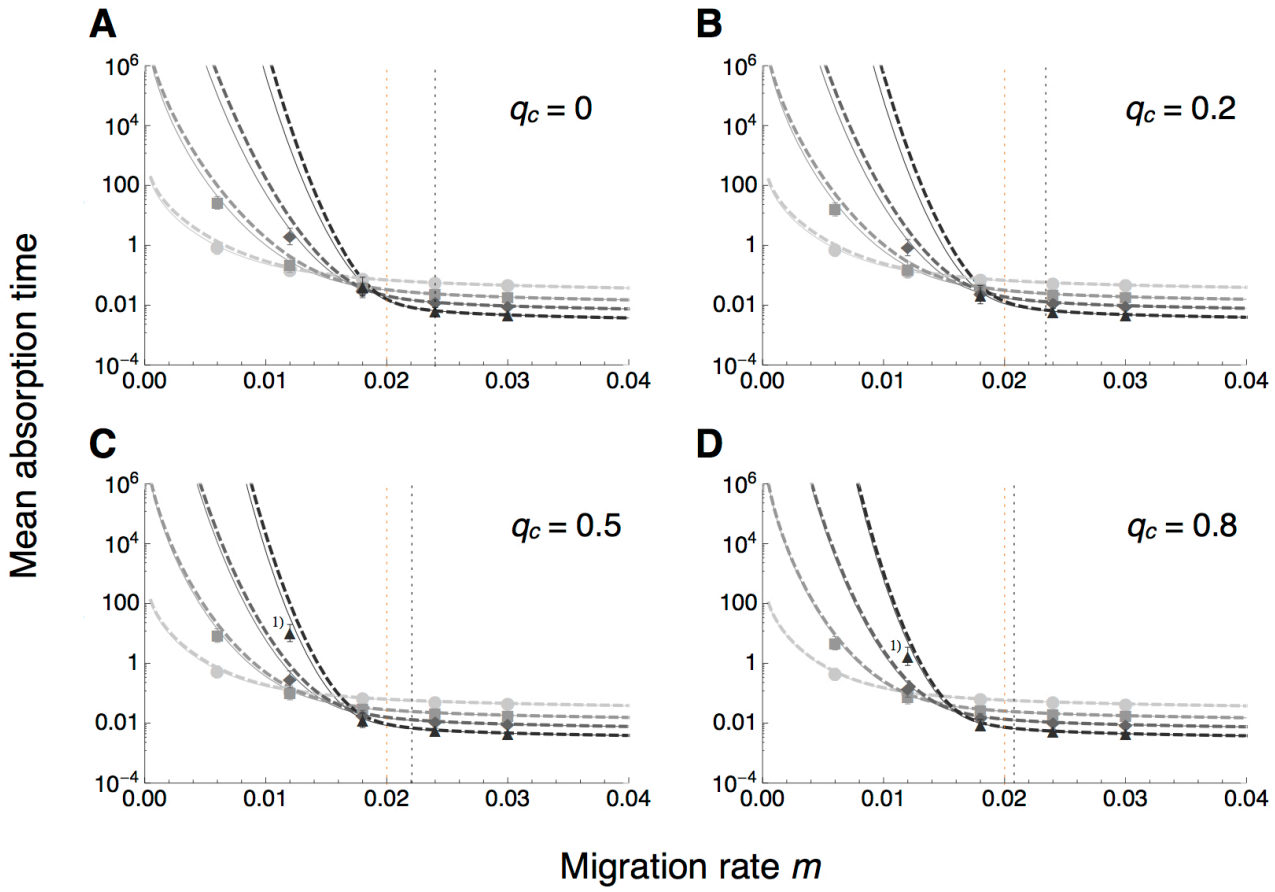

Figure S14 Mean absorption time of A_1 as a function of the migration rate. Two approximations derived under the assumption of quasi-linkage eqilibrium (QLE) are compared. Solid curves show $\bar{t}_{\mathrm{QLE}}$ (Eq. 8) and thick dashed curves $\bar{t}_{\mathrm{QLE},\rho\gg 0}$ (Eq. 114 in File S1). The effective population size N_e increases from light to dark grey, taking values of 100, 250, 500, and 1000. The vertical dotted lines denote the critical values of m below which A_1 can invade in the deterministic one-locus (orange) and two-locus (black) model. Dots and whiskers show the mean and 95% empirical interquantile range across 1000 runs of the mean absorption time in 1000 replicates per run. Where points and whiskers are missing, simulations could not be completed within the time limit of 72 hours per replicate on the computer cluster. Data points labelled by 1) are from parameter combinations for which fewer than 1000 replicates per run could be realised, because some took longer than the limit of 72 hours. (A) Monomorphic continent: $q_c = 0$. (B)–(D) Polymorphic continent with continental frequency of B_1 equal to $q_c = 0.2$, $q_c = 0.5$, and $q_c = 0.8$, respectively. Other parameters are $a = 0.02$, $b = 0.04$, $r = 0.1$, and $p_0 = 1/(2N)$ (we assumed $N_e = N$). Time is in multiples of $2N_e$ generations and plotted on the log scale.

