Adapt globally, act locally: The effect of selective sweeps on bacterial sequence diversity

J. Majewski, McGill University
Frederick M Cohan, Wesleyan University
Adapt Globally, Act Locally: The Effect of Selective Sweeps on Bacterial Sequence Diversity

Jacek Majewski1 and Frederick M. Cohan
Department of Biology, Wesleyan University, Middletown, Connecticut 06459-0170
Manuscript received September 20, 1996
Accepted for publication April 23, 1999

ABSTRACT

Previous studies have shown that genetic exchange in bacteria is too rare to prevent neutral sequence divergence between ecological populations. That is, despite genetic exchange, each population should diverge into its own DNA sequence-similarity cluster. In those studies, each selective sweep was limited to acting within a single ecological population. Here we postulate the existence of globally adaptive mutations, which may confer a selective advantage to all ecological populations constituting a metapopulation. Such adaptations cause global selective sweeps, which purge the divergence both within and between populations. We found that the effect of recurrent global selective sweeps on neutral sequence divergence is highly dependent on the mechanism of genetic exchange. Global selective sweeps can prevent populations from reaching high levels of neutral sequence divergence, but they cannot cause two populations to become identical in neutral sequence characters. The model supports the earlier conclusion that each ecological population of bacteria should form its own distinct DNA sequence-similarity cluster.

It is becoming increasingly clear that a full accounting of ecological diversity in the bacterial world requires a molecular approach. Molecular techniques have demonstrated that only a small fraction of bacterial species are culturable (Amman et al. 1995; Huber et al. 1995; Ohkuma and Kudo 1996), so our best hope of identifying the full scope of bacterial biodiversity is to characterize the sequence diversity of genes that can be amplified directly from natural habitats (Knight et al. 1992; Pace 1997). Such surveys typically yield clusters of organisms with similar sequences, and each sequence-similarity cluster is typically interpreted as a distinct ecological population (Britschgi and Giovannoni 1991; Murray and Stackebrandt 1995; Boivin-Jahns et al. 1996).

This interpretation is justified because in studies of more familiar and cultivable taxa, bacterial systematists have found an empirical correspondence between ecologically distinct populations and sequence-similarity clusters. That is, groups of bacteria known to be ecologically different generally fall into separate sequence-similarity clusters (Vandamme et al. 1996; Palys et al. 1997); conversely, ecologically uncharacterized strains that fall into separate sequence clusters have subsequently been found to have different ecological properties (Balmelli and Pifferetti 1996; Normand et al. 1996). Sequence surveys appear to be an efficient method for discovering the ecological diversity of culturable as well as non-culturable bacteria (Vandamme et al. 1996; Palys et al. 1997).

While ecologically distinct groups of bacteria are frequently distinguishable as separate sequence-similarity clusters, it is important to find a strong theoretical basis for this observation. If there are times when multiple ecological populations of bacteria fall together into the same sequence cluster, molecular approaches may severely underestimate bacterial biodiversity (Cohan 1994a,b, 1995, 1996, 1999).

Recent theory has shown why ecological populations should correspond to sequence clusters (Cohan 1994a,b; Palys et al. 1997). In this theory, ecological populations are defined so that (1) each adaptive mutation confers a benefit only in the genetic background of its original population, and (2) mutant cells bearing an adaptive mutation can outcompete only members of their own population. Natural selection favoring adaptive mutants within a particular population purges that population of genetic diversity at all loci, owing to the low rate of recombination in bacteria. [Each such purging event is called a “selective sweep” (Guttman and Dykhuizen 1994b); we refer to recurrent selective sweeps as “periodic selection” (Atwood et al. 1951; Koch 1974; Levin 1981).] Because an adaptive mutant does not outcompete cells from other populations, periodic selection purges only the diversity within populations and not the divergence between populations. Each round of periodic selection thereby enhances the distinctness of ecological populations at all loci and fosters the divergence
of different ecological populations into separate sequence-similarity clusters.

The tendency for bacterial populations to form separate sequence clusters is opposed by recombination between populations (Cohan 1994a). Depending on the rates of interpopulation recombination and the intensity of periodic selection, the model has shown three possible classes of outcomes of neutral sequence divergence between populations: (1) under extremely low rates of recombination, populations will diverge without bound, so that every nucleotide site that can be substituted harmlessly will eventually become substituted; (2) under higher rates of recombination, populations will reach an equilibrium level of divergence, so that populations fall into distinct sequence clusters, but divergence between them never becomes saturated; and (3) under yet higher rates of recombination, ecologically distinct populations will not be distinguishable by neutral sequence data, as the levels of divergence within and between populations will be nearly equal. Given the low rates of recombination estimated thus far for bacteria (Selander and Musser 1990; Maynard-Smith et al. 1993; Whittam and Ake 1993; Guttman and Dykhuizen 1994a; Roberts and Cohan 1995), the model predicts that each population should be distinct as a separate sequence-similarity cluster, as described in cases 1 and 2 above (Cohan 1994a, 1995; Palys et al. 1997).

Nevertheless, it is not clear that the existing model adequately predicts the degree of sequence divergence between ecological populations. Here we present an alternative and more general model for periodic selection, in which some mutations may be adaptive outside of the context of their original populations. In this model, the domain of competitive superiority of an adaptive mutant (i.e., the cell) is still its own ecological population, but the adaptive mutation (i.e., the allele) can be recombined into other populations, where it can confer higher fitness and cause a local selective sweep within each recipient population (Figure 1). This process may homogenize the populations for any segment that is cotransferred between populations along with the adaptive mutation. We have hypothesized that globally adaptive mutations could homogenize populations for neutral sequence diversity at all gene loci, provided that the size of fragments recombined is large enough and that universally adaptive mutations recur throughout the genome.

In this article, we present a coalescence model to explore the conditions under which universally adaptive mutations can homogenize neutral sequence diversity across ecological populations. We tested whether different ecological populations might fail to diverge into separate sequence-similarity clusters under the rates of recombination observed in bacteria. We also tested whether universally adaptive mutations may prevent populations with low recombination rates from diverging without bound.

**Figure 1.** Transmission of a globally adaptive mutation from population 1, where the mutation originally occurred, to population 2. The asterisk represents an adaptive mutation; the B locus is a gene of interest, whose allelic diversity we are monitoring. The allele at this locus was originally associated with the adaptive mutation in population 1. The diagram shows a possible scenario for coalescence of gene segments from different populations. A selective sweep has just been completed in population 1: most of the local diversity at locus B has been purged and the population contains mostly the allele originally associated with the adaptive mutation (allele B1). The remaining diversity in population 1 is represented byb, the allele is then cotransferred along with the adaptive allele into population 2. The fraction of the genome that is cotransferred is h, and y is the distance between the target of selection and the gene of interest. When the selective sweep is complete in population 2, most of the diversity at locus B has been purged by the B allele (provided that recombination rates are low enough). Hence, the selective sweep leads to coalescence of neutral gene segments within the same population and also in different populations.

**THE MODEL**

**Ecological populations and adaptive mutations:** A metapopulation consists of n closely related ecological populations, each containing N cells (Table 1). Each population is adapted to a different ecological niche. Recombination occurs rarely within and between these populations, and the metapopulation is closed outside of recombination with other such metapopulations.

