

WILEY



Neutral Genetic Diversity in a Metapopulation with Recurrent Local Extinction and Recolonization

Author(s): John R. Pannell and Brian Charlesworth

Source: *Evolution*, Vol. 53, No. 3 (Jun., 1999), pp. 664-676

Published by: Society for the Study of Evolution

Stable URL: <http://www.jstor.org/stable/2640708>

Accessed: 05-09-2016 07:47 UTC

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://about.jstor.org/terms>



Wiley, Society for the Study of Evolution are collaborating with JSTOR to digitize, preserve and extend access to *Evolution*

NEUTRAL GENETIC DIVERSITY IN A METAPOPULATION WITH RECURRENT LOCAL EXTINCTION AND RECOLONIZATION

JOHN R. PANNELL¹ AND BRIAN CHARLESWORTH

Institute for Cell, Animal and Population Biology, University of Edinburgh, Edinburgh EH9 3JT, United Kingdom

Abstract.—Many species exist as metapopulations in balance between local population extinction and recolonization, processes that may strongly affect the distribution of neutral genetic diversity within demes and in the metapopulation as a whole. In this paper we use both the infinite-alleles and the infinite-sites models to reframe Slatkin's propagule-pool and migrant-pool models in terms of mean within-deme and among-deme genetic diversity; the infinite-sites model is particularly relevant to DNA sequence data. Population turnover causes a major reduction in neutral genetic diversity within demes, π_S , and in the metapopulation as a whole, π_T . This effect is particularly strong for propagule-pool colonization, in which colonists are drawn from a single extant deme. Because metapopulation dynamics affect both within-deme and total metapopulation diversity similarly, comparisons between species with different ecologies on the basis of ratios such as F_{ST} are difficult to interpret and absolute measures of divergence between populations should be used as well. Although the value of F_{ST} in a metapopulation with local extinction depends strongly on the mode of colonization, this has almost no effect on the numerator of the F_{ST} ratio, $\pi_T - \pi_S$, so that F_{ST} is influenced mainly by the effect of the colonization mode on the denominator (π_T). Our results also indicate that it is inappropriate to use measures of average within-deme diversity in species with population turnover to estimate the scaled mutation rate, θ , because extinction can greatly reduce π_S . Finally, we discuss the effect of population turnover on the effective size of a metapopulation.

Key words.—Geographic invariance, infinite-alleles model, infinite-sites model, population differentiation, population structure.

Received October 6, 1998. Accepted January 8, 1999.

Many species consist of geographically isolated populations, or demes, linked by gene flow through migration. In recent years, a good deal of attention has been directed toward theoretical predictions of the effect of such population structure on neutral genetic diversity, and these models have been used to infer migration rates from data on allele frequencies or variation at nucleotide sites (Slatkin 1985; Slatkin and Barton 1989; Bossart and Pashley Prowell 1998). Most genetic models of subdivided populations assume that demes persist at a constant size over time. However, local populations in some species may be ephemeral, resulting in the persistence of a metapopulation in a dynamic balance between local extinctions and the recolonization of available habitat by individuals from extant demes (Harrison and Taylor 1997). In such a metapopulation, we would like to know how recurrent extinction will affect neutral genetic diversity within demes and in the metapopulation as a whole.

Slatkin (1977) provided the first theoretical study of the effects of recurrent local extinction and recolonization on genetic diversity. His model II extended Maruyama's (1970) finite-island/infinite-alleles model to derive equations for the effective number of alleles in a metapopulation with two contrasting modes of recolonization: a "propagule-pool" model, in which all colonists are derived from a single extant deme in the metapopulation, and a "migrant-pool" model, in which the colonists of a new deme are drawn randomly from the whole metapopulation in the same way as migrants are sampled. Using a similar approach, Maruyama and Kimura (1980) derived expressions for the within- and between-population diversity in a metapopulation with local extinction

and recolonization, following Slatkin's propagule-pool model.

Slatkin (1977) gave exact recurrence equations for the identity in allelic state under the infinite-alleles model (Kimura and Crow 1964) within and between demes and provided approximations for the effective number of alleles in the metapopulation for two cases: (1) when the extinction and migration rates are small and of the same order of magnitude; and (2) when the extinction rate exceeds the migration rate by an order of magnitude or more. He did not, however, give explicit formulae for the probabilities of identity within and between demes at equilibrium. Maruyama and Kimura (1980) presented approximations for these probabilities for a modified propagule-pool model of colonization with low extinction rates, but they did not examine the effects of migrant-pool colonization, nor did they model a metapopulation in which the rate of extinction is high. Several studies have examined modifications to Slatkin's (1977) models (Wade and McCauley 1988; Whitlock and McCauley 1990; Whitlock 1992a; Ingvarsson 1997), but these have all estimated genetic differentiation among demes using F_{ST} , and none gives a partitioning of expected diversity into between- and within-deme components.

Models of subdivided populations without local extinction have shown that the average nucleotide diversity within a population under the infinite-sites model is invariant with respect to the migration rate, as long as migration is conservative (i.e., it does not change the sizes of individual demes), the matrix describing the migration process is irreducible (i.e., genes can ultimately move between any two demes, directly or via other demes), and at least some individuals remain in each deme from one generation to the next (Slatkin 1987; Strobeck 1987). When these conditions are met, the expected equilibrium within-deme diversity un-

¹ Present address: Department of Plant Sciences, University of Oxford, Oxford OX1 3RB, United Kingdom; E-mail: john.pannell@plants.ox.ac.uk.

der the infinite-sites model is given by $\theta = 4N_Tu$, where N_T is the total breeding size of the metapopulation. A useful generalization of this result to nonconservative migration has recently been derived for the infinite-sites model by Nagylaki (1998b). This suggests that an appropriately weighted measure of within-population nucleotide diversity might provide an approximate estimator of θ under fairly general conditions. But recurrent population extinction and colonization clearly violate the conditions that are sufficient for the invariance principle to hold, and it is therefore important to determine the extent to which the average within-deme diversity deviates from θ in a metapopulation with population turnover.

Recently, Whitlock and Barton (1997) have investigated the effects of population turnover on the "eigenvalue effective size" of a metapopulation. This is defined as the size of an idealised Fisher-Wright population whose rate of increase in probability of identity by descent between a pair of genes is the same as its asymptotic rate of increase in a metapopulation with a specified structure and dynamics. They also defined the "mutation effective size," which determines the total genetic diversity that is maintained in a population at equilibrium (Whitlock and Barton 1997). The mutation effective size predicts the genetic diversity maintained under the infinite-sites model of mutation (Kimura 1971), according to the relation $\pi_T = 4N_eu$ (Tajima 1983), where π_T is the total nucleotide diversity maintained in the population (see below), u is the mutation rate per nucleotide site, and N_e is the mutation effective size of the population. This mutation effective size will in general differ from the eigenvalue effective size, although the two may be equivalent under special circumstances (Whitlock and Barton 1997; see also the Discussion).

To date, studies of the genetic effects of population turnover in a metapopulation have thus focussed either on genetic differentiation among demes or on the eigenvalue effective size of the metapopulation; little attention has been given to predicting the effects of extinction and colonization on within-deme diversity (but see Maruyama and Kimura 1980). In view of the increasing interest in studies of DNA sequence variability in groups such as flowering plants, which are expected often to have subdivided populations that are subject to extinction and recolonization (Husband and Barrett 1996; Valverde and Silvertown 1997; Barrett and Pannell 1998; Giles et al. 1998; Pannell and Barrett 1998), it is essential to develop theoretical models to predict patterns of within- and between-population variability.

This paper has three main aims. First, we reframe Slatkin's (1977) propagule-pool and migrant-pool models in terms of within- and among-deme diversities and investigate the effects of extinction and colonization on these quantities. We stress that the model is entirely the work of Slatkin (1977); our contribution is to present its results in terms that may be relevant to data on DNA sequence diversity sampled from within and among populations. Second, we derive approximations for within-deme and total metapopulation diversities for both the infinite-alleles and infinite-sites models; the latter are more appropriate for the interpretation of data on DNA sequence variation. We compare them with exact solutions to Slatkin's recurrence equations for a range of parameter values to indicate the extent to which our approximations are

robust. We also relate our approximations for the infinite-sites model to the corresponding expressions for expected pairwise coalescent times. Finally, we discuss our results in relation to the effective size of a metapopulation, the validity of the invariance principle, and the more recent extensions of Slatkin's (1977) metapopulation model that use F_{ST} . We argue that descriptions of population differentiation solely by means of relative measures such as F_{ST} sacrifice important information about the effects of metapopulation processes on genetic diversity within and among populations, and that, in general, no easy method of estimating θ exists if extinction rates are higher than migration rates.

