The Evolutionary Dynamics of a Genetic Barrier to Gene Flow: From the Establishment to the Emergence of a Peak of Divergence

Takahiro Sakamoto and Hideki Innan[1]
SOKENDAI, The Graduate University for Advanced Studies, Hayama, Kanagawa 240-0193, Japan

ABSTRACT Divergent selection works when an allele establishes in the subpopulations in which it is adaptive, but not in the ones in which it is deleterious. While such a locally adaptive allele is maintained, the target locus of selection works as a genetic barrier to gene flow or a barrier locus. The genetic divergence (or F_{ST}) around the barrier locus can be maintained, while in other regions of the genome, genetic variation can be mixed by gene flow or migration. In this work, we consider theoretically the evolutionary process of a barrier locus, from its birth to stable preservation. Under a simple two-population model, we use a diffusion approach to obtain analytical expressions for the probability of initial establishment of a locally adaptive allele, the reduction of genetic variation due to the spread of the adaptive allele, and the process to the development of a sharp peak of divergence (genomic island of divergence). Our results will be useful to understanding how genomes evolve through local adaptation and divergent selection.

KEYWORDS speciation; population genetics; diffusion theory; migration; gene flow; divergent selection

A genomic island of divergence could arise when a locally adapted allele establishes in a certain subpopulation (*e.g.*, Wu 2001; Turner *et al.* 2005; Nosil 2012). This local establishment could be stably maintained by divergent selection if the allele confers sufficient benefit in the subpopulations in which it is adaptive, but not in the ones in which it is deleterious. Such a locus works as a genetic barrier to gene flow, or a barrier locus, because migrants are maladaptive. Due to recombination, the genomic region that is affected by divergent selection is limited, thereby creating a peak of divergence along the chromosome (*i.e.*, a genomic island of divergence). Further development of multiple barrier loci in the genome might initiate ecological speciation (Turner *et al.* 2005; Nosil 2012). Here, we are interested in the evolutionary dynamics of a barrier locus, from its establishment via a partial local sweep, through the emergence of a peak of divergence, to its stable preservation.

We consider the process theoretically by dividing into three phases—establishment, consolidation and equilibrium—as illustrated in Figure 1. We consider a simple situation with two subpopulations: I and II. Assuming a relatively high migration rate between them, the levels of polymorphism within the two subpopulations are similar to each other (measured by the heterozygosities, h_{w1} and h_{w2}, for subpopulations I and II). In the meantime, the population divergence (measured by h_b, the heterozygosity between the two subpopulations) is very low (Figure 1A). Then, a *de novo* mutation (star in Figure 1A) arises in subpopulation I, in which the mutation is advantageous, whereas it is maladaptive (or deleterious) in subpopulation II. In the establishment phase, the mutation spreads in subpopulation I and nearly fixes (Figure 1B), but its frequency in subpopulation II is low because it should be selected against if migrated into subpopulation II. In a strict sense, this is not a fixation that can be treated mathematically as an absorbing state, because migration keeps providing maladaptive alleles. Therefore, after Kimura (1954), we hereafter use the terminology of "quasi-fixation" for this nearly fixed state. The quasi-fixation should occur quickly, and a partial local selective sweep occurs in subpopulation I (Figure 1B), thereby establishing a barrier locus. Around the barrier locus, it is typical to observed a "block" of region with low genetic variation in

doi: https://doi.org/10.1534/genetics.119.302311
Manuscript received April 19, 2019; accepted for publication May 16, 2019; published Early Online June 6, 2019.
Supplemental material available at FigShare: https://doi.org/10.25386/genetics.8188019.
[1]Corresponding author: Graduate University for Advanced Studies, Hayama, Kanagawa 240-0193, Japan. E-mail: innan_hideki@soken.ac.jp

subpopulation I, with a slightly elevated genetic divergence (F_{ST}). The consolidation phase starts after the initial establishment of the barrier locus, during which the block of low genetic variation gradually shrinks in length over time by recombination and migration, while new mutations accumulate and the divergence between two subpopulations increases particularly near the barrier locus (Figure 1C). Then, at the end, a stable sharp peak of divergence arises in the equilibrium phase (Figure 1D). The equilibrium shape of the peak of divergence is determined mainly by the balance between selection intensity and the rates of recombination and migration.

The scope of this work is to provide a unified and comprehensive theoretical understanding of the evolution of a new peak of divergence, from its birth to stable preservation in equilibrium. We use a simple two-population model, where migration is allowed between subpopulations I and II. Suppose a *de novo* mutation arises that confers a selective advantage specific to subpopulation I, which is the initial state of our system. Under this model, we derive the following: for the establishment phase,

1. The establishment probability of the *de novo* mutation, that is, the probability that the mutation quasi-fixes in subpopulation I.
2. The expected reduction of genetic variation within subpopulations I and II after the quasi-fixation (*i.e.*, partial local sweep).

for the consolidation phase,

3. the evolutionary dynamics at both the barrier locus and the linked neutral sites since the quasi-fixation.

and for the equilibrium phase,

4. The expected shape of the peak of divergence at equilibrium.

Several theoretical works have focused on a specific part of these aspects. For (1) the establishment probability, perhaps the most flexible, useful theoretical framework was introduced by Barton (1987) in a general multiple-island-model. By using a diffusion approximation, Barton (1987) derived a partial differential equation for the establishment probability. Essentially the same result was obtained by Pollak (1966), who used a branching process, and the establishment probability was derived from a probability generating function. Barton's differential equation was solved, and closed forms of the establishment probability have been available only in several specific situations in *continuous* habitat models. In a one-dimensional continuous habitat model, Barton (1987) solved his partial differential equation analytically, assuming two forms of fitness gradient (linear and pocket). Kirkpatrick and Peischl (2013) used a branching process, from which they obtained a partial differential equation that is similar to that of Barton (1987). Then, the authors successfully incorporated changes in fitness gradient along time, and derived an approximate establishment probability.

In *discrete* population models, Barton's general formula (and also Pollak's one) is difficult to handle, and has not been fully explored even in a simple two-population model with symmetric migration. Therefore, the currently available theoretical results are not based on Barton's differential equation, and have some limitations. In a continent-island model with *unidirectional* migration, Aeschbacher and Bürger (2014) solved the establishment probability of a locally beneficial mutation linked to another locally beneficial mutation that was already established, where mathematical treatment is facilitated by unidirectional migration (see also Yeaman *et al.* 2016). Yeaman and Otto (2011) obtained an approximate establishment probability by using a heuristic approach that is a combination of the leading eigenvalue of the transition matrix of deterministic process and Kimura's formula of fixation probability (Kimura 1962). As shown in their paper, this formula well describes the establishment probability when a *de novo* mutation arises in the adapted subpopulation (*i.e.*, subpopulation I in our model), but it does not work when it arises in the maladapted subpopulation (*i.e.*, subpopulation II in our model). Recently, Tomasini and Peischl (2018) provided an approximate establishment probability by assuming a slightly supercritical branching process. Their formula works well under the assumption of slightly supercritical approximation, namely, the leading eigenvalue of the transition matrix of deterministic model is not large, but it may not work well when the selection intensity in the adapted subpopulation is very large.

In this work, we derive a closed form formula of the establishment probability in a two-population model with bidirectonal migration along the formulation of Barton (1987). We extend Barton's derivation with simultaneous quadratic equations and solve them allowing unequal subpopulation sizes. Our formula is more general than previous ones (Yeaman and Otto 2011; Tomasini and Peischl 2018); it works with strong selection and it allows that a *de novo* mutation can arise either subpopulation I or II.

To the best of our knowledge, there is no theoretical work on the hitch-hiking process of a partial local sweep in a two-population model. With regard to a single population model, many studies investigated the reduction of polymorphism theoretically due to a selective sweep. These studies considered a selected site and a linked neutral site, and assumed that a sufficiently advantageous mutation arises and goes to fixation in the population. Along this fixation, they derived how much polymorphism can be reduced at the linked site. Maynard Smith and Haigh (1974) first obtained the reduction of polymorphism, where the stochastic effect of genetic drift at the linked site was ignored. The model was extended to include the stochastic effect by using a coalescent approach (Kaplan *et al.* 1989) and by using a diffusion method (Stephan *et al.* 1992; see also Barton 1998; Etheridge *et al.* 2006). Durrett and Schweinsberg (2004) used a different approach for a faster approximate simulation of a selective sweep and derived some analytical expressions (see also Schweinsberg and Durrett 2005).

