doi:10.1111/j.1558-5646.2011.01277.x

ESTABLISHMENT AND MAINTENANCE OF ADAPTIVE GENETIC DIVERGENCE UNDER MIGRATION, SELECTION, AND DRIFT

Sam Yeaman[1,2,3] and Sarah P. Otto[1]

[1]*Department of Zoology, University of British Columbia, Vancouver, British Columbia, V6T 1Z4, Canada*

[2]*Université de Neuchâtel, Institute of Biology, 11 Rue Emile-Argand, Neuchâtel 2000, Switzerland*

[3]*E-mail: yeaman@zoology.ubc.ca*

Received October 6, 2010
Accepted February 21, 2011

There is a long tradition in population genetics of exploring the maintenance of variation under migration–selection balance using deterministic models that assume infinite population size. With finite population size, stochastic dynamics can greatly reduce the potential for the maintenance of polymorphism, but this has yet to be explored in detail. Here, classical two-patch models are extended to predict: (1) the probability of a locally beneficial mutation rising in frequency in the patch where it is favored and (2) the critical threshold migration rate above which the maintenance of polymorphism is much less likely. Individual-based simulations show that these approximations provide accurate predictions across a wide range of parameter space.

KEY WORDS: **Dispersal, environmental heterogeneity, gene flow, genetic drift, local adaptation, migration–selection balance.**

Understanding how evolutionary processes affect the establishment and maintenance of genetic divergence between populations is critical to interpreting empirical evidence for local adaptation and speciation. Most theoretical research on population divergence under migration–selection balance has focused on deterministic models (reviewed in Felsenstein 1976; Karlin 1982; Lenormand 2002; Nagylaki and Lou 2008; also see Nagylaki and Lou 2007; Star et al. 2007; Bürger 2009a,b; Nagylaki 2009). The earliest studies by Haldane (1930) and Wright (1931) considered a continent-island model, showing that a locally adapted allele would be deterministically maintained in an island patch as long as the rate of immigration, m, of the alternate allele (fixed on the continent) was lower than the selection coefficient, s, favoring the locally adapted allele ($m/s < 1$). Moran (1962), Maynard Smith (1966), and Bulmer (1972) generalized this result to a two-patch model with reciprocal migration (m_{ij} from patch j into patch i) with population size regulated independently in each patch (soft selection). A mutant allele spreads deterministically when $m_{12}/s_1 + m_{21}/s_2 < 1$, where the fitnesses of the resident homozygote in the two patches are defined as $1 - s_1$ and $1 - s_2$ relative to the mutant-bearing heterozygotes, with s_1 and s_2 assumed to have opposite signs (Bulmer 1972). Counterintuitively, this result predicts the maintenance of polymorphism across a wide region of parameter space when selection coefficients and migration rates are closely balanced. In the extreme case, where $s_1 = -s_2$ and $m_1 = m_2$, a polymorphism is maintained regardless of the migration rate, even when selection is very weak. We would expect polymorphism to be readily lost in finite populations with very weak selection, however.

Although several studies have explored the effects of population size on the probability of a single mutant allele fixing in both patches of a finite two-patch model (Tachida and Iizuka 1991; Gavrilets and Gibson 2002; Whitlock and Gomulkiewicz 2005), previous work has not yet determined the conditions under which a polymorphism is likely to be maintained. Billiard and Lenormand (2005) approximated equilibrium allele frequencies under selection, drift, and low rates of migration, but their approach breaks down with moderate levels of migration and so

cannot be used to determine when local adaptation is likely to persist. Alternatively, the allele frequency dynamics under selection, drift, and migration can be described using a diffusion model in the case where the migrant pool arriving at a patch never changes (yielding, e.g., Wright's 1931 allele frequency distribution), but this assumes that migrants carry an allele, A, into the patch at a fixed frequency and that migration out of the patch has negligible effect on the dynamics elsewhere.

The aim of this article is to understand the effect of drift on the maintenance of polymorphism between two interconnected patches of finite size. Specifically, we develop approximations for the probability that locally adapted alleles increase when rare and the critical migration threshold below which adaptive divergence among patches is likely to persist. These predictions are then tested using individual-based simulations.

Analytical Model

We briefly rederive the conditions for a protected polymorphism for the deterministic two-allele model with two patches connected by migration (Bulmer 1972), which is the foundation for our results with finite patch sizes. We use the following notation for parameters in patch i: the allele frequencies of the two competing alleles (A and a) are p_i and q_i; W_i represents the fitness of each genotype; m_{ij} represents the proportion of gametes in patch i that have arrived from patch j; finally, $\bar{W}_i$ represents the mean fitness. A generation consists of random mating, followed by selection on diploids, gamete production, and gamete migration. The allele frequencies in the next generation then become:

$$\begin{aligned} p_i' &= (1 - m_{ij})\, p_i^s + m_{ij}\, p_j^s \\ q_i' &= (1 - m_{ij})\, q_i^s + m_{ij}\, q_j^s, \end{aligned} \tag{1}$$

where,

$$\begin{aligned} p_i^s &= p_i(p_i W_{i,AA} + q_i W_{i,Aa})/\bar{W}_i \\ q_i^s &= q_i(q_i W_{i,aa} + p_i W_{i,Aa})/\bar{W}_i. \end{aligned} \tag{2}$$

We assume throughout that selection favors allele a in patch 1 and allele A in patch 2; to emphasize this assumption, we will use the following alternative fitness definitions where convenient: $W_{1,AA} = 1 - s$, $W_{1,Aa} = 1 - hs$, $W_{1,aa} = 1$, $W_{2,AA} = 1$, $W_{2,Aa} = 1 - kt$, and $W_{2,aa} = 1 - t$.

The standard approach to finding the regions of parameter space that allow the maintenance of polymorphism is to consider the stability of the monomorphic equilibria when either allele is fixed (e.g., Moran 1962; Maynard Smith 1966; Bulmer 1972). If both fixation states are unstable, it follows that a polymorphism will be maintained. The stability of the equilibrium where A is fixed can be found by evaluating the Jacobian matrix at $q_i = 0$. The leading eigenvalue of this matrix ($\lambda_{l,a}$) represents the rate of change in frequency of allele a when rare:

$$\lambda_{l,a} = \frac{1}{2}\psi + \frac{1}{2}\sqrt{\psi^2 - 4(1 - m_{12} - m_{21})\frac{W_{1,Aa}}{W_{1,AA}}\frac{W_{2,Aa}}{W_{2,AA}}}, \tag{3}$$

where,

$$\psi = (1 - m_{12})\,\frac{W_{1,Aa}}{W_{1,AA}} + (1 - m_{21})\,\frac{W_{2,Aa}}{W_{2,AA}}. \tag{4}$$

A similar result holds for the invasion of A when rare ($\lambda_{l,A}$, interchanging A and a as well as m_{12} and m_{21}). In an infinite population, we expect the population to remain polymorphic as long as both $\lambda_{l,a}$ and $\lambda_{l,A}$ are greater than one because both alleles increase in frequency when rare.

To simplify the presentation, we now assume that migration is symmetric between the two patches ($m_{ij} = m_{ji}$); analogous results can be obtained for unequal migration using the approach that follows. To begin, we find the critical migration rates below which each allele invades by solving for when $\lambda_{l,a} = 1$ and $\lambda_{l,A} = 1$. For the invasion of allele a when rare, the critical migration threshold is

$$m_{crit,a}^{\infty} = \frac{1}{\dfrac{W_{1,Aa}}{W_{1,Aa} - W_{1,AA}} - \dfrac{W_{2,Aa}}{W_{2,AA} - W_{2,Aa}}}, \tag{5}$$

where, the ∞ superscript indicates that the population size is assumed to be infinite. Equation (5) is equivalent to the threshold found by Bulmer (1972) using his fitness definitions ($W_{1,AA}/W_{1,Aa} = 1 - s_1$ and $W_{2,AA}/W_{2,Aa} = 1 - s_2$). We have written equation (5) so that the two fractions in the denominator are positive under the assumption that allele a is favored in patch 1 and allele A in patch 2. If the difference between these two fractions is negative ($m_{crit,a}^{\infty} < 0$) or is zero, then allele a can invade for any migration rate. If there is intermediate dominance with $h = 0.5$ and $k = 0.5$, equation (5) reduces to:

$$m_{crit,a}^{\infty} = \frac{st}{2t - 2s}. \tag{6}$$

Similarly, $m_{crit,A}^{\infty}$ can be obtained by interchanging the allelic designations, A and a, in equation (5).