METHODS

Assumptions of the Model

We describe here Slatkin's (1977) model II in full, reframing his expressions in terms of diversity measures rather than in terms of identity probabilities. Consider a metapopulation consisting of n occupied habitat patches or demes, each containing N monoecious diploid individuals that reproduce by random mating, with a random proportion of selfing, N^{-1} . We define h_0 and h_1 as the probabilities that two genes sampled from within a deme and from two different demes, respectively, are distinct alleles.

Slatkin (1977) assumed an infinite-alleles model of mutation (Kimura and Crow 1964), whereby each gene mutates to a new allele with probability u per generation. Alternatively, we can assume an infinite-sites model of mutation (Kimura 1971), such that u is the mutation rate per nucleotide base. In this case, we use π_S and π_B for within- and between-population diversity, respectively (Charlesworth 1998). π_S and π_B are then the respective probabilities that two genes, each drawn from the same or different demes, differ at a random nucleotide site. The infinite-sites model can be derived as the limiting case of the infinite-alleles model as u tends to zero, so that the same basic recursion relations apply to both models.

In our model, generations are discrete. At the beginning of each generation, a proportion e of extant demes goes extinct; if n demes are extant and the metapopulation is at equilibrium, ne new demes are reestablished. Thus, e measures the probability that a given deme was founded during the current generation. Note that it is not necessary to assume that the same habitat patches are immediately recolonized after extinction, as is often stated in models of this type; one requires only that the metapopulation is at a demographic steady state, with the total number of extant demes remaining constant from one generation to the next. Thus, there may be any number of unoccupied habitat patches in the metapopulation, as long as n of them are occupied after colonization has taken place.

Colonization and migration may be assumed to occur concurrently after extinction. For both processes, we assume that individuals can come only from populations that have not just gone extinct. New colonies are founded by k individuals, either originating all from a single source population (the propagule-pool model) or randomly sampled from the entire metapopulation (the migrant-pool model). Migration involves the replacement of a proportion m of the individuals

within each extant deme by immigrants drawn randomly from the rest of the metapopulation. Note that recolonization and migration are assumed to be in general quantitatively different because they involve the movement of different numbers of individuals into demes. In the propagule-pool model, the processes of migration and recolonization are also qualitatively different because the source pools from which colonists and immigrants are drawn differ. Wade and McCauley (1988) have discussed the biological implications of these differences.

After colonization and migration have taken place, random mating occurs in each deme. Demes with N individuals maintain that size, whereas demes that have just been recolonized by k individuals increase their size to N individuals in a single generation. The assumption of instant population growth can be generalized by letting $\varphi(N)$ be the probability that an extant population is of size N ; we then replace terms in $(2N)^{-1}$ by $\int \varphi(N)(2N)^{-1} dN$, the inverse of the harmonic mean population size; such a formulation would imply, somewhat unrealistically, that local extinction is independent of the population size. Departures from the assumptions of an idealized Wright-Fisher population can be included by replacing N by the inbreeding effective size of the local population.

Recursion Relations

We wish to estimate h_0 and h_1 immediately after the population growth phase of the metapopulation cycle. With these assumptions, and assuming that $(1 - u)^2 \approx (1 - 2u)$ and $n^{-1}(n - 1) \approx 1$ (i.e., assuming a low mutation rate and a large number of demes), the equations of Slatkin (1977) give:

$$h'_0 \approx 1 - (1 - 2u) \left\{ (1 - e) \left[a \left(\frac{1}{2N} + \left(1 - \frac{1}{2N} \right) (1 - h_0) \right) + (1 - a)(1 - h_1) \right] + eA \right\} \quad (1)$$

and

$$h'_1 \approx 1 - (1 - 2u) \left\{ (1 - e)^2 \left[b \left(\frac{1}{2N} + \left(1 - \frac{1}{2N} \right) (1 - h_0) \right) + (1 - b)(1 - h_1) \right] + 2e(1 - e)B_1 + e^2B_2 \right\}, \quad (2)$$

where the primes indicate diversities in the subsequent generation. Here, $(1 - a)$ and $(1 - b)$ represent the probability that the two genes sampled were not both in the same deme the previous generation; A is the probability that two genes sampled from the same newly founded population are identical; B_1 is the probability that two identical genes are sampled from different populations, one from a newly founded population and the other from an existing population; and B_2 is the probability that two identical genes are sampled from different newly founded populations. Expressions for a , b , A , B_1 , and B_2 are given in the Appendix.

Expressions for the probabilities of identity by descent, as opposed to identity in state, for genes sampled from within

and between populations can of course be retrieved from equations (1) and (2) by substituting F_0 and F_1 for $1 - h_0$ and $1 - h_1$, respectively, and dividing by $(1 - 2u)$. Thus,

$$F'_0 \approx (1 - e) \left[a \left(\frac{1}{2N} + \left(1 - \frac{1}{2N} \right) (1 - F_0) \right) + (1 - a)(1 - F_1) \right] + eA \quad (3)$$

and

$$F'_1 \approx (1 - e)^2 \left[b \left(\frac{1}{2N} + \left(1 - \frac{1}{2N} \right) (1 - F_0) \right) + (1 - b)(1 - F_1) \right] + 2e(1 - e)B_1 + e^2B_2. \quad (4)$$

Identity by descent can be interpreted from the perspective of the coalescent process as gene lineages are followed backward in time until they coalesce at their most recent common ancestor (Hudson 1990). Thus, the probability of identity by descent of two genes sampled from within the same deme is simply the probability that they coalesce one generation back in time. Similarly, two genes sampled randomly from different demes in the metapopulation will be identical by descent with a probability equal to the probability of their coalescence in the previous generation. Mean coalescent times for genes sampled from the same deme, T_0 , and from different demes, T_1 , at equilibrium can be found using the method of Nagylaki (1998b) by solving:

$$T_0 \approx (1 - e) \left[a \left(\frac{1}{2N} + \left(1 - \frac{1}{2N} \right) (1 + T_0) \right) + (1 - a)(1 + T_1) \right] + eA \quad (5)$$

and

$$T_1 \approx (1 - e)^2 \left[b \left(\frac{1}{2N} + \left(1 - \frac{1}{2N} \right) (1 + T_0) \right) + (1 - b)(1 + T_1) \right] + 2e(1 - e)B_1 + e^2B_2, \quad (6)$$

where parameters h_0 and h_1 in expressions for A , B_1 , and B_2 in the Appendix are substituted for $(-T_0)$ and $(-T_1)$, respectively.

The total allelic diversity in the metapopulation is given by

$$\bar{h} = \frac{h_0}{n} + \left(1 - \frac{1}{n} \right) h_1. \quad (7)$$

The expected coalescent time for two genes sampled randomly from the whole metapopulation at equilibrium, \bar{T} , is calculated similarly by weighting T_0 and T_1 by n^{-1} and $(1 - n^{-1})$, respectively, and the total nucleotide diversity in the metapopulation under the infinite-sites model, π_T , is calculated by weighting π_S and π_B in the same way. When n is large, \bar{h} , \bar{T} , and π_T are well approximated by h_1 , T_1 , and π_B , respectively. For conciseness, we thus give all approximations in terms of within-deme and total metapopulation measures only. Under the infinite-sites model, $\pi_{ij} = 2\bar{T}_{ij}u$, where π_{ij} and \bar{T}_{ij} are, respectively, the expected pairwise site diversity and time to coalescence of two genes sampled from demes i and j (Hudson 1990). Results on coalescent times and nucleotide diversities can thus be interchanged.