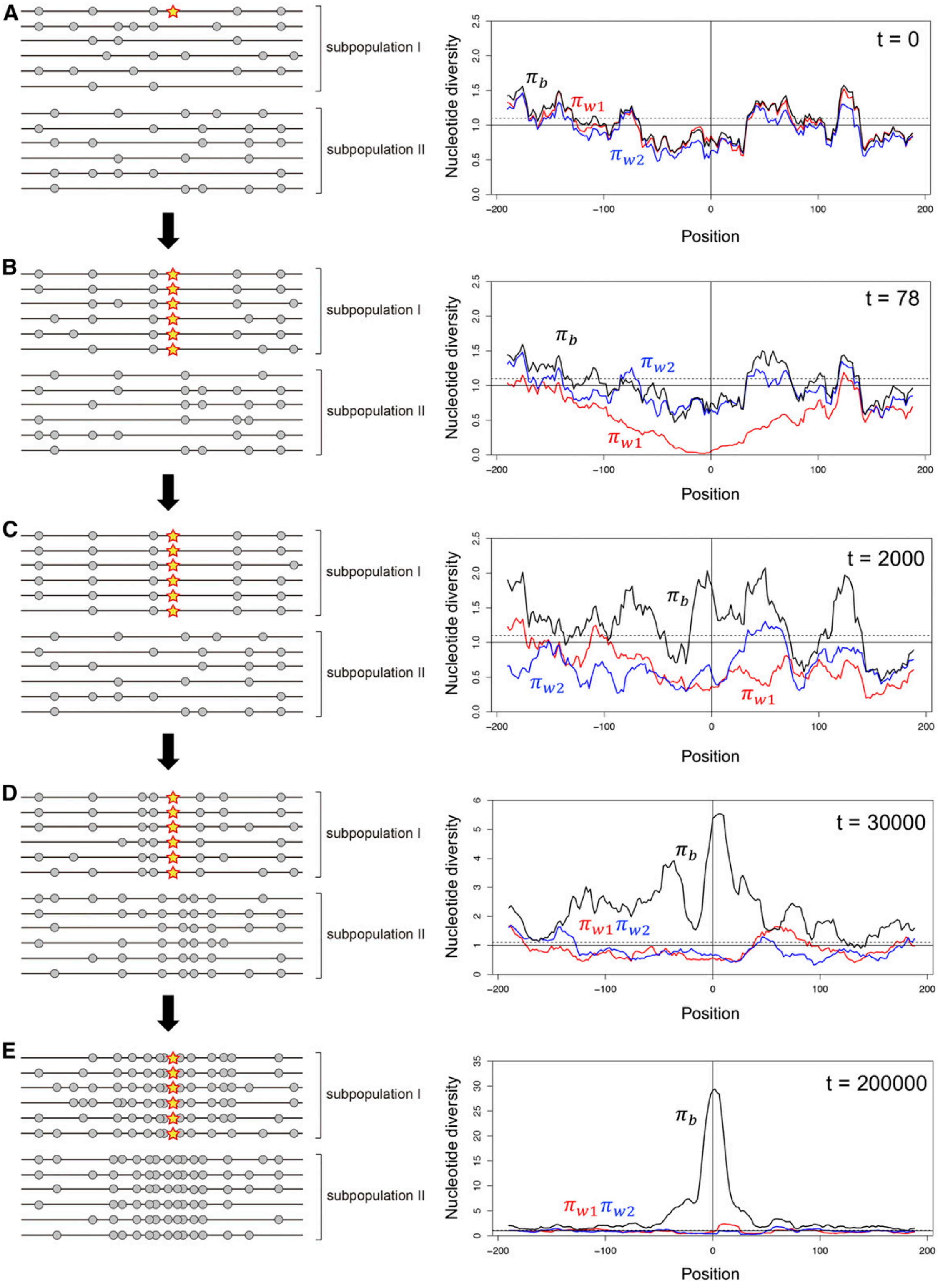

Figure 1 Illustrating the evolution of a barrier locus in a simple two-population model with fairly high migration between them. (A) A locally adaptive *de novo* mutation arises in subpopulation I at position 0. A typical pattern of polymorphism is shown in left. The star is the locally adaptive mutation, and gray circles are neutral polymorphism in the surrounding region. The right panel shows the spacial distributions of nucleotide diversity obtained by a simulation. The simulation considers two subpopulations with population sizes are $2N_1 = 2N_2 = 2000$, between which symmetric migration is allowed at rate $4N_1m_1 = 4N_2m_2 = 5.0$. We assume selection intensity $s_1 = 0.2$ and $s_2 = -0.2$. The entire simulated region is 400 kb if a population recombination rate of $4Nr$ = 0.001 per site is assumed. See Appendix A for details about the simulation. The polymorphism levels within the two populations (π_{w1} and π_{w2}) are in red and blue, and divergence between the two populations (π_b) is in black. π_{w1}, π_{w2}, and π_b can be considered as the averages of h_{w1}, h_{w2}, and h_b in a 20-kb window. The *y*-axis is adjusted such that $E(\pi_{w1}) = E(\pi_{w2}) = 1$ under neutrality (the solid line) and the

There are several theoretical studies on a sweep in multi-population models available, but these considered a fixation across multiple subpopulations, not a local fixation. In a model with multiple subpopulations, Slatkin and Wiehe (1998) and Santiago and Caballero (2005) considered the process where a beneficial mutation fixes in the entire population through weak migration. Kim and Maruki (2011) allowed stronger migration, and derived an analytical expression in a two-population model. Our interest is different from these studies in that we consider a locally beneficial mutation that can quasi-fix only in the subpopulation in which it is beneficial (not in the entire population). We here extend the theory of Stephan's diffusion model (Stephan *et al.* 1992) to a two-population model, and consider how much polymorphism can be reduced locally at a linked site after a partial local sweep.

We then turn to the evolutionary dynamics at both the barrier locus and the linked neutral sites after the completion of the partial local sweep. We here consider this process after a local sweep as described in Figure 1. A local sweep creates a "block" of a fairly long region with almost no genetic variation within the subpopulation in which the new mutation is adaptive (*i.e.*, subpopulation I in our model). In this work, given an arbitrary configuration of genetic variation after a local sweep, we obtain, analytically, the moments of allele frequency at a linked site, with which we describe how a genomic island decays. Yeaman *et al.* (2016) investigated a similar problem in a different situation, where an genomic island evolves due to the clustering of two barrier loci. In their model, considering a secondary contact, erosion starts when there already are a large number of fixed sites that spread over the genome, and islands appear because selection works to maintain divergence at selected site(s), while losing divergence in other regions through homogenization by migration. By using the structured coalescent, they obtained the expected spatial distribution of F_{ST} (in terms of relative coalescent time) around selected sites as a function of the time since the secondary contact. They also considered the scenario where a *de novo* mutation broadens a genomic island that has been created by a barrier locus, and revealed the final shape of a two-barrier island is the same as the genomic island under the secondary-contact scenario. However, their derivation did not consider the effect of selective sweep of the *de novo* mutation. It should be noted that, because our derivation accepts any arbitrary initial allele frequency at a linked site, it can be applied to any situation, not only that after a secondary contact but also that after a local sweep.

In the equilibrium phase, the balance between selection, migration, recombination, and mutation holds. Theoretical treatment at equilibrium is relatively straightforward, and there are several theoretical studies on the spatial distribution of F_{ST} (Charlesworth *et al.* 1997; Akerman and Bürger 2014; Yeaman *et al.* 2016). Under our framework for the consolidation phase, essentially the same result can be provided as a special case, with time going to infinity.

Model and Results

We consider a two-population model with discrete generations and monoecious diploid individuals that mate at random. The diploid population sizes of subpopulations I and II are assumed to be constant at N_1 and N_2, respectively. As illustrated in Figure 1, we are specifically interested in selection for local adaptation in subpopulation I. We consider a genomic region encompassing a selected site at position 0, which is referred to as locus A (Figure 2). At locus A, two alleles (A/a) are allowed with no recurrent mutation between them. Allele A confers a selection coefficient s_1 in subpopulation I and s_2 in subpopulation II (we assume $s_1 > 0$ and $s_2 < 0$). Additive selection is assumed so that the fitness of individuals with AA, Aa, and aa are given by $1 + 2s_1$, $1 + s_1$, and 1 in subpopulation I, and $1 + 2s_2$, $1 + s_2$ and 1 in subpopulation II. Selection works only at this selected site, and all remaining sites are assumed to be neutral. For the following derivation under a two-locus model, we consider a secondary neutral site (locus B), at which two alleles (B/b) are allowed with recurrent mutation between them (Figure 2). The mutation rate from allele B to allele b is u, and that from allele b to allele B is v. The recombination rate between the two loci, A and B, is r.

The system starts when a *de novo* mutation (allele A) arises in a single individual either in population I or II, where allele a is fixed in both subpopulations. Therefore, the initial state is $(x_1, x_2) = (1/2N_1, 0)$ or $(0, 1/2N_2)$, where x_1 and x_2 are frequencies of the new allele A in subpopulations I and II, respectively. Throughout this article, we assume strong selection and weak migration so that maladapted individuals are rare in each subpopulation once the initial establishment is achieved.

Establishment probability

We derive the establishment probability of a new *de novo* allele using the general framework of Barton (1987), who derived a simultaneous quadratic equation from the diffusion theory. This section focuses only on the selected locus A (see Figure 2), at which we are interested in the probability that allele A quasi-fixes in subpopulation I. Following Haldane (1927), we approximate the establishment probability by the probability that the new mutation increases in frequency and escapes from immediate extinction. This is because,

broken line exhibits $E(\pi_b)$, (*i.e.*, = the genomic average). (B) The mutation quasi-fixes in subpopulation I, causing a drastic reduction in π_{w1}. (C) Migration shuffles polymorphisms in the two subpopulations, while selection works to maintain the quasi-fixation of the mutation. (D) The divergence gradually increases around the mutation, and (E) a clear peak of divergence arises. It should be noted that the star is fixed in the sample (left panes in B–E), but it does not necessarily mean that it is fixed in subpopulation I because there should be maladaptive immigrant alleles at a low frequency.

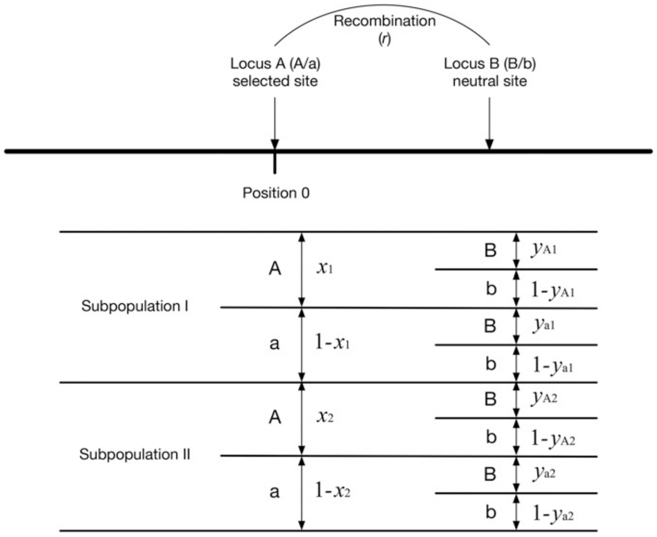

Figure 2 Two-locus model used in this work. Locus A targeted by divergent selection is placed at position 0, and a linked neutral locus B can be placed at an arbitrary position. The frequencies of allele A and allele a at locus A, and those of allele B and allele b at locus B in the two subpopulations are illustrated.

under the assumption of strong selection, the behavior of such a mutation is almost deterministic once it escapes from extinction by genetic drift.