Following Cohan (1994a), we define an ecological population as the domain of competitive superiority of an adaptive mutant. Thus, an adaptive mutant would outcompete to extinction all other strains from the same population (because they are adapted to the same niche) but would not drive to extinction strains from other populations. While an adaptive mutant (i.e., the cell) has a competitive advantage only within its own ecological population, an adaptive mutation (i.e., the mutant allele) may be either locally or globally adaptive.

A locally adaptive mutation confers a benefit only in the genetic background of its original population, whereas a globally adaptive mutation can confer a benefit in any genetic background within the metapopulation. A global selective sweep occurs when a globally adaptive mutation recombines from its original population into other populations: any cell receiving the adaptive muta-
tion from the original population is then able to out-
compete other members of its own population (Figure 1).
Whereas a globally adaptive mutation confers fitness
globally to all cells in the metapopulation, natural selec-
tion acts only locally to favor the adaptive genotype
within each population.

We assume that selective sweeps are rare events and
that the duration of the sweep is short relative to the
time between sweeps.

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>Definitions of parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>Number of populations in the metapopulation</td>
</tr>
<tr>
<td>N</td>
<td>Number of cells in each population</td>
</tr>
<tr>
<td>c</td>
<td>Rate (per gene segment per genome per generation) at which individuals of a given population integrate DNA from any other individual in the metapopulation</td>
</tr>
<tr>
<td>c1</td>
<td>Rate (per gene segment per genome per generation) at which individuals of a given population integrate DNA from other individuals of the same population</td>
</tr>
<tr>
<td>c2</td>
<td>Rate (per gene segment per genome per generation) at which individuals of a given population integrate DNA from other populations</td>
</tr>
<tr>
<td>c3</td>
<td>Rate (per gene segment per genome per generation) at which individuals of a given population integrate DNA from one particular population (other than their own)</td>
</tr>
<tr>
<td>σp0, σl</td>
<td>Per population rate of global and local selective sweeps, respectively</td>
</tr>
<tr>
<td>μs, μl</td>
<td>Per capita rate of globally and locally adaptive mutations, respectively</td>
</tr>
<tr>
<td>z</td>
<td>Selective advantage of an adaptive mutant</td>
</tr>
<tr>
<td>h</td>
<td>Average size of a recombining DNA fragment, measured as a fraction of the genome</td>
</tr>
<tr>
<td>y</td>
<td>Distance between the adaptive mutation and the gene segment of interest, measured as fraction of the genome, for a particular selective sweep</td>
</tr>
<tr>
<td>p1</td>
<td>Probability that the local selective sweep results in coalescence of two segments from the same population</td>
</tr>
<tr>
<td>p2</td>
<td>Probability that a global selective sweep results in coalescence of two segments from the same population</td>
</tr>
<tr>
<td>p3</td>
<td>Probability that a global selective sweep results in coalescence of two segments from different populations</td>
</tr>
<tr>
<td>P</td>
<td>Average of p over many selective sweeps, representing all possible values of y; P, P0, P1, P3 are the average values of p1, p2, p3, respectively</td>
</tr>
<tr>
<td>q</td>
<td>Probability of corecombination of the adaptive allele and the segment of interest</td>
</tr>
<tr>
<td>μ0</td>
<td>Neutral mutation rate per third-base site</td>
</tr>
<tr>
<td>Δ0, Δ4</td>
<td>Number of nucleotide substitutions per third-base site, for two gene segments chosen randomly from the same population and different populations, respectively</td>
</tr>
<tr>
<td>π0, π4</td>
<td>Nucleotide divergence over all sites within and between populations, respectively</td>
</tr>
</tbody>
</table>

Rate of fixation of adaptive mutations: Following Co
an (1994a), adaptive mutation is modeled as a one-
step process that occurs randomly over time at a rate
μs (for global adaptations) or μl (for local adaptations)
per capita per generation. Each adaptive mutation con-
fers a selective advantage z. Taking into account that
only the fraction 2z of adaptive mutations is expected
to become fixed (assuming that the population size N

Recombination within and between populations: Re-
combination in bacteria is unidirectional and the seg-
ment recombined is usually a small fraction of the
gene (Smith 1988). We therefore model recombi-
nation as a gene conversion process in which a seg-
ment of the recipient DNA is replaced with the homolog
of the donor. The model is concerned with recombi-
nation at two loci: a “gene of interest,” whose sequence
divergence we wish to predict, and a selected gene,
whose adaptive mutation is favored by selection. Each
gene is assumed to be short enough so that it is not
split by recombination. A single recombination event
can involve one or both of the genes, depending on the
size of the recombining fragment (h) and the distance
between the loci (y).

Recombination follows a modified island model,
where c is the rate (per gene segment per genome per
generation) at which individuals integrate (as recipi-
ents) DNA at a gene segment of interest from other
individuals of the same ecological population; c is the
rate at which individuals integrate DNA from any other
ecological population; c is the total rate of recombi-
nation at which an individual integrates DNA from any
other individual in the metapopulation; thus c = c +
c. The value c is the rate at which individuals integrate
DNA from a particular ecological population (other than
their own). In a metapopulation consisting of n ecologi-
cal populations, c = c/ (n − 1).

Probability that a selective sweep leads to coales-
cence: Our model determines the expected time (going
backward from the present) to coalescence into a com-
mon ancestor for two homologous gene segments oc-
curring today in two different individuals. These indi-
viduals may be cells from the same or different ecological
populations of the metapopulation.

We define p as the probability that a selective sweep
leads to coalescence at a gene segment of interest. This
is the probability that two cells chosen from the meta-
population immediately following a selective sweep are
identical by descent for the gene segment of interest.
Whether a selective sweep results in coalescence at a
gene of interest depends on the relative magnitudes of the selective advantage of the adaptive mutation, the rate at which recombination separates the gene of interest from the selected gene, and the population size. If the rate of recombination is high and the selective advantage low, the event is unlikely to lead to coalescence.

We consider several instances of the variable $p$, corresponding to the probabilities of coalescence within and between populations, for globally and locally adaptive mutations: $p_l$ is defined as the probability that a local selective sweep within a population leads to coalescence of segments from that population; $p_g$ is the probability that a global selective sweep leads to coalescence of segments from the same population; and $p_{gd}$ is the probability that a global selective sweep leads to coalescence of segments from different populations of the metapopulation.

In appendix a, we derive a method (adapted from Kaplan et al. 1989) for calculating $p_l$, $p_g$, and $p_{gd}$, for the special case of two ecological populations, i.e., $n = 2$. The variables as $p_l$, $p_g$, and $p_{gd}$ are functions of $C_l$, $C_g$, $N$, and $q$, where $q$ is the probability that a recombination event results in corecombination of the adaptive allele with the segment of interest. This probability is a function of the length $h$ of the DNA taken up by a recipient cell during a recombination event and of the distance $y$ between the adaptive mutation and the segment of interest (both $h$ and $y$ are measured as fractions of the genome):

$$q = \begin{cases} \frac{(h-y)}{h} & \text{if } h \leq 1/2; \ y \leq h, \\ 0 & \text{if } h \leq 1/2; \ y > h, \\ \frac{(h-y)}{h} & \text{if } h > 1/2; \ y \leq 1 - h, \\ \frac{(2h-1)}{h} & \text{if } h > 1/2; \ y > 1 - h. \end{cases}$$

Because the coalescence of homologous segments from different populations requires that the transfer of the adaptive mutation from population 1 to population 2 includes the segment of interest (Figure 1), $p_{gd}$ will be highly dependent on the probability of cotransfer.