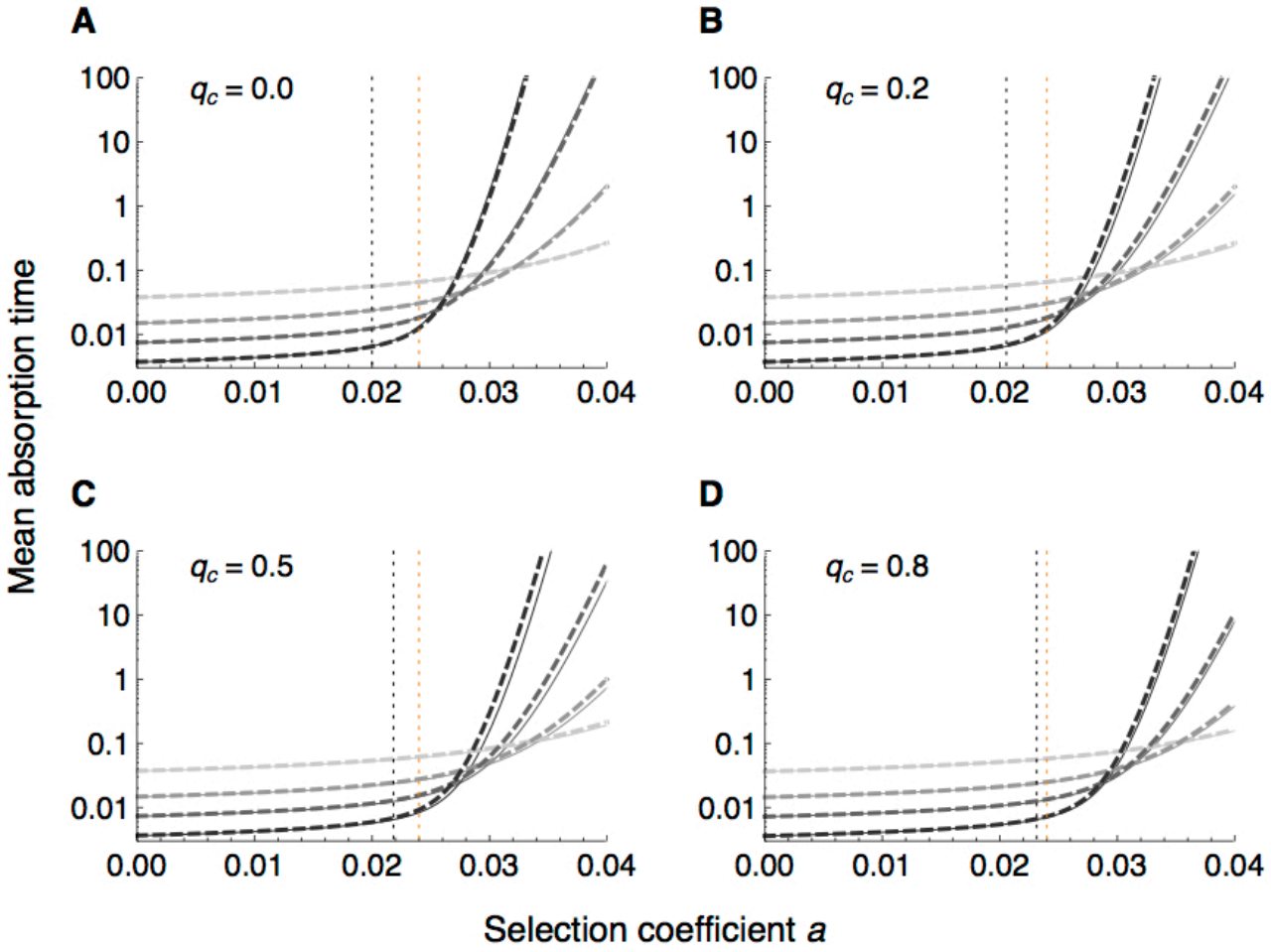

Figure S15 Mean absorption time of A_1 as a function of its selective advantage. Two approximations derived under the assumption of quasi-linkage eqilibrium (QLE) are compared. Solid curves show $\bar{t}_{\mathrm{QLE}}$ (Eq. 8) and thick dashed curves $\bar{t}_{\mathrm{QLE},\rho\gg 0}$ (Eq. 114 in File S1). The effective population size N_e increases from light to dark grey, taking values of 100, 250, 500, and 1000. The vertical dotted lines denote the critical values of a above which A_1 can invade in the deterministic one-locus (orange) and two-locus (black) model. (A) Monomorphic continent ($q_c = 0$). (B)–(D) Polymorphic continent with q_c equal to 0.2 in (B), 0.5 in (C) and 0.8 in (D). Other parameters are $b = 0.04$, $m = 0.024$, $r = 0.1$, and $p_0 = 1/(2N)$ (we assumed $N_e = N$). Time is in multiples of $2N_e$ generations and plotted on the log scale.

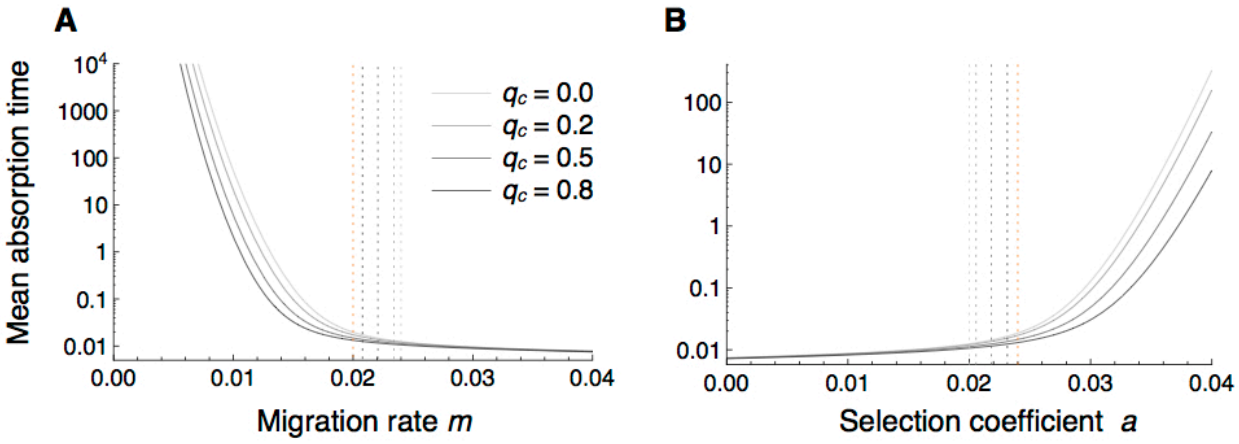

Figure S16 Effect of the continental frequency q_c of B_1 on the mean absorption time of A_1. Curves show the diffusion approximation $\bar{t}_{\mathrm{QLE}}$ (Eq. 8), derived under the assumption of quasi-linkage equilibrium. The continental frequency q_c of B_1 increases from light to dark grey, taking values of 0, 0.2, 0.5, and 0.8. (A) The mean absorption time is given in multiples of $2N_e$ generations as a function of the migration rate. Vertical dotted lines denote the critical values of m below which A_1 can invade in the respective deterministic two-locus model (grey) and, as a reference, in the one-locus model (orange). The selection coefficient in favour of A_1 is $a = 0.02$. (B) As in (A), but as a function of the selective advantage of allele A_1. The migration rate is $m = 0.024$. In both panels, $b = 0.04$, $r = 0.1$, $N_e = 500$, and $p_0 = 1/(2N)$ (we assumed $N_e = N$).