Let $F(x_1, x_2)$ be the establishment probability when the frequencies of allele A are x_1 and x_2 in the two subpopulations. By using an analogous procedure to Barton (1987), we derive $p_1 = F(1/2/N_1, 0)$ and $p_2 = F(0, 1/2N_2)$, the establishment probability when the new allele arises in subpopulations I and II, respectively. According to the diffusion theory, F satisfies the Kolmogorov backward equation:

$$\begin{aligned} 0 = \frac{x_1}{4N_1}\frac{\partial^2 F}{\partial x_1^2} + \frac{x_2}{4N_2}\frac{\partial^2 F}{\partial x_2^2} + \{s_1x_1 + m_1(x_2 - x_1)\}\frac{\partial F}{\partial x_1} \\ + \{s_2x_2 + m_2(x_1 - x_2)\}\frac{\partial F}{\partial x_2}, \end{aligned} \tag{1}$$

where $m_1(m_2)$ is the proportion of immigrant individuals just after migration in subpopulation I (II). To keep the subpopulation sizes constant, we assume $N_1m_1 = N_2m_2$, and we ignore higher order terms of $o(x_i)$ ($i.e., x_1^2, x_2^2$). This is reasonable because of the assumption that the establishment probability is determined mainly at low frequencies. Because the extinction probabilities of individual mutations are independent, we can write F as

$$F(x_1, x_2) = 1 - \exp(-2N_1x_1\psi_1 - 2N_2x_2\psi_2) \tag{2}$$

where $\exp(-\psi_i)$ is the extinction probability of a new mutant in subpopulation i; therefore, p_i is determined as $p_i = 1 - \exp(-\psi_i)$. After substitution of Equation 2 into Equation 1, one can show that solutions to Equation 1 can be obtained by solving the following system of equations:

$$\begin{aligned} \psi_1^2 &= 2(s_1 - m_1)\psi_1 + 2\frac{N_2}{N_1}m_2\psi_2 \\ \psi_2^2 &= 2\frac{N_1}{N_2}m_1\psi_1 + 2(s_2 - m_2)\psi_2, \end{aligned} \tag{3}$$

which corresponds to equation 4b in Barton (1987). Equation 3 can be rearranged to

$$\psi_1\left(\psi_1^3 - 2a\psi_1^2 + \left(a^2 - bd\right)\psi_1 + b(ad - bc)\right) = 0 \tag{4}$$

$$\psi_2 = \frac{\psi_1^2 - a\psi_1}{b}, \tag{5}$$

where $a = 2(s_1 - m_1)$, $b = 2\frac{N_2}{N_1}m_2 = 2m_1$, $c = 2\frac{N_1}{N_2}m_1 = 2m_2$, and $d = 2(s_2 - m_2)$. Equation 4 can be solved by using the solution of a cubic equation. Equations 4 and 5 have, at most, one solution that fulfills $p_1 > 0$ and $p_2 > 0$. The condition where Equations 4 and 5 have such a solution is $a + d > 0$ or $ad - bc < 0$, which corresponds to the situation where the deterministic growth rate of the mutant allele is positive (see Appendix B for details).

Figure 3 shows the establishment probability from Equations 4 and 5 as a function of migration rate. We first consider a symmetric model ($N_1 = N_2 = 1000$), and two selection intensities ($s_1 = 0.02$ and $s_1 = 0.1$) are assumed, while $s_2 = -0.01$ is fixed (Figure 3, A and B). The establishment probability can be computed when a locally adaptive mutation arises either in subpopulation I or II, represented as $F(1/2N_1, 0)$ and $F(0, 1/2N_2)$, respectively. We performed a forward simulation to check the performance of our analytical result (Appendix A). For each parameter set, we ran 1,000,000 independent replications of the simulation, and counted the number of replications where the new allele A was preserved in 10,000 generations. The establishment probability was then obtained as the proportion of such replications. Therefore, it includes replications where two alleles (A and a) coexisted (case C) and those where allele A is completely fixed in both subpopulations (case F). The proportion of case C in the established replications (Pc) decreases with increasing migration rate (see below).

Our result (red in Figure 3) is in excellent agreement with the simulation result: $F(1/2N_1, 0)$ is approximately $F_{m=0} = \frac{1 - \exp(-2s_1)}{1 - \exp(-4N_1s_1)}$ when the migration rate is very low, consistent with the prediction in a single population model (Kimura 1957). As the migration rate increases, $F(1/2N_1, 0)$ decreases and $F(0, 1/2N_2)$ increases, and they become similar to each other. With a very high migration rate ($m \sim 0.5$), the two subpopulations can be considered as a single random-mating population, and the fixation probability of a single mutation is mainly determined by the average selection coefficient, $\bar{s} = \frac{s_1N_1 + s_2N_2}{N_1 + N_2}$, namely, $F_{m=0.5} = \frac{1 - \exp(-2\bar{s})}{1 - \exp(-4N_T\bar{s})}$ where $N_T = N_1 + N_2$ (Nagylaki 1980). Indeed, in our simulations, allele A was fixed in both populations in almost all established cases ($Pc = 1$). In each panel in Figure 3, a gray region is placed such that $Pc > 0.9$ in the

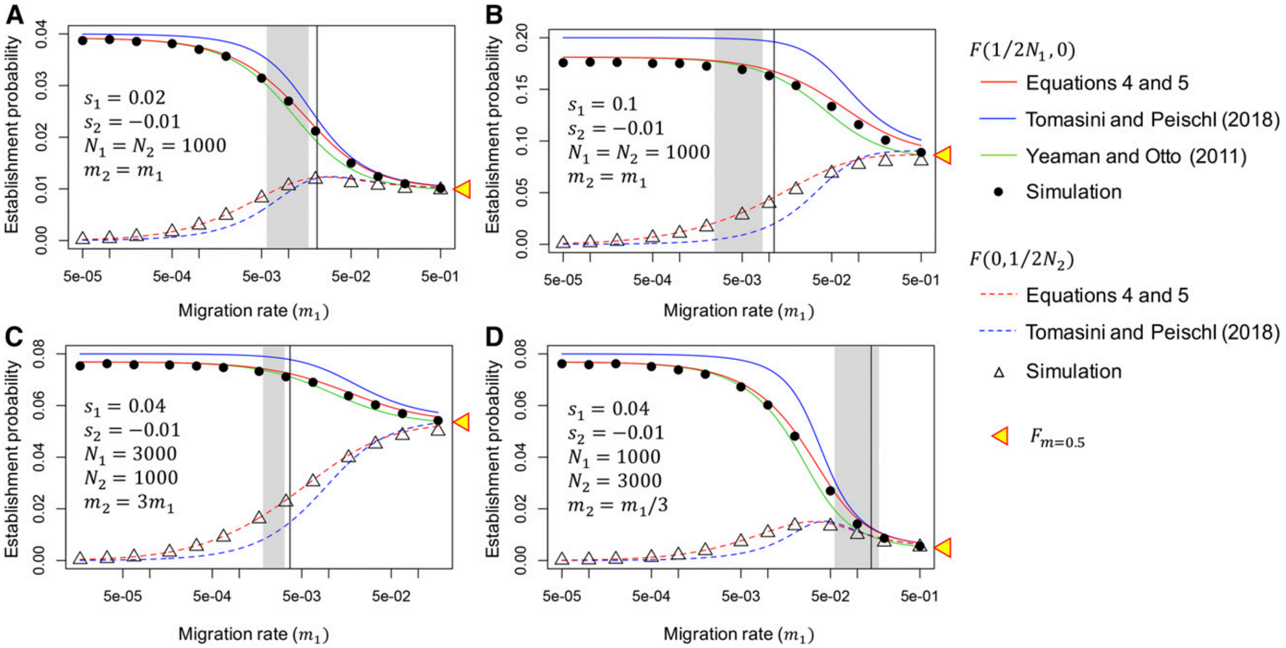

Figure 3 Establishment probability as a function of migration rate. (A) Weak selection ($s_1 = 0.02$ and $s_2 = -0.01$) and (B) strong selection ($s_1 = 0.1$ and $s_2 = -0.01$) are assumed in a symmetric model ($N_1 = N_2$). (C and D) Asymmetric population settings are considered ($N_1 = 3N_2$ in C and $N_1 = N_2/3$ in D). Our result in red is compared with those of Yeaman and Otto (2011) and Tomasini and Peischl (2018), together with the result of our forward simulation. The establishment probability for a mutation that arises in subpopulation I ($F(1/2N_1, 0)$) is shown by solid lines and closed circles, and that for a mutation that arises in subpopulation II ($F(0, 1/2N_2)$) is shown by broken lines and open triangles. The establishment probability at the high migration limit ($m = 0.5$) is shown by a yellow triangle. In each panel, a gray region is placed such that the proportion of the replications where two alleles (A and a) coexisted (Pc) > 0.9 in the left, while $Pc < 0.1$ in the right. The vertical line presents the critical migration rate, above which allele A fixes in the entire population, obtained by Equation 6 (see text for detail).

left, while $Pc < 0.1$ in the right. It has been demonstrated that, under a deterministic model, the condition where two alleles (A and a) coexist is $s_1 s_2 < 0$ and

$$\left| \frac{m_1(1+s_1)}{s_1} + \frac{m_2(1+s_2)}{s_2} \right| < 1 \tag{6}$$

(Nagylaki and Lou 2008). The critical migration rate predicted by this equation is shown by the vertical lines in Figure 3, which roughly agrees with the line of $Pc = 0.9$ (see Yeaman and Otto 2011). This indicates that the pattern dramatically changes in a short range of m_1, and the left side is the scope of this article. Similar results were also obtained in asymmetric models ($N_1 = 3N_2$ in Figure 3C and $N_1 = N_2/3$ in Figure 3D).

Figure 3 quantitatively compares our analytical results with those of previous studies (Yeaman and Otto 2011; Tomasini and Peischl 2018). It is found that $F(1/2N_1, 0)$ from Yeaman and Otto (2011) is almost as good as ours, but unfortunately $F(0, 1/2N_2)$ was not provided by Yeaman and Otto (2011). It seems that Tomasini and Peischl (2018) overestimates $F(1/2N_1, 0)$ and underestimates $F(0, 1/2N_2)$.

Reduction of genetic variation due to a selective sweep

When a new locally adaptive mutation (a → A) arises and quasi-fixes in subpopulation I, genetic variation in the surrounding region in subpopulation I should be reduced dramatically due to the hitch-hiking effect. In this section, we consider a two-locus model as defined in Figure 2. We derive the degree of reduction in heterozygosity at a linked neutral site (locus B) in subpopulation I, D_{LS}, by extending the diffusion approach of Stephan *et al.* (1992), who investigated the effect of hitch-hiking in a single population model with no population structure.