The critical threshold for the invasion of allele a, given by equation (5), is more restrictive than the equivalent threshold for the invasion of A whenever their selection and dominance coefficients satisfy:

$$\frac{W_{1,Aa}}{W_{1,aa} - W_{1,Aa}} + \frac{W_{1,Aa}}{W_{1,Aa} - W_{1,AA}} > \frac{W_{2,Aa}}{W_{2,Aa} - W_{2,aa}} + \frac{W_{2,Aa}}{W_{2,AA} - W_{2,Aa}}. \tag{7}$$

Without loss of generality, we assume that the alleles are labeled such that equation (7) is satisfied, and we define the patches

such that patch 1 is where allele a is favored. For example, in the additive case, equation (7) reduces to $t > s$, indicating that we should label allele a as the one that is more deleterious in the patch where it is locally maladapted. Defining the alleles such that equation (7) is satisfied, it is then sufficient to ask whether $m < m^{\infty}_{crit,a}$ to determine whether a polymorphism will be maintained in an infinite population, and we refer to this deterministic critical threshold as m^{∞}_{crit} for simplicity.

For some combinations of parameters, the critical threshold rises above one, implying that both alleles will be maintained deterministically for any migration rate. It can be shown that for m^{∞}_{crit} to be greater than one, the geometric mean fitness of the heterozygotes across patches must be higher than that of the resident homozygotes:

$$\sqrt{\frac{W_{1,Aa}}{W_{1,AA}}\frac{W_{2,Aa}}{W_{2,AA}}} > 1. \quad (8)$$

Equation (8) is most likely to be satisfied when the dominance coefficients (h and k) are small, causing the heterozygotes to have a high relative fitness.

In this deterministic model, we expect a polymorphism to be maintained as long as the smaller of the two leading eigenvalues ($\lambda_{l,a}$ or $\lambda_{l,A}$ from eq. 3) rises above one, as illustrated in Figure 1A for the simple additive case ($h = 0.5$, $k = 0.5$) with symmetric migration ($m_{ij} = m_{ji} = m$). When selection is much stronger in one patch than the other, the criterion for polymorphism approaches the threshold in the continent-island model (Haldane 1930; Wright 1931; $m/(s/2) < 1$ when $t \gg s$, $m/(t/2) < 1$ when $s \gg t$, selection is divided by two to account for the fitness definitions used here) because the patch with strong selection exhibits little polymorphism, as assumed for the continent (dashed lines in Fig. 1A). Similar behavior is seen in the continuous-time model of Gavrilets and Gibson (2002; their Fig. 1).

With finite population size, the magnitude of $\lambda_l - 1$ is important, not just its sign, as alleles experiencing very weak diversifying selection will be easily lost due to the combined effects of migration and drift. Here, we splice together the above deterministic results with standard diffusion results for a single patch to predict when both locally adapted alleles are likely to be maintained in finite populations. For a mutation experiencing selection of constant strength S in a single population of size N, Kimura (1962) showed that the probability of rising from a single copy to fixation is

$$\Pr[fix] = \frac{1 - e^{-2S}}{1 - e^{-4NS}}. \quad (9)$$

Because $\lambda_l - 1$ describes the asymptotic rate of increase of a rare allele in a deterministic analysis, we define the "diversification coefficient" to be $\delta = \lambda_l - 1$ and use δ in place of S in equation (9)

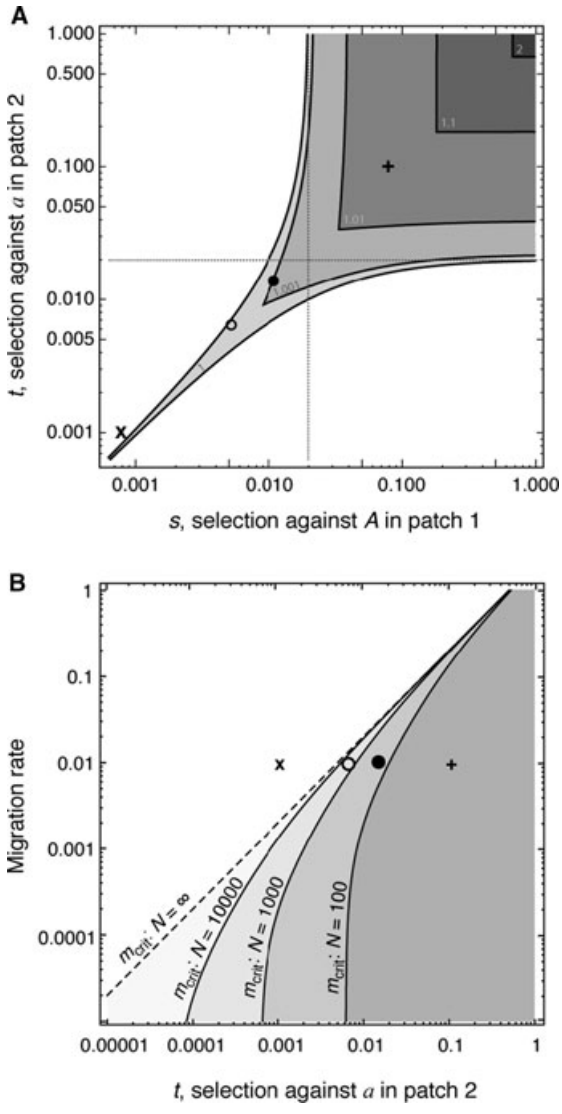

Figure 1. Regions maintaining a balanced polymorphism as a function of the selection coefficients (s, t) and migration rate (m). (A) Contour plot showing the smaller of the two leading eigenvalues ($\lambda_{l,a}$ or $\lambda_{l,A}$; contour values in gray) from a deterministic analysis with $h = k = 0.5$ and $m = 1/100$. Within the white region, deterministic selection does not maintain a balanced polymorphism, and we expect loss of one of the alleles. (B) The critical migration rate, below which polymorphism will be maintained using equation (5) for infinitely large populations and equation (11) for finite populations, assuming $s = 0.8\ t$ and $h = k = 0.5$. The four points correspond to the same parameter sets in both panels and illustrate the fact that even though $\lambda_l > 1$ in a deterministic analysis (points "○" "●" and "+"), the evolutionary forces can be too weak to maintain a polymorphism in the face of drift in finite populations (when $N < 2643$ for "○" and when $N < 220$ for "●").

to approximate the probability that a locally adapted mutation rises to high frequency, given that it is introduced into the patch where it is favored (hereafter, invasion probability). Here, N is assumed to equal the local patch size, and the allele is assumed

to be nearly additive (see Supporting information for the general case with dominance).

Alternatively, the "diversification coefficient" could be defined based on branching process results, setting $\delta = P/2$, where P is the probability of establishment of a rare allele in the patch where it is favored in a two-patch model (see Supporting information; Barton 1987). The branching process approach also assumes that the total population size is infinitely large, but it accounts for chance loss of the allele while rare. As shown in the Supporting information (Fig. S6), defining δ using the asymptotic rate of increase, $\lambda_l - 1$, or the probability of establishment, $P/2$, yields qualitatively similar results, and both approaches accurately describe the invasion probability in finite populations once δ is plugged into equation (9). Because the λ_l is more easily solved, however, we focus the text on the former approach. The branching process approach can, however, be easily adapted to give the invasion probability when the mutation is first introduced into the patch where it is disfavored (see Supporting information).

We next adjust the critical migration rate to account for the role of drift by solving for the migration rate that results in $2\delta > 1/(2N)$, which is roughly where the invasion probability based on equation (9) transitions from being dominated by drift ($\sim 1/2N$) to being dominated by selection ($\sim 2\,\delta$) (Crow and Kimura 1970). In essence, this approach finds the boundary where the net evolutionary force acting deterministically on the allele frequencies (migration and selection) has the same probability of explaining the spread of the rare allele as the stochastic noise generated by sampling error in a finite population. Throughout, we assume that the fitness differences are not so small that drift would overwhelm selection and hamper the spread of the alleles where they are locally favored, even without the swamping effects of migration. Specifically, we assume that

$$2\left(\frac{W_{1,Aa}}{W_{1,AA}} - 1\right) \gg \frac{1}{2N} \text{ and } 2\left(\frac{W_{2,Aa}}{W_{2,aa}} - 1\right) \gg \frac{1}{2N}. \quad (10)$$

Solving for when $2\delta > 1/(2N)$, we obtain a new threshold for the migration rate below which the system is expected to maintain polymorphism

$$m^N_{crit} = \frac{1}{\frac{W_{1,Aa}}{W_{1,Aa} - W_{1,AA}\left(1 + \frac{1}{4N}\right)} - \frac{W_{2,Aa}}{W_{2,AA}\left(1 + \frac{1}{4N}\right) - W_{2,Aa}}}. \quad (11)$$

Decreasing the population size makes it more difficult to maintain a polymorphism, both because m^N_{crit} is more likely to be positive and because m^N_{crit}, when positive, tends to be smaller (recall that negative values of m^N_{crit} allow polymorphism to be maintained for any migration rate, provided assumption eq. 10 is satisfied).