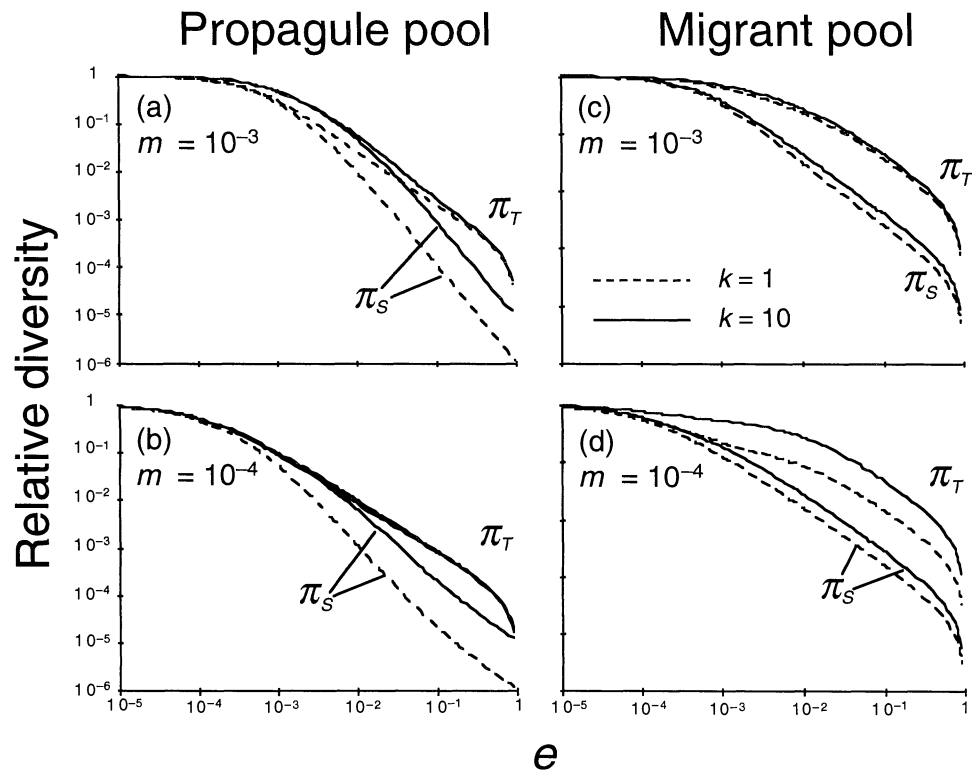


FIG. 1. Numerical calculations of the exact solution to equations (1) and (2) at equilibrium for selected parameter values, assuming an infinite-sites model of mutation. Diversities are plotted against the extinction rate and expressed relative to the case with no extinction. π_T and π_S are given for the propagule-pool (a and b) and the migrant-pool models (c and d). Panels (a) and (c): $m = 10^{-3}$; (b) and (d): $m = 10^{-4}$. Curves for $k = 1$ and $k = 10$ are plotted. Other parameter values: $u = 10^{-8}$, $N = 10^3$, and $n = 500$. Note that both axes are on a log scale.

RESULTS

Exact Solutions for Within-Deme and Total Diversities

The exact solutions to equations (1) and (2) are complicated. However, we have calculated the equilibrium solutions to these recursions numerically over a range of extinction rates, expressing the results relative to the corresponding cases with no extinction. Figure 1 shows some examples for the infinite-sites model. In both the propagule-pool and migrant-pool models, local population extinction always causes a substantial decrease in π_T . In the propagule-pool model, extinction invariably leads to a decrease in π_S , whereas in a metapopulation with migrant-pool colonization, intermediate values of e may cause an increase in π_S relative to the case of no extinction if k is much higher than Nm . This is because colonies that are founded by several individuals drawn from the metapopulation as a whole will tend to have diversities that approach the total metapopulation diversity. However, under propagule-pool colonization, founding individuals are drawn from a single extant deme and so the extinction-recolonization process does not contribute to the mixing of genes from different demes; such mixing occurs only through migration among extant demes (i.e., after colonization has taken place). These results can be understood in the light of approximate solutions to the recursion equations, which we now discuss.

Algebraic Approximations

Substantial algebraic manipulations are required to find approximate equilibrium solutions for within- and between-population diversities from equations (1) and (2). In all the approximations presented here, it is assumed that $u \ll 1$, $m \ll 1$, $n \gg 1$, and $k \ll N$. Recall that in a metapopulation with a large number of demes, h_1 and π_B approximate \bar{h} and π_T , respectively. Table 1 presents approximations for h_1 and \bar{h} assuming an infinite-alleles model of mutation under a propagule-pool and a migrant-pool model of colonization. Corresponding expressions for π_S and π_T are given in Table 2, assuming an infinite-sites model of mutation.

For the infinite-sites model (Table 2), we have expressed our results conventionally in terms of $\theta = 4N_Tu$, the scaled mutation rate for a panmictic population of total size $N_T = Nn$. It is assumed that $\theta \ll 1$. For both the infinite-alleles and the infinite-sites models, we follow Slatkin (1977) in giving expressions for two ranges of e : $e \leq m$ and $e \gg m$. In the former instance, it is assumed that $(1 - e)^2 \approx (1 - 2e)$; in the latter, terms of order me^{-1} can be dropped and higher-order terms in e are retained. Table 3 presents expressions for F_{ST} based on approximations for the infinite-sites model and following its definition by Hudson et al. (1992) as:

$$F_{ST} = \frac{\pi_T - \pi_S}{\pi_T}. \quad (8)$$

TABLE 1. Approximations for h_0 and \bar{h} for the propagule-pool and migrant-pool models. The equivalent expressions from Slatkin (1977) are also given by equating his approximation for n_e with $(1 - \bar{h})^{-1}$ and solving for \bar{h} .

	h_0	\bar{h}	Slatkin's equivalent \bar{h}
Propagule pool			
$e \leq m$	$\frac{4Nnum}{e + m + 4Nnum + (Nnue/k) + un}$	$\frac{4Nnum + (Nnue/k) + un}{e + m + 4Nnum + (Nnue/k) + un}$	$\frac{4Nnum + un}{e + m + 4Nnum + un}$
$e \gg m$	$\frac{4Nku[2mn - e(2 - e)(2mn - 1)]}{(2 - e)e^2N + 2nu(1 - e)(k + eN)}$	$\frac{2(1 - e)(k + eN)un}{(2 - e)e^2N + 2nu(1 - e)(k + eN)}$	$\frac{2(1 - e)un}{2(1 - e)un + 2e - e^2}$
Migrant pool			
$e \leq m$	$\frac{4Nnum + 2Nnue[1 - (1/2k)]}{e + m + 4Nnum + 2Nnue + un}$	$\frac{4Nnum + 2Nnue + un}{e + m + 4Nnum + 2Nnue + un}$	$\frac{4Nnum + 2Nnue + un}{e + m + 4Nnum + 2Nnue + un}$
$e \gg m^*$	$\frac{2eNnu[1 - (1/2k)]}{E[1 - e + (eN/k)] + 2eNnu + un}$	$\frac{2eNnu + un}{E[1 - e + (eN/k)] + 2eNnu + un}$	(see text)

* $E = [1 - (1 - e)^2]/2(1 - e)$.

This formulation is equivalent to Nei's (1973) G_{ST} (Charlesworth 1998).

Reduction in π_S and π_T and Breakdown of the Invariance Principle.

Models of subdivided populations without local extinction have shown that the average allelic and nucleotide diversity within populations is invariant with respect to the migration rate, given the assumptions identified in the introduction (Slatkin 1987; Strobeck 1987; Nagylaki 1998b). This principle of invariance no longer holds for a metapopulation with population turnover, whether colonization follows a propagule-pool or a migrant-pool model, unless $e \ll m$ (in which case π_S reduces to θ ; see Table 2). Increases in e reduce within-deme diversity below θ , particularly when extinction is followed by recolonization from a propagule pool. Thus,

for example, if $e \leq m$ under the propagule-pool model, $\pi_S \approx \theta m(e + m)^{-1}$.

In a metapopulation without extinction, π_S and π_T are proportional to the total number of individuals, and thus to both n and N . This is also true for a metapopulation with extinction under both the propagule-pool and the migrant-pool models when $e \leq m$ and when $e \gg m$ under the migrant-pool model. However, when $e \gg m$ in the propagule-pool model, π_S and π_T become independent of N , although π_T continues to depend on n . Note that if $m \ll n^{-1}$ and $e \gg m$, our approximation for π_S reduces to $4kue^{-1}$ under a propagule-pool model of colonization, a result that is independent of both N and n and that is a small fraction $k(eNn)^{-1}$ of what it would be in a metapopulation without local extinction. If $kN^{-1} \ll e \ll 1$, then π_T reduces to une^{-1} (Table 2).

These results can be understood intuitively by recognizing

TABLE 2. Approximations for π_S , π_T , and $\pi_T - \pi_S$ for the propagule-pool and migrant-pool models.