Overview of Stephan's diffusion approach: We first briefly introduce the approach of Stephan *et al.* (1992), which provides the basis of our derivation below. The expected reduction of heterozygosity at locus B for a single population model with diploid size N is denoted by D_0. With the assumption of strong selection, Stephan *et al.* (1992) assumed that the behavior of the frequency (x) of the beneficial allele A with selection coefficient, s, follow a deterministic function:

$$\frac{dx}{dt} = sx(1-x), \tag{7}$$

where selection is additive. This deterministic treatment works once the frequency of allele A exceeds a certain threshold such that it escapes from immediate extinction by genetic drift, as mentioned in the previous section. This treatment makes the following derivation much easier because the dynamics can be described by a two-dimensional diffusion equation. It should be noted that x with no subscript denotes the frequency of allele A in the single population model, whereas, in our two-population model, the frequencies of allele A in subpopulations I and II are denoted by x_1 and x_2, respectively (see Figure 2). We consider another biallelic neutral locus (B/b), and the recombination rate between this neutral locus and the selected locus is assumed to be r. y_A is the frequency of allele B among A-chromosomes, and y_a is the frequency of allele B among a-chromosomes. Then, the expected changes of an arbitrary function $f(y_A, y_a)$ is described as the following ordinary differential equation:

$$\frac{d}{dt}E(f) = E(L(f)), \tag{8}$$

where L is a differential operator of the Kolmogorov backward equation:

$$\begin{aligned} L = &\frac{y_A(1-y_A)}{4Nx}\frac{\partial^2}{\partial y_A^2} + r(1-x)(y_a - y_A)\frac{\partial}{\partial y_A} \\ &+ \frac{y_a(1-y_a)}{4N(1-x)}\frac{\partial^2}{\partial y_a^2} + rx(y_A - y_a)\frac{\partial}{\partial y_a}. \end{aligned} \tag{9}$$

By using this formula, Stephan *et al.* (1992) solved the first and second moments of y_A and y_a after a sweep, from which the expected reduction of heterozygosity at the linked site can be computed numerically. With some approximation, Stephan *et al.* (1992) further obtained a nice closed form of the solution:

$$D_0 = \frac{2r}{s}(2Ns)^{-2r/s}\Gamma\left(-\frac{2r}{s}, \frac{1}{2Ns}\right). \tag{10}$$

In this work, we found that this equation somehow undervalues the effect of random genetic drift at the linked neutral locus, perhaps due to the approximation of Stephan *et al.* (1992). We noted that, in Equation 10, D_0 goes to $\exp(-1/2Ns)$ in the limit of $r \to \infty$, whereas heterozygosity should decrease by genetic drift by a factor of $1/2N$ per generation, even in the absence of the hitch-hiking effect. Here, we consider the expected reduction of heterozygosity along the quasi-fixation as $\exp(-\log(2N)/Ns)$, because the fixation time is approximately given by

$$T = \int_{1/2N}^{1-1/2N} \frac{dx}{sx(1-x)} \approx \frac{2\log(2N)}{s}.$$

This equation means that the expected reduction of heterozygosity due to genetic drift, $\exp(-\log(2N)/Ns)$, is not negligible compared to $\exp(-1/2Ns)$. To correct for this factor, we add this into Equation 10:

$$D_0' = \frac{2r}{s}(2Ns)^{-2r/s}\Gamma\left(-\frac{2r}{s}, \frac{1}{2Ns}\right)\exp\left(-\frac{\log(2N)}{Ns}\right). \tag{11}$$

We found that this heuristic approach is in very good agreement with the numerical solution obtained by directly computing Equation 9.

Local sweep in the two-population model: In this work, we extend Stephan derivation (Stephan *et al.* 1992) to the two-population model defined above. We first consider the dynamics of the new mutant allele frequency (x_1) at the selected locus (position 0) in subpopulation I. The major difference from the corresponding formula in Stephan *et al.* (1992) (*i.e.*, Equation 7) is that the effect of migration should be considered in the two-population model. Because allele A is very rare in subpopulation II under the assumption of strong selection and low migration, we can ignore migrants with A allele from subpopulations II to I. Then, the dynamics of x_1 can be approximated by a deterministic function:

$$\frac{dx_1}{dt} = s_1x_1(1-x_1) - m_1x_1. \tag{12}$$

We set the time such that $t = 0$ when the mutation arises, and $t = \tau$ when the mutation quasi-fixes. We next consider the neutral locus B (B/b). As illustrated in Figure 2, y_{A1} (y_{A2}) is the frequency of haplotype A-B among A-chromosomes in subpopulation I (II), and y_{a1} (y_{a2}) is the frequency of haplotype a-B among a-chromosomes in subpopulation I (II). We assume that y_{A2} is very small throughout the sweep process. Then, the expected changes of an arbitrary function $f(y_{A1}, y_{a1}, y_{a2})$ is described as the following ordinary differential equation:

$$\frac{d}{dt}E(f) = E(L(f)), \tag{13}$$

where L is a differential operator of the Kolmogorov backward equation. Following Ohta and Kimura (1969), we obtain L for our model as

$$\begin{aligned} L = &\frac{y_{A1}(1-y_{A1})}{4N_1x_1(t)}\frac{\partial^2}{\partial y_{A1}^2} + r(1-x_1(t))(y_{a1} - y_{A1})\frac{\partial}{\partial y_{A1}} \\ &+ \frac{y_{a1}(1-y_{a1})}{4N_1(1-x_1(t))}\frac{\partial^2}{\partial y_{a1}^2} + \left\{rx_1(t)(y_{A1} - y_{a1}) + \frac{m_1}{(1-x_1(t))(1-m_1)+m_1}(y_{a2} - y_{a1})\right\}\frac{\partial}{\partial y_{a1}} \\ &+ \frac{y_{a2}(1-y_{a2})}{4N_2}\frac{\partial^2}{\partial y_{a2}^2} + \left\{x_1(t)m_{e,1\to 2}(y_{A1} - y_{a2}) + (1-x_1(t))m_2(y_{a1} - y_{a2})\right\}\frac{\partial}{\partial y_{a2}}. \end{aligned} \tag{14}$$

This equation is derived such that several terms are added to Equation 9 for incorporating random genetic drift within subpopulation II (first term on the third line) and the effect of migration. The second term of $\partial/\partial y_{a1}$ (second line) is for migration from subpopulation II to subpopulation I, and the term of $\partial/\partial y_{a2}$ (third line) is for migration from subpopulation I to subpopulation II. Due to the assumption of strong selection, migrant A-chromosomes from subpopulation I to subpopulation II should be selected out immediately. Therefore, the migration rate of locus B can be *effectively* considered as the product of migration rate and the probability that at least one recombination event occurs before selection purges allele A, $m_{e,1\to 2}$:

$$m_{e,1\to 2} = \frac{(1+s_2)r}{1-(1+s_2)(1-r)} m_2 \tag{15}$$

(Bengtsson 1985). Then, Equation 13 directly allows us to compute the first and second moments of y_{A1} and y_{a2} after the quasi-fixation of allele A (*i.e.*, $y_{A1}(\tau)$ and $y_{a2}(\tau)$. We obtain heterozygosity within each subpopulation (h_{w1} and h_{w2}) and between them (h_b) at $t=\tau$ as

$$\begin{aligned} h_{w1}(\tau) &= 2E(y_{A1}(\tau)) - 2E\left(y_{A1}(\tau)^2\right), \\ h_{w2}(\tau) &= 2E(y_{a2}(\tau)) - 2E\left(y_{a2}(\tau)^2\right), \\ h_b(\tau) &= E(y_{A1}(\tau)) + E(y_{a2}(\tau)) - 2E(y_{A1}(\tau)y_{a2}(\tau)), \end{aligned} \tag{16}$$

from which the expected reduction of heterozygosity is obtained as

$$D_{LS} = h_{w1}(\tau)/h_{w1}(0). \tag{17}$$

Generally, D_{LS} involves the initial frequencies, $y_{a1}(0)$ and $y_{a2}(0)$. However, it should be noted that their quantitative effect on D_{LS} is not large unless $y_{a1}(0)$ and $y_{a2}(0)$ are not very similar.

Figure 4 shows the effect of migration on the reduction in heterozygosity. The plot in red is the case of no migration, where our result is essentially identical to that of Stephan *et al.* (1992), and the plots in blue and green are for migration cases. We consider three pairs of population sizes, $N_1 = N_2 = 1000$ in A, $N_1 = 1000, N_2 = 5000$ in B, and $N_1 = 5000, N_2 = 1000$ in C. For each parameter set, filled circles represent the average over 100,000 replications of forward simulation (see Appendix A). In Figure 4, $h_{w1}(\tau)$, $h_{w2}(\tau)$, and $h_b(\tau)$ are plotted such that $h_{w1}(0) = h_{w2}(0) = 1$ before the sweep, so that $h_{w1}(\tau)$ directly corresponds to D_{LS}. In all cases, our theoretical result from Equation 14 is in excellent agreement with the simulation results. It is found that the effect of a local partial sweep seems to be only on subpopulation I, and there is almost no effect on the variation in subpopulation II. Moving away from the selected site at position 0, D_{LS} is larger for a higher migration rate. This is because that migration brings standing variation maintained in subpopulation II into subpopulation I, thereby increasing the polymorphism level in subpopulation I. We observed $h_b(\tau)$ is slightly elevated around the selected site at position 0. If we assume $1 - h_w(\tau)/h_{all}(\tau)$ roughly approximates $\mathrm{F_{ST}}$, where h_{all} is heterozygosity when the two subpopulations are merged together, it can be said that a local sweep creates a relatively wide block of region with elevated $\mathrm{F_{ST}}$, which can be considered as an initial peak of divergence.

Consolidation of a barrier locus with a peak of divergence

When a new locally adaptive mutation (a → A) quasi-fixes in subpopulation I, a block of region with elevated $\mathrm{F_{ST}}$ arises, where genetic variation in subpopulation I is dramatically reduced (Figure 1B). In this section, by using the two-locus model defined in Figure 2, we consider the process after this state, but our derivation is flexible enough to plug in any initial state.