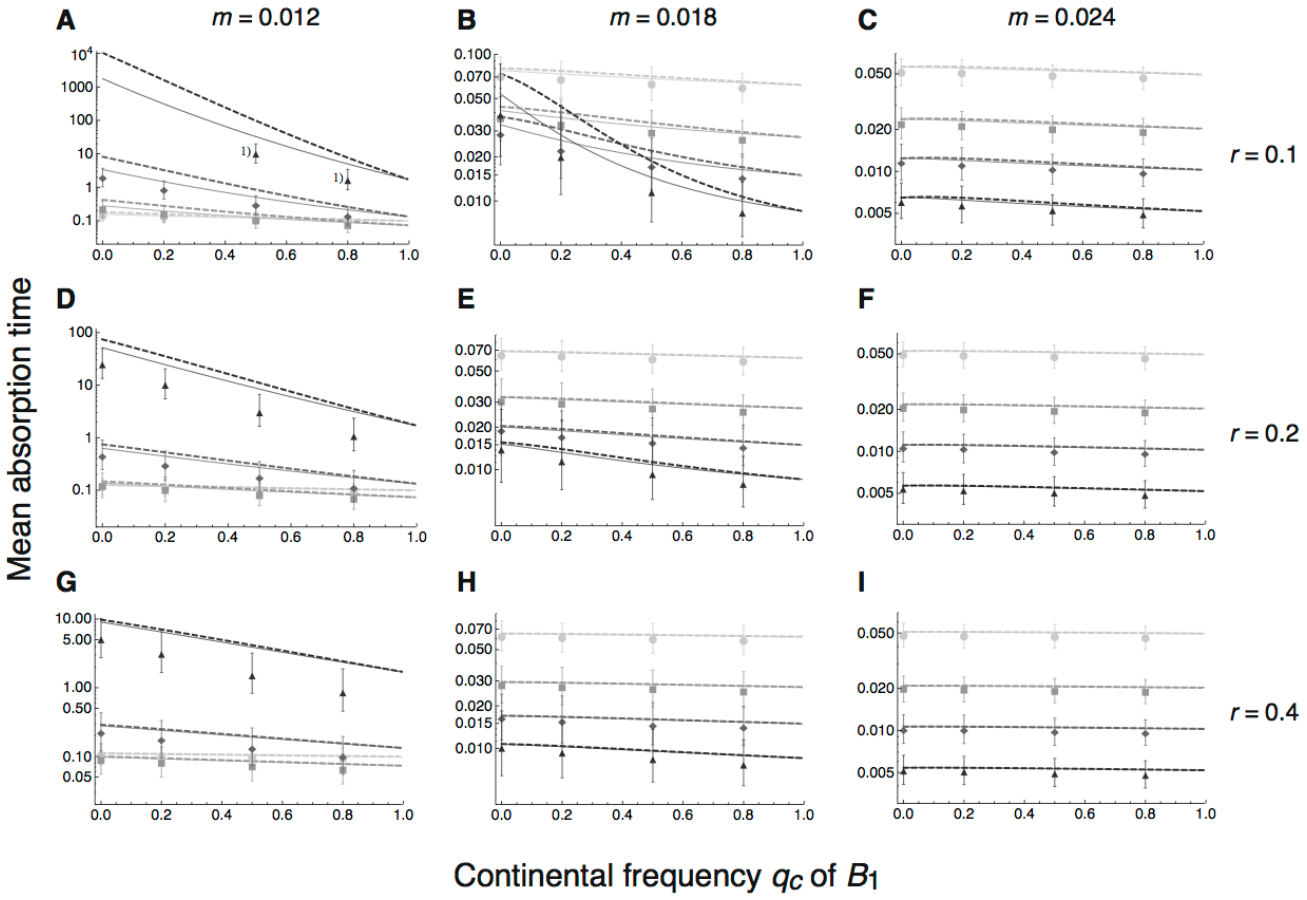

Figure S17 Mean absorption time of A_1 as a function of the continental frequency B_1. Two approximations derived under the assumption of quasi-linkage eqilibrium (QLE) are compared. Solid curves show $\bar{t}_{\mathrm{QLE}}$ (Eq. 8) and thick dashed curves $\bar{t}_{\mathrm{QLE},\rho\gg 0}$ (Eq. 114 in File S1). The effective population size N_e increases from light to dark grey, taking values of 100, 250, 500, and 1000. Dots and whiskers show the mean and 95% empirical interquantile range across 1000 runs of the mean absorption time in 1000 replicates per run. Where points and whiskers are missing, simulations could not be completed within the time limit of 72 hours per replicate on the computer cluster. Data points labelled by 1) represent parameter combinations for which fewer than 1000 replicates per run could be realised, because some took longer than the limit of 72 hours. The migration rate m increases from the left to the right column, taking values of 0.012, 0.018, and 0.024. The recombination rate r increases from the top to the bottom row, taking values of 0.1, r=0.2, and r=0.4. Other parameters are $a = 0.02$, $b = 0.04$, and $p_0 = 1/(2N)$ (we assumed $N_e = N$). Time is in multiples of $2N_e$ generations and plotted on the log scale.

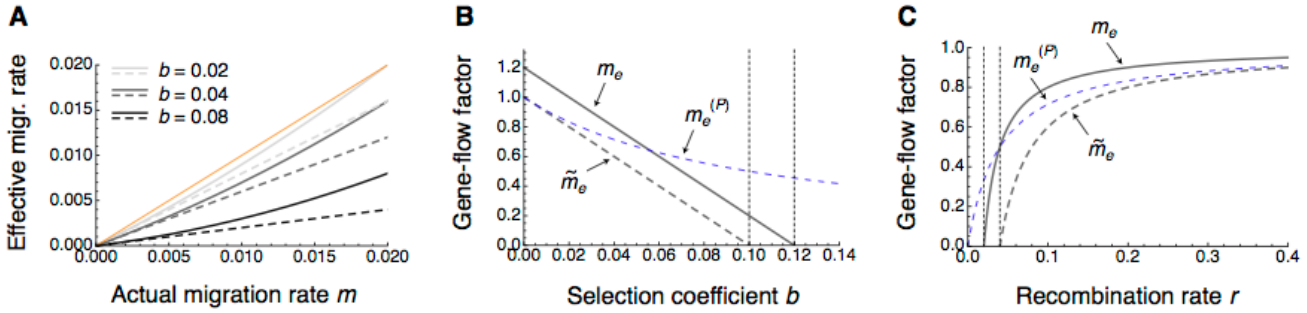

Figure S18 Effective migration rate at a weakly beneficial mutation arising in linkage to a migration–selection polymorphism. (A) The effective migration rate under the QLE approximation up to second (m_e, Eq. 19, solid) and first ($\tilde{m}_e$, Eq. 20, dashed) order of the actual migration rate m. The orange curve has a slope of 1 and represents the marginal case of linkage to a neutral background ($b = 0$). Parameter values are $b = 0.02$ (light grey), $b = 0.04$ (medium grey), $b = 0.08$ (black), and $r = 0.1$. (B) The gene-flow factor (ratio of effective to actual migration rate, Bengtsson 1985) as a function of the selective advantage b of the beneficial background allele B_1. Grey solid and dashed curves show the gene-flow factor computed using m_e and $\tilde{m}_e$, respectively. The curves cross the horizontal axis at $b = m + r$ and $b = r$, respectively (vertical lines). The blue dashed curve gives the gene-flow factor for Petry's (1983) $m_e^{(P)}$ in Eq. (21). Parameters are $m = 0.02$ and $r = 0.1$. (C) As in panel (B), but as a function of the recombination rate. Vertical dotted lines indicate $r = b - m$ and $r = b$ for m_e/m and $\tilde{m}_e/m$, respectively. Parameters are $b = 0.04$ and $m = 0.02$.

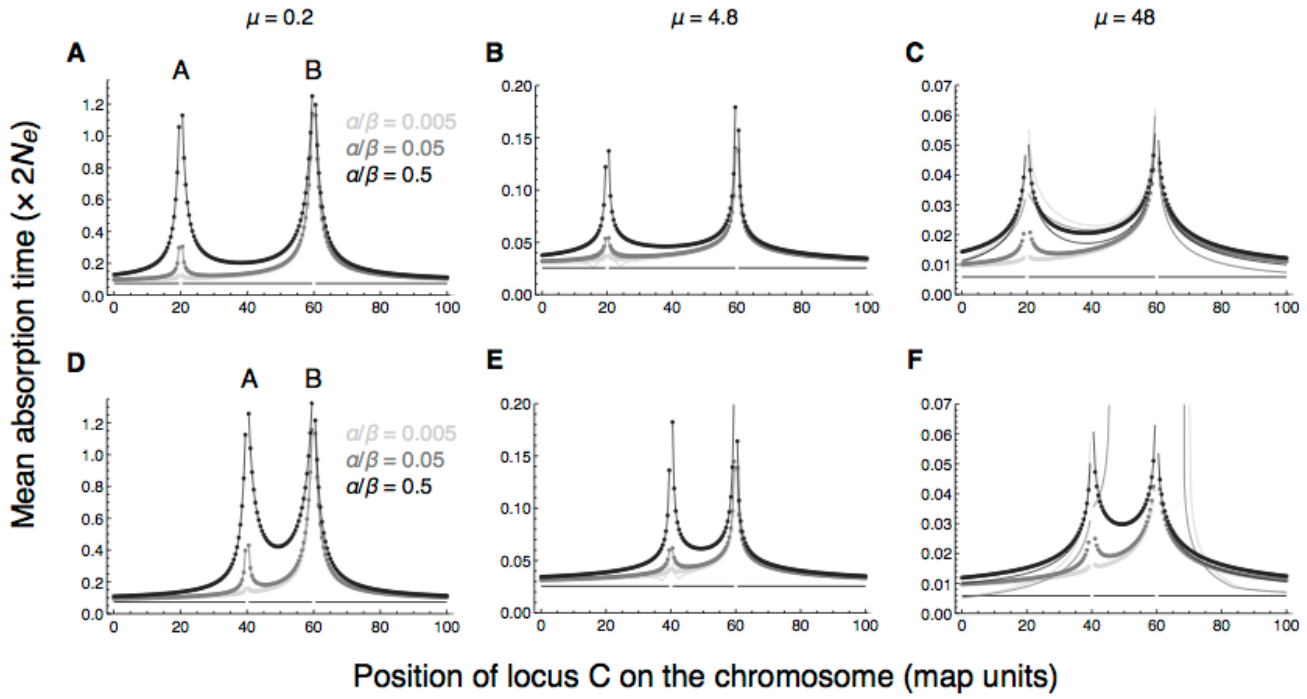

Figure S19 Effect of linkage to two selected sites on the absorption time of a neutral mutation. The mutation occurs at the neutral locus C. The loci A and B under selection are located 20 and 60 map units from the left end of the chromosome in panels (A)–(C), whereas locus A is located 40 map units from the left end of the chromosome in panels (D)–(F). One map unit (centimorgan) corresponds to a recombination rate of $r = 0.01$ and the effective population size is $N_e = 100$. The scaled selection coefficient in favour of B_1 is $\beta = 80$ and the scaled migration rate increases from left to right from $\mu = 0.2$ in (A) and (D) to $\mu = 4.8$ in (B) and (E) and $\mu = 48$ in (C) and (F). From light to dark, α/β takes values of 0.005, 0.05, and 0.5, where α is the scaled selection coefficient in favour of A_1. Points show values computed using the approximate effective migration rates in Eq. (23) and curves are based on numerically computed exact effective migration rates (Procecure S9). For μ large and α small (light grey curves in F), the latter were affected by numerical errors causing strong deviation. The horizontal black line denotes the baseline for free recombination between locus C and the selected sites.