	π_S	π_T			$\pi_T - \pi_S$
		$e \ll kN^{-1}$	$e \approx o(N^{-1})$	$e \gg kN^{-1}$	
Propagule pool					
$e \leq m$	$\frac{\theta m}{(e + m)}$	$\frac{\theta}{(e + m)}\left(m + \frac{1}{4N}\right)$	$\frac{\theta}{(e + m)}\left(m + \frac{1}{4N} + \frac{e}{4k}\right)$	$\frac{\theta}{(e + m)}\left(m + \frac{e}{4k}\right)$	$\frac{\theta}{(e + m)}\left(\frac{1}{4N}\right)^{**}$
$e \gg m$	$\frac{4ku[2mn - e(2 - e)(2mn - 1)]}{(2 - e)e^2}$		$\frac{2(1 - e)(k + eN)un}{(2 - e)e^2N}$	$\frac{2(1 - e)un}{(2 - e)e}$	$\left[\frac{2(1 - e)un}{(2 - e)e}\right]^{***, \ddagger}$
Migrant pool					
$e \leq m$	$\frac{\theta}{(e + m)}\left(m + \frac{e}{2} - \frac{e}{4k}\right)$	$\frac{\theta}{(e + m)}\left(m + \frac{1}{4N}\right)$	$\frac{\theta}{(e + m)}\left(m + \frac{1}{4N} + \frac{e}{2}\right)$	$\frac{\theta}{(e + m)}\left(m + \frac{e}{2}\right)$	$\frac{\theta}{(e + m)}\left(\frac{1}{4N}\right)^{**}$
$e \gg m^*$	$\frac{\theta}{2E[1 - e + (eN/k)]}\left(e - \frac{e}{2k}\right)$		$\frac{\theta}{2E[1 - e + (eN/k)]}\left(e + \frac{1}{2N}\right)$	$\frac{\theta e}{2E[1 - e + (eN/k)]}$	$\frac{\theta}{2E[1 - e + (eN/k)]}\left(\frac{e}{2k}\right)^{**}$

* $E = [1 - (1 - e)^2]/2(1 - e)$.

** $e \ll kN^{-1}$.

*** $e \gg kN^{-1}$.

\ddagger Approximation assumes further that $mn \ll 1$ and $k \ll n$.

TABLE 3. Approximations for F_{ST} for the propagule-pool and migrant-pool models.

	$e \leq m$	$e \gg m$
Propagule pool	$\frac{1}{4Nm + 1}$	1^*
Migrant pool		
$e \ll kN^{-1}$	$\frac{1}{4Nm + 1}$	
$e \approx o(N^{-1})$	$\frac{1 + (eN/k)}{4Nm + 1 + 2eN}$	$\frac{k + eN}{k(1 + 2eN)}$
$e \gg kN^{-1}$		$\frac{1}{2k}$

* It is assumed here that $k \ll n$.

that as e increases, the inbreeding effective size of each deme approaches k , the initial number of colonists. The dependence of π_T on n rather than on Nn when $e \gg m$ in the propagule-pool model reflects the almost total absence of any genetic diversity within demes and the fact that almost all diversity will be found between demes. This is because, under the propagule-pool model, there can be no mixing of individuals from different sources during colonization and because the low expected life span of a colony, e^{-1} , when the rate of population turnover is high precludes the mixing of individuals through migration among extant demes. The situation approaches that described by the random replacement of clonal lines with no migration/recombination between extant lines, as modeled by Maruyama and Kimura (1980) (see below). It is also evident that when the extinction rate is much higher than the migration rate and colonization follows a propagule-pool model, the term $(\pi_T - \pi_S)$ is always small if n is large, so that F_{ST} approaches unity (Table 3, and see below).

In contrast to the propagule-pool model, when $e \gg m$ in the migrant-pool model π_S and π_T depend on both n and N (Table 2). This is because, even in the absence of migration among extant demes, colonization from a migrant pool causes the mixing of genes from across the metapopulation. Note that if $m \ll e \ll 1$ and if $k \ll N$, E in Table 2 reduces to e and within-deme and total metapopulation diversities under the migrant-pool model are approximated, respectively, by:

$$\pi_S \approx \frac{\theta}{2(1 + eNk^{-1})} \left(1 - \frac{1}{2k} \right) \quad (9)$$

and

$$\pi_T \approx \frac{\theta}{2(1 + eNk^{-1})} \left(1 + \frac{1}{2eN} \right). \quad (10)$$

Effect of Population Turnover on F_{ST}

The most widely used measure of genetic differentiation in subdivided populations is Wright's (1951) F_{ST} , or modifications thereof (see Charlesworth 1998). In Figure 2, we give examples of the effect of population turnover on F_{ST}

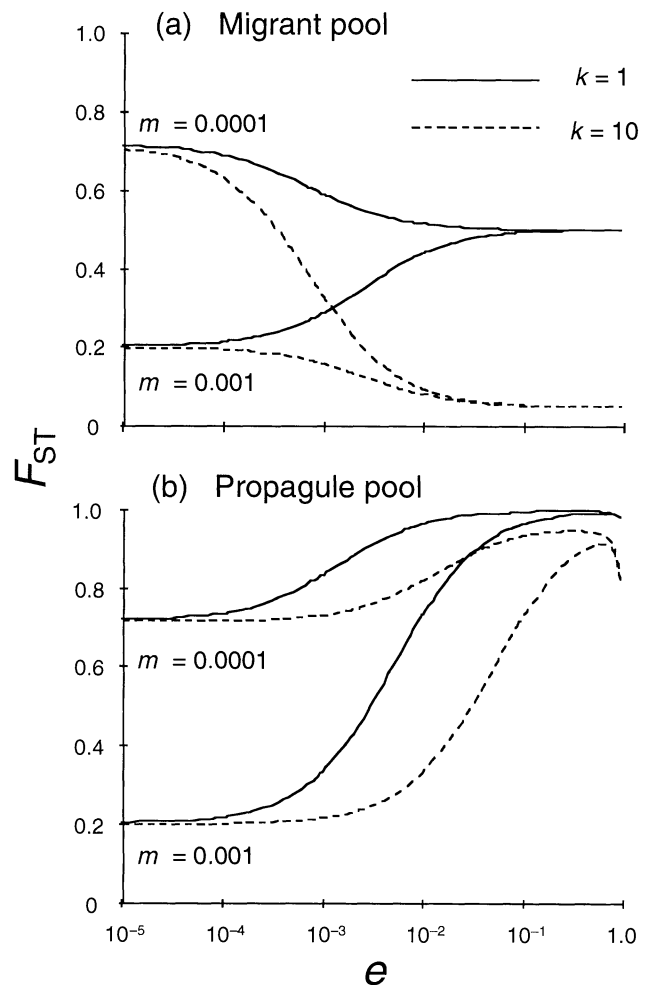


FIG. 2. Numerical calculations of the exact solutions to $F_{ST} = (\pi_T - \pi_S)/\pi_T$, with π_T and π_S given by the equilibrium solutions to equations (1) and (2): (a) the migrant-pool; and (b) the propagule-pool models. Results are given for $k = 1$, $k = 10$, $m = 10^{-3}$, and $m = 10^{-4}$. Other parameter values: $u = 10^{-8}$, $N = 10^3$, and $n = 500$. Note that the x-axis is on a log scale.

for several parameter values. Under propagule-pool colonization, increasing e always leads to an increase in F_{ST} when $e > m$ (see below). (The reduction in F_{ST} as e approaches unity is due to the loss of all diversity from the metapopulation; see Fig. 1.) In contrast, with migrant-pool colonization, whether F_{ST} increases or decreases with increasing e depends on the relative size of Nm and k , as observed by Wade and McCauley (1988) and Whitlock and McCauley (1990). In general, population turnover under the propagule-pool model brings about much greater increases in F_{ST} than it does under migrant-pool assumptions, which lead to a decrease in F_{ST} if $k > Nm + 0.5$ (Wade and McCauley 1988).

For the propagule-pool model with $e \leq m$ or $e \gg m$, our approximation for F_{ST} is independent of the extinction rate (see also Fig. 2). When $e \leq m$, F_{ST} is given by Wright's (1951) familiar approximation $(4Nm + 1)^{-1}$; when $e \gg m$, $F_{ST} \approx 1$ (Table 3). For the migrant-pool model, if e is much smaller than both m and kN^{-1} , then $F_{ST} \approx (4Nm + 1)^{-1}$, as for propagule-pool colonization and for Wright's island mod-