Again, the critical threshold can rise above one for some parameters, implying that both alleles will be maintained for any migration rate when:

$$\sqrt{\frac{W_{1,Aa}}{W_{1,AA}}\frac{W_{2,Aa}}{W_{2,AA}}} > \left(1 + \frac{1}{4N}\right). \quad (12)$$

Drift in smaller populations makes equation (12) more restrictive.

Individual-based Simulations

To assess the accuracy of the above splicing approach, we used a modification of the Nemo platform (Guillaume and Rougemont 2006) to run individual-based simulations (as per Yeaman and Whitlock 2011) based on the above lifecycle and parameters. We measured: (1) the probability of a novel mutation rising to high frequency in the patch where it is favored; (2) the persistence time of a polymorphism (Supporting information); and (3) the fraction of time in which populations exhibited local adaptation. For the latter, we measured whether populations in each patch had a higher fitness than they would have had if transplanted to the other patch.

PROBABILITY OF A NEW MUTATION RISING TO HIGH FREQUENCY

We first tested the accuracy of the splicing approach in predicting the invasion probability of a locally adapted allele. In these simulations, all individuals in both patches were initially *AA* homozygotes and a single *a* mutant was introduced in one randomly chosen individual in the patch where it was beneficial. No mutations occurred thereafter and simulations were run for 5000 generations, with between 20,000 and 1,000,000 replicates for each parameter set, adjusted so that the standard errors were roughly the size of the symbols on the plots. The observed invasion probability was calculated from the fraction of simulations in which the *a* mutant either rose to a frequency above 0.5 (in the entire meta-population) or persisted for 5000 generations (we note that for cases with strongly asymmetrical selection and/or dominance, invasion may yield stable polymorphism at frequencies below 0.5; in such cases, as long as the allele persists for at least 5000 generations, it would still be recorded as a successful invasion).

Substituting the diversification coefficient δ for S in Kimura's equation (9) provided an accurate prediction for the invasion probability for a range of selection coefficients when mutations were codominant ($h = k = 0.5$; Fig. 2B). Alleles introduced where they were locally favored had a much higher invasion probability when selection was stronger, and in all cases, invasion probability increased with decreasing migration. Although it is intuitive that strongly selected mutations should have a higher establishment probability, this result runs counter to the

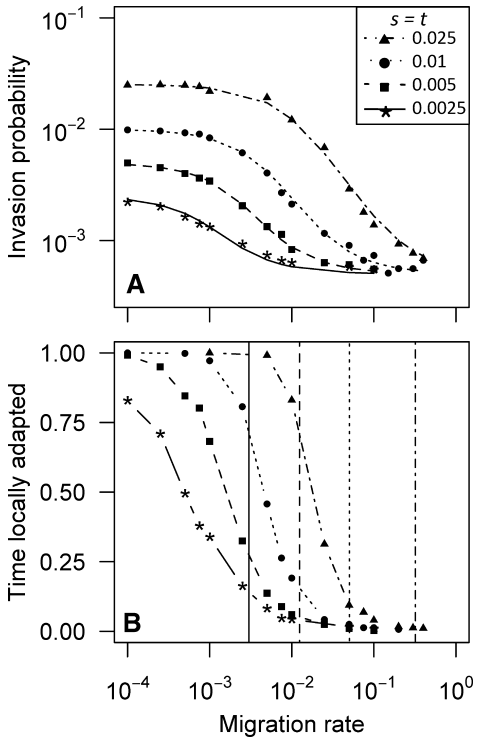

Figure 2. Invasion probability of a new mutation in a finite population (A) and fraction of time points at which both populations were locally adapted under recurrent mutation (B). For the four selection regimes shown, the curves in (A) correspond to predictions from Kimura's fixation probability using δ for S in equation (9), and the vertical lines in (B) correspond to m^N_{crit}. Other parameters: $h = k = 0.5$ and $N = 1000$.

deterministic result that locally adapted alleles should always invade when selection coefficients and migration rates are symmetrical (Bulmer 1972; Spichtig and Kawecki 2004). For very low migration rates, the fixation probability of the introduced allele was, as expected, nearly twice its selective advantage ($2(W_{1,Aa}/W_{1,AA} - 1)$). The fixation probability dropped with increasing migration, because the rare allele spent an increasing amount of time in the patch where it was disfavored.

We also find similar results with a smaller patch size ($N = 100$; Fig. S1), with discrepancies only when selection is so weak that inequality (10) begins to break down. The splicing approach using equation (9) is less accurate, however, when the alleles are not codominant (Fig. S2), but we found that using Kimura's general solution for arbitrary dominance provided good agreement in these cases (Fig. S3), even when selection and dominance coefficients were asymmetrical (Fig. S4). As selection tends toward neutrality, the success of the invading allele will increasingly depend upon dynamics in the patch where it is disfavored, so that using the single patch size for N in equation (9) may yield inaccurate results (but this will usually be in the region of parameter space that fails to satisfy inequality (10)).

CRITICAL MIGRATION THRESHOLD AND MAINTENANCE OF POLYMORPHISM

We next tested the accuracy of the critical migration threshold, m^N_{crit}, below which we expect local adaptation to persist, asking whether populations were more likely to exhibit local adaptation for $m < m^N_{crit}$. Simulations were initialized with all Aa individuals and run for 20,000 generations. Because our interest here was on whether polymorphic populations exhibited local adaptation, we included bidirectional mutation (from A to a or vice versa) at a rate of $\mu = 10^{-4}$ to facilitate the maintenance of both alleles. Populations were censused every 50 generations for the final 5000 generations, with at least 60 independent replicates per parameter combination. At each census, populations were considered to be locally adapted if, in both patches, the mean fitness of individuals in a patch was higher than had those individuals been transplanted to the other patch.

The critical threshold m^N_{crit} provided an accurate prediction for the migration rate below which populations were locally adapted for a substantial fraction of time in each patch (Fig. 2B; $h = k = 0.5$). In all cases, m^∞_{crit} was infinite and, therefore, failed to predict this transition between maintenance and loss of local adaptation. Similar results were obtained for asymmetrical selection and dominance coefficients ($h \neq k$ and $s \neq t$; Fig. S9B) and for other population sizes (Fig. S10). In Figures S8–S10, we show that the persistence time of locally adapted alleles rises dramatically for migration rates below m^N_{crit}.

Discussion

If two alleles experience opposing selection pressures in different parts of the range of a species, divergent selection and limited migration may favor the maintenance of both alleles. Because alleles can be lost stochastically, however, drift can reduce the potential for the maintenance of polymorphism, especially in small populations. The current study illustrates how the net effect of deterministic evolutionary processes (here, migration and selection) can be accurately approximated using the rate of change in frequency of a rare allele within an infinitely large population ($\delta = \lambda_l - 1$). This diversification coefficient can then be spliced into Kimura's equation for the fixation probability from a single-patch model to predict the invasion probability of an allele introduced into the patch where it is favored (Fig. 2A) and the critical migration rate below which populations tend to exhibit local adaptation (Fig. 2B). When the diversification coefficient is strong relative to drift (i.e., $2\delta \gg 1/(2N)$), locally favored alleles have a high probability of invading when rare and a long persistence time, and sampled populations are likely to exhibit local adaptation. As migration rates increase, however, the net diversification coefficient declines, resulting in less local adaptation.

Although the leading eigenvalue of the Jacobian matrix, λ_1, integrates the effects of selection and migration to describe the asymptotic rate of increase of a rare allele, it is not obvious that splicing this into Kimura's equations for fixation probability would work as well as we observed in our simulations. Similarly, despite the coarseness of the splicing approach for calculating m^N_{crit}, the stochastic migration threshold functions agree well over a range of symmetrical and asymmetrical selection and dominance coefficients. In the extreme case of additive alleles with equal and opposite effects in the two patches, the deterministic analysis (m^∞_{crit}) predicts that polymorphism can be maintained regardless of the strength of selection (Fig. 1). This prediction is not met, however, in finite populations. Instead, we find that polymorphism is likely to be maintained only when the migration rate is less than m^N_{crit}, a more restrictive condition, especially when selection is weak.