We use a similar diffusion approach to the previous section but we focus on the behavior of y_{A1} and y_{a2}. The expected changes of an arbitrary function $f(y_{A1}, y_{a2})$ is described as the following ordinary differential equation:

$$\frac{d}{dt}E(f) = E(L(f)), \tag{18}$$

where L is a differential operator of the Kolmogorov backward equation, which is given by

$$\begin{aligned} L = {} & \frac{y_{A1}(1-y_{A1})}{4N_1}\frac{\partial^2}{\partial y_{A1}^2} + \frac{y_{a2}(1-y_{a2})}{4N_2}\frac{\partial^2}{\partial y_{a2}^2} + \big[v-(u+v)y_{A1} \\ & + m_{e,2\to 1}(y_{a2}-y_{A1})\big]\frac{\partial}{\partial y_{A1}} + \big[v-(u+v)y_{a2} \\ & + m_{e,1\to 2}(y_{A1}-y_{a2})\big]\frac{\partial}{\partial y_{a2}}. \end{aligned} \tag{19}$$

The two terms in the first line of Equation 19 are for random genetic drift within subpopulations I and II, and the terms in the second line describes the deterministic change of the frequency of allele B due to mutation and migration. As well as the previous section, we use the effective migration rate (Bengtsson 1985):

$$m_{e,2\to 1} = \frac{(1+\tilde{s}_1)r}{1-(1+\tilde{s}_1)(1-r)} m_1, \tag{20}$$

where $\tilde{s}_1 = 1/(1+s_1) - 1$ is the relative selection coefficient of maladapted individuals in subpopulation I. $m_{e,1\to 2}$ is defined by Equation 15. We consider the dynamics of the first and second order moments, and put $\boldsymbol{y} = (E(y_{A1}), E(y_{a2}), E(y_{A1}^2), E(y_{A1}y_{a2}), E(y_{a2}^2))^T$. By using Equation 18, we derive a differential equation for $\boldsymbol{y}$ as follows:

$$\frac{d\boldsymbol{y}}{dt} = \boldsymbol{Q}\boldsymbol{y} + \boldsymbol{e}, \tag{21}$$

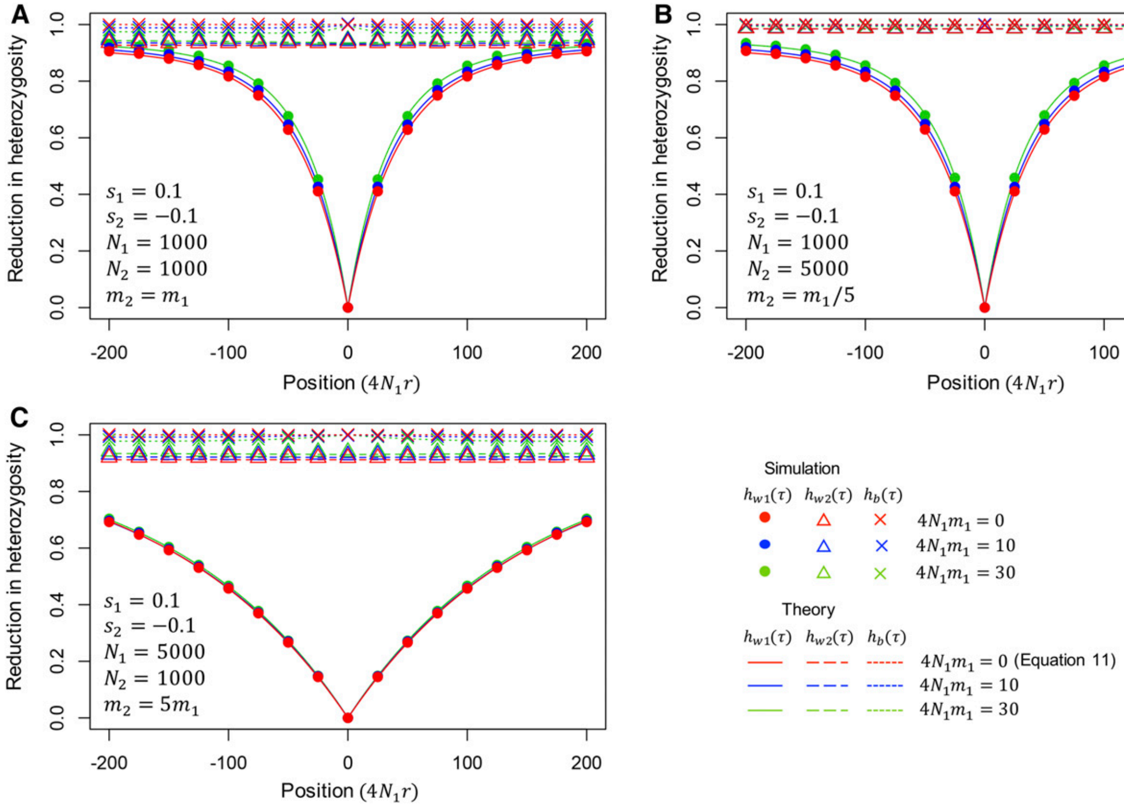

Figure 4 The expected reduction of heterozygosity after a local partial sweep in the two-population model. Position is shown in $4N_1r$ from the selected site. Theoretical results for $h_{w1}(\tau)$ $h_{w2}(\tau)$ and $h_b(\tau)$ computed from (14)–(17) by assuming $y_{a1}(0) = y_{a2}(0) = 0.3$ for convenience, but very similar results were obtained for other values of $y_{a1}(0)$ and $y_{a2}(0)$. In the case of no migration (red), our results is identical to Stephan *et al.* (1992) (*i.e.*, Equation 11). Results for three parameter sets are shown: (A) s_1=0.1, s_2=-0.1,N_1=1000, N_2=1000, m_2=m_1, (B) s_1=0.1, s_2=-0.1,N_1=1000, N_2=5000, m_2=m_1/5, (C) s_1=0.1, s_2=-0.1,N_1=5000, N_2=1000, m_2=5m_1.

where $\boldsymbol{Q}$ is the 5 × 5 matrix given by

$$\boldsymbol{Q} = \begin{pmatrix} -(u+v+m_{e,2\to1}) & m_{e,2\to1} & 0 & 0 & 0 \\ m_{e,1\to2} & -(u+v+m_{e,1\to2}) & 0 & 0 & 0 \\ 2v+\dfrac{1}{2N_1} & 0 & -2(u+v+m_{e,2\to1}+\dfrac{1}{4N_1}) & 2m_{e,2\to1} & 0 \\ v & v & m_{e,1\to2} & -(2u+2v+m_{e,2\to1}+m_{e,1\to2}) & m_{e,2\to1} \\ 0 & 2v+\dfrac{1}{2N_2} & 0 & 2m_{e,1\to2} & -2(u+v+m_{e,1\to2}+\dfrac{1}{4N_2}) \end{pmatrix} \tag{22}$$

and $\boldsymbol{e} = (v, v, 0, 0, 0)^T$. See Appendix C for details. By solving Equation 21, $\boldsymbol{y}$ is given by

$$\boldsymbol{y}(t) = \exp(t\boldsymbol{Q})\boldsymbol{y}(0) + \boldsymbol{Q}^{-1}(\exp(t\boldsymbol{Q}) - \boldsymbol{I})\boldsymbol{e} \tag{23}$$

where $\boldsymbol{I}$ is the identity matrix of size 5 (Appendix C). $\boldsymbol{y}$ at equilibrium is given by $\tilde{\boldsymbol{y}} = -\boldsymbol{Q}^{-1}\boldsymbol{e}$. Our solution at equilibrium is well consistent with previous studies (Charlesworth *et al.* 1997; Yeaman *et al.* 2016) that used the coalescent approach (see Appendix D for a proof).

Figure 5 compares our theoretical results from Equation 23 (broken lines) with simulation results (closed circles). $N_1 = N_2 = 1000, s_1 = -s_2 = 0.05, u = v = 2.5 \times 10$ $-6, m_1 = m_2 = 1.25 \times 10^{-3}$ are assumed to represent a strong selection case. As the initial condition $(t = 0)$, we set $h_{w1} = 0, h_{w2} = 0.18$, and $h_b = 0.1$, representing a situation after a local sweep in subpopulation I. Equation 23 describes how a

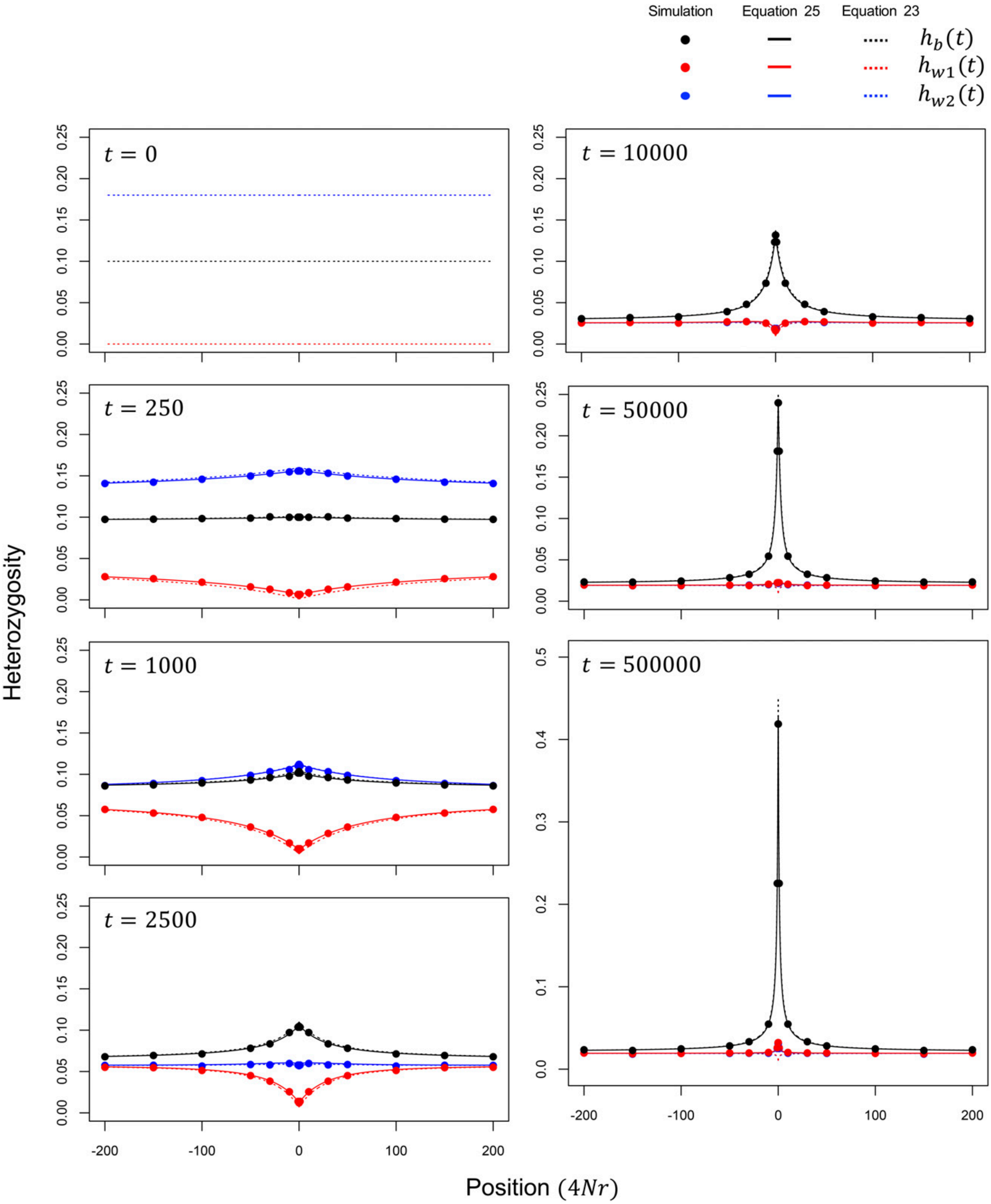

Figure 5 Temporal change of heterozygosity (h_{w1}, h_{w2}, h_b) after a local sweep in subpopulation I. The spacial distributions of h_{w1}, h_{w2}, and h_b are shown for seven time points (t = 0, 250, 1000, 2500, 10,000, 50,000, and 500,000 generations after a sweep). Position in the simulated regions is shown in $4Nr$ from the selected site. $N_1 = N_2 = 1000, s_1 = -s_2 = 0.05, u = v = 2.5 \times 10^{-6}, m_1 = m_2 = 1.25 \times 10^{-3}$, $y_1(0) = 0.0$ and $y_2(0) = 0.1$ are assumed. Theoretical results from Equations 23 and 25 are shown by broken and solid lines, respectively. Simulation results (closed circles) are the averages over 50,000 replications of forward simulations.