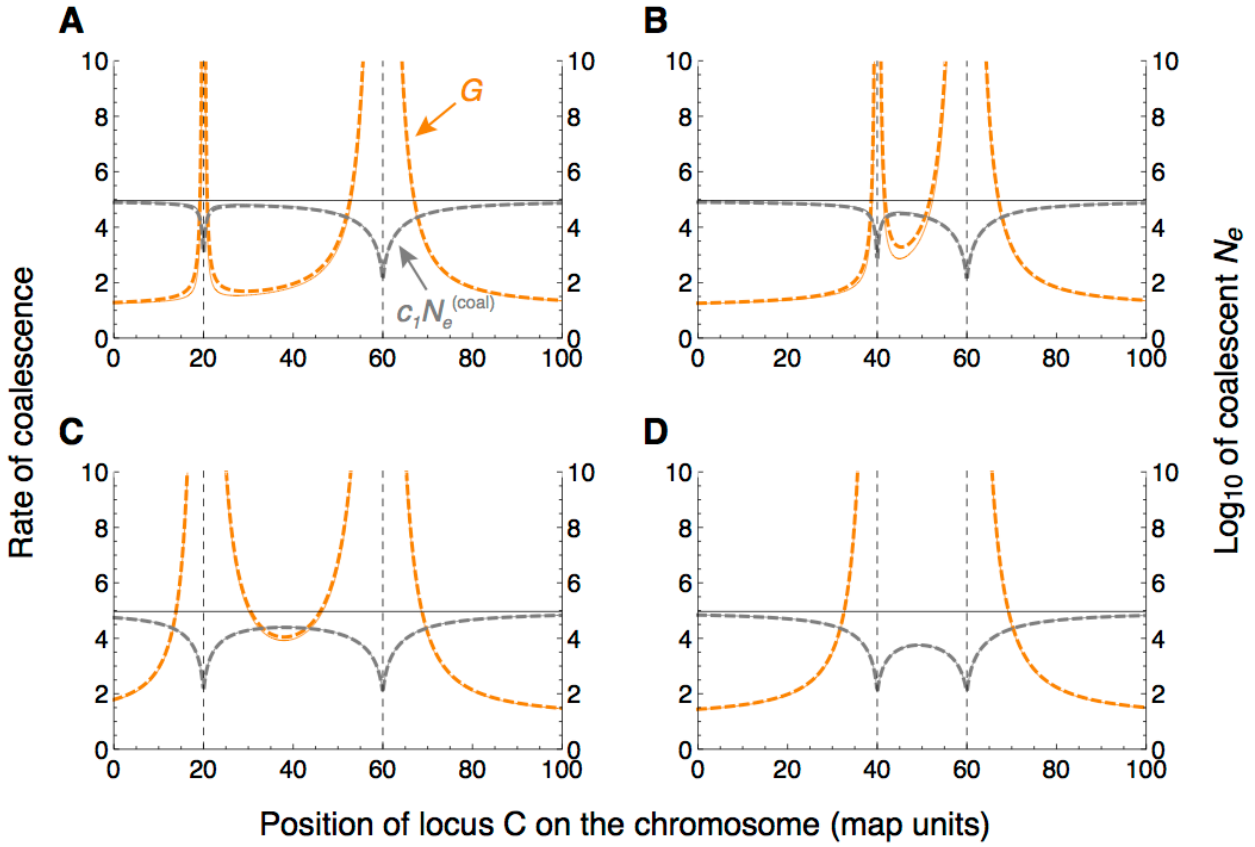

Figure S20 Effect on neutral coalescence of linkage to two sites at migration–selection balance. The rate of coalescence G (orange, see Eq. 25) and the coalescent effective size of the island population, $c_1\bar{N}/G$, are given as a function of the position (in map units) of the neutral locus C. Solid and thick dashed curves are for values computed using the exact and approximate (Eq. 23) effective migration rate, respectively (they overlap almost completely). One map unit (centimorgan) corresponds to a recombination rate of $r = 0.01$ and the position of the sites under selection is indicated by vertical dashed lines. The total population size is $\bar{N} = 10^8$, the fraction of the island is $N_1/\bar{N} = c_1 = 0.01$ and the selection coefficient at locus B (position 60) is $b = 0.4$. (A) and (B) The migration rate to the island is of the same order of magnitude as selection at locus A: $a = 0.02$, $m_1 = 0.024$. (C) and (D) Immigration is weak compared to selection at locus A: $a = 0.2$, $m_1 = 0.024$. Throughout, $m_1/m_2 = c_2/c_1 = (1 - c_1)/c_1$, so actual migration is conservative (Wakeley 2009, p. 194). The horizontal black line gives the baseline-effective population size at the neutral locus in the absence of linked selection. For alternative parameter combinations, see File S9.

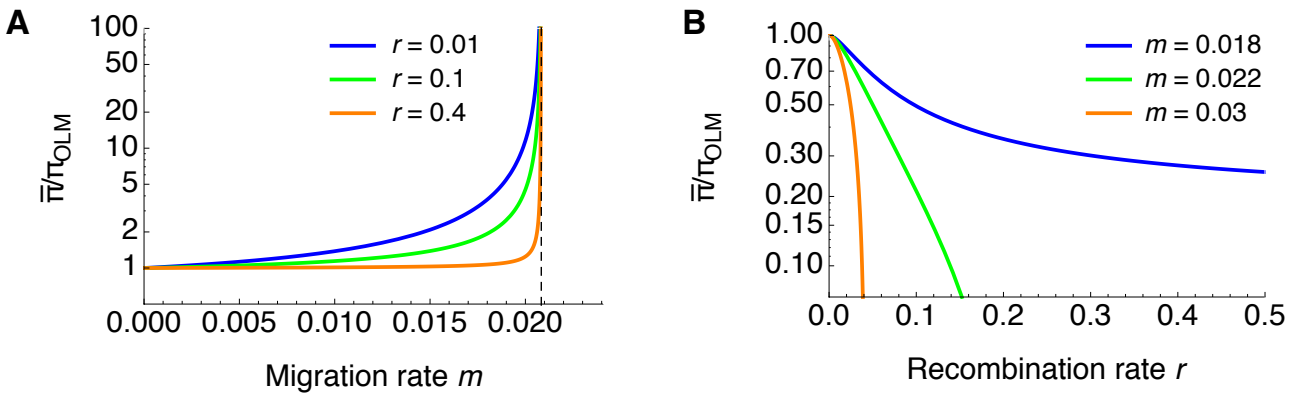

Figure S21 Mean invasion probability of A_1 with linkage to a background polymorphism compared to no linkage. Curves show the ratio of the weighted mean invasion probability, $\bar{\pi}$, divided by that of the one-locus model, π_{OLM} ($r = 0.5$). The ratio was computed from numerical solutions to the branching process (Eq. 3) and is shown as a function of the migration (m) and recombination (r) rate in panels (A) and (B), respectively. The vertical dashed line in panel (A) shows the critical migration rate $a/(1-b)$, beyond which allele A_1 cannot be established under the deterministic one-locus model. In panel (B), for $m = 0.018$ (blue curve), allele A_1 can be established independently of r. For stronger migration (green and orange curves), A_1 can be established only if r is below a critical value (where the green and orange curves cross the x-axis, respectively). Other parameter values are $a = 0.02$, $b = 0.04$, and $q_c = 0$. Compare to Figure 7 for the relative effect of m on mean extinction time.