We show that the interplay between migration, selection, and drift tends to result in threshold behavior of the system, with rapid transitions at the critical migration rate (m^N_{crit}) between regions of parameter space that do and do not maintain polymorphism. Drift in finite populations can dramatically change the conditions maintaining a polymorphism with locally adapted alleles because when migration and selection nearly balance, the net effect of these deterministic processes can be very weak and overwhelmed by drift, even when selection is not itself weak.

ACKNOWLEDGMENTS

We would like to thank M. Whitlock for extensive discussion in the development of this manuscript. We would also like to thank N. Barton for suggesting the branching process approach, M. Kirkpatrick, D. Irwin, R. Bürger, and anonymous reviewers for constructive criticism, and F. Guillaume for providing extensive help with the individual-based simulations. Individual-based simulations were run on the WestGrid computing cluster. Funding for this project was provided by an NSERC PGS-D to SY, an NSERC Discovery Grant to SPO, an NSERC Discovery Grant to Mike Whitlock, and a grant to Laurent Lehmann (PP00P3–123344; Swiss NSF).

LITERATURE CITED

Billiard, S., and T. Lenormand. 2005. Evolution of migration under kin selection and local adaptation. Evolution 59:13–23.

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

Bulmer, M. G. 1972. Multiple niche polymorphism. Am. Nat. 106:254–257.

Bürger, R. 2009a. Multilocus selection in subdivided populations II. Maintenance of polymorphism under weak or strong migration. J. Math. Biol. 58:979–997.

———. 2009b. Polymorphism in the two-locus levene model with nonepistatic directional selection. Theor. Popul. Biol. 76:214–228.

Crow, J. F., and M. Kimura, eds. 1970. An introduction to population genetics theory. Harper and Row, New York.

Felsenstein, J. 1976. The theoretical population genetics of variable selection and migration. Annu. Rev. Genet. 10:253–280.

Gavrilets, S., and N. Gibson. 2002. Fixation probabilities in a spatially heterogeneous environment. Popul. Ecol. 44:51–58.

Guillaume, F., and J. Rougemont. 2006. Nemo: an evolutionary and population genetics programming framework. Bioinformatics 22:2556–2557.

Haldane, J. B. S. 1930. A mathematical theory of natural and artificial selection. Part VI. Isolation. Proc. Camb. Philos. Soc. 26:220–230.

Karlin, S. 1982. Classifications of selection migration structures and conditions for a protected polymorphism. Evol. Biol. 14:61–204.

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

Lenormand, T. 2002. Gene flow and the limits to natural selection. Trends Ecol. Evol. 17:183–189.

Maynard Smith, J. 1966. Sympatric speciation. Am. Nat. 100:637–650.

Moran, P. A. P. 1962. The statistical processes of evolutionary theory. Clarendon Press, Oxford.

Nagylaki, T. 2009. Polymorphism in multiallelic migration-selection models with dominance. Theor. Popul. Biol. 75:239–250.

Nagylaki, T., and Y. Lou. 2007. Evolution under multiallelic migration-selection models. Theor. Popul. Biol. 72:21–40.

———. 2008. The dynamics of migration-selection models. Tutorials in Mathematical Biosciences IV: evolution and ecology. Lecture Notes in Mathematics. 1922:117–170.

Spichtig, M., and T. J. Kawecki. 2004. The maintenance (or not) of polygenic variation by soft selection in heterogeneous environments. Am. Nat. 164:70–84.

Star, B., R. J. Stoffels, and H. G. Spencer. 2007. Single-locus polymorphism in a heterogeneous two-deme model. Genetics 176:1625–1633.

Tachida, H., and M. Iizuka. 1991. Fixation probability in spatially changing environments. Genet. Res. 58:243–251.

Whitlock, M. C., and R. Gomulkiewicz. 2005. Probability of fixation in a heterogeneous environment. Genetics 171:1407–1417.

Wright, S. 1931. Evolution in Mendelian populations. Genetics 16:97–159.

Yeaman, S., and M. C. Whitlock. 2011. Genetic architecture of adaptation under migration-selection balance. Evolution. doi:10.1111/j.1558-5646.2011.01269.x [Epub ahead of print].

Associate Editor: N. Barton

Supporting Information

The following supporting information is available for this article:

Figure S1. Even in very small populations ($N = 100$), the predicted invasion probability is accurate as long as selection is strong relative to the inverse of the population size (purple and green curves).

Figure S2. Invasion probability of a new mutation in a finite population ($N = 1000$ in each patch).

Figure S3. Local invasion probability for dominant and recessive mutations.

Figure S4. Invasion probability for mutations with unequal dominance and selection coefficients ($s \neq t$, $h \neq k$; see legend), with curves giving the numerical solutions to equation (S1) for patches of size $N = 1000$.

Figure S5. Eigenvectors and the invasion of locally adapted alleles.

Figure S6. Predicted invasion probabilities according to the branching process solution (P_1; blue curves) versus the splicing approximation with either $\delta = \lambda_l - 1$ (red curves) or $\delta = P_1/2$ (black curves).

Figure S7. Predicted invasion probabilities according to the branching process solution when the mutation is introduced into the patch where it is disfavored (P_2; dashed lines) versus simulation results (points).

Figure S8. Persistence time of a pair of divergent alleles under migration and selection.

Figure S9. (A) The persistence time and (B) the fraction of time points at which both populations are locally adapted with asymmetrical selection ($s \neq t$) and dominance ($h \neq k$) coefficients, as shown in the legend.

Figure S10. (A) The persistence time of a pair of divergent alleles and (B) the fraction of time points at which both populations were locally adapted under recurrent mutation, both increase with increasing population size.

Figure S11. Change in the diversification coefficient (δ) with migration for the cases considered in Figures S8 (dashed black curves) and S9 (gray solid curves), ordered in the same manner as in those figures.

Supporting Information may be found in the online version of this article.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

SUPPLEMENTARY MATERIALS

Invasion probability for dominant or recessive mutations

While equation (9) provides an accurate approximation for the local invasion probability of an additive allele in a finite population with selection and migration (substituting δ for *S*), even in small populations (Figure S1), it tends to over- or underestimate the invasion probability of dominant or recessive mutations, respectively (Figure S2). More generally, for a mutation with a dominance coefficient of *H* and a selection coefficient of *S*, Kimura (1962) showed that the probability of fixation of an allele introduced at frequency *p* in a diploid population equals:

$$u(p) = \frac{\int_0^p e^{-2cDx(1-x)-2cx}dx}{\int_0^1 e^{-2cDx(1-x)-2cx}dx}, \qquad \text{(S1)}$$

where $c = NS$ and $D = 2H - 1$. As δ represents the rate of increase of a mutation when rare, which is equal to the rate of increase of a heterozygote, we can substitute δ for all *HS* terms and $\delta/(1 - h)$ for all *S* terms, to approximate the invasion probability of a mutation with arbitrary dominance in the patch where it is favored. (Note that *h* in this paper refers to the dominance of the locally deleterious allele, whereas *H* in Kimura's formulation refers to the dominance of the beneficial allele; hence, with weak selection, we replace *H* with $1 - h$.) Figure S3 shows the predicted invasion probabilities for both dominant and recessive mutations with $s = t = 0.0025$ over a range of migration rates. This modified approach yields very accurate predictions when $h = k = 0.25$, correcting the discrepancy shown in Figure S2C. The modified approach is also very accurate for low migration rates when $h = k = 0.75$; once migration rates rise above m_{crit}^N (solid vertical line), however, the local invasion probability predicted by Kimura's equation

differs from the observations in the simulations, likely because the simulations track the invasion to high frequency, while Kimura's equation measures a rise to fixation. The match to simulations continues to be good even when dominance and selection coefficients are not symmetrical in the two patches (Figure S4).

We note that when the deterministic analysis predicts invasion (i.e., there is a leading eigenvalue greater than one), the leading eigenvalue is the eigenvalue that describes the rate of spread when the rare allele is predominantly in the patch where it is favored (i.e., this is the direction in which the associated eigenvector points). There is always a second eigenvalue that is less than one that describes the rate of decline of the allele when it occurs primarily in the patch where it is disfavored (i.e., along the direction of the second eigenvector, see Figure S5). If we introduce a single mutation in the patch where it is selected against, the initial frequency vector will be closer to this second eigenvector and loss will occur with high probability (see next section). We thus focus our attention on cases where an allele is introduced in the patch where it is favored, yielding an initial frequency vector closer to the eigenvector associated with the leading eigenvalue. In this case, the dynamics of the rare allele can be well approximated by focusing on the leading eigenvalue.