We assume that the size ($h$) of the recombining DNA fragment is constant, while adaptive mutations occur randomly throughout the genome. Because we are interested in modeling the consequences of multiple selective sweeps, we need to calculate the mean probability ($P$) that a selective sweep leads to coalescence, averaged over all possible distances ($y$) between the neutral marker and the adaptive mutations (i.e., between 0 and $1/2$ because the bacterial chromosome is circular):

$$P = 2 \int_0^{1/2} p(y) dy.$$  

Both the probabilities $p$ and the above integral were evaluated numerically.

**The coalescence model:** Our coalescence model calculates the expected time that two homologous gene segments (occurring in different organisms) have diverged since their last common ancestor. These gene segments are postulated to be short enough so that they are not split by recombination. The following are the expected times to coalescence for two strains from the same and different ecological populations, $E(t_s)$ and $E(t_d)$ (derived in appendix b):

$$E(t_s) = \frac{1 + \sigma_s(1 - P_s)E(t_s) + \sigma_g(1 - P_g)E(t_s) + 2c_sE(t_s)}{1/N + \sigma_s + \sigma_g + 2c_s}$$  

$$E(t_d) = \frac{1 + \sigma_s(1 - P_{gd})E(t_d) + 2c_sE(t_d)}{\sigma_g + 2c_s}.$$  

These equations were solved using Mathematica (Wolfram 1991). The solutions are too long to present here. Numerical representations of the probability density functions, $P_{1s}(t)$ and $P_{1d}(t)$, of the respective times to coalescence were also calculated using a method outlined in appendix b.

**Calculation of the expected nucleotide divergence:** The expected nucleotide sequence divergence is predicted using the probability density functions for $t_s$ and $t_d$ following Cohan (1994a). Nucleotide substitutions are postulated to consist only of synonymous mutations, and every third base substitution is taken to be synonymous, with no synonymous substitutions allowed at the first or second bases of codons. The number of neutral substitutions per third base site ($D$) is then obtained by multiplying the time to coalescence by twice the per third base site rate of mutation ($\mu$),

$$D = 2\mu d,$$  

where $t_s$ or $t_d$ are the times until coalescence of segments in the same population or different populations, respectively.

The nucleotide sequence divergence over all sites, $\pi$, may be calculated by correcting for multiple substitutions per site (Jukes and Cantor 1969) and correcting for substitutions occurring only at third base sites:

$$\pi = \frac{1}{3} \cdot \frac{3}{4} (1 - e^{-4D/3}).$$  

We used the probability density functions $P_{1s}(t)$ and $P_{1d}(t)$ to calculate the expected divergence

$$E(\pi) = \frac{1}{3} \cdot \frac{3}{4} \int_0^t (1 - e^{-4D/3}) \cdot P_{1s}(t) dt.$$  

The integration was carried out numerically. The expected nucleotide divergence between segments from the same population, $E(\pi_s)$, and different populations, $E(\pi_d)$, was calculated using $P_{1s}(t)$ and $P_{1d}(t)$, respectively. The neutral nucleotide divergence after infinite time reaches a level of $\pi$, which we refer to as “unbounded divergence.”
The ratio of globally to locally adaptive mutation rates: We next explore the effect of recurrent adaptive mutations on population structure. We focus on the significance of the ratio of globally to locally adaptive mutations. We maintain the total frequency of adaptive mutations constant, while allowing the ratio of global:local adaptations to vary. We consider three relative frequencies of global:local events, 1:0, 1:1, and 0:1 (Figure 3).

Figure 3 shows that globally adaptive mutations reduce neutral sequence divergence between populations compared to the case with only local adaptations. This effect is most pronounced at low recombination rates. When only local selective sweeps are possible, the model shows that a recombination rate of \(10^{-10}\) leads to unbounded neutral divergence between populations (i.e., \(\pi_d = 1/4\)). Increasing the global:local ratio decreases the divergence between populations by up to 50-fold.

The divergence within populations also decreases, but to a much lower extent. Thus, increasing the proportion of globally adaptive mutations makes the populations less distinct in neutral characters.

Consider next whether globally adaptive mutations can prevent different ecological populations from diverging into separate sequence clusters. We define populations as falling into separate sequence-similarity clusters when \(E(\pi_d) > 2E(\pi_c)\) (Palys et al. 1997). Using this criterion, Figure 4 shows that the critical recombination rates necessary for populations to diverge into separate clusters are nearly the same whether or not globally adaptive mutations occur.

Analysis of a simplified model with no locally adaptive mutations: We concentrated on the effect of globally adaptive mutations by considering the special case of a two-component metapopulation in which all the adaptive mutations are global (Figure 5). For this special case we treated the coalescence equations analytically to gain further insight into the behavior of the sequence divergence functions presented in Figure 3. Noting that bacterial populations are always large enough so that the probability of coalescence by drift is negligible relative to coalescence by periodic selection (i.e., \(1/N \ll \sigma_P\)), Equations 3 and 4 reduce to

\[
E(t_c) = \frac{4c_c + \sigma_g p_{gs}}{(2c_c + \sigma_g p_{gs})(2c_c + \sigma_g p_{gs}) - 4c_c^2}, \tag{8}
\]

\[
E(t_d) = \frac{4c_d + \sigma_g p_{gs}}{(2c_d + \sigma_g p_{gs})(2c_d + \sigma_g p_{gs}) - 4c_d^2}. \tag{9}
\]

We may consider \(\sigma_g p_{gs}\) and \(\sigma_g p_{gs}\) as pseudoparameters, representing the diversity-purging effect of periodic selection (i.e., the rate of selective sweeps times the probability of coalescence within each sweep). The times to coalescence are then determined by only three factors: \(c_c\), the rate of recombination between populations; \(\sigma_g p_{gs}\), the within-population diversity-purging effect of global

![Figure 2](image-url)
periodic selection; and \( \sigma_gP_g \), the between-population diversity-purging effect of global periodic selection.

Consider the relative magnitudes of \( \sigma_gP_g \) and \( \sigma_dP_d \).

We used Equation A4 of appendix a to calculate the
values of \( P_g \) and \( P_d \) across the range of frequency of recombination (c) and selective advantage (z) considered in this article, and we found that \( P_g \geq P_d / h \). We assume that the size of the recombination fragment \( h \) is usually <10\% for the genome (see discussion). Therefore \( \sigma_gP_g \gg \sigma_dP_d \). This leaves four regions of magnitude for \( c_h; c_h \gg \sigma_gP_g, c_h \sim \sigma_gP_g, c_h \ll \sigma_gP_g, \) and \( c_h \ll \sigma_dP_d \). These regions correspond to regions I through IV, respectively, of Figure 5.