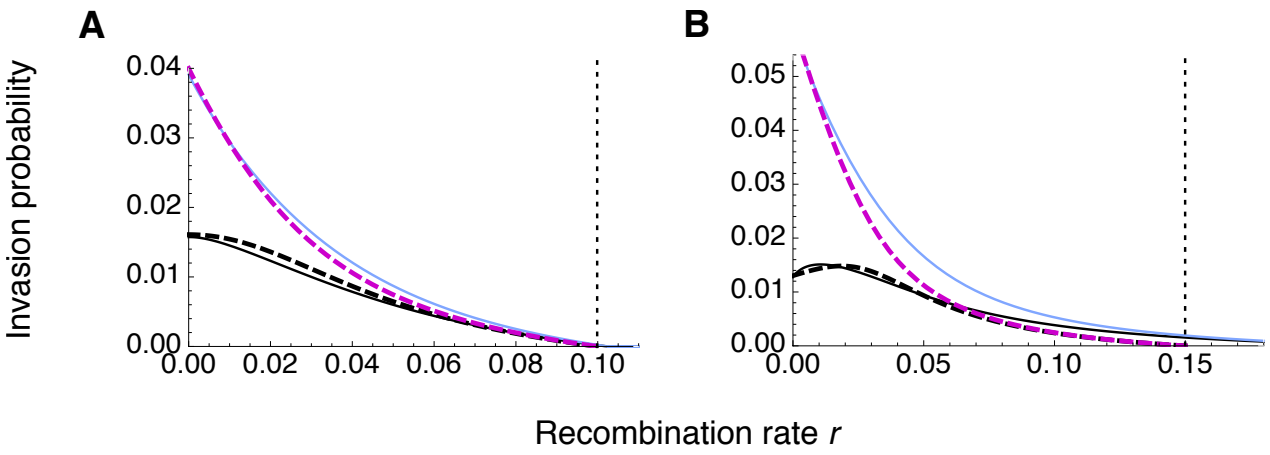

Figure S22 Comparison of branching-proccess and `splicing' approximations to the invasion probability of A_1 as a function of the recombination rate. Black curves represent the branching-process solution averaged across the two backgrounds (B_1 and B_2). The solid curve gives the exact numerical solution and the dashed curve the analytical approximation for a slighty-supercritical process (based on Eq. 12). The dashed purple curve represents the approximation based on the `splicing approach' as proposed by Yeaman (2013). As a reference, the thin blue curve gives the numerical branching-process solution conditional on A_1 arising on the beneficial background B_1. (A) A case where $r_{\rm opt} = 0$; $a = 0.02$, $m = 0.024$. (B) A case where $r_{\rm opt} > 0$; $a = 0.03$, $m = 0.032$. In both panels, $b = 0.04$, and the vertical dotted line gives the critical recombination rate below which A_1 can invade according to deterministic continuous-time theory.

File S1
Additional Methods

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 $x_i \geq 0$ $(i = 1, \ldots, 4)$ and $\sum_{i=1}^{4} x_i = 1$, which are equivalent to $0 \leq p, q \leq 1$ and

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

The matrix of relative fitnesses on the island is

$$\mathbf{W} = \begin{array}{c} \\ A_1A_1 \\ A_1A_2 \\ A_2A_2 \end{array} \begin{array}{c} \begin{array}{ccc} B_1B_1 & B_1B_2 & B_2B_2 \end{array} \\ \begin{pmatrix} w_{11} & w_{12} & w_{22} \\ w_{13} & w_{14} = w_{23} & w_{24} \\ w_{33} & w_{34} & w_{44} \end{pmatrix} \end{array}, \tag{27}$$

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,

$$x_1' = (1-m)(x_1w_{1\bullet} - rw_{14}D)/\bar{w}, \tag{28a}$$
$$x_2' = (1-m)(x_2w_{2\bullet} + rw_{14}D)/\bar{w}, \tag{28b}$$
$$x_3' = (1-m)(x_3w_{3\bullet} + rw_{14}D)/\bar{w} + mq_c, \tag{28c}$$
$$x_4' = (1-m)(x_4w_{4\bullet} - rw_{14}D)/\bar{w} + m(1-q_c), \tag{28d}$$

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,

$$x_1' = (1-m)\left[w_1x_1 + rw_{14}x_2\hat{q}_B - rw_{14}x_1(1-\hat{q}_B)\right]/\bar{w}, \tag{29a}$$
$$x_2' = (1-m)\left[w_2x_2 - rw_{14}x_2\hat{q}_B + rw_{14}x_1(1-\hat{q}_B)\right]/\bar{w}, \tag{29b}$$

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

$$w_1 = w_{13}\hat{q}_B + w_{14}(1-\hat{q}_B), \tag{30a}$$
$$w_2 = w_{24}(1-\hat{q}_B) + w_{14}\hat{q}_B. \tag{30b}$$

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

$$\bar{w} = \hat{q}_\mathrm{B}^2 w_{33} + 2\hat{q}_\mathrm{B}(1-\hat{q}_\mathrm{B})w_{34} + (1-\hat{q}_\mathrm{B})^2 w_{44}. \tag{31}$$

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

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

with

$$\lambda_{11} = (1-m)\left[w_1 - r(1-\hat{q}_\mathrm{B})w_{14}\right]/\bar{w}, \tag{33a}$$
$$\lambda_{12} = (1-m)r(1-\hat{q}_\mathrm{B})w_{14}/\bar{w}, \tag{33b}$$
$$\lambda_{21} = (1-m)r\hat{q}_\mathrm{B}w_{14}/\bar{w}, \tag{33c}$$
$$\lambda_{22} = (1-m)\left[w_2 - r\hat{q}_\mathrm{B}w_{14}\right]/\bar{w}. \tag{33d}$$

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 $\hat{q}_\mathrm{B}$. We note that Eqs. (29) to (33) are valid for both a monomorphic and a polymorphic continent. The difference comes in only via $\hat{q}_\mathrm{B}$, 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.: 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 $\bar{\bar{w}}$.

3 Marginal one-locus migration–selection model

We denote the marginal one-locus migration–selection equilibrium by $E_\mathrm{B} = (p = 0, q = \hat{q}_\mathrm{B}, D = 0)$. This equilibrium is assumed to be realised on the island before occurrence of the A_1 mutation. The equilibrium frequency $\hat{q}_\mathrm{B}$ 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 $\hat{q}_\mathrm{B}$ is obtained by solving

$$q'_\mathrm{B} = (1-m)\frac{\tilde{w}_1}{\tilde{\bar{w}}}q_\mathrm{B} + mq_c = q_\mathrm{B} \tag{34}$$

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

$$\tilde{\bar{w}} = q_\mathrm{B}^2 w_{33} + 2q_\mathrm{B}(1-q_\mathrm{B})w_{34} + (1-q_\mathrm{B})^2 w_{44} \tag{35}$$

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

$$\hat{q}_\mathrm{B} = \frac{w_{34}(1-m) - \tilde{\bar{w}} + \sqrt{4(1-m)mq_c(w_{34}-w_{33})\tilde{\bar{w}} + \left[\tilde{\bar{w}} - (1-m)\right]^2}}{2(1-m)(w_{34}-w_{33})}, \tag{36}$$

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

$$m < \frac{w_{34} - \tilde{\bar{w}}}{w_{34}}. \tag{37}$$

We note that $\tilde{\bar{w}}$ 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. $0 < \hat{q}_\mathrm{B} < 1$), if the migration rate is below a critical value,

$$m < \frac{b}{1-a} =: m_\mathrm{B}. \tag{38}$$

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 $m = m_{\mathsf{B}} > m^*$, 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. $m < \min(b, m_{\mathsf{B}})$), 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

$$\hat{q}_{\mathsf{B}} = \frac{b - (1-a)m + 2bmq_c + \sqrt{R}}{2b(1+m)}, \tag{39}$$

where

$$R = 4b(1-a-b)m(1+m)q_c + \left[b - (1-a)m + 2bmq_c\right]^2 \geq 0. \tag{40}$$

In contrast to the case of a monomorphic continent, where E_{B} exists only if $m < m_{\mathsf{B}}$, 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