Approximations using branching processes

It is also possible to define the diversification coefficient using a branching process approach with migration (Barton 1987). Letting P_i equal the probability that a single allele introduced in patch i will become established and assuming a Poisson distributed number of offspring, solutions for P_i are obtained following the logic of Haldane (1927),

who noted that for the allele to be ultimately lost, each offspring copy must also be ultimately lost:

$$1 - P_1 = \sum_j Poisson(1 + s_1)\left(1 - P_1^*\right)^j$$
$$1 - P_2 = \sum_j Poisson(1 + s_2)\left(1 - P_2^*\right)^j$$

where s_i is the selection coefficient favoring the rare allele in the i^{th} patch. In the terminology used in the main text, $s_1 = \frac{W_{1,Aa}}{W_{1,AA}} - 1 = \frac{1 - hs}{1 - s} - 1$ and $s_2 = \frac{W_{2,Aa}}{W_{2,AA}} - 1 = -kt$ when considering the invasion of allele *a*. Here, $P_1^* = \left(1 - m_{12}\right)P_1 + m_{12}P_2$ and $P_2^* = \left(1 - m_{21}\right)P_2 + m_{21}P_1$ describe the average establishment probability for an offspring copy of the allele, given that the offspring might have migrated to the other patch, where m_{ij} is the migration rate between the i^{th} and j^{th} patches. This logic requires that the fate of each offspring copy is independent of all others, effectively requiring an infinitely large population size without competition in each patch.

Assuming weak selection and migration and evaluating the sums, the fixation probabilities satisfy:

$$\begin{matrix} m_{21}(P_2 - P_1) + s_1 P_1 - \frac{P_1^2}{2} \approx 0 \\ m_{12}(P_1 - P_2) + s_2 P_2 - \frac{P_2^2}{2} \approx 0 \end{matrix},$$

These equations can be solved for the P_i, but the solutions are cumbersome roots to a cubic polynomial. Alternatively, they can be numerically solved.

When $m << m_{crit}^N$, the predicted invasion probability in the patch where the allele is favoured, P_1, is quite close to the probability predicted using $\delta = \lambda_l - 1$ for *S* in Kimura's diffusion equation, as in the main text (Figure S6; red vs. blue lines). As

migration approaches m^N_{crit}, however, the branching process breaks down because the net effect of selection and migration become weak relative to drift in a finite population. Similarly, when the mutation is introduced into the patch where it is disfavored, the branching process approach yields accurate predictions of the probability of invasion when $m << m^N_{crit}$, but underestimates the invasion probability when migration is above this threshold (Figure 7). We can, however, define the diversification coefficient from $P_1 = 2\delta$ (recall that the expected fixation probability in a single large population is twice the selection coefficient, assuming weak selection), and use this value of δ for S in Kimura's diffusion equation to approximate the fixation probability in finite populations (black curves). Both splicing approaches, with $\delta = \lambda_l - 1$ or $\delta = P_1/2$, yield similarly accurate approximation to simulations (compare Figure S6 to Figure 2).

We can also derive the critical migration threshold by solving for the migration rate that satisfies $P_1 = 2\ \delta = 1/2N$. This yields values for the critical threshold that are quite similar to the approach described in the main body of the manuscript (m^N_{crit} for branching process approach vs. the splicing approach is {0.318 vs. 0.320},{0.052 vs. 0.050}, {0.0133 vs. 0.0124} and {0.00335 vs. 0.00300} for $s = t = 0.025$, 0.01, 0.005, and 0.0025, respectively, when $N = 1000$ and $h = k = 0.5$).

We note that the splicing approach based on the eigenvalues of a deterministic analysis, with $\delta = \lambda_l - 1$, can be described more easily without resorting to solutions to cubic equations and so is used in the main text. That said, if we wanted to know the relative invasion probability of allele a when introduced into patch 1 versus patch 2, this is more naturally accomplished using the branching process solutions for P_1 and P_2, respectively.

Critical migration threshold and the persistence time of a polymorphism

To explore the utility of the m_{crit}^{N} threshold for predicting regions of parameter space within which adaptive divergence can be maintained for long periods of time, each patch was initially fixed for the locally favored allele, and the system was allowed to evolve without any subsequent mutation. Persistence time was measured as the elapsed time before loss of either allele, averaged over 60 independent simulation replicates. Because the fixation time becomes very long with decreasing *m*, simulations were run for a maximum of 100,000 generations.

With high migration rates ($m > m_{crit}^{N}$), the waiting time before the loss of either one of the alleles was approximately described by the neutral expectation (equation 8.9.3, Crow and Kimura 1970), as shown in Figure S8. As *m* fell below the m_{crit}^{N} threshold, persistence times increased dramatically. Because m_{crit}^{N} is considerably higher for alleles with larger selection coefficients, simulations with stronger selection maintained polymorphism over a much larger region of parameter space (Figure S8). In all cases, m_{crit}^{∞} was infinite and failed to predict the transition to longer persistence times, which occurred near m_{crit}^{N}.

The critical migration thresholds also provided reasonably accurate predictions of the migration rate below which polymorphism tends to be maintained for a long period of time when selection and dominance coefficients were asymmetrical (Figure S9A). In these cases again, long-term persistence is typically only seen with migration rates lower than m_{crit}^{N} (left of vertical lines), although polymorphism persists for migration rates slightly larger than this cutoff whenever one allele is more strongly protected from loss than the other allele (i.e., $\lambda_{l,a}$ and $\lambda_{l,A}$ differ substantially). For example, in Figure S9 at

$m = m_{crit}^{N}$, the diversification coefficient (δ) favoring the spread of allele *A* when rare is 323 (triangles), 360 (squares) and 14 (circles) times larger than that favoring the spread of allele *a*, so that polymorphism is more likely to persist whenever *A* becomes rare. Results for the fraction of time points at which both populations exhibited local adaptation (i.e., higher mean fitness when raised in the patch of origin than when transplanted to the other patch) follow similar patterns to those seen for persistence time (Figure S9B). Furthermore, across a range of population sizes, m_{crit}^{N} captures the transition between when local adaptation persists and when it rapidly disappears (Figure S10).

As migration rates vary, we note that the width of the transition region between when local adaptation is consistently maintained and when it is rare varies from case-to-case. As shown in Figure S10, differences in the width of this transition can be explained by how rapidly the diversification coefficient, δ, drops with increasing migration.

Literature Cited in Supplementary Materials

Barton, N.H. 1987. The probability of establishment of an advantageous mutation in a subdivided population. Genetical Research. 50:35-40.

Haldane, J.B.S. 1927. A mathematical theory of natural and artificial selection, Part V: Selection and mutation. Mathematical Proceedings of the Cambridge Philosophical Society. 23:838-844.

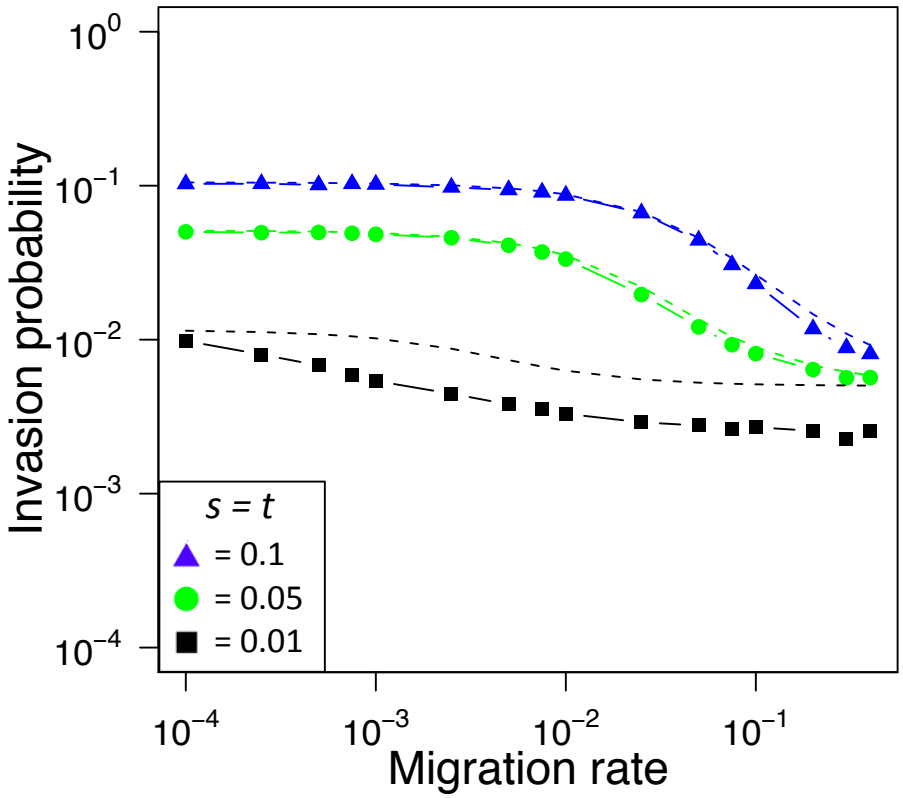

Figure S1. Even in very small populations ($N = 100$), the predicted invasion probability is accurate as long as selection is strong relative to the inverse of the population size (purple and green curves). The accuracy becomes poor with weaker selection (black curves), because drift overwhelms selection even in the patch where the allele is locally favored (i.e., inequality (10) is no longer satisfied). Simulations as in Figure 2A, but with $N = 100$ and $h = k = 0.5$. Dashed lines correspond to predictions from Kimura's fixation probability using δ for S in equation (9).