Region I of Figure 5, \( c_h \gg \sigma_gP_g \), corresponds to very high recombination rates, yielding the following approximation of Equations 8 and 9:

\[
E(t_s) \approx E(t_d) \approx 2\sigma_gP_g. \tag{10}
\]

Region I of the graph corresponds to the case where high rates of recombination within populations (\( c_h \)) diminish the diversity-purging effect of periodic selection (\( P_g \)), while high values of interpopulation recombination (\( c_h \)) further homogenize the populations, making them indistinguishable (such that expected divergence levels within and between populations are equal).

The conditions in region II, \( c_h \sim \sigma_gP_g \gg \sigma_dP_d \), yield

\[
E(t_s) = \frac{2}{\sigma_gP_g} \tag{11}
\]

\[
E(t_d) = \frac{2}{\sigma_gP_g} + \frac{1}{2c_h}. \tag{12}
\]

In region II of Figure 5, recombination is no longer sufficient to prevent populations from diverging. The divergence between populations is greater than that within populations and is determined by the equilibrium between recombination (which acts to homogenize the populations) and local diversity-purging events (which tend to keep the populations distinct).
The conditions of region III, $\alpha_p^P_{gs} \gg \alpha_g^P_{gd} \sim \alpha_c$, yield

$$E(t_c) = \frac{4c_6 + \alpha_g^P_{gd}}{\alpha_g^P_{gd}(2c_6 + \alpha_g^P_{gd})}$$

$$E(t_d) = \frac{1}{2c_6 + \alpha_g^P_{gd}}.$$  

Region III reflects the increasing significance of global periodic selection. Divergence between populations is determined by the combined homogenizing effects of recombination ($c_6$) and global periodic selection ($\alpha_g^P_{gd}$).

Region IV corresponds to the case of extremely rare recombination, $\alpha_p^P_{gs} \gg \alpha_g^P_{gd} \gg c_6$, yielding

$$E(t_c) = \frac{1}{\alpha_g^P_{gs}}$$

$$E(t_d) = \frac{1}{\alpha_g^P_{gd}}.$$  

This is the limiting case, where recombination between populations becomes so infrequent that its effects are entirely overwhelmed by periodic selection. In this limit, the divergence within populations is determined solely by the intensity of local purging of diversity, while the divergence between populations is only limited by the intensity of global purging of diversity.

Under the conditions of rare recombination (region IV), the ratio of the times to coalescence (i.e., $E[t_c]/E[t_d]$) approaches $1/h$. This follows from two consequences of rare recombination. First, because the locus of interest and the adaptive mutation are rarely separated by recombination, a selective sweep almost certainly leads to coalescence of gene segments from the same population (i.e., $P_{gs} \approx 1$). Second, the transmission of the adaptive mutation from population 1 to population 2 is likely to be the result of a single transfer event. Hence, the probability of coalescence of two gene segments from different populations ($P_{gd}$) approaches the probability that the transfer event was a cotransfer of the adaptive mutation and the segment of interest (averaged over all distances between the two loci). That is, $P_{gd} \approx h$, and $E[t_c]/E[t_d] = 1/h$.

**Effect of recombination fragment size on population divergence:** We consider next the effect of the recombination fragment size ($h$) on the distinctness of ecological populations. In general, larger recombination fragments increase the probability ($q$) that a gene of interest will cotransfer across populations with an adaptive mutation (Equation 1), thus fostering coalescence of segments between populations (Figure 2). Hence, larger sizes of recombination fragments tend to make ecological populations appear less distinct in neutral characters (A, Figure 5). The effect of $h$ on the distinctness of populations is most important at low between-population recombination rates (Figure 5).

The effect of $h$ on population distinctness (quantified as $E[\tau_s]/E[\tau_{cd}]$) is shown explicitly in Figure 6. Under very low rates of between-population recombination, the distinctness ratio approaches $1/h$ for large fragment sizes (i.e., $h > 10\%$; Figure 6). Thus, global periodic selection alone (i.e., with little recombination between populations) cannot reduce the distinctness ratio of populations to 1 (so that $E[\tau_s]/E[\tau_{cd}] = E[\tau_{cd}]$ unless the recombination fragment size reaches 100$\%$ of the genome.

We explored in more detail the effect of $h$ on population distinctness for the case when within-population divergence levels are 1$\%$ (i.e., $E[\tau_{cd}] = 0.01$), because this is the divergence level frequently observed within bacterial sequence-similarity clusters (Palys et al. 1997).
Figure 6.—The distinctness of populations as a function of the size of recombination fragment \( h \). The distinctness of populations is quantified as the ratio of between-within population divergence \( E(p_d) / E(p_s) \). Note that for large fragment sizes \( (h > 10\%) \), the distinctness ratio approaches \( 1/h \). All selective sweeps are global \( (\mu_l = 0, \mu_g = 2 \times 10^{-12}) \); other parameters are as in Figure 3. The rate of between-population recombination is set to \( c_s = c_b = 10^{-12} \), a value too low to produce a homogenizing effect on the metapopulation, while sufficiently high to allow globally adaptive mutations to transfer across populations. Divergence levels within and between populations are based on the solution of Equation 7; \( P_{gs} \) and \( P_{gd} \) are calculated from \( N, c_s, q, \) and \( z \), using Equations 2 and A4.

With this level of divergence, global periodic selection is quite ineffective in reducing divergence between populations when recombination fragments are small (Figure 7). For example, global periodic selection with a recombination fragment of 1% of the genome cannot reduce the between-population divergence by \( >4\% \); however, with larger recombination fragments \( (e.g., h = 5\%) \), global periodic selection may significantly reduce the between-population divergence from unbounded neutral divergence \( (\pi_d = 0.25) \) to a much more limited level of divergence \( (\pi_d = 0.09; \) Figure 7).

**DISCUSSION**

This study presents a coalescence model for investigating the effect of globally adaptive mutations on neutral sequence divergence in bacteria. We used this model to test whether interpopulation transfer of globally adaptive mutations might prevent neutral sequence divergence between ecologically distinct populations of bacteria.

**Assumptions of the model:** If globally adaptive mutations are to reduce divergence between ecological populations at every locus in the genome, we must assume that every gene locus has the opportunity to hitchhike from population to population along with globally adaptive mutations (Figure 1). We therefore assume that globally adaptive mutations that confer benefits in more than one population exist, that they are numerous, and that they appear throughout the genome. The latter two assumptions are required because only a limited fraction of the genome can be cotransferred (and subsequently homogenized) across populations with any given adaptive mutation: the segments transferred in bacterial recombination are generally small (Smith 1988), and the transfer of large segments across populations is probably disfavored by natural selection (Cohan 1994b; Zawadzki and Cohan 1995).

Consider next the central premise of the model, that globally adaptive mutations exist and are numerous. The likelihood of globally adaptive mutations must de-
pend on the degree of ecological divergence between populations. In the early stages of population divergence, a mutation that is adaptive in one population is likely to be adaptive in others. As the populations become progressively more finely tuned to their respective niches, accumulating many niche-specific adaptations, we should see fewer adaptive mutations that can benefit more than one population. We therefore expect globally adaptive mutations to prevent neutral sequence divergence genome-wide only between the most closely related populations.