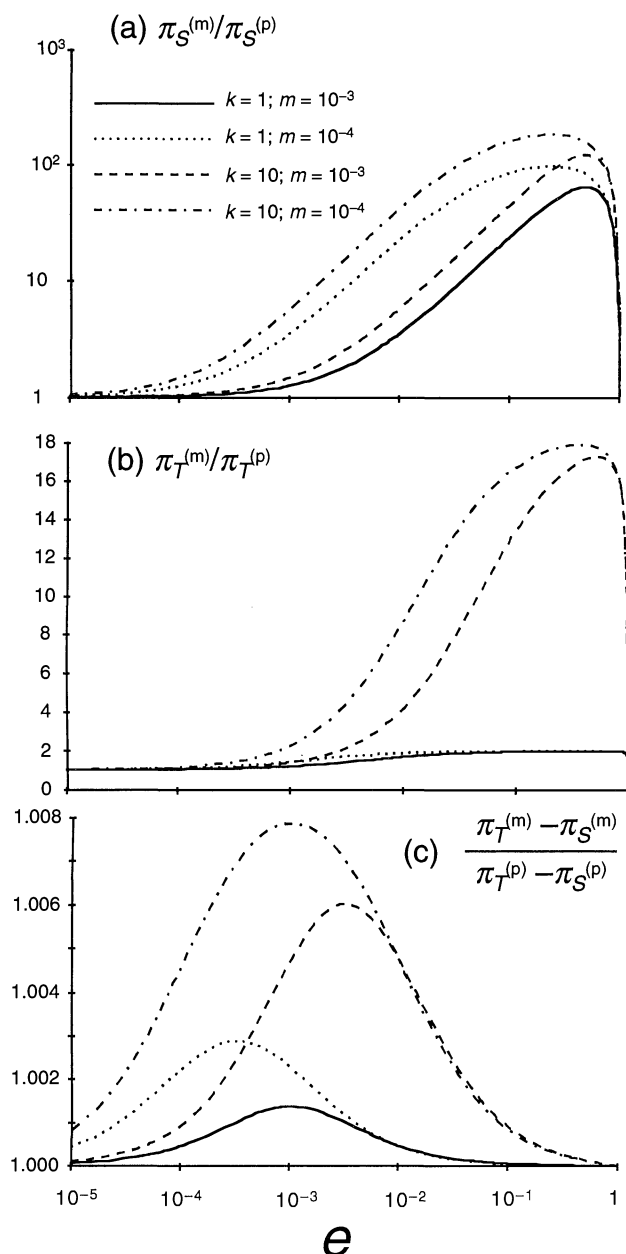


FIG. 3. Numerical comparisons of exact solutions of the propagule-pool and migrant-pool models for selected parameter values. Curves for (a) π_S , (b) π_T , and (c) $(\pi_T - \pi_S)$ are plotted as a ratio of the solution for the migrant-pool model to that for the propagule-pool model with the given parameter values. Results are given for $k = 1$, $k = 10$, $m = 10^{-3}$, and $m = 10^{-4}$. Other parameter values: $u = 10^{-8}$, $N = 10^3$, and $n = 500$. Note that the x-axis is on a log scale.

el without extinction. If e is much greater than both m and kN^{-1} , $F_{ST} \approx (2k)^{-1}$, which is substantially lower than its value of unity under propagule-pool colonization with high e .

These differences between the effects of propagule-pool and migrant-pool colonization on F_{ST} are difficult to interpret because F_{ST} is a function of both within-deme and total metapopulation genetic diversity (see eq. 8). In Figure 3 we therefore compare the effects on measures of genetic diver-

sities of population turnover with propagule-pool versus migrant-pool colonization. A metapopulation with relatively high extinction rates and migrant-pool colonization maintains a great deal more genetic variation at equilibrium than the corresponding case with propagule-pool colonization, both within individual demes (Fig. 3a) and in the metapopulation as a whole (Fig. 3b). However, the absolute value of $(\pi_T - \pi_S)$ is insensitive to the mode of colonization, even for relatively high values of $(k - Nm)$ (Fig. 3c). The differences in F_{ST} between propagule-pool and migrant-pool colonization can thus be attributed almost entirely to differences in their effect on total metapopulation diversity.

Effective Size of a Metapopulation

Following Whitlock and Barton (1997), we define the mutation effective size, N_e , as the size of a Wright-Fisher population in which diversity π_T is maintained in a balance between drift and mutation at equilibrium. Setting $\pi_T = 4N_e u$ in the expression for the nucleotide diversity of a metapopulation with propagule-pool colonization with $e \leq m$ (Table 2), we obtain:

$$N_e \approx \frac{n}{4(e + m)F_{ST}}. \quad (11)$$

This is equivalent to the approximate equations (23) and (24) of Whitlock and Barton (1997), which describe the eigenvalue effective size of a metapopulation, *i.e.*, the size of a Wright-Fisher population without mutation for which the asymptotic rate of increase in identity by descent is equal to that expected in a metapopulation with extinction.

Thus, for propagule-pool colonization with an extinction rate $e \leq m$, the mutation effective size is approximately equal to the eigenvalue effective size of a metapopulation as derived by Whitlock and Barton (1997). When $e \gg m$ and $e \gg kN^{-1}$, our approximation $2(1 - e)un(2e - e^2)^{-1}$ for π_T (Table 2) yields $N_e \approx n(4e)^{-1}$ by substitution into $\pi_T = 4N_e u$. This is equivalent to Maruyama and Kimura's (1980) expression for the effective size of a diploid metapopulation with propagule-pool colonization and no migration between extant demes. Whitlock and Barton's equation (23) also reduces to Maruyama and Kimura's expression for $e \ll m$. Under these conditions, the mutation effective size depends only on the number of demes and on the extinction rate, and is independent of the number of colonists and the size of each colony. Moreover, it appears that Whitlock and Barton's (1997) eigenvalue effective size is a good approximation of the mutation effective size under both propagule-pool and migrant-pool colonization, but only as long as $nNu \ll 1$. Therefore, there is no general approximate equivalence of the mutation and eigenvalue effective sizes (Whitlock and Barton 1997). This is perhaps not surprising, in view of the fact that such an equivalence does not necessarily hold even in the case of the conventional island model, unless n is large (compare equations 34 and 38a of Nagylaki [1986]).

Comparison of Approximate and Exact Solutions for π_S and π_T

We compared our approximations with the exact equilibrium solutions for π_T and π_S for a wide range of parameter

values, making no assumptions regarding the relative sizes of e and kN^{-1} . There was typically good correspondence between the approximations and the exact solutions for both propagule-pool and migrant-pool models for $e < m$ and $e \gg m$. However, for intermediate values of e relative to m , our expressions in Table 2 may be very poor approximations to the exact values and should not be used.

Comparison of the Approximations with Those of Slatkin

Slatkin (1977) assumed an infinite-alleles model of mutation and presented approximate solutions to his recurrence equations in terms of n_e , the effective number of alleles in a metapopulation with recurrent local extinctions, where $n_e = (1 - \bar{h})^{-1}$. Our approximations for \bar{h} are equivalent to Slatkin's n_e for migrant-pool colonization when $e \leq m$ (Table 1). However, Slatkin's estimate for n_e for the case where $e \gg m$ in the migrant-pool model appears to be in error. Our approximation for this case is:

$$n_e = 1 + \frac{2eNnu + un}{E(1 - e + eN/k)}. \quad (12)$$

For the propagule pool model, Slatkin's (1977) expressions imply somewhat simpler approximations for \bar{h} than ours (Table 1). These differences are due to Slatkin's implicit assumption that $e \ll kN^{-1}$ for the case where $e \leq m$ and that $e \gg kN^{-1}$ for the case where $e \gg m$. These further assumptions are probably reasonable for small e and large e , respectively, but when e is of order N , the more complex approximations presented in Table 1 are better.

DISCUSSION

Total Genetic Diversity in a Metapopulation

The process of population turnover in a metapopulation, where recurrent local extinction is followed by the recolonization of unoccupied sites by individuals from extant demes, causes a major reduction in the genetic diversity maintained within demes and in the total metapopulation. This effect is apparent for both propagule-pool colonization, where k colonists all come from the same source population, and for migrant-pool colonization, where colonists are drawn randomly from the metapopulation as a whole. Slatkin (1977) first showed that recurrent extinction causes a sharp reduction in this measure of diversity, and Maruyama and Kimura (1980) used a modification of Slatkin's propagule-pool model to demonstrate a reduction in the mutation effective size of a metapopulation due to recurrent local extinction. More recently, Whitlock and Barton (1997) have shown that recurrent extinction reduces the eigenvalue effective size of a metapopulation, and, thus, by implication, its diversity. However, as we have discussed, the two measures of effective population size are not, in general, equivalent, nor are they equivalent to Nagylaki's (1980, 1982, 1998b) migration effective size. Therefore, predictions of diversity or coalescent times require explicit calculations.

Whitlock and Barton (1997) attributed the expected decrease in effective population size of a metapopulation to a dramatic increase in the variance in reproductive success among individuals brought about by the process of extinction

and recolonization. In a metapopulation without extinction, each individual has a mean of one successful progeny, with a variance of one. In contrast, extinction affords individuals that recolonize sites an expected reproductive output of Nk^{-1} progeny, which will typically be much higher than one, whereas those individuals in demes on the verge of extinction fail to contribute any progeny to the metapopulation. This increased variance in reproductive success due to recurrent local extinction is responsible for the reduction in total diversity predicted by our analysis and the studies cited above.