sharp peak of divergence grows along time. As time goes, h_{w1} and h_{w2} become closer to each other, and eventually reaches their equilibrium values (t≫10,000). h_b also decreases except for a short region surrounding the selected site. The rate of erosion (decrease of h_b) increases moving away from the selected site. At the selected site, h_b gradually increases and eventually develops a sharp peak, and simultaneously h_{w1} and h_{w2} also exhibit a small peak that can be created by migration between two subpopulations. It reaches an equilibrium after a significant amount of time, where the selection-migration balance holds so that the shape of the peak does not change much.

Figure 5 shows that Equation 23 (broken lines) is consistent with the simulation results, but the agreement could be further improved if we account for the presence of locally maladapted allele, *i.e.*, allele A (a) in subpopulation I (II). At migration-selection equilibrium, alleles A and a are present in subpopulation I and II at an expected frequency of $1-x_1 \approx -m_1/\tilde{s}_1$ and $x_2 \approx -m_2/s_2$, respectively. Even though these frequencies are small under our assumption of weak migration relative to selection, we show in the following that the approximation in Equation 23 can be improved by accounting for them. Let us focus on the fate of a single neutral allele at the neutral locus linked to an immigrant locally maladaptive allele. We ask how long such a neutral immigrant allele survives on the locally maladaptive background. The linked neutral allele will either be eliminated by selection against the locally maladapted allele in its background, or it recombines off its deleterious background onto a locally beneficial background. The expected time until elimination by selection or recombination in subpopulations I and II are, respectively, given by

$$\begin{aligned} t_{2\to1} &= \sum_{i=0}^{\infty}\{(1+\tilde{s}_1)(1-r)\}^i = \frac{1}{1-(1+\tilde{s}_1)(1-r)}, \\ t_{1\to2} &= \sum_{i=0}^{\infty}\{(1+s_2)(1-r)\}^i = \frac{1}{1-(1+s_2)(1-r)}. \end{aligned} \tag{24}$$

Therefore, the expected numbers of neutral alleles from the other subpopulation with the maladapted allele is $N_1m_1t_{2\to1}$ and $N_2m_2t_{1\to2}$ in subpopulations I and II, respectively.

Let the frequencies of B in subpopulations I and II including those on the locally maladapted background $\tilde{y}_1$ and $\tilde{y}_2$. Accounting for the presence of locally maladaptive alleles, the first- and second-order moments of $\tilde{y}_i$ are:

$$\begin{aligned} E(\tilde{y}_1) &= (1-m_1t_{2\to1})E(y_{A1}) + m_1t_{2\to1}E(y_{a2}) \\ E(\tilde{y}_2) &= m_2t_{1\to2}E(y_{A1}) + (1-m_2t_{1\to2})E(y_{a2}) \\ E(\tilde{y}_1^2) &= (1-m_1t_{2\to1})^2E(y_{A1}^2) + m_1^2t_{2\to1}^2E(y_{a2}^2) + 2m_1t_{2\to1}(1-m_1t_{2\to1})E(y_{A1}y_{a2}) \\ E(\tilde{y}_1\tilde{y}_2) &= (1-m_1t_{2\to1})m_2t_{1\to2}E(y_{A1}^2) + m_1t_{2\to1}(1-m_2t_{1\to2})E(y_{a2}^2) \\ &\quad + \{(1-m_1t_{2\to1})(1-m_2t_{1\to2}) + m_1t_{2\to1}m_2t_{1\to2}\}E(y_{A1}y_{a2}) \\ E(\tilde{y}_2^2) &= m_2^2t_{1\to2}^2E(y_{A1}^2) + (1-m_2t_{1\to2})^2E(y_{a2}^2) + 2m_2t_{1\to2}(1-m_2t_{1\to2})E(y_{A1}y_{a2}). \end{aligned} \tag{25}$$

See Appendix C for details. Figure 5 shows that Equation 25 fits the simulation results better than Equation 23. A notable improvement is seen in h_{w1} for a narrow region around the selected site. Because Equation 23 ignores the presence of maladaptive alleles (assuming their immediate death), Equation 23 predicts a small dip, but our simulation demonstrated rather that a small peak arises. This small peak of h_{w1} is well described by the improved Equation 25.

Discussion

In the early stages of ecological speciation with gene flow, divergent selection is required to maintain phenotypes that are adaptive to each niche (Wu 2001; Turner *et al.* 2005; Nosil 2012). Each target locus of divergent selection works as a barrier locus to migration, because maladaptive migrants should be selected out in a short time. Such a barrier locus can be formed if a locally adaptive mutation arises and becomes established in subpopulations where it is adaptive. This quasi-fixation of a locally adaptive mutation causes a local partial sweep, thereby creating a block of region with elevated F_{ST}. Then, while divergent selection maintains the mutation, recombination shuffles genetic variation in the linked regions and mutations accumulate around the barrier locus. Through this process, a sharp peak of divergence develops in a narrow region around the barrier locus.

This article considers theoretically the evolutionary behavior of a barrier locus, from its initial establishment to stable preservation. The process was divided into three phases: establishment, consolidation and equilibrium (Figure 1). We obtained (1) the establishment probability of a locally adaptive mutation, (2) the expected reduction of genetic variation within subpopulations I and II after a partial local sweep, (3) the evolutionary dynamics at both the barrier locus and the linked neutral sites since the sweep, and (4) the expected shape of the peak of divergence around the barrier locus at equilibrium.

For (1), we derived a closed-form formula of the establishment probability along the formulation of Barton (1987). Our simulations showed that our theoretical results for $F(1/2N_1, 0)$ and $F(0, 1/2N_2)$ outperform the previous approximations, although Yeaman and Otto (2011)'s heuristic approach is almost as good as ours. Because we focused on divergent selection so that allele A is quasi-fixed in subpopulation I, whereas allele a is quasi-fixed in subpopulation II, we assumed $s_1 > 0$ and $s_2 < 0$. However, as shown in Figure 3, it is possible that either allele A or a could fix in the entire population, even if $s_1 > 0$ and $s_2 < 0$ hold, although it might take an extremely long time. In contrast, Gavrilets and Gibson (2002) and Whitlock and Gomulkiewicz (2005) obtained the probability of such eventual fixation in the entire population. These studies and ours can be understood in a single framework as follows. Assuming $s_1 > 0$ and $s_2 < 0$, the establishment of allele A first occurs and is maintained quite stably for a long time, but with time going toward infinity, allele A could fix in the entire population most likely when the average selection coefficient $\bar{s}$ is positive, while allele a could likely fix when $\bar{s}$ is negative. This is why our formula of the establishment probability (Equation 2) is the same as the numerator of the fixation probability when $\bar{s}$ is positive (equations 7 and 8 in Gavrilets and Gibson 2002 and equation 6 in Whitlock and Gomulkiewicz 2005). On the other hand, the establishment probability significantly differs from the fixation probability of Gavrilets and Gibson (2002) and Whitlock and Gomulkiewicz (2005) when $\bar{s}$ is negative because such a mutation hardly goes to eventual fixation, although it can be maintained as a quasi-fixed state for a sufficiently long time.

For (2), we extended the diffusion method of Stephan *et al.* (1992) to our two-population model. Because the beneficial allele A quasi-fixes only in one subpopulation, the process is very similar to that of a single population model (Stephan *et al.* 1992), except that migration between two subpopulations has some effect. Our theoretical result (see Figure 4) demonstrated a relatively minor effect of migration; with an increasing migration rate, the level of polymorphism in subpopulation I increases because migration brings genetic variation from subpopulation II.

For (3) and (4), we considered the evolutionary dynamics at both the barrier locus and the linked neutral sites since the quasi-fixation, followed by the development of a stable peak of divergence around the barrier locus. This process to equilibrium can be described by a single Equation 25. Furthermore, Equation 25 is flexible enough to plug in any initial state, such as a secondary contact of already diverged subpopulation. To demonstrate this, in Supplemental Material, Figure S1, we compare the pattern after a local sweep (left panels) and that after a secondary contact (right panels) (see also Appendix E for details). After a secondary contact, h_b is already high across the genome, and h_b gradually decreases, but selection works to keep divergence around the selected site, thereby creating a peak of divergence. After a very long time (*i.e.*, in equilibrium), the shape of the peak becomes identical to that after a sweep, as pointed out by Yeaman *et al.* (2016). We further performed simulations to investigate how robust our derivation is when the selection intensity is reduced (although we assumed strong selection). The results with 10 times lower selection intensity are shown in Figure S2. This selection intensity is fairly weak, and close to the lower limit to maintain the quasi-fixation state of the two alleles. Yet, Equation 25 is in fairly good agreement with the simulation results, although the performance of Equation 23 is not very good. This is because the frequency of maladaptive alleles is not negligible with a reduced selection intensity.