Does a typical adaptive mutation confer a benefit in more than one population? Recently, Guttman and Dykhuizen (1994b) provided evidence that one adaptive mutation precipitated selective sweeps in all the ecological populations included within Escherichia coli. A selective sweep apparently purged sequence diversity within a small chromosomal region from all the various sequence clusters of E. coli, while these clusters retained their distinctness for all other chromosomal regions studied. This is exactly the pattern expected soon after a global selective sweep. As shown in Figure 2, for genes that are closely linked to the adaptive mutation \( q = 1 \), there is nearly total purging of diversity both within and between populations; for genes that are not linked to the adaptive mutation \( q = 0 \), there is purging of diversity within populations but none between populations. Provided that each of the E. coli sequence clusters is actually a separate ecological population (Cohan 1994a,b, 1999; Palys et al. 1997), the selective sweep demonstrated by Guttman and Dykhuizen (1994b) appears to have been driven by a globally adaptive mutation.

Globally adaptive mutations as a homogenizing force in neutral sequence evolution: Analysis of our model has shown that, in general, globally adaptive mutations tend to make populations less distinct. Especially under extremely low recombination rates, globally adaptive mutations severely depress neutral sequence divergence between populations while having only a minor effect on within-population diversity (Figure 3). Populations that would diverge without bound in the absence of global periodic selection may be prevented from diverging without bound in the presence of global periodic selection.

Nevertheless, global periodic selection does not homogenize neutral sequence divergence to the extent that populations become indistinguishable. Consider, for example, ecological populations whose average within-population sequence divergence is \( \sim 1\% \), a value typical for sequence-similarity clusters in bacteria (Palys et al. 1997). In the absence of global periodic selection, such populations diverge into separate sequence-similarity clusters whenever the between-population recombination rate is \( \sim 10^{-6} \); in the presence of global periodic selection, even for a large recombination fragment \( h = 10\% \), the critical recombination rate decreases only slightly to \( 10^{-7.8} \) (Figure 4). Recombination rates between most bacterial populations are unlikely to exceed either critical value (Whittem and Ake 1993; Roberts and Cohan 1995; Palys et al. 1997; Cohan 1999). We therefore conclude that in spite of the homogenizing effect of global periodic selection, ecological populations should diverge into separate sequence-similarity clusters.

Analysis of the model has shown that the effect of global adaptations on between-population divergence is highly dependent on the size of the fragment recombined (Figure 7). If the recombination fragment is small (\(< 1\% \) of the genome), global periodic selection is virtually ineffective in reducing between-population divergence; however, if the recombining fragment is large (e.g., 5\% of the genome), global periodic selection may significantly reduce the between-population divergence (Figure 7).

The effect of global periodic selection on sequence divergence may therefore depend on the mode of genetic transfer between populations, because the various modes of transfer differ greatly in the length of DNA recombined. In naturally competent taxa, such as Streptococcus and Bacillus, transformation may be the predominant mode of DNA exchange. The average fragment of DNA incorporated in both Streptococcus and Bacillus transformation is \(< 1\% \) of the genome (Humbert et al. 1995; Zawadzki and Cohan 1995). To the extent that transformation is the primary mode of transferring adaptive mutations across populations in these taxa, global periodic selection should have virtually no effect on sequence divergence (Figure 7).

Other modes of recombination, such as transduction and conjugation, can transfer much larger segments of DNA. A generalized transducing phage can, in principle, transfer segments as large as the phage’s own genome, which could be \( \sim 10\% \) of the bacterium’s genome (Fraenkel-Conrat 1985; Arber 1994). Conjugating plasmids can transfer even larger segments: in the case of the Hfr plasmid of E. coli, most of the genome can be transferred (but there may be additional fitness constraints on the size of large transferred fragments; see above). Therefore, when transduction and conjugation are the principal means of transfer of adaptive mutations, global periodic selection can have an important role in reducing divergence between populations.

In summary, global periodic selection can limit the sequence divergence between ecological populations. The effect of global periodic selection is most pronounced for groups of populations with low between-population recombination, such that global periodic selection is the only constraint on divergence between populations. Global periodic selection is unlikely to prevent the divergence of ecological populations into separate sequence clusters. A quantitative prediction of the homogenizing effect of global periodic selection would require more information about the rate of mutations that confer adaptations in multiple populations, infor-
mation about how evenly globally adaptive mutations are distributed throughout the genome, and information about the size of fragments that can be transferred between populations and then successfully accommodated by the receiving population.

We thank Michael Feldgarden for suggesting that we explore a model of global periodic selection and Richard Hudson for suggesting important improvements to the model. This work was supported by Environmental Protection Agency grants R82-1388-010 and R82-5348-010 and by research funds from Wesleyan University.

LITERATURE CITED


APPENDIX A: Probability That a Periodic Selection Event Leads to Coalescence

We consider the special case of a metapopulation consisting of two ecological populations. The adaptive mutation driving the periodic selection event begins in population 1 and is subsequently passed into population 2 by recombination. We use a two-locus, four-allele model. A is the locus under selection, while B is the segment of interest whose neutral sequence divergence we are investigating. Alleles in population 1 are designated by subscript 1; those in population 2 are designated by subscript 2. Within population 1, the advantageous allele is designated as A1, and all other alleles at the selected locus are designated a1. A2 designates all the other alleles at the selected locus in this population. An allele
at locus B can be attached to any of the four A alleles, i.e., A₁, a₁, A₂, a₂. The frequencies of the alleles A₁ and A₂ in their respective populations are x₁ and x₂.

Let gᵢ(Y,t) be the conditional probability that if a randomly selected gene B from generation t of the meta-population is attached to the allelic type Y (at locus A), its ancestor in generation (t − 1) was attached to allelic type X. [gᵢ(Y,t) is equivalent to the quantity fᵢ(Y,t)/f(Y,t) of Hudson and Kaplan (1988).] Following Hudson and Kaplan (1988), and keeping only the highest order terms in N (where N is the population size), 8 of the 16 gᵢ probabilities are

\[
g₁(A₁,t) = 1 + O\left(\frac{1}{N}\right)
\]

\[
g₁(a₁,t) = \frac{x₁(t-1)R₁₁}{N} + O\left(\frac{1}{N^{2}}\right)
\]

\[
g₁(A₂,t) = \frac{x₁(t-1)R₁₂}{2N} + \frac{x₂(t-1)R₂₁}{N} + O\left(\frac{1}{N^{2}}\right)
\]

\[
g₁(a₂,t) = \frac{x₁(t-1)R₁₂}{2N} + O\left(\frac{1}{N^{2}}\right)
\]

\[
g₁(a₁,t) = 1 + O\left(\frac{1}{N}\right)
\]

\[
g₁(A₁,t) = \frac{(1-x₁(t-1))R₁₁}{N} + O\left(\frac{1}{N^{2}}\right)
\]

\[
g₁(A₂,t) = \frac{(1-x₁(t-1))R₁₂}{2N} + O\left(\frac{1}{N^{2}}\right)
\]

\[
g₁(a₂,t) = \frac{(1-x₁(t-1))R₁₁}{2N} + \frac{(1-x₁(t-1))R₂₁}{N} + O\left(\frac{1}{N^{2}}\right)
\]

where \(R₁₁ = R₂₂ = 2Nc₁(1-q)\), the per-population rate at which the two loci are separated by recombination within a population; \(R₁₂ = R₂₁ = 2Nc₂(1-q)\), the per-population rate at which the two loci are separated by recombination between populations; and \(R₁ = R₂ = Nc₃q\), the per-population rate at which a DNA segment containing both loci is transferred between populations. The remaining 8 gᵢ values may be obtained by substituting 2 for 1 and 1 for 2 in all the indices of the above equations, e.g.,

\[
gᵢ(a₁,t) = \frac{x₂(t-1)R₂₂}{N} + O\left(\frac{1}{N^{2}}\right).
\]