Within-Population Diversity

Models of subdivided populations in which each deme maintains a constant size have shown that, as long as genes in each deme may eventually reach any other deme through migration, at least some individuals remain in each deme from one generation to the next, and migration does not change the size of individual demes (see introduction), then the average within-deme diversity at genetic equilibrium is independent of the details of the migration process under the infinite-sites model (Slatkin 1987; Strobeck 1987). Nagylaki (1998b) has shown that this result can be generalized to the case of a nonconservative migration model, if the mean diversity is calculated by weighting the contribution from each deme by the corresponding component of the left leading eigenvector of the migration matrix, replacing N_T , the total breeding size of the population, with Nagylaki's migration effective population size (Nagylaki 1980, 1982, 1998b); this parameter is less than or equal to N_T . This result is useful because it suggests that, at least in principle, an analogue of the scaled mutation rate, θ , can be obtained from a weighted estimate of within-population diversity in the case of a subdivided population. Total diversity is, of course, inflated over the within-population diversity as a result of subdivision and cannot be used to estimate θ unless there is so little evidence of population subdivision that panmixia is a good approximation.

Our results show that the invariance principle does not hold for a metapopulation with recurrent extinction and recolonization. Even when $e \leq m$, Table 2 shows that π_S is always less than or equal to θ . It is thus inappropriate to use measures of average within-deme diversity to estimate θ in species with a metapopulation structure and population turnover. This essentially follows from the fact that the simple migration-matrix approach used in the standard models of geographically structured populations cannot be applied in a straightforward manner to a metapopulation with recurrent local extinction. In general, it will be very hard to use measures of DNA sequence diversity for other than descriptive purposes when there is reason to suspect extinction-recolonization population dynamics.

Population Differentiation in a Metapopulation

Most previous studies of genetic diversity in a metapopulation have been concerned with predicting the effects of local extinction and recolonization on F_{ST} (Slatkin 1985; McCauley 1987, 1989, 1991, 1993; Wade and McCauley 1988; Whitlock and McCauley 1990; Whitlock 1992a; Ingvarsson 1997; Whitlock and Barton 1997; Le Corre and Kremer

1998). However, recently several authors have suggested that F_{ST} may be a poor measure of population differentiation generally, because it can be particularly sensitive to the total genetic diversity across the species, irrespective of differentiation between demes (Charlesworth et al. 1997; Charlesworth 1998; Nagylaki 1998a). The problem is most acute when levels of diversity are high. For example, analyses will tend to yield low estimates of F_{ST} if within-deme diversity is high, even if different demes have no alleles in common (and are thus completely differentiated; Nagylaki 1998a). This suggests that for loci with high levels of diversity, such as microsatellites, F_{ST} should be avoided as a measure of population differentiation. Nevertheless, because total diversity in a metapopulation with recurrent local extinctions should be low, the problem of high within-deme diversity is less likely to be encountered in species with a colonizing habit.

A more general problem associated with relative measures of genetic differentiation such as F_{ST} is due to the fact that metapopulation processes affect both within-population and total metapopulation diversities, potentially making comparisons among species difficult to interpret (McCauley 1995). For example, Wade and McCauley (1988) and Whitlock and McCauley (1990) have discussed the relative effects on F_{ST} (or G_{ST}) of propagule-pool versus migrant-pool colonization: F_{ST} is always increased by population turnover if sites are recolonized following the propagule-pool model, whereas the effect of extinction on F_{ST} in a metapopulation with migrant-pool colonization depends on the relative values of k and Nm . By considering the effects of population turnover on both within-deme and total metapopulation diversity, we find that the value of the between-population diversity ($\pi_T - \pi_S$) is insensitive to the mode of colonization (Fig. 3). Thus, differential effects of the colonization mode are due chiefly to its effect on the denominator of the F_{ST} ratio, π_T . Therefore, the much greater increase in F_{ST} with extinction under the propagule pool model when $e \geq m$ (Fig. 2) mainly reflects the effect of e on total diversity rather than an effect on between-population diversity. Conversely, for $e \leq m$, Table 2 shows that the between-population diversity is reduced by increased e , whereas F_{ST} is nearly independent of e . Thus, differences in F_{ST} between species with different ecologies do not necessarily reflect differences in absolute measures of population differentiation.

Wade and McCauley (1988) have discussed the difference between the propagule-pool and migrant-pool models of colonization from a biological perspective. In particular, they noted that in most cases colonization is likely to be intermediate between the two extreme models originally proposed by Slatkin (1977). Whitlock and McCauley (1990) included this possibility in their model of within-population diversity in an infinite metapopulation by introducing a parameter, ϕ , for the probability that two immigrant genes came from the same source population; $\phi = 0$ and $\phi = 1$ correspond to migrant-pool and propagule-pool colonization, respectively. Wade and McCauley (1988) cited McCauley's (1987) study of a milkweed beetle metapopulation, *Tetraopes tetraophthalmus*, as an example "truly intermediate between migrant- and propagule-pool modes of colony formation." In that study, it was found that "sites were often visited by only 1–4 in-

seminated females of unknown but most probably independent origin" (Wade and McCauley 1988). Using demographic and genetic data from a metapopulation of the forked fungus beetle, *Bolitotherus cornutus*, Whitlock (1992b) estimated ϕ to be about 0.5. Ingvarsson et al. (1997) also estimated ϕ to be about 0.5 for a metapopulation of the mycophagous beetle *Phalacrus substriatus*, but they noted that this value was probably conservatively low. It seems likely that, in highly structured metapopulations, colonization may often follow the propagule-pool model, with colonists tending to originate from neighboring colonies (in a stepping-stone fashion), rather than as a random sample of the metapopulation as a whole. This is particularly likely when k is low, with the colonizing propagule comprising either a single self-fertile individual, a previously mated female, or a group of related individuals such as seeds in the same fruit (Wade and McCauley 1988).

What values of e are we likely to find in species characterized by metapopulation dynamics, and to what extent may we use our expressions for F_{ST} to detect population turnover in a metapopulation? There are two results that are relevant to these questions. First, we have seen that, for $e \leq m$ and $e \ll N^{-1}$, F_{ST} is insensitive to population turnover. Thus, if we are to use F_{ST} to detect metapopulation dynamics, we should expect to find evidence for population turnover only if $e > m$ (see Table 3). Second, in situations in which there is rapid population turnover (i.e., $e > m$), we may use an estimate of F_{ST} to predict the likely mode of colonization. If F_{ST} is less than about 0.5, colonists most likely originate from a migrant pool; for $F_{ST} > 0.5$, propagule-pool colonization seems more probable (Table 3). The metapopulation studies cited above have tended to find F_{ST} in the lower of these two ranges. This would suggest either that the extinction rate is of the order of the migration rate or, perhaps less likely, that colonization has been from a migrant pool.

Rates of Equilibration

The results we have presented in this paper have all assumed that an equilibrium has been reached in the distribution of genetic variation as a result of its accumulation through mutation and its loss through drift and due to its redistribution among demes through migration and recolonization following local extinction. Perturbations in parameter values, such as the size of demes, the migration and extinction rates, or in the total number of demes occupied in the metapopulation, will cause levels of within-deme and total metapopulation diversities to deviate from their equilibrium values, and substantial time may be required before a new equilibrium is established. One advantage that F_{ST} has over absolute measures of diversity is its relatively rapid equilibration following perturbations (Crow and Aoki 1984; McCauley 1995). This is because F_{ST} is generally determined by the redistribution of the extant genetic variation in the metapopulation by gene flow among local populations, whereas absolute levels of diversity depend on the rates of mutation and drift at the metapopulation level, which will tend to be very small (McCauley 1995). Nevertheless, equilibration times for average within-deme diversity, total metapopulation diversity, and F_{ST} depend strongly on the nature of the perturbation. In Table 4 we present the results of deterministic iterations

TABLE 4. The number of generations for π_S , π_T , and F_{ST} to change from one equilibrium value to a value halfway toward the new equilibrium after a perturbation in the number of demes, n . Calculations were performed by iterating Slatkin's (1977) recursions. Equilibrium values prior to and after the perturbation were determined using the exact solutions to equations (1) and (2). We traced the change in π_T , π_S , and F_{ST} for perturbations in: (a) the number of individuals per deme, N (from $N = 10$ to $N = 1000$ and vice versa); (b) the extinction rate, e (from $e = 0$ to $e = 0.001$ and vice versa); and (c) the number of demes, n (from $n = 10$ to $n = 1000$ and vice versa). Results are shown for the propagule-pool and the migrant-pool models and for $k = 1$ and $k = 10$. Other parameter values used, unless stated otherwise: $u = 10^{-8}$, $N = 1000$, $n = 500$, $e = 0.001$, $m = 0.001$.