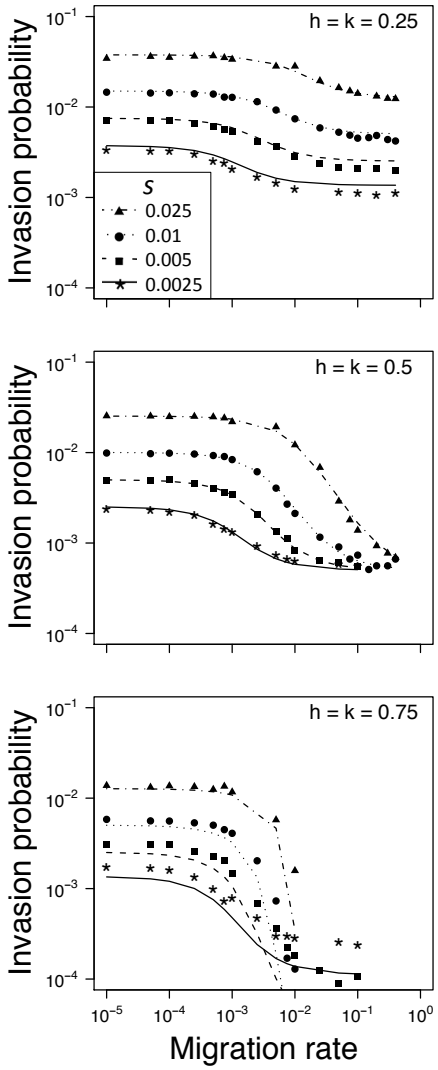

Figure S2. Invasion probability of a new mutation in a finite population ($N = 1000$ in each patch). Symbols correspond to symmetrical selection coefficients ($s = t$; see legend). Dashed lines correspond to predictions from Kimura's fixation probability using $\delta = \lambda_1 - 1$ for S in equation (9). Results for $s = t = 0.025$ and 0.01 and $h = k = 0.75$ are not shown for high migration rates, as the low frequency of invasion precluded an accurate estimate of its probability. The middle panel repeats results from Figure 2A for reference.

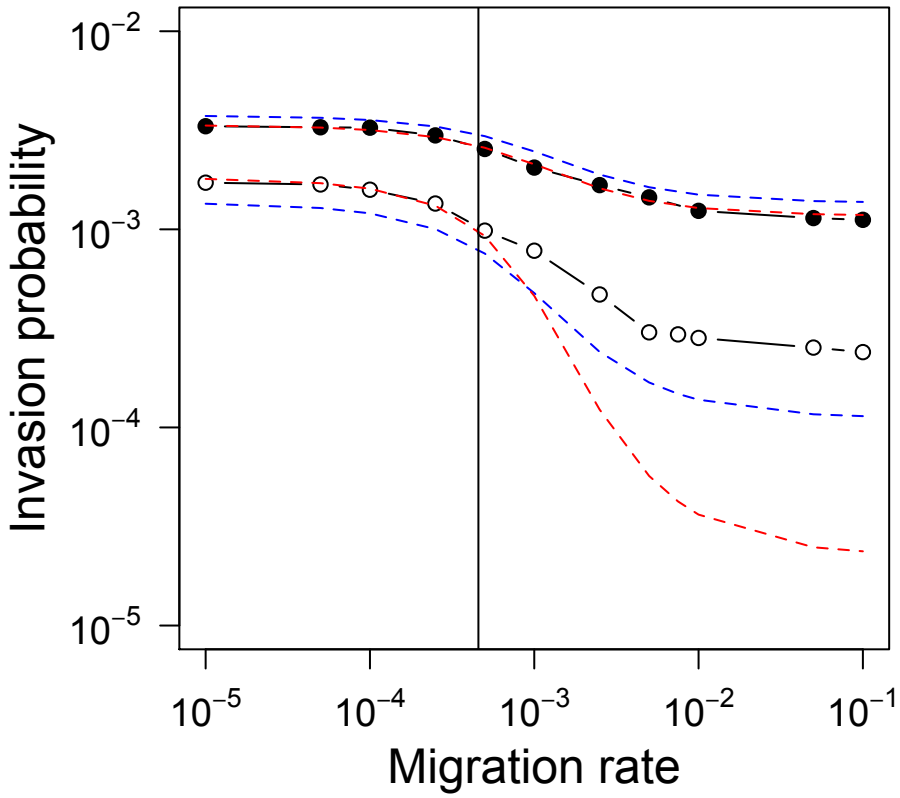

Figure S3. Local invasion probability for dominant and recessive mutations. Results for simulations are shown as solid lines with filled circles ($h = k = 0.25$) and open circles ($h = k = 0.75$) for $s = t = 0.0025$ and $N = 1000$; predictions based on the fixation probability for an additive allele (equation 9) are shown with blue dashed curves, while numerical solutions of equation (S1) for arbitrary dominance are shown with red dashed curves. The vertical solid line indicates the critical migration threshold, m_{crit}^{N}, for $h = k = 0.75$; m_{crit}^{N} is undefined for $h = k = 0.25$, as these parameters satisfy inequality (12), implying that any migration rate can permit divergence.

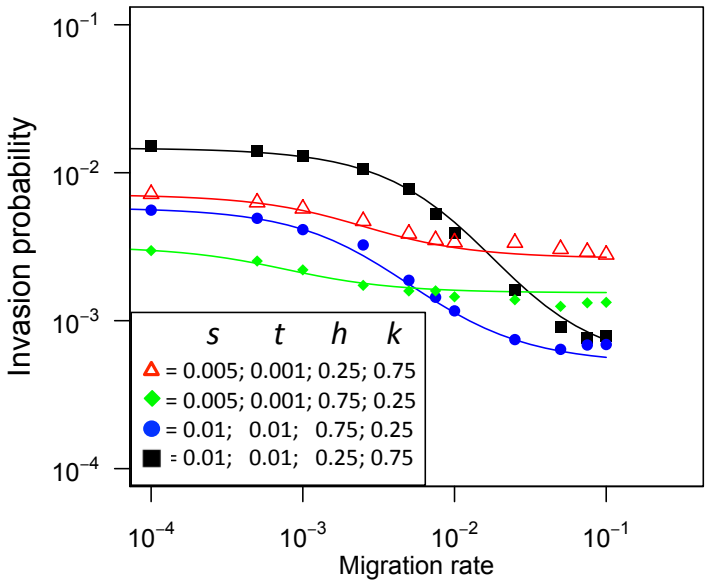

Figure S4. Invasion probability for mutations with unequal dominance and selection coefficients ($s \neq t$, $h \neq k$; see legend), with curves giving the numerical solutions to equation (S1) for patches of size $N = 1000$.