We now define the Q process. Suppose that m B genes are selected at random at the end of the selective sweep (time t = 0). Let \(Q(0) = (i, j, k, l)\), where i, j, k, l represent the number of B genes attached to A₁, a₁, A₂, a₂, respectively. Going back in time, Q(t) describes the number of ancestral B genes attached to each A allele at time t (i.e., t generations before time 0). The total number of ancestral B genes in generation t is denoted by \(|Q(t)|\). Note that \(|Q(t)|\) never increases, because the number of ancestral alleles can only stay constant or decrease (if two or more of the sampled alleles had a common ancestor in the previous generation). We are interested in the cases where Q(t) changes states, i.e., \(Q(t-1) \neq Q(t)\). There are two possible cases.

**Case 1.** \(|Q(t-1)| = |Q(t)|\): The only possible state changes allowed by this condition result from recombination between parental genes. Given \(Q(t) = (i, j, k, l)\), there are 12 possible states of \(Q(t-1)\):

\[
\{(i + 1, j - 1, k, l); (i + 1, j, k - 1, l); (i + 1, j, k, l - 1); (i, j + 1, k, l - 1); (i, j + 1, k - 1, l); (i, j, k + 1, l); (i, j, k, l + 1); (i, j, k, l + 1); (i - 1, j, k, l + 1); (i - 1, j, k - 1, l + 1)\}.
\]

Note that all other jumps would require more than a single recombination event and their probabilities are therefore of the order \(1/N^2\) and are negligible. We are interested in the probabilities that the process jumps from \((i, j, k, l)\), to any of the above states, e.g.,

\[
P[Q(t-1) = (i + 1, j - 1, k, l) | Q(t) = (i, j, k, l)] = j \cdot gᵢ(a₁,t).
\]

The probability that a selected gene B from generation t is attached to a₁ while its ancestor was attached to A₁ is given by \(gᵢ(a₁,t)\). Because we are sampling j a₁ alleles, \(P[Q(t-1) = (i + 1, j - 1, k, l) | Q(t) = (i, j, k, l)]\) is given by the probability of coalescence by drift, i.e.,

\[
\frac{1}{N xᵢ(t-1)} + O\left(\frac{1}{N^{2}}\right).
\]

where \(xᵢ(t-1)\) is the frequency of the allelic type aᵢ in the parental generation. Thus, the probability that two B alleles attached to A₁ at t have a common ancestor at \(t-1\) is

\[
\frac{1}{N xᵢ(t-1)} + O\left(\frac{1}{N^{2}}\right).
\]

Because i B genes attached to A₁ are being sampled,

\[
P[Q(t-1) = (i - 1, j, k, l) | Q(t) = (i, j, k, l)]
\]
where if \( i < 2 \), \( \frac{1}{2} \) is interpreted as 0. Similarly,

\[
P[Q(t - 1) = (i, j, k, l) | Q(t) = (i, j, k, l)] = \left(\frac{1}{2}\right) \left[ \frac{1}{N} \right]^{l-1} + O\left(\frac{1}{N^2}\right).
\]

and because the chance of more than one coalescence event per generation is of the order of \( 1/N^2 \), jumps of greater than 1 state are ignored. We have now defined the probabilities of every possible state change of \( Q(t) \). The probability that the \( Q \) process does not change state can thus be written as

\[
P(Q(t - 1) = Q(t) | Q(t) = (i, j, k, l)) = 1 - h_{ijt}(x_i(t - 1), x_j(t - 1)) + O\left(\frac{1}{N^2}\right).
\]

where \( h_{ijt} \) is the total probability of the \( Q \) process changing states:

\[
h_{ijt}(x_i(t - 1), x_j(t - 1)) = \left(\frac{1}{2}\right) \left[ \frac{1}{N} \right]^{l-1} + \left(\frac{1}{2}\right) \left[ \frac{1}{N} \right]^{k-1} + \left(\frac{1}{2}\right) \left[ \frac{1}{N} \right]^{i-1} + \left(\frac{1}{2}\right) \left[ \frac{1}{N} \right]^{j-1} + i \left( g_i(x_i,t) + g_i(x_i,t) + g_i(x_i,t) \right) + j \left( g_j(x_j,t) + g_j(x_j,t) + g_j(x_j,t) \right) + k \left( g_k(x_k,t) + g_k(x_k,t) + g_k(x_k,t) \right) + l \left( g_l(x_l,t) + g_l(x_l,t) + g_l(x_l,t) \right).
\]

**Calculation of \( p_{gsr} \) and \( p_{gs} \):** At the end of the selective sweep we sample two \( B \) alleles from the metapopulation. We want to know the probability that the two alleles had a common ancestor during the selective sweep. We consider two cases:

Case 1: the two genes are sampled from different ecological populations. The probability of coalescence during the sweep is \( p_{gsr} \).

Case 2: the two genes are sampled from the same ecological population. The probability of coalescence during the sweep is \( p_{gs} \).

We begin by considering case 1, calculation of \( p_{gsr} \). We follow the \( Q \) process back in time, going from \( \tau_f \) (end of the selective sweep) to \( \tau_0 \) (beginning of selective sweep; Figure B). We can write \((1 - p_{gsr})\) as the probability of escaping coalescence, i.e., leaving the selective sweep \((t = \tau_f)\) with one \( B \) gene in population 1 (attached to \( A_1 \)) and one \( B \) gene in population 2 (attached to \( A_2 \)) and entering it \((t = \tau_0)\) with two ancestral \( B \) genes:

\[
1 - p_{gsr} = P[Q(\tau_0) = (i, j, k, l) | Q(\tau_f) = (1,0,1,0)]
\]

\[
(i + j + k + l) = 2. \tag{A1}
\]

We also need to define \( P_{ijkl}(t) \), the probability of finding the \( Q \) process in the state \((i, j, k, l)\) at time \((t)\), given that at the end of the selective sweep \( Q(\tau_f) = (1,0,1,0) \).

\[
P_{ijkl}(t) = P[Q(t) = (i, j, k, l) | Q(\tau_f) = (1,0,1,0)].
\]

Hence, we can rewrite Equation A1 as

\[
1 - p_{gsr} = \sum_{i+j+k+l=2} P_{ijkl}(\tau_0). \tag{A2}
\]