We thus developed analytical expressions for the evolutionary behavior of a barrier locus, from its emergence to development of a peak of divergence. In the early stages of ecological speciation, it is possible that multiple barrier loci develop and genomic islands of divergence arise, but this does not necessarily mean that the emergence of genomic islands of divergence always results in speciation. It is possible that genomic islands of divergence could disappear due to environmental changes, or by chance, and no speciation occurs. To achieve speciation, many other forces would be necessary, including emergence of additional islands (Feder *et al.* 2012a,b; Via 2012; Aeschbacher and Bürger 2014; Yeaman *et al.* 2016), further divergence on a genomic-scale possible due to a reduction in migration rate, and environmental changes. More theoretical studies are needed to fully understand the process to ecological speciation.

Acknowledgments

We thank S. Takuno and three anonymous reviewers for helpful comments. This work was supporter by JSPS (Japan Society for the Promotion of Science) and SOKENDAI (The Graduate University for Advanced Studies).

Literature Cited

Aeschbacher, S., and R. Bürger, 2014 The effect of linkage on establishment and survival of locally beneficial mutations. Genetics 197: 317–336. https://doi.org/10.1534/genetics.114.163477

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. https://doi.org/10.1007/s00285-013-0660-z

Barton, N. H., 1987 The probability of establishment of an advantageous mutant in a subdivided population. Genet. Res. 50: 35–40. https://doi.org/10.1017/S0016672300023314

Barton, N. H., 1998 The effect of hitch-hiking on neutral genealogies. Genet. Res. 72: 123–133. https://doi.org/10.1017/S0016672398003462

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

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. https://doi.org/10.1017/S0016672397002954

Durrett, R., and J. Schweinsberg, 2004 Approximating selective sweeps. Theor. Popul. Biol. 66: 129–138. https://doi.org/10.1016/j.tpb.2004.04.002

Etheridge, A., P. Pfaffelhuber, and A. Wakolbinger, 2006 An approximate sampling formula under genetic hitchhiking. Ann. Appl. Probab. 16: 685–729. https://doi.org/10.1214/105051606000000114

Feder, J. L., S. P. Egan, and P. Nosil, 2012a The genomics of speciation-with-gene-flow. Trends Genet. 28: 342–350. https://doi.org/10.1016/j.tig.2012.03.009

Feder, J. L., R. Gejji, S. Yeaman, and P. Nosil, 2012b Establishment of new mutations under divergence and genome hitchhiking. Philos. Trans. R. Soc. Lond. B Biol. Sci. 367: 461–474. https://doi.org/10.1098/rstb.2011.0256

Gavrilets, S., and N. Gibson, 2002 Fixation probabilities in a spatially heterogeneous environment. Popul. Ecol. 44: 51–58. https://doi.org/10.1007/s101440200007

Haldane, J. B. S., 1927 A mathematical theory of natural and artificial selection, part v: selection and mutation. Proc. Camb. Philos. Soc. 23: 838–844. https://doi.org/10.1017/S0305004100015644

Kaplan, N. L., R. R. Hudson, and C. H. Langley, 1989 The "hitchhiking effect" revisited. Genetics 123: 887–899.

Kim, Y., and T. Maruki, 2011 Hitchhiking effect of a beneficial mutation spreading in a subdivided population. Genetics 189: 213–226. https://doi.org/10.1534/genetics.111.130203

Kimura, M., 1954 Process leading to quasi-fixation of genes in natural populations due to random fluctuation of selection intensities. Genetics 39: 280–295.

Kimura, M., 1957 Some problems of stochastic processes in genetics. Ann. Math. Stat. 28: 882–901. https://doi.org/10.1214/aoms/1177706791

Kimura, M., 1962 On the probability of fixation of mutant genes in a population. Genetics 47: 713–719.

Kirkpatrick, M., and S. Peischl, 2013 Evolutionary rescue by beneficial mutations in environments that change in space and time. Philos. Trans. R. Soc. Lond. B Biol. Sci. 368: 20120082. https://doi.org/10.1098/rstb.2012.0082

Maynard Smith, J. M., and J. Haigh, 1974 The hitch-hiking effect of a favourable gene. Genet. Res. 23: 23–35. https://doi.org/10.1017/S0016672300014634

Nagylaki, T., 1980 The strong-migration limit in geographically structured populations. J. Math. Biol. 9: 101–114. https://doi.org/10.1007/BF00275916

Nagylaki, T., and Y. Lou, 2008 The dynamics of migration-selection models, pp. 117–170 in *Tutorials in Mathematical Biosciences IV: Evolution and Ecology* (Lecture Notes in Mathematics). Springer, Berlin. https://doi.org/10.1007/978-3-540-74331-6_4

Nosil, P., 2012 *Ecological Speciation*. Oxford University Press, Oxford. https://doi.org/10.1093/acprof:osobl/9780199587100.001.0001

Ohta, T., and M. Kimura, 1969 Linkage disequilibrium at steady state determined by random genetic drift and recurrent mutation. Genetics 63: 229–238.

Pollak, E., 1966 On the survival of a gene in a subdivided population. J. Appl. Probab. 3: 142–155. https://doi.org/10.2307/3212043

Santiago, E., and A. Caballero, 2005 Variation after a selective sweep in a subdivided population. Genetics 169: 475–483. https://doi.org/10.1534/genetics.104.032813

Schweinsberg, J., and R. Durrett, 2005 Random partitions approximating the coalescence of lineages during a selective sweep. Ann. Appl. Probab. 15: 1591–1651. https://doi.org/10.1214/105051605000000430

Slatkin, M., and T. Wiehe, 1998 Genetic hitch-hiking in a subdivided population. Genet. Res. 71: 155–160. https://doi.org/10.1017/S001667239800319X

Stephan, W., T. H. Wiehe, and M. W. Lenz, 1992 The effect of strongly selected substitutions on neutral polymorphism: analytical results based on diffusion theory. Theor. Popul. Biol. 41: 237–254. https://doi.org/10.1016/0040-5809(92)90045-U

Tomasini, M., and S. Peischl, 2018 Establishment of locally adapted mutations under divergent selection. Genetics 209: 885–895.

Turner, T. L., M. W. Hahn, and S. V. Nuzhdin, 2005 Genomic islands of speciation in anopheles gambiae. PLoS Biol. 3: e285. https://doi.org/10.1371/journal.pbio.0030285

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. https://doi.org/10.1098/rstb.2011.0260

Wakeley, J., 2009 *Coalescent Theory: An Introduction*. Roberts & Company, Greenwood Village, CO.

Whitlock, M. C., and R. Gomulkiewicz, 2005 Probability of fixation in a heterogeneous environment. Genetics 171: 1407–1417. https://doi.org/10.1534/genetics.104.040089

Wu, C.-I., 2001 The genic view of the process of speciation. J. Evol. Biol. 14: 851–865. https://doi.org/10.1046/j.1420-9101.2001.00335.x

Yeaman, S., and S. P. Otto, 2011 Establishment and maintenance of adaptive genetic divergence under migration, selection, and drift. Evolution 65: 2123–2129. https://doi.org/10.1111/j.1558-5646.2011.01277.x

Yeaman, S., S. Aeschbacher, and R. Bürger, 2016 The evolution of genomic islands by increased establishment probability of linked alleles. Mol. Ecol. 25: 2542–2558. https://doi.org/10.1111/mec.13611

Communicating editor: G. Coop

Appendix

Appendix A: Forward Simulation

Here, we describe the setting and assumptions of our forward simulations. A model with two subpopulations (I and II) is used. Subpopulations I and II consist of $2N_1$ and $2N_2$ haploids. We are interested in how DNA sequence evolves at the population level around a selected locus. We considered a genomic region encompassing a selected locus at the center, and assumed the infinite-site model for simulating patterns of nucleotide polymorphisms (*e.g.*, Figure 1). For other simulations, we consider a two-locus model with the selected locus and a linked neutral locus. The recombination rate between the two loci is r. The fitness of an individual is determined by the allelic state at the selected locus: The fitness of an individual with allele A and a are, respectively, $1 + s_1$ and 1 in subpopulation I, and $1 + s_2$ and 1 in subpopulation I. Every generation, migration is allowed such that $2Nm$ individuals are swapped between the two subpopulations, Then, to construct a new population in the next generation, $2N_1$ and $2N_2$ individuals are chosen randomly from the current subpopulations I and II, respectively, where their fitness is taken into account. No recurrent mutation is allowed at this site in order to trace the fate of the mutation (unless otherwise mentioned). In contrast, at the linked neutral locus, recurrent mutation is allowed at rate μ per generation. Heterozygosities within and between (h_w and h_b) subpopulations can be scored at any arbitrary time point.

Appendix B: The Solution of Equations 4 and 5

First, we present a proof that there is, at most, one solution that fulfills $p_1 > 0$ and $p_2 > 0$, and the condition on which such a solution exists is $a + d > 0$ or $ad - bc < 0$. Then, we give a closed expression of the solution.

For ψ_1 and ψ_2 to satisfy $p_1 > 0$ and $p_2 > 0$, $\psi_1 > 0$ and $\psi_2 > 0$ are needed. Note that $b, c > 0$ because the migration rate and population size are always positive. Although, in this work, we consider only the case of $d < 0$, Equations 4 and 5 may also work in the case of $d \geq 0$. Therefore, here, we present the proof that allows $d \geq 0$. We set $f(x) = x^3 - 2ax^2 + (a^2 - bd)x + (abd - b^2c)$, and note that the first derivative of $f(x)$ is $f'(x) = 3x^2 - 4ax + (a^2 - bd)$. We discuss the complementary following three cases.

1. $a \geq 0$
   From Equation 5, $\psi_1 > a$, then $x > a$ is needed. Because the x-coordinate of the vertex of $f'(x)$, $\frac{2}{3}a$, is not greater than a, $f'(x)$ increases monotonically when $x > a$. Noting that $f(a) = -b^2c < 0$, there is only one solution to $f(x) = 0$.

2. $a < 0$ and $d \leq 0$
   From Equation 5, $\psi_1 > 0$, then $x > 0$ is needed. Because $f'(0) = a^2 - bd > 0$ and the x-coordinate of the vertex of $f'(x)$, $\frac{2}{3}a$, is smaller than 0, $f'(x) > 0$ when $x > 0$. Therefore, whether $f(x) = 0$ has a solution or not in $(0, \infty)$ depends on the sign of $f(0)$. If $f(0) \geq 0$, *i.e.*, $b(ad - bc) \geq 0$, there is no solution. Otherwise, there is only one solution.