	$k = 1$		$k = 10$	
	from $N = 1000$ to $N = 10$	from $N = 10$ to $N = 1000$	from $N = 1000$ to $N = 10$	from $N = 10$ to $N = 1000$
(a)				
Propagule pool				
π_S	14	107,132	14	257,548
π_T	89,531	258,087	89,571	397,426
F_{ST}	13	231	14	271
Migrant pool				
π_S	14	154,480	14	408,043
π_T	90,363	300,596	91,155	542,951
F_{ST}	13	198	13	198
(b)				
	from $e = 0.0$ to $e = 0.001$	from $e = 0.001$ to $e = 0.0$	from $e = 0.0$ to $e = 0.001$	from $e = 0.001$ to $e = 0.0$
Propagule pool				
π_S	164,085	758,900	375,814	823,948
π_T	257,986	844,762	397,232	844,784
F_{ST}	231	277	270	276
Migrant pool				
π_S	227,720	771,971	667,305	958,869
π_T	299,971	846,062	540,281	847,166
F_{ST}	205	272	183	289
(c)				
	from $n = 1000$ to $n = 10$	from $n = 10$ to $n = 1000$	from $n = 1000$ to $n = 10$	from $n = 10$ to $n = 1000$
Propagule pool				
π_S	5272	512,169	8336	785,421
π_T	5185	512,080	8196	785,273
F_{ST}	23,426	2363	636	665
Migrant pool				
π_S	6017	596,081	10,931	1,069,260
π_T	5976	596,038	10,892	1,069,217
F_{ST}	32,248	4071	35,297	1864

of equations (1) and (2) after three types of perturbation: an increase or decrease in N , the number of individuals per deme; an increase or decrease in e , the extinction rate; and an increase or decrease in n , the total number of demes in the metapopulation.

New equilibrium values for π_S and F_{ST} are quickly reached after a decrease in the size of N , but π_T requires orders of magnitude more time to equilibrate (Table 4a). This is because the rate of reduction in π_S is determined by the size of N^{-1} , whereas that of π_T depends on the much smaller size of $(nN)^{-1}$. Thus, when n is large, an increase in F_{ST} after a drop in N will occur about as rapidly as the corresponding decrease in within-deme diversity. However, when N increases, equilibration of π_S takes place at a rate of the order of that for π_T (Table 4a). Note that in this case F_{ST} equilibrates very much more rapidly than either measure of absolute diversity.

When there is a perturbation in the extinction rate (Table 4b), F_{ST} equilibrates much more rapidly than π_S or π_T . Note that, again, both π_S and π_T take longer to change when diversity has to be generated by new mutations (i.e., when the mutation effective size of the metapopulation increases as a result of decreasing e) than when equilibration involves the loss of diversity through drift (i.e., when e increases). This difference will of course depend on the relative values of u and N_e^{-1} ; if $N_e > u^{-1}$ after a perturbation, the loss of diversity

through drift will take longer than its increase through mutation.

Rates of equilibration after perturbations in the total number of occupied sites appear to be less straightforward (Table 4c). In particular, when n suddenly decreases, it may take a great deal longer for F_{ST} to equilibrate than it does for π_S and π_T , although it should be noted that this was not so for $k = 10$ in our iterations of the propagule-pool model. The results presented in Table 4c are the number of generations required for each of the three diversity measures to change halfway toward their new equilibrium value after a perturbation in n . We also plotted the trajectories over time for this process of equilibration and found that, whereas π_S and π_T approached their new equilibrium along a smooth asymptote, F_{ST} altered relatively quickly at first, but then continued to change much more slowly (results not shown). This was reflected by the fact that F_{ST} changed 10% toward its new equilibrium more rapidly than did π_S and π_T , but it took orders of magnitude longer than the absolute measures of diversity to go 90% of the way toward its new equilibrium. The finding that F_{ST} can take longer to equilibrate than either π_S and π_T when n decreases with constant N is a further illustration of the strong dependence of F_{ST} on π_T , the denominator of the F_{ST} ratio, as noted above. However, alteration of n does not in fact lead to large changes in F_{ST} (for the parameter values presented in Table 4c, even the change

in n between 10 and 1000 caused the equilibrium value of F_{ST} to alter by less than 5%). Thus, for perturbations leading to measurable changes in F_{ST} , its equilibration will generally be much more rapid than that of absolute measures of diversity (Crow and Aoki 1984). In metapopulations that are in a state of recovery after a perturbation, it is likely that π_T and π_S will not be near their equilibrium values, whereas F_{ST} is much more likely to be so. It should also be noted that whereas π_T and π_S tend to be changed substantially by perturbations in n , N , and e , albeit slowly, F_{ST} is sensitive to perturbations from equilibrium conditions in N and e , but is almost unaffected by perturbations in n , such as may occur as a result of large-scale range expansions or contractions in the distribution of a species.

Relation between Breeding System and Measures of Diversity

In many plant species, seed dispersal may be highly localized, with very low seed immigration rates between colonies (Levin and Kerster 1974; Levin 1981). Gene flow between extant populations may be particularly restricted in highly self-fertilizing plant species in which flowers are typically small and pollen production and dispersal are limited (Morgan and Barrett 1989; Fenster et al. 1995). If such species experience population turnover in the context of a metapopulation with propagule-pool colonization, we may expect to find very low within-deme diversity. In a recent study of within- and among-deme nucleotide diversity in highly self-fertilizing populations of *Leavenworthia* species, Liu et al. (1998) found no genetic variability within populations, but substantial diversity between them. In discussing these results, the authors proposed that repeated extinction and recolonization events may have contributed to the observed patterns, noting, however, that "in inbreeders extinction and recolonization probably reduces species-wide diversity" (Liu et al. 1998).

Based on the results presented here, this assessment appears to be sound. Whereas population turnover, whether with propagule-pool or migrant-pool colonization, may drastically reduce within-population diversity, especially in the absence of migration between extant demes, total species diversity will be reduced by frequent recurrent extinction to an extent not evident in the *Leavenworthia* data. Indeed, total diversity among selfing populations was found to be similar to that among outcrossing populations (about 4%; Liu et al. 1998). We find that it is possible to account for the low within- and relatively high among-population nucleotide diversity observed by Liu et al. (1998) in *Leavenworthia* only if we assume a sufficiently low extinction rate combined with an absence of migration between extant demes. For instance, under a propagule-pool mode of colonization, with $u = 10^{-8}$, $n = 500$, $N = 10^3$, $k = 1$, and $m = 0$, an extinction rate of 10^{-4} yields diversity values similar to those found in the highly selfing *Leavenworthia* species ($\pi_S = 3.6 \times 10^{-5}$; $\pi_T = 4.8 \times 10^{-2}$). Increasing m above zero yields diversity measures in line with observations of outcrossing *Leavenworthia* populations (e.g., with $m = 10^{-3}$: $\pi_S = 1.6 \times 10^{-2}$; $\pi_T = 2.1 \times 10^{-2}$). However, although it is possible that selfing and outcrossing species may indeed differ in their rates of

migration due to differences in pollen flow, such a low extinction rate, as well as absolutely no migration between demes in selfing species, seems to be an implausible scenario for a metapopulation at equilibrium, as discussed above.

It is possible that the selfing populations of *Leavenworthia* are simply not at genetic and demographic equilibrium, but rather are the result of a recent range expansion of the species from a more genetically diverse source not sampled by Liu et al. (1998). Large differences in within-deme allozyme diversity have been often found in inbreeding species, with some populations maintaining relatively high levels of diversity and others devoid of any variation at all (Schoen and Brown 1991). In the highly selfing species *Avena barbata*, for example, small roadside populations generally lacked variation at enzyme loci, whereas larger field populations, which may have been the source of the individuals that colonized the roadside habitat, were more highly diverse (Jain et al. 1981). The observation of large differences in diversity among populations of selfing species certainly suggests a need for diversity estimates from a larger sample of populations than was presented for *Leavenworthia*. It also suggests a need for models that predict not only the mean within-deme diversity in a metapopulation, but also its distribution. Even in idealized models, where each deme is assumed to be identical, variance in diversity is likely to be high because of the age structure established by population turnover in the metapopulation (Wade and McCauley 1988). In a metapopulation in which demes differ from one another in their rates of extinction, colonization, and migration, such variance will be even higher (Whitlock and Barton 1997). The population bottlenecks associated with colonization events would be expected to produce departures from standard neutral expectation under this type of scenario; if sufficient within-deme variability can be detected, tests such as Tajima's D could be used to detect these effects (Charlesworth et al. 1993), although their power in such situations remains to be determined.