To calculate all the relevant \( P_{ijkl}(\tau_0) \) values, we use the differential equations governing their behavior:

\[
\frac{dP_{1100}}{dt} = -h_{1100}P_{1100} + 2P_{1100} \cdot g_i(A_1,t) + 2P_{0200} \cdot g_i(a_1,t) + P_{1010} \cdot g_i(A_2,t) + P_{0101} \cdot g_i(a_2,t)
\]

\[
\frac{dP_{1000}}{dt} = -h_{1000}P_{1000} + 2P_{1000} \cdot g_i(A_1,t) + 2P_{0200} \cdot g_i(A_1,t) + P_{1100} \cdot g_i(A_2,t) + P_{0101} \cdot g_i(a_2,t)
\]

\[
\frac{dP_{0100}}{dt} = -h_{0100}P_{0100} + 2P_{0100} \cdot g_i(A_1,t) + 2P_{0020} \cdot g_i(a_2,t) + P_{1100} \cdot g_i(A_2,t) + P_{0101} \cdot g_i(a_2,t)
\]

\[
\frac{dP_{0101}}{dt} = -h_{0101}P_{0101} + 2P_{0101} \cdot g_i(A_1,t) + 2P_{0020} \cdot g_i(a_2,t) + P_{1100} \cdot g_i(A_2,t) + P_{0101} \cdot g_i(a_2,t)
\]

\[
\frac{dP_{0010}}{dt} = -h_{0010}P_{0010} + 2P_{0010} \cdot g_i(a_1,t) + 2P_{0002} \cdot g_i(a_2,t) + P_{1100} \cdot g_i(A_2,t) + P_{0011} \cdot g_i(a_2,t)
\]

\[
\frac{dP_{0011}}{dt} = -h_{0011}P_{0011} + 2P_{0011} \cdot g_i(a_1,t) + 2P_{0002} \cdot g_i(a_2,t) + P_{1100} \cdot g_i(A_2,t) + P_{0011} \cdot g_i(a_2,t)
\]
We also need the equations describing changes in the frequencies of the adaptive alleles $A_1$ and $A_2$, $x_1(t)$ and $x_2(t)$:

$$\frac{dx_1}{dt} = -zx_1(1 - x_1)$$

$$\frac{dx_2}{dt} = -zx_2(1 - x_2).$$

To be able to treat the model deterministically, we follow Kaplan et al. (1989) and limit $t$ to

$$\tau_1(1 - \varepsilon) \leq t \leq \tau_1(\varepsilon),$$

where $\tau(\varepsilon)$ corresponds to the time in the early phase of the selective sweep, where $x_1 = \varepsilon$, and $\tau_1(1 - \varepsilon)$ corresponds to a time near the end of the selective sweep, where $x_2 = 1 - \varepsilon$. We take $\varepsilon = 5/Nz$.

It remains for us to establish the boundary conditions

$$P_{1010}(\tau_1(1 - \varepsilon)) = 1,$$

while all other $P_{ij}(\tau_1(1 - \varepsilon))$ are zero. Also,

$$x_1(\tau_2(1 - \varepsilon)) = 1 - \varepsilon.$$

Because the adaptive mutation enters population 2 later than it enters population 1, we must determine the value of $x_1$ at the time when $x_2 = 1 - \varepsilon$. For that purpose we need to first run the frequency equations forward in time (starting at $x_1 = \varepsilon$, $x_2 = 0$). We then allow for the transfer of a single adaptive mutation to
population 2. This takes place at time $E(\tau)$, which is the expectation of the transfer of the first adaptive allele that will become fixed in population 2. (This is in fact equal to the time at which 1/2z alleles have crossed over.) We run the equations forward in time until $\tau_i(1-\varepsilon)$ to establish the end boundary conditions for $x_i$; then we can run both $x_i$ and $x_i$ backward, along with the equations for $P_{ijkl}$ following Equation A3 (see Figure 8).

The transfer of the first adaptive allele from population 1 to population 2 is a stochastic event and hence introduces a discontinuity in the solution to differential equations for $P_{ijkl}$. At the time immediately preceding the initial transfer of the adaptive allele into population 2 (time $\tau_i+$), no $A_i$ alleles existed. Therefore, at time $\tau_i+$ the probabilities of a $B$ allele being attached to an $A_i$ allele ($P_{ijkl}(\tau_i+)$ for $k \neq 0$) must be zero. However, these $P$ terms might have nonzero values at $\tau_i(\varepsilon)$. Following Kaplan et al. (1989), we may neglect the contribution of $P_{ijkl}(\tau_i(\varepsilon))$ to $P_{ijkl}$. However, a nonzero value of $P_{ijkl}(\tau_i(\varepsilon))$ implies that immediately after the transfer event, one of the sampled alleles is attached to an $A_i$ allele. In this case, there exist two possible states of the $Q$ process immediately preceding the transfer: $Q(\tau_i+) = (i+1,j,0,1)$ if the transfer was a core combination of $A$ and $B$ (with a probability $q_1$); or $Q(\tau_i+) = (i,j,0,1)$ if only the $A$ allele was transferred (with a probability $q-1$). Hence, the additional boundary conditions at $\tau_i+$ needed to account for the transfer are

$$
P_{ijkl}(\tau = \tau_i) = P_{ijkl}(\tau = \tau_i) = 0$$

Using the above boundary conditions, we can obtain the values $P_{ijkl}(\tau_i(\varepsilon))$ at the beginning of the selective sweep. Then, using the approximation $P_{ijkl}(\tau_i(\varepsilon)) = P_{ijkl}(\tau_i)$, Equation A2 becomes

$$1 - P_{gs} = P_{ijkl}(\tau_i(\varepsilon)) + P_{ijkl}(\tau_i(\varepsilon))$$

$$+ P_{ijkl}(\tau_i(\varepsilon)) + P_{ijkl}(\tau_i(\varepsilon)) + P_{ijkl}(\tau_i(\varepsilon)).$$

The above analysis also applies to case 2, the calculation of $P_{ijkl}$. We only need to alter the initial conditions, i.e., allowing $Q(\tau) = (2,0,0,0)$ (if we are sampling two alleles from the original population where the adaptive mutation first occurred) or $Q(\tau) = (0,0,2,0)$ (if we are sampling two alleles from the population to which the adaptive mutation has been transferred). Because the probabilities of choosing either population are equal, we calculate $P_{ijkl}$ as the average of the two values.

**APPENDIX B: The Expected Time to Coalescence**

**Coalescence within populations**

Following Cohan (1994a), we first consider the time to coalescence $t_s$ for two segments currently residing in cells of the same ecological population. The statistical properties of $t_s$ are investigated by dividing $t_s$ into two constituent quantities: the time $t_r$ necessary to go back into the past of the two lineages to reach a “key event” (defined below) and the additional time necessary to go beyond the key event to reach a coalescence of the two lineages into a common ancestor (if the key event did not result in coalescence). A key event is any event that changes the expected time to coalescence. In the case of segments from the same ecological population, a key event may be any of the following: coalescence by drift acting on $N$ cells of the population, a local selective sweep, a global selective sweep, or a genetic exchange event in which one of the segments is transferred into its present population from another ecological population.