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

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

$$\hat{q}_{\mathsf{B}} \approx 1 - \frac{m}{b}, \tag{42}$$

$$m_{\mathsf{B}} \approx b, \tag{43}$$

and

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

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

$$\mathbb{E}\left[x_1'\right] = \lambda_{11}x_1 + \lambda_{21}x_2, \tag{45a}$$

$$\mathbb{E}\left[x_2'\right] = \lambda_{12}x_1 + \lambda_{22}x_2, \tag{45b}$$

or, in matrix form

$$\mathbb{E}\left[\mathbf{x}'\right] = \mathbf{x}\mathbf{L}, \tag{46}$$

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

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

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

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

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

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

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

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

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

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

$$\lambda_{11} = E + Fr, \tag{51a}$$
$$\lambda_{12} = -Fr, \tag{51b}$$
$$\lambda_{21} = Hr, \tag{51c}$$
$$\lambda_{22} = J - Hr, \tag{51d}$$

where

$$E = \frac{1 + b + am}{1 - a + b}, \tag{52a}$$
$$F = -\frac{m}{b}, \tag{52b}$$
$$H = \frac{b - (1 - a)m}{b(1 - a + b)}, \tag{52c}$$
$$J = \frac{1 + m(a - b)}{1 - a + b}. \tag{52d}$$

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

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

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

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

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

where

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

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

$$R_2 = b^2 + 2br - 4mr + r^2. \tag{55}$$

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

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

(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

$$f_1(s_1, s_2) = e^{(E+Fr)s_1 - Frs_2 - E} = s_1 \tag{57a}$$
$$f_2(s_1, s_2) = e^{Hrs_1 + (J-Hr)s_2 - J} = s_2 \tag{57b}$$

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,

$$\lambda_{11} = \tilde{E} + \tilde{F}r, \tag{58a}$$

$$\lambda_{12} = \tilde{G}r, \tag{58b}$$

$$\lambda_{21} = \tilde{H}r, \tag{58c}$$

$$\lambda_{22} = \tilde{J} + \tilde{I}r, \tag{58d}$$

with

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

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

$$\lambda_{11} \approx \frac{1}{2}\left(2+2a+b-m-\sqrt{R_3}\right) - \frac{b+m-\sqrt{R_3}}{2b}r, \qquad \lambda_{12} \approx \frac{b+m-\sqrt{R_3}}{2b}r,$$

$$\lambda_{21} \approx \frac{b-m+\sqrt{R_3}}{2b}r, \qquad \lambda_{22} \approx \frac{1}{2}\left(2+2a-b-m-\sqrt{R_3}\right) - \frac{b-m+\sqrt{R_3}}{2b}r,$$

where

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

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

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

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

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

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

(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

$$\nu = \nu(\xi) = 1 + \xi, \tag{62}$$

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 $\xi \to 0$, $q_i(\xi)$ converges to 1 and

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

Here, $v_k = i$ is the ith 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

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

where u_i is the ith 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

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

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 maint text. After some algebra using *Mathematica* (File S5), we obtain the appropriately normalised left and right eigenvectors of $\mathbf{L}$ as

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

and

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

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 $\hat{q}_\mathrm{B}$ 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

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

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 $q_c \to 0$ and assuming $m < m_\mathrm{B}$ (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 ($r_{\mathrm{opt}} > 0$) in some cases, whereas it is maximised at $r_{\mathrm{opt}} = 0$ in other cases (Figures 1A and 1B). To distinguish between these two regimes, we note that $r_{\mathrm{opt}} > 0$ 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

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

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 $\hat{q}_{\mathrm{B}}$ 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

$$\begin{aligned} 1 - \pi_1 &= e^{-\lambda_{11}\pi_1} \cdot e^{-\lambda_{12}\pi_2} \\ 1 - \pi_2 &= e^{-\lambda_{21}\pi_1} \cdot e^{-\lambda_{22}\pi_2}. \end{aligned}$$

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

$$\ln\left[1 - \pi_1(r)\right] = -\lambda_{11}(r)\pi_1(r) - \lambda_{12}(r)\pi_2(r) \tag{69a}$$
$$\ln\left[1 - \pi_2(r)\right] = -\lambda_{21}(r)\pi_1(r) - \lambda_{22}(r)\pi_2(r). \tag{69b}$$

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

$$\ln\left[1 - \pi_1(r)\right] = -\frac{1-m}{\bar{w}}\Big\{\big[w_1 - (1-\hat{q}_{\mathrm{B}})rw_{14}\big]\pi_1(r) + (1-\hat{q}_{\mathrm{B}})rw_{14}\pi_2(r)\Big\} \tag{70a}$$
$$\ln\left[1 - \pi_2(r)\right] = -\frac{1-m}{\bar{w}}\Big\{\hat{q}_{\mathrm{B}}rw_{14}\pi_1(r) + \left(w_2 - \hat{q}_{\mathrm{B}}rw_{14}\right)\pi_2(r)\Big\}. \tag{70b}$$

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

$$\frac{\pi_1'(0)}{1-\pi_1^\circ} = (1-m)\frac{w_1\pi_1'(0) - (1-\hat{q}_{\mathrm{B}})w_{14}\left(\pi_1^\circ - \pi_2^\circ\right)}{\bar{w}} \tag{71a}$$
$$\frac{\pi_2'(0)}{1-\pi_2^\circ} = (1-m)\frac{w_2\pi_2'(0) + \hat{q}_{\mathrm{B}}w_{14}\left(\pi_1^\circ - \pi_2^\circ\right)}{\bar{w}}, \tag{71b}$$

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

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

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

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

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

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

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

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

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

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

After some algebra (File S6), we find that $\bar{\pi}'(0) > 0$, and hence $r_\mathrm{opt} > 0$, 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

$$\ln Q_1(r) = (E+Fr)\,Q_1(r) - FrQ_2(r) - E \tag{77a}$$
$$\ln Q_2(r) = HrQ_1(r) + (J-Hr)Q_2(r) - J, \tag{77b}$$

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 $Q_i'(0) = \frac{dQ_1(r)}{dr}\big|_{r=0}$. Here, we used $Q_i^\circ = Q_i(0)$ for the extinction probability of A_1 condidtional on initial occurrence on background B_i ($i \in \{1,2\}$). Solving for $Q_1'(0)$ and $Q_2'(0)$ yields

$$Q_1'(0) = \frac{FQ_1^\circ(Q_1^\circ - Q_2^\circ)}{1-EQ_1^\circ} \tag{78a}$$
$$Q_2'(0) = \frac{HQ_2^\circ(Q_1^\circ - Q_2^\circ)}{1-JQ_2^\circ}. \tag{78b}$$

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

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

where

$$Z_1 := E = \frac{1+b+am}{1-a+b}, \tag{80a}$$
$$Z_2 := J = \frac{1+m(a-b)}{1-a+b}. \tag{80b}$$

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 $Z_i \le 1$, 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

$$Z_1 = 1 + \epsilon \tag{81}$$

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

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

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

$$Q_1^\circ = 1 - \pi_1^\circ \approx 1 - 2\epsilon. \tag{83}$$

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

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

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 $\hat{q}_\mathrm{B}$ and $1 - \hat{q}_\mathrm{B}$, 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 $\hat{q}_\mathrm{B}$. For additive fitnesses, $\hat{q}_\mathrm{B}$ 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 $w_1/\bar{w} = (1+b+am)/\left[(1 - a + b)(1 - m)\right]$ increases quickly with m (File S6). This explains why π_1° increaes 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

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

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

valid for a small relative to m and b. Noting that $\bar{\pi}'(0) = -\left[\hat{q}_\mathrm{B}Q_1'(0) + (1 - \hat{q}_\mathrm{B})Q_2'(0)\right]$ and using $\hat{q}_\mathrm{B}$ 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