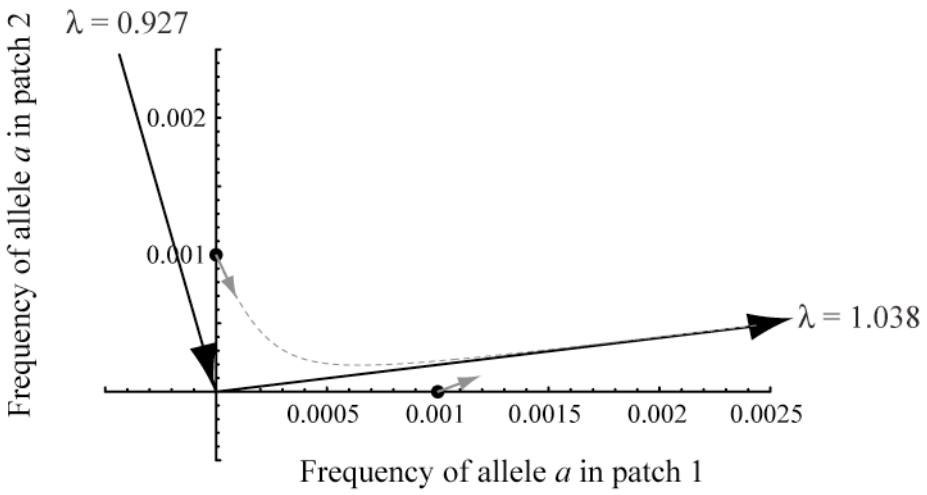

Figure S5. Eigenvectors and the invasion of locally adapted alleles. The frequency of an allele, *a*, in patches 1 and 2 can be represented as a vector {x,y}, with movement in the xy-plane representing evolutionary change (here with $h = k = 0.5$, $s = t = 0.1$, and $m = 0.02$). While the *a* allele is rare, the dynamical system can be described by stretching by a factor $\lambda = 1.038$ along the eigenvector associated with the leading eigenvalue (the long, nearly horizontal, arrow) and by shrinking by a factor $\lambda = 0.927$ along the second eigenvector (the long, nearly vertical, arrow). Thus, if the *a* allele starts primarily in patch 1 where it is favored (the dot at {0.001,0}), it will spread at a rate that is well predicted by the leading eigenvalue and in the direction of the associated eigenvector (grey arrow indicates the change over five generations). In contrast, if the *a* allele starts primarily in patch 2 where it is disfavored (the dot at {0,0.001}), the dynamics will initially be dominated by the second eigenvalue, and allele *a* declines in frequency. In an infinite population, the *a* allele would eventually spread even if introduced into the patch where it is disfavored (dashed grey curve), but in a finite population, the initial decline when introduced into patch 2 will often lead to the loss of the rare allele.

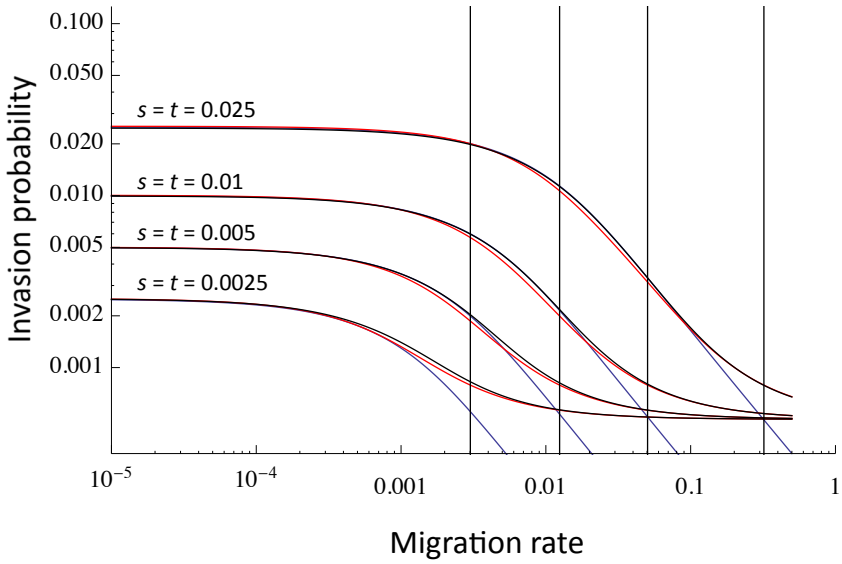

Figure S6. Predicted invasion probabilities according to the branching process solution (P_1; blue curves) versus the splicing approximation with either $\delta = \lambda_l - 1$ (red curves) or $\delta = P_1/2$ (black curves). Figure corresponds to the data presented in Figure 2A, with $s = t$, $h = k = 0.5$ and $N = 1000$. Critical migration thresholds (m_{crit}^{N}) are shown as vertical lines (from left to right: $s = t$ equal to 0.0025, 0.005, 0.01, 0.025, respectively).

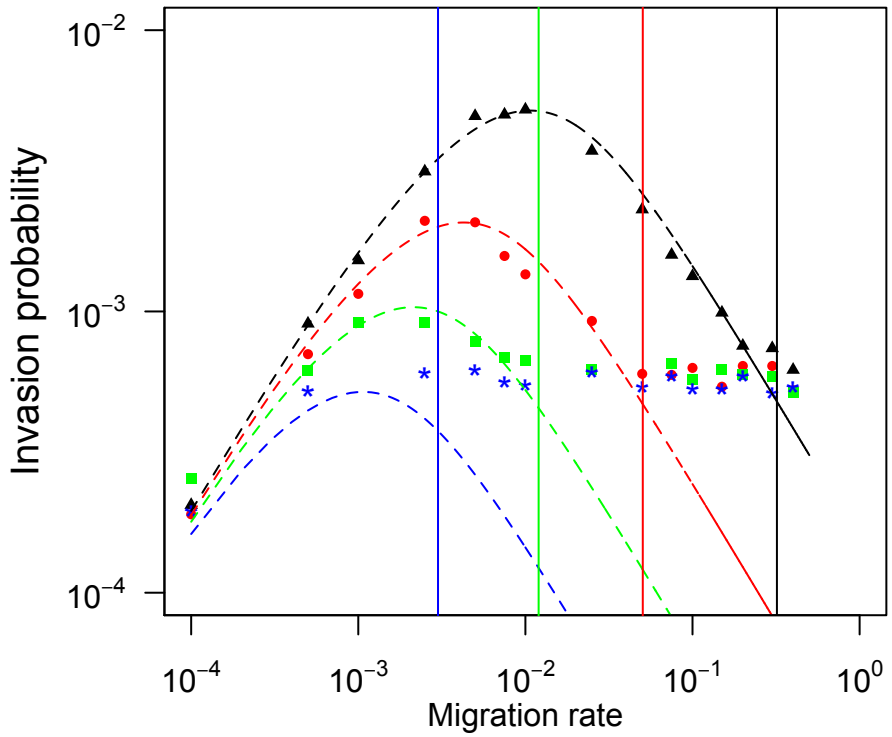

Figure S7. Predicted invasion probabilities according to the branching process solution when the mutation is introduced into the patch where it is disfavoured (P_2; dashed lines) versus simulation results (points). Figure corresponds to the data presented in Figure 2A, with $s = t$, $h = k = 0.5$ and $N = 1000$. Critical migration thresholds (m_{crit}^{N}) are shown as vertical lines (from left to right: $s = t$ equal to 0.0025 (blue), 0.005 (green), 0.01 (red), 0.025 (black), respectively).

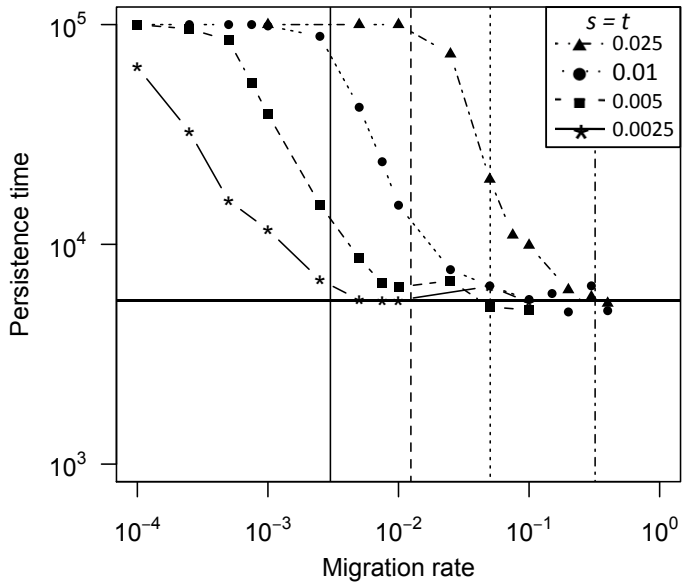

Figure S8. Persistence time of a pair of divergent alleles under migration and selection. Plots show simulation results for $N = 1000$, $h = k = 0.5$, $s = t$, and the same selection coefficients and symbols as in Figure 2. Vertical lines correspond to m_{crit}^{N} for the four selection regimes; the thick horizontal line indicates the expected persistence time under neutrality. Recall that the simulations are terminated after 10^5 generations, so points near the top of the graph are underestimates of the actual persistence times.