The variable $t_s$ is thus defined below, where the first term $t_r$ represents the time to reach the most recent key event, and the other three terms represent the additional time necessary to go beyond the key event to reach a coalescence (Table 2):

$$t_s = k_s + x_{gs}(x_{gs}^{d} + x_{gs}^{d})$$

$$+ x_{gs}(x_{gs}^{d} + x_{gs}^{d}) + x_{gs}^{d}.$$
variable $\chi_{at}$ indicates whether the key event was a local selective sweep ($\chi_{at} = 1$ if the key event was a local selective sweep; $\chi_{at} = 0$ else). Two lineages that are in the same population at the end of the selective sweep may begin the sweep in one of three states: a single lineage (i.e., the lineages coalesce); two lineages in the same population; or they may begin as two lineages in different populations. The random variable $\chi_{at}$ indicates whether the lineages escape coalescence and begin in the same population (1 if yes, 0 else, as above), and $t'_c$ is the additional time to coalescence in this case. The random variable $\chi_{at}$ indicates whether the lineages begin the selective sweep in different populations (1 if yes, 0 else), and $t'_c$ is the additional time to coalescence in this case. Accordingly, the sum ($\chi_{at}t'_c + \chi_{at}t'_c$) represents the additional time to coalescence when the key event is a local sweep.

**Global periodic selection as the key event:** The random variable $\chi_{gs}$ indicates whether the key event was a global selective sweep (1 if yes, 0 else). The random variable $\chi_{gs}$ indicates whether the lineages escape coalescence and begin in the same population (1 if yes, 0 else, as above), and $t'_c$ is the additional time to coalescence in this case. The random variable $\chi_{gs}$ indicates whether the lineages begin the selective sweep in different populations (1 if yes, 0 else), and $t'_c$ is the additional time to coalescence in this case. Accordingly, the sum ($\chi_{gs}t'_c + \chi_{gs}t'_c$) represents the additional time to coalescence when the key event is a global selective sweep.

**Recombination between populations as the key event:** The random variable $\chi_{cd}$ indicates whether the key event is a recombination event between populations, in which one of the lineages enters the current population (1 if yes, 0 else). The random variable $t'_c$ indicates the additional time to coalescence beyond this key event.

**Coalescence between populations**

Next consider the time to coalescence, $t_c$, for segments that are now in two cells belonging to different ecological populations. As above, we divide $t_c$ into the time $k_c$, to go back to the most recent key event (a global selective sweep or a genetic exchange event in which one of the lineages enters its current population from the population of the other lineage) and the additional time required to go beyond the key event to reach coalescence:

$$t_c = k_c + \chi_{gs}(t'_c + t'_c) + \chi_{cd}t'_c.$$

**Global selective sweep as the key event:** The random variable $\chi_{gs}$ indicates whether the two lineages presently in different populations escape coalescence and begin the sweep in different populations, and $t'_c$ is a random variable for the additional time to coalescence in this case. The random variable $\chi_{gs}$ indicates whether the lineages escape coalescence and begin the sweep in the same population, and $t'_c$ represents the additional time to coalescence in this case. The sum ($\chi_{gs}t'_c + \chi_{gs}t'_c$) represents the additional time to coalescence when the key event is a global selective sweep.

**Between-population recombination as the key event:** The random variable $\chi_{cd}$ indicates whether the key event was a recombination event between populations, such that before the event the lineages were in the same population and afterward were in different populations (1 if yes, 0 else). The random variable $t'_c$ represents the additional time to coalescence in this case.

**Expected values of $t_c$ and $t_c$**

The expected values of all indicator variables were calculated following Cohan (1994a), except when the key event is a selective sweep. The expectation of $\chi_{gs}$ is the probability that two segments from the same population escape coalescence and begin the sweep in the same population (1 – $P_I$) and likewise for the expected values of other indicator variables. $E(\chi_{gs}) = 1 - P_{gs}$, $E(\chi_{gs}) = 1 - P_{gs}$, $E(\chi_{gs}) = 1 - P_{gs}$, $E(\chi_{gs}) = 1 - P_{gs}$. All the relevant expected values are calculated in appendix A.

The expected values of $t_c$ and $t_c$ are as follows:

$$E(t_c) = [1 + \alpha_p(1 - P_{gs})E(t_c) + P_{gs}E(t_c)] + \sigma_p(1 - P_{gs} - P_{gs})E(t_c) + P_{gs}E(t_c) + \sigma_pE(t_c).$$

It can be shown that if the selective sweeps are rapid and infrequent (i.e., duration of a sweep is much greater than the interval between sweeps) the above equations reduce to

$$E(t_c) = \frac{1 + \alpha_p(1 - P_{gs})E(t_c) + \sigma_p(1 - P_{gs})E(t_c)}{1 + \alpha_p + \sigma_p + 2\alpha_p}$$

$$E(t_c) = \frac{1 + \alpha_p(1 - P_{gs})E(t_c) + \sigma_p(1 - P_{gs})E(t_c)}{1 + \alpha_p + \sigma_p + 2\alpha_p}.$$

**Probability density functions of $t_c$ and $t_c$**

We define $P(t_c)$ as the probability that the time $t_c$ to coalescence for two segments currently in the same population is equal to $t$. Similarly, $P_{gs}(t)$ is the probability that the time $t_c$ to coalescence for two segments currently in different populations is equal to $t$. The values $P(t_c)$ and $P_{gs}(t)$ can then be expressed as the probability that the most recent key event occurred at exactly time $t$ and led to coalescence or, if the key event occurred at any other time $t < t$ and did not lead to coalescence, the additional time necessary to reach coalescence was $t - t$. We can consider the key events to be exponentially distributed (Hudson and Kaplan 1988). Hence,

$$P(t_c) = \left\{ \begin{array}{ll}
\frac{1}{\lambda} & \text{if } t < \lambda \\
0 & \text{otherwise}
\end{array} \right.$$

$$P(t_c) = \left\{ \begin{array}{ll}
\frac{1}{\lambda} & \text{if } t < \lambda \\
0 & \text{otherwise}
\end{array} \right.$$

$$P_{gs}(t_c) = \left\{ \begin{array}{ll}
\frac{1}{\lambda} & \text{if } t < \lambda \\
0 & \text{otherwise}
\end{array} \right.$$
\[
\begin{align*}
P_{ts}(t) &= \sigma_g \sigma_{gd} e^{-(\sigma_g + 2\sigma_d)t} \\
&+ \int_0^t e^{-(1/N + \sigma_g + 2\sigma_d)} P_{ts}(t - \tau) d\tau \\
&+ 2\sigma_g \int_0^t e^{-(1/N + \sigma_g + 2\sigma_d)} P_{sd}(t - \tau) d\tau \\
&+ \sigma_g (1 - P_{sd}) \int_0^t e^{-(1/N + 2\sigma_d)} P_{sd}(t - \tau) d\tau \\
&+ 2\sigma_g \int_0^t e^{-(1/N + 2\sigma_d)} P_{sd}(t - \tau) d\tau.
\end{align*}
\]

The above recursions were solved numerically to yield numerical representations of the two probability distribution functions.