3. $a < 0$ and $d > 0$
   From Equation 5, $\psi_1 > 0$, then $x > 0$ is needed. Because the x-coordinate of the vertex of $f'(x)$, $\frac{2}{3}a$, is smaller than 0, $f'(x)$ increases monotonically when $x > 0$. Noting that $f(0) = b(ad - bc) < 0$, there is only one solution.

Noting that $b, c > 0$ and $ad - bc$ is negative when $ad \leq 0$, the condition on which one solution exists is rearranged to $a + d > 0$ or $ad - bc < 0$. This is the same as the condition where a deterministic model,

$$\frac{d}{dt}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \tag{B1}$$

has a positive growth rate. In other words, the matrix in Equation B1 has at least one positive eigenvalue.

Next, we present a closed form of ψ_1. From the above proof, if there is a nonzero real root of $f(\psi_1) = 0$, which fulfills $p_1 > 0$ and $p_2 > 0$, the root is the largest real root of $f(\psi_1) = 0$. Therefore, by using the solution of cubic equation, ψ_1 can be expressed as

$$\psi_1 = \begin{cases} 0 & \text{when } a + d \leq 0 \text{ and } ad - bc \geq 0 \\ \sqrt[3]{-\frac{Q}{2} + \sqrt{R}} + \sqrt[3]{-\frac{Q}{2} - \sqrt{R}} - \frac{A_2}{3} & \text{when } R > 0 \text{ and } (a + d > 0 \text{ or } ad - bc < 0) \\ 2\,S\cos\left(\frac{1}{3}\arccos\left(\frac{T}{2S}\right)\right) - \frac{A_2}{3} & \text{when } R \leq 0 \text{ and } (a + d > 0 \text{ or } ad - bc < 0) \end{cases}, \tag{B2}$$

where $A_0 = abd - b^2c, A_1 = a^2 - bd, A_2 = -2a, P = A_1 - \frac{A_2^2}{3}, Q = A_0 - \frac{A_1A_2}{3} + \frac{2}{27}A_2^3, R = (\frac{P}{3})^3 + (\frac{Q}{2})^2, S = \sqrt{-\frac{P}{3}}, T = -\frac{Q}{S^2}$. In the above expression, we assume the range of principal value of $y = \arccos(x)$ as $0 \leq y \leq \pi$.

Appendix C: Derivation of Equations 21, 23, and 25

Here, we describe the derivation of Equations 21, 23, and 25 in more detail. By applying Equation 18 to $f = y_{A1}, y_{a2}, y_{A1}^2, y_{A1}y_{a2}$, and y_{a2}^2, we can derive the time derivative of moments of y_{A1} and y_{a2} as follows:

$$\begin{aligned}
\frac{dE(y_{A1})}{dt} &= v - (u + v + m_{e,2\to1})E(y_{A1}) + m_{e,2\to1}E(y_{a2}) \\
\frac{dE(y_{a2})}{dt} &= v - (u + v + m_{e,1\to2})E(y_{a2}) + m_{e,1\to2}E(y_{A1}) \\
\frac{dE(y_{A1}^2)}{dt} &= \left(2v + \frac{1}{2N_1}\right)E(y_{A1}) - 2\left(u + v + m_{e,2\to1} + \frac{1}{4N_1}\right)E(y_{A1}^2) + 2m_{e,2\to1}E(y_{A1}y_{a2}) \\
\frac{dE(y_{A1}y_{a2})}{dt} &= vE(y_{A1}) + vE(y_{a2}) + m_{e,1\to2}E(y_{A1}^2) - (2u + 2v + m_{e,2\to1} + m_{e,1\to2})E(y_{A1}y_{a2}) + m_{e,2\to1}E(y_{a2}^2) \\
\frac{dE(y_{a2}^2)}{dt} &= \left(2v + \frac{1}{2N_2}\right)E(y_{a2}) - 2\left(u + v + m_{e,1\to2} + \frac{1}{4N_2}\right)E(y_{a2}^2) + 2m_{e,1\to2}E(y_{A1}y_{a2}).
\end{aligned} \tag{C3}$$

By setting $\boldsymbol{y} = (E(y_{A1}), E(y_{a2}), E(y_{A1}^2), E(y_{A1}y_{a2}), E(y_{a2}^2))^T$, $\boldsymbol{e} = (v, v, 0, 0, 0)^T$ and defining $\boldsymbol{Q}$ as Equation 22, Equation C3 can be rearranged in a matrix form (Equation 21). Then, by using the solution of a linear differential equation with constant coefficients, the solution of Equation 21 is given by

$$\begin{aligned}
\boldsymbol{y}(t) &= \exp(t\boldsymbol{Q})\boldsymbol{y}(0) + \int_0^t \exp((t-s)\boldsymbol{Q})\boldsymbol{e}ds \\
&= \exp(t\boldsymbol{Q})\boldsymbol{y}(0) + \boldsymbol{Q}^{-1}(\exp(t\boldsymbol{Q}) - \boldsymbol{I})\boldsymbol{e}.
\end{aligned} \tag{C4}$$

The solution 23 is further improved by accounting for neutral immigrant alleles linked to maladaptive alleles. To do so, we derive the expected time of a neutral immigrant allele until its elimination by selection or recombination as Equation 24. The expected frequencies of such an allele are $m_1t_{2\to1}$ and $m_2t_{1\to2}$ in subpopulations I and II, respectively. $\tilde{y}_1$ and $\tilde{y}_2$, denote the frequencies of B in subpopulations I and II including those on the locally maladapted background. Then, $\tilde{y}_1$ and $\tilde{y}_2$ can be approximated by

$$\begin{aligned}
\tilde{y}_1 &= (1 - m_1t_{2\to1})y_{A1} + m_1t_{2\to1}y_{a2}, \\
\tilde{y}_2 &= (1 - m_2t_{1\to2})y_{a2} + m_2t_{1\to2}y_{A1}.
\end{aligned} \tag{C5}$$

By using Equation C5 and taking expectations, the first and second-order moments of $\tilde{y}_1$ and $\tilde{y}_2$ are given by Equation 25.

Appendix D: Comparison Between Diffusion and Coalescent at Equilibrium Phase

In the main text, we show that replacing the migration rate in the neutral diffusion equation by the effective migration rate well approximates the effect of linkage with the locus under divergent selection. In a neutral model, heterozygosity at equilibrium in a structured population is already well studied by the coalescent theory under the infinite-site model (reviewed in Wakeley 2009). In this work, we alternatively used the forward diffusion approach because the diffusion approach can be applied to more general conditions. In this Appendix, we show our diffusion result at equilibrium is consistent with that of the coalescent theory.

We attempt to derive the expected heterozygosity under the infinite-site setting along our diffusion-based derivation. In practice, we first consider a K-allele model, and then the results will be transformed to the infinite-site model. Let B allele be one of the alleles at the locus. We put y_1 and y_2 as frequency of allele B in subpopulation I and II, respectively. In the following derivation, we assume $N_1 = N_2 = N$ and $m_1 = m_2 = m$. The differential operator of the Kolmogorov backward equation is as follows,

$$
\begin{aligned}
L = & \frac{y_1(1-y_1)}{4N}\frac{\partial^2}{\partial y_1^2} + \frac{y_2(1-y_2)}{4N}\frac{\partial^2}{\partial y_2^2} \\
& + [v-(u+v)y_1 + m(y_2-y_1)]\frac{\partial}{\partial y_1} + [v-(u+v)y_2 + m(y_1-y_2)]\frac{\partial}{\partial y_2},
\end{aligned} \tag{D1}
$$

At the equilibrium, we derive the moments up to the second order as

$$
\begin{aligned}
E(y_1) = E(y_2) &= \frac{V}{U+V}, \\
E(y_1^2) = E(y_2^2) &= \frac{V(V+1)(U+V+M) + V^2M}{(U+V)(U+V+M+1)(U+V+M) - M^2(U+V)}, \\
E(y_1y_2) &= \frac{MV(V+1) + V^2(U+V+M+1)}{(U+V)(U+V+M+1)(U+V+M) - M^2(U+V)},
\end{aligned} \tag{D2}
$$

where $U = 4Nu$, $V = 4Nv$, and $M = 4Nm$. In the limit to the infinite-allele model, that is, $v = \frac{u}{K-1}$ and $K \to \infty$, the expected heterozygosity within and between subpopulation goes to

$$
\begin{aligned}
h_w &= 1 - KE(y_1^2) \to \frac{U^2 + 2UM}{(U+M+1)(U+M) - M^2}, \\
h_b &= 1 - KE(y_1y_2) \to \frac{U(U+2M+1)}{(U+M+1)(U+M) - M^2}.
\end{aligned} \tag{D3}
$$

This result under the infinite-allele setting can be transformed to the infinite-site mode: If we put $U = \frac{\theta}{n}$ and n goes to ∞, π_w and π_b are described as

$$
\begin{aligned}
\pi_w &= nh_w \to 2\theta \\
\pi_b &= nh_b \to \theta(2 + \frac{1}{M}),
\end{aligned} \tag{D4}
$$

which is identical with the result from the coalescent theory (Charlesworth *et al.* 1997; Yeaman *et al.* 2016).

Appendix E: Comparing the Scenarios of Local Partial Sweep and Secondary Contact

We compute Equation 23 for a scenario of secondary contact, and compared with the results of a local partial sweep shown in Figure 5. For a secondary-contact scenario, we assume that already diverged two subpopulation have merged so that there are a number of fixed sites between the two subpopulations. To make a realization of this situation, we set $y_1(0) = 0.1$ and $y_2(0) = 0.9$, and the other parameters are identical to those used for Figure 5 (*i.e.*, $4N_1s_1 = -4N_2s_2 = 200$ and $4N_1m_1 = 4N_2m_2 = 5$). Figure S1 compares the patterns after a local sweep (left panels) and after a secondary contact (right panels). After a secondary contact, h_b is already high across the genome, and h_b gradually decreases but selection works to keep divergence around the selected site, thereby creating a peak of divergence. In equilibrium, the shape of the peak becomes identical to that after a sweep, in agreement with Yeaman *et al.* (2016).

Figure S2 shows the results with identical parameter sets to those for Figure S1 except for the selection intensity. The purpose is to check how robust our derivation is when the selection intensity is reduced. We here set $4N_1s_1 = -4N_2s_2 = 20$ and $4N_1m_1 = 4N_2m_2 = 5$. The results with 10 times lower selection intensity are shown in Figure S1. This selection intensity is fairly weak, and it is close to the lower limit to maintain the quasi-fixation state of the two alleles.