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

After some algebra, one can show that $\bar{\pi}'(0) > 0$, and hence $r_\mathrm{opt} > 0$, 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 $r_\mathrm{opt} > 0$ (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 $r_\mathrm{opt} > 0$. Analytical and numerical computations in File S6 suggest that if we start with a monomorphic continent ($q_c = 0$) in a constellation where $r_\mathrm{opt} > 0$ holds, and then increase q_c, the maximum in $\bar{\pi}(r)$ shifts to 0 ($r_\mathrm{opt} \to 0$). There must be a critical value of q_c at which the shift from $r_\mathrm{opt} > 0$ to $r_\mathrm{opt} = 0$ 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 (~) 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 ($i = 1, \ldots, 4$) 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 $\lim_{\Delta t \to 0} \frac{\Delta x}{\Delta t}$ for $x \in \{p, q, D\}$ results in

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

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

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

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

$$\hat{\tilde{q}}_\mathsf{B} = 1 - \frac{m}{b} \tag{88}$$

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 $m < b = \tilde{m}_\mathsf{B}$. 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 $\tilde{E}_\mathsf{B}$ 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 $\tilde{E}_\mathsf{B}$ 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 $\hat{\tilde{q}}_\mathsf{B}$ of B_1 at the marginal one-locus migration–selection equilibrium $\tilde{E}_\mathsf{B}$ as

$$\hat{\tilde{q}}_\mathsf{B} = \frac{b - m + \sqrt{R_3}}{2b}, \tag{89}$$

with $R_3 = (b-m)^2 + 4bmq_c > 0$ as previously encountered in Eq. (59) in section 4. Equilibrium $\tilde{E}_\mathsf{B}$ 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 $\tilde{E}_\mathsf{B}$ under the two-locus dynamics. The Jacobian matrix evaluated at $\tilde{E}_\mathsf{B} = (p = 0, q = \hat{\tilde{q}}_\mathsf{B}, D = 0)$ is

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

and its leading eigenvalue is

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

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

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

(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 $\tilde{E}_\mathrm{B}$ is unstable and A_1 can invade whenever $r < \tilde{r}_\mathrm{B}$, where

$$\tilde{r}_\mathrm{B} = \begin{cases} \infty & \text{if } 0 \le m \le a, \\ \tilde{r}^{**} & \text{if } m > a. \end{cases} \tag{93}$$

Figure S7 shows the division of the (m, r)-parameter space into areas where $\tilde{E}_\mathrm{B}$ 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 $\tilde{E}_\mathrm{B}$ is not hyperbolic. These are given by

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

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

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

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

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

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

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

Figure S8 shows the partition of the (m, q_c)-parameter space into areas where $\tilde{E}_\mathrm{B}$ is asymptotically stable (blue) and unstable (orange), respectively. There are parameter combinations such that $\tilde{E}_\mathrm{B}$ 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 $\tilde{E}_{\mathrm{B}}$ is unstable if $q_c < \tilde{q}_{c+}^{**}$. If recombination is intermediate, i.e. $\tilde{r}^* < r < b^2/[4(m-a)]$, then $\tilde{E}_{\mathrm{B}}$ 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 $\tilde{E}_{\mathrm{B}}$ 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 $\tilde{E}_{\mathrm{B}}$ is asymptotically stable (blue) and unstable (orange), respectively. As just shown, there are parameter combinations such that $\tilde{E}_{\mathrm{B}}$ 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 $\tilde{E}_{\mathrm{B}}$ under the two-locus dynamics could be obtained in terms of a critical migrationr rate m^{**} at which $\tilde{E}_{\mathrm{B}}$ 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 $\tilde{E}_{\mathrm{B}}$ 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 $\tilde{E}_{\mathrm{B}}$ 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 $\tilde{E}_{\mathrm{B}}$ (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

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

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

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

(cf. Eq. 4.3 in Bürger and Akerman 2011). As $r \to \infty$, Eq. (97) converges to the case of no linkage, where $\hat{\tilde{p}}_+ = 1 - m/a$, $\hat{\tilde{q}}_+ = 1 - m/b$, and $\hat{\tilde{D}}_+ = 0$. 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

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

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

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

Setting $q_c = 0$ and recalling that $m < \tilde{m}^* = a\left(1 + \frac{b-a}{r}\right)$ 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) $\tilde{E}_{+\mathrm{QLE}}$. Graphical exploration in File S7 confirms that $\tilde{E}_{+\mathrm{QLE}}$ is asymptotically stable whenever it exists under the QLE regime.

Finally, we ask when $\tilde{E}_{+\mathrm{QLE}}$ exists in the admissible state space. We note that $\hat{\tilde{p}}_{+\mathrm{QLE}}$ is a strictly decreasing function of the recombination rate r, independently of the migration rate m. In contrast, $\hat{\tilde{q}}_{+\mathrm{QLE}}$ 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 $r_{\hat{\tilde{p}}^0_{+\mathrm{QLE}}}$ and $r_{\hat{\tilde{p}}^1_{+\mathrm{QLE}}}$ the recombination rates at which $\hat{\tilde{p}}_{+\mathrm{QLE}}$ equals 0 and 1, respectively. Analogously, we use $r_{\hat{\tilde{q}}^0_{+\mathrm{QLE}}}$ and $r_{\hat{\tilde{q}}^1_{+\mathrm{QLE}}}$ for the recombination rates at which $\hat{\tilde{q}}_{+\mathrm{QLE}}$ equals 0 and 1, respectively. These critical recombination rates are found to

be

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

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

and

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

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

As shown in File S7, if $m < a$, $\tilde{E}_{+\mathrm{QLE}}$ exists in the admissible state space if and only if $r > \max\left(r_{\hat{\tilde{p}}^1_{+\mathrm{QLE}}}, r_{\hat{\tilde{q}}^1_{+\mathrm{QLE}}}\right)$. If $m \geq a$, $\tilde{E}_{+\mathrm{QLE}}$ exists in the admissible state space if and only if $\max\left(r_{\hat{\tilde{p}}^1_{+\mathrm{QLE}}}, r_{\hat{\tilde{q}}^1_{+\mathrm{QLE}}}\right) < r < r_{\hat{\tilde{p}}^0_{+\mathrm{QLE}}}$. At a first glance, it may seem surprising to obtain an upper limit on r. However, as is easily verified, $r_{\hat{\tilde{p}}^0_{+\mathrm{QLE}}}$ is also the critical value at which $\tilde{E}_{+\mathrm{QLE}}$ coincides with the QLE approximation of $\tilde{E}_{\mathrm{B}}$, 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 developped (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, $q = \hat{\tilde{q}}_{\mathrm{B}}$ constant, where $\hat{\tilde{q}}_{\mathrm{B}} = 1 - m/b$ is the frequency of B_1 at the one-locus migration–selection equilibrium in continuous time (Eq. 88). Equation (87) is therefore approximated by

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

$$\dot{q} = \frac{dq}{dt} = 0, \tag{101b}$$

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

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

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

In the limit of $r \to \infty$, we recover the one-locus migration-selecion 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

$$V(p) = p(1-p) \tag{103}$$

(Karlin and Taylor 1981, p. 159).

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

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

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

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

Inserting Eq. (104), we find,

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

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 $\tilde{E}_{\mathrm{B}}$, it is required that $m < \tilde{m}^* = a\left(1+\frac{b-a}{r}\right)$ 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

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

To make the assumption of quasi-linkage equilibrium explicit, we will add the subscript QLE to relevant quantities from now on. The densities $t_{i,\mathrm{QLE}}(p;p_0)$ 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 $t_{1,\mathrm{QLE}}(p;p_0) = 2\psi(y)^{-1}(1-p)^{-1}p^{-1}\int_0^p \psi(y)dy$ and $p^{-1}\int_0^p \psi(y)dy \to 1$ as $p \to 0$ (File S7), we approximate $t_{1,\mathrm{QLE}}(p;p_0)$ by

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

whenever p is small. Recall from Eq. (107) that $t_1(p;p_0)$ is needed only if $0 \le p \le p_0$. 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, $p \le p_0$ automatically implies that p is small whenever $t_{1,\mathrm{QLE}}(p;p_0)$ is employed. The approximation in Eq. (108) is therefore valid for our purpose.