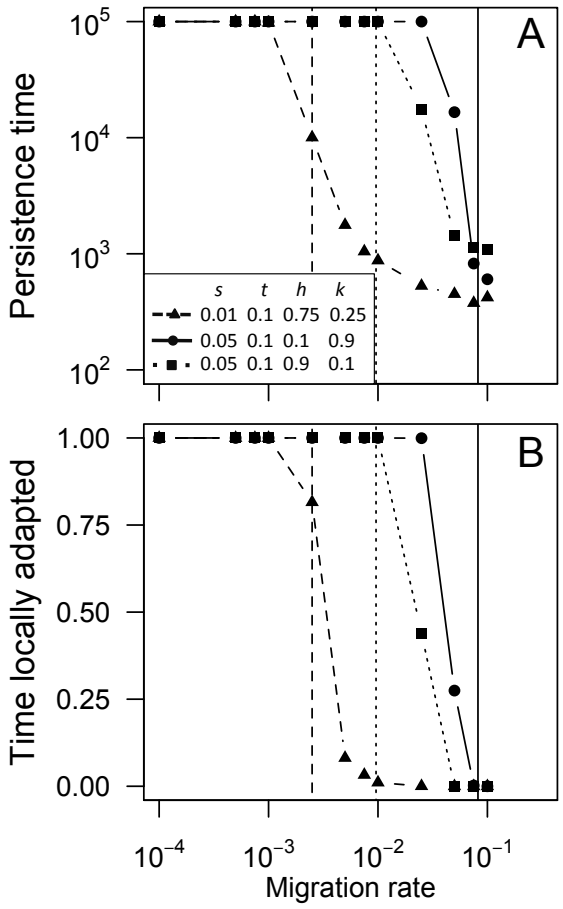

Figure S9. (A) The persistence time and (B) the fraction of time points at which both populations are locally adapted with asymmetrical selection ($s \neq t$) and dominance ($h \neq k$) coefficients, as shown in the legend. All other parameters are as in Figure S8.

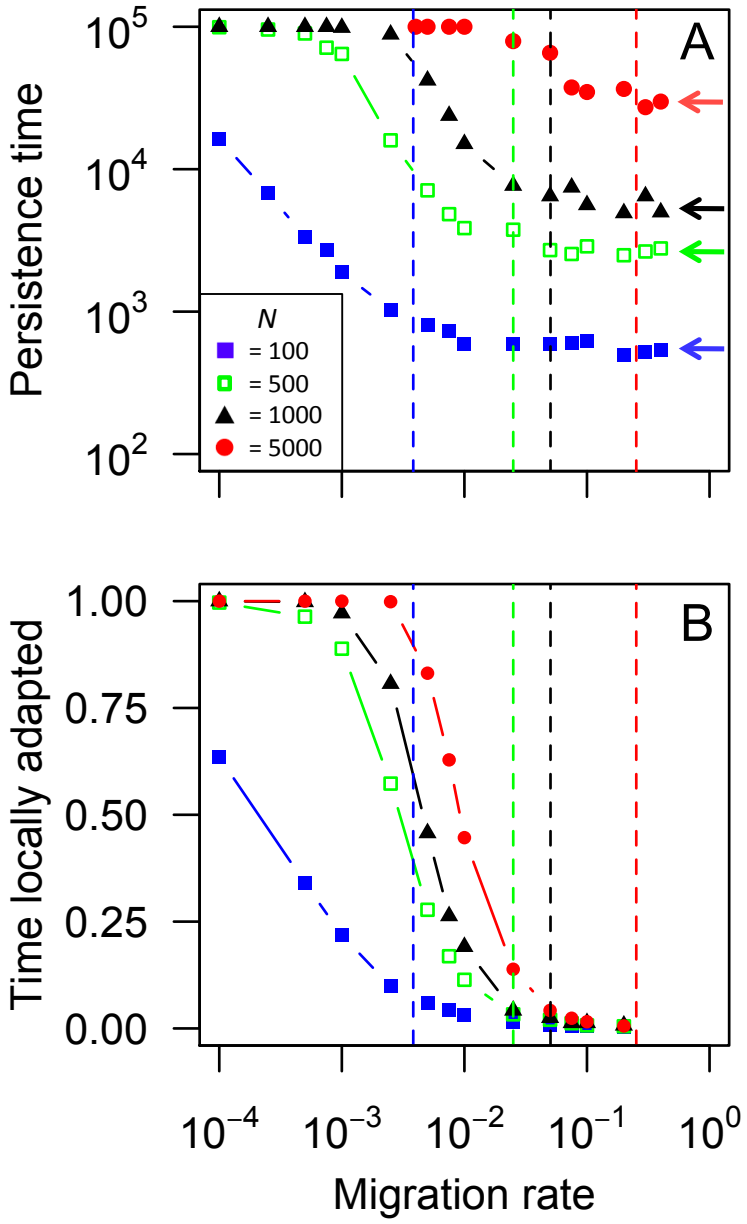
A
Persistence time
10^5
10^4
10^3
10^2
N
= 100
= 500
= 1000
= 5000
B
Time locally adapted
1.00
0.75
0.50
0.25
0.00
10^{-4}
10^{-3}
10^{-2}
10^{-1}
10^0
Migration rate

Figure S10. (A) The persistence time of a pair of divergent alleles and (B) the fraction of time points at which both populations were locally adapted under recurrent mutation both increase with increasing population size. Simulations followed the approach used to generate the data in Figure 2B with $h = k = 0.5$ and $s = t = 0.01$ but explored a range of population sizes: $N = 100$ (blue squares), $N = 500$ (green squares), $N = 1000$ (black triangles), and $N = 5000$ (red circles). Dashed lines correspond to m_{crit}^{N} for the four population sizes while the arrows in (A) indicate the expected persistence times under neutrality for these values of N (from Crow and Kimura 1970; equation 8.9.3). The deterministic critical threshold predicts the maintenance of both alleles for all cases considered in this figure because $m_{crit}^{\infty} \rightarrow \infty$, whereas m_{crit}^{N} accurately predicts that there should be a threshold migration rate below which local adaptation is likely to be observed. Simulations in panel A were stopped after a maximum of 10^5 generations.

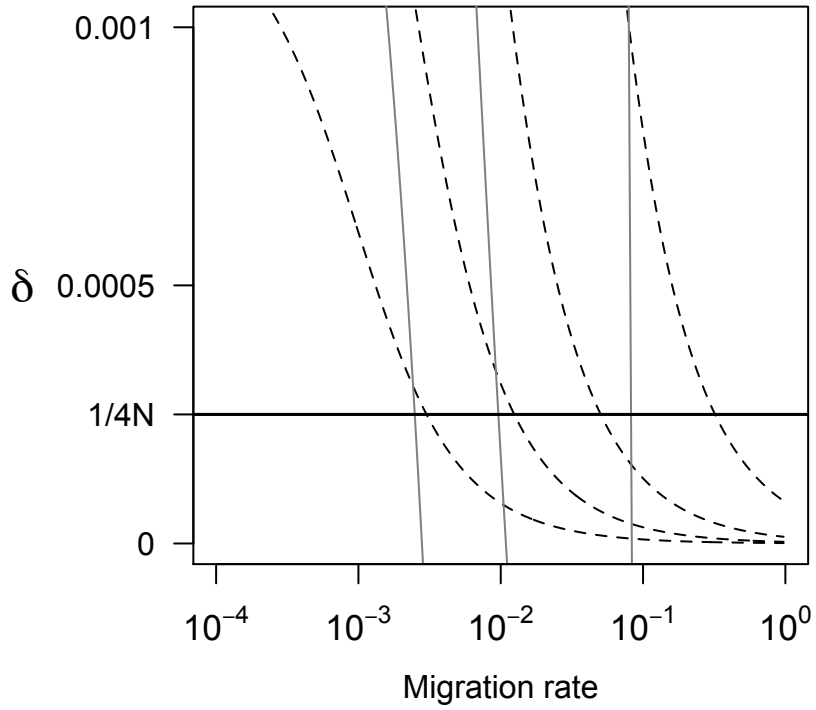

Figure S11. Change in the diversification coefficient (δ) with migration for the cases considered in Figures S8 (dashed black curves) and S9 (grey solid curves), ordered in the same manner as in those figures. A thick horizontal line is plotted at 1/4*N*, so that each curve crosses this line at the critical migration threshold, m_{crit}^{N}. When the slope of δ at this threshold is steeper (grey curves), we observe a narrower transition zone between when local adaptation is consistently maintained and when it is rare (note the rapid transition in Figure S9 relative to Figure S8). In this comparison, the width of the transition zone is smaller when selection and dominance are asymmetrical (Figure S9), but this need not be the case depending on the parameters.