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

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

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

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

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

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

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 $p_0 \to 0$.

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

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

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

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

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

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

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

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

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

$$\tilde{t}_{2,\mathrm{QLE},\rho\gg 0}(p;p_0) = \frac{2p_0}{V(p)\psi_{\rho\gg 0}(p)} = 2p_0e^{2p\alpha}p^{-1}(1-p)^{\frac{2\mu(\mu-\beta+\rho)}{\rho}-1}$$

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

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

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

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

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

$$\tilde{t}_{1,\mathrm{OLM}}(p;p_0) = 2e^{2p\alpha}(1-p)^{2\mu-1} \qquad \text{if } 0 \le p \le p_0,$$

$$\tilde{t}_{2,\mathrm{OLM}}(p;p_0) = 2p_0e^{2p\alpha}p^{-1}(1-p)^{2\mu-1} \qquad \text{if } p_0 \le p \le 1$$

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 polymorhpic 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

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

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

$$M(p) \approx M_{\rho\gg0}(p) = \alpha p(1-p) - \mu p + \frac{\mu\left(\beta-\mu-2\beta q_c+\sqrt{R_5}\right)}{2\rho}\,p \tag{117}$$

The infinitesimal variance is the same as for a monomorphic continent, $V(p) = p(1-p)$. Inserting $M_{\rho\gg0}(p)$ from Eq. (117) and $V(p)$ into the definition of $\psi(p)$ in Eq. (105), we obtain

$$\psi_{\rho\gg0}(p) = e^{-2\alpha p}(1-p)^{\frac{\mu\left(\beta-\mu-2\rho-2\beta q_c+\sqrt{R_5}\right)}{\rho}}. \tag{118}$$

The STDs $t_{1,\mathrm{QLE},\rho\gg0}(p;p_0)$ and $t_{2,\mathrm{QLE},\rho\gg0}(p;p_0)$ are found by insertion of $\psi_{\rho\gg0}(p)$ from Eq. (118) into Eq. (113). Exploiting the fact that $x^{-1}\int_0^x \psi_{\rho\gg0}(p)dp$ converges to 1 as x approaches 0, the STDs can be approximated by

$$\tilde{t}_{1,\mathrm{QLE},\rho\gg0}(p;p_0) = 2e^{2p\alpha}(1-p)^{\frac{\mu\left(\mu-\beta+2\beta q_c+2\rho-\sqrt{R_5}\right)}{\rho}-1}, \tag{119a}$$

$$\tilde{t}_{2,\mathrm{QLE},\rho\gg0}(p;p_0) = 2p_0e^{2p\alpha}p^{-1}(1-p)^{\frac{\mu\left(\mu-\beta+2\beta q_c+2\rho-\sqrt{R_5}\right)}{\rho}-1} \tag{119b}$$

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 $\tilde{t}_{i,\mathrm{QLE},\rho\gg0}(p;p_0)$ 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 $D_{\mathsf{AB}} = D$, D_{AC} and D_{CB} the linkage disequilibria between the indicated loci, and by $D_{\mathsf{ACB}} = y_1 - pqn - pD_{\mathsf{CB}} - qD_{\mathsf{AC}} - nD_{\mathsf{AB}}$ 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

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

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

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

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

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

(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 $r_{\mathsf{AB}} = r_{\mathsf{AC}} + r_{\mathsf{CB}}$. 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 $D_{\mathsf{AC}} = D_{\mathsf{CB}} = D_{\mathsf{ACB}} = 0$ 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 $(n, D_{\mathsf{AC}}, D_{\mathsf{CB}}, D_{\mathsf{ACB}})$ to $(n_c, 0, 0, 0)$. In the limit of weak migration, i.e. $m \ll (a, b, r)$, the latter is given by

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

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

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

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 $r_{\mathsf{AC}} = r_{\mathsf{AB}} + r_{\mathsf{BC}}$). In Eq. (87c), r must be replaced by r_{AC}. Then, the Jacobian $\mathbf{J}_N^{\mathsf{ABC}}$ approximating convergence of $(n, D_{\mathsf{AC}}, D_{\mathsf{BC}} = D_{\mathsf{CB}}, D_{\mathsf{ABC}} = D_{\mathsf{ACB}})$ to $(n_c, 0, 0, 0)$ in the limit of weak migration is equal to $\mathbf{J}_N^{\mathsf{ACB}}$ with the last entry of the last row replaced by $-a - b - r_{\mathsf{AC}} + \frac{m(a+b+3r_{\mathsf{AB}})}{a+b+r_{\mathsf{AB}}}$. To first order in m, the leading eigenvalue of $\mathbf{J}_N^{\mathsf{ABC}}$ is

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

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,

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

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

Recall that the Jacobian matrices $\mathbf{J}_N^{\mathsf{ACB}}$ and $\mathbf{J}_N^{\mathsf{ABC}}$ 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 $\mathsf{C}-\mathsf{A}-\mathsf{B}$. 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 agaist 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.

File S2

Deterministic analysis of a diploid two-locus continent–island model in discrete time.

File S2 is a PDF version of the *Mathematica* Notebook 2LocContIsland_Determ_Discr.nb (see File S10) and available for download at http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.114.163477/-/DC1.

File S3

Branching-process approximation of the invasion probability of a weakly beneficial mutation linked to an established polymorphism at migration–selection balance.

File S3 is a PDF version of the *Mathematica* Notebook 2LocContIsland_Stoch_Discr.nb (see File S10) and available for download at http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.114.163477/-/DC1.

File S4

Comparison of the Jacobian of the marginal one-locus migration–selection equilibrium (E_{B}) to the mean matrix of the corresponding branching process.

File S4 is a PDF version of the *Mathematica* Notebook 2LocContIsland_Compare_JacobianVsMeanMatrix.nb (see File S10) and available for download at http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.114.163477/-/DC1.

File S5

Analytical approximation of the invasion probability for a slightly supercritical branching process.

File S5 is a PDF version of the *Mathematica* Notebook 2LocContIsland_Stoch_Discr_SlightlySupercritBP.nb (see File S10) and available for download at http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.114.163477/-/DC1.

File S6

Derivative of the weighted mean invasion probability $\bar{\pi}$ at recombination rate $r = 0$.

File S6 is a PDF version of the *Mathematica* Notebook 2LocContIsland_Stoch_Discr_OptRecombRate.nb (see File S10) and available for download at http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.114.163477/-/DC1.

File S7

Diffusion approximation of sojourn and absorption times assuming quasi-linkage disequilibrium.

File S7 refers to the *Mathematica* Notebook 2LocContIsland_Stoch_DiffusionApprox_QLE.nb archived in File S10. File S10 is available for download at http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.114.163477/-/DC1.

File S8

The effective migration rate experienced by a neutral site linked to two loci at migration–selection balance.

File S8 is a PDF version of the *Mathematica* Notebook 2LocContIsland_Determ_effMigRate.nb (see File S10) and available for download at http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.114.163477/-/DC1.

File S9

The effect on neutral variation of migration and selection at two linked sites.

File S9 is a PDF version of the *Mathematica* Notebook 2LocContIsland_Stoch_NeutralLinkedMut.nb (see File S10) and available for download at http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.114.163477/-/DC1.

File S10

Archive of *Mathematica* Notebooks in the NB format (Files S2–S9).

File S10 is available for download as a ZIP file at http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.114.163477/-/DC1.

File S11

Archive of *Java* source code, binaries, and JAR files for simulations as described in Methods.

File S11 is available for download as a ZIP file at http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.114.163477/-/DC1.