It is also possible that the low within-deme and high between-deme diversities observed in selfing species such as *Leavenworthia* are the result of the homogenizing influences of local adaptation and/or background selection on the genome within demes (Liu et al. 1998). This may be particularly plausible for highly selfing species in which the whole genome is tightly linked and neutral genetic diversity could repeatedly be lost through local selective sweeps or background selection against deleterious mutations (Charlesworth et al. 1993; Nordborg 1997). Models investigating the effects on neutral genetic diversity of selection at linked loci in a metapopulation context are needed to assess the likelihood of this possibility. Finally, the extremely protracted equilibration of π_T after a reduction in local population size, in contrast with the rapid approach to equilibrium of π_S (see Table 4a), suggests that the high values of π_T in the *Leavenworthia* selfers may reflect a history in which local effective population sizes were once much larger than they are today, and that π_T is yet to equilibrate. Notwithstanding this possibility, it should be borne in mind that the time scales over which such equilibration takes place may greatly exceed the age of a metapopulation in any given landscape, except under the most stable geoclimatic conditions. This has im-

portant implications for the interpretation of diversity data using models that assume equilibrium conditions, like those analyzed here.

It is clear that an assessment of the nature of recolonization in species with a metapopulation structure must be made on a case-by-case basis; there may even be variation in extinction-recolonization dynamics within a single species. In a recent study of the population structure in the perennial plant species *Silene dioica*, Giles et al. (1998) found that genetic differentiation among small patches occupied by the species on islands was about twice as high as it was between individuals sampled from different islands. They interpreted this result by suggesting that the colonization of patches followed a propagule-pool model, whereas gene flow between islands was more akin to a migrant-pool model. The effects of processes acting at different spatial and temporal scales are likely to be common in species with highly structured populations, and further empirical data from such species are greatly needed.

ACKNOWLEDGMENTS

We thank the Royal Society for financial support; T. Nagylaki and B. Haubold for helpful discussions; and N. Barton, D. McCauley, M. Slatkin, and M. Whitlock for valuable comments on the manuscript.

LITERATURE CITED

- BARRETT, S. C. H., AND J. R. PANNELL. 1998. Metapopulation dynamics and mating-system evolution in plants. In P. Hollingsworth, R. Bateman, and R. Gornell, eds. *Molecular systematics and plant evolution*. Chapman and Hall, London. *In press*.
- BOSSART, J. L., AND D. PASHLEY PROWELL. 1998. Genetic estimates of population structure and gene flow: limitations, lessons and new directions. *Trends Ecol. Evol.* 13:202–206.
- CHARLESWORTH, B. 1998. Measures of divergence between populations and the effect of forces that reduce variability. *Mol. Biol. Evol.* 15:538–543.
- CHARLESWORTH, B., M. T. MORGAN, AND D. CHARLESWORTH. 1993. The effect of deleterious mutations on neutral molecular variation. *Genetics* 134:1289–1303.
- CHARLESWORTH, B., M. NORDBORG, AND D. CHARLESWORTH. 1997. The effects of local selection, balanced polymorphism and background selection on equilibrium patterns of genetic diversity in subdivided populations. *Genet. Res.* 70:155–174.
- CROW, J. F., AND K. AOKI. 1984. Group selection for a polygenic behavioral trait: estimating the degree of population subdivision. *Proc. Natl. Acad. Sci., USA* 81:6073–6077.
- FENSTER, C. B., P. K. DIGGLE, S. C. H. BARRETT, AND K. RITLAND. 1995. The genetics of flora development differentiating two species of *Mimulus* (Scrophulariaceae). *Heredity* 74:258–266.
- GILES, B. E., E. LUNDQVIST, AND J. GOUDET. 1998. Restricted gene flow and subpopulation differentiation in *Silene dioica*. *Heredity* 80:715–723.
- HARRISON, S., AND D. TAYLOR. 1997. Empirical evidence for metapopulation dynamics. Pp. 27–42 in I. Hanski and M. E. Gilpin, eds. *Metapopulation biology: ecology, genetics, evolution*. Academic Press, San Diego, CA.
- HUDSON, R. R. 1990. Gene genealogies and the coalescent process. *Oxford Surv. Evol. Biol.* 7:1–45.
- HUDSON, R. R., D. D. BOOS, AND N. L. KAPLAN. 1992. A statistical test for detecting geographic subdivision. *Mol. Biol. Evol.* 9:138–151.
- HUSBAND, B. C., AND S. C. H. BARRETT. 1996. A metapopulation perspective in plant population biology. *J. Ecol.* 84:461–469.
- INGVARRSSON, P. K. 1997. The effect of delayed population growth on the genetic differentiation of local populations subject to frequent extinctions and recolonizations. *Evolution* 51:29–35.
- INGVARRSSON, P. K., K. OLSSON, AND L. ERICSON. 1997. Extinction-recolonization dynamics in the mycophagous beetle *Phalacrus substriatus*. *Evolution* 51:187–195.
- JAIN, S. K., K. N. RAI, AND R. S. SINGH. 1981. Population biology of *Avena* XI. Variation in peripheral isolates of *A. barbata*. *Genetica* 56:213–215.
- KIMURA, M. 1971. Theoretical foundation of population genetics at the molecular level. *Theor. Pop. Biol.* 2:174–208.
- KIMURA, M., AND J. F. CROW. 1964. The number of alleles that can be maintained in a finite population. *Genetics* 49:725–738.
- LATTER, B. D. H. 1973. The island model of population differentiation: a general solution. *Genetics* 73:147–157.
- LE CORRE, V., AND A. KREMER. 1998. Cumulative effects of founding events during colonisation on genetic diversity and differentiation in an island and stepping-stone model. *J. Evol. Biol.* 11:495–512.
- LEVIN, D. A. 1981. Dispersal versus gene flow in plants. *Ann. Missouri Bot. Gard.* 68:233–253.
- LEVIN, D. A., AND H. W. KERSTER. 1974. Gene flow in seed plants. *Evol. Biol.* 7:139–220.
- LIU, F., L. ZHANG, AND D. CHARLESWORTH. 1998. Genetic diversity in *Leavenworthia* populations with different inbreeding levels. *Proc. R. Soc. London, B* 265:293–301.
- MARUYAMA, T. 1970. Effective number of alleles in a subdivided population. *Theor. Pop. Biol.* 1:273–306.
- MARUYAMA, T., AND M. KIMURA. 1980. Genetic variability and effective population size when local extinction and recolonization of subpopulations are frequent. *Proc. Natl. Acad. Sci. USA* 77:6710–6714.
- MCCAULEY, D. E. 1987. Population genetic consequences of local colonization: evidence from the milkweed beetle *Tetraopes tetraophthalmus*. *Flor. Entomol.* 70:21–30.
- . 1989. Extinction, colonization, and population structure: a study of a milkweed beetle. *Amer. Nat.* 134:365–376.
- . 1991. Genetic consequences of local population extinction and recolonization. *Trends Ecol. Evol.* 6:5–8.
- . 1993. Evolution in metapopulations with frequent local extinction and recolonization. *Oxford Surv. Evol. Biol.* 10:109–134.
- . 1995. Effects of population dynamics on genetics in mosaic landscapes. Pp. 178–198 in L. Hansson, L. Fahrig, and M. Gray, eds. *Mosaic landscapes and ecological processes*. Chapman and Hall, London.
- MORGAN, M. T., AND S. C. H. BARRETT. 1989. Reproductive correlates of mating system variation in *Eichhornia paniculata* Spreng. Solms (Pontederiaceae). *J. Evol. Biol.* 2:183–204.
- NAGYLAKI, T. 1980. The strong-migration limit in geographically structured populations. *J. Math. Biol.* 9:101–114.
- . 1982. Geographical invariance in population genetics. *J. Theor. Biol.* 99:159–172.
- . 1986. Neutral models of geographical variation. Pp. 216–237 in P. Tautu, ed. *Stochastic spatial processes. Lecture Notes in Mathematics* no. 1212. Springer-Verlag, Berlin.
- . 1998a. Fixation indices on subdivided populations. *Genetics* 148:1325–1332.
- . 1998b. The expected number of heterozygous sites in a subdivided population. *Genetics* 149:1599–1604.
- NEI, M. 1973. Analysis of gene diversity in subdivided populations. *Proc. Natl. Acad. Sci. USA* 70:3321–3323.
- NORDBORG, M. 1997. Structured coalescent processes on different time scales. *Genetics* 146:1501–1514.
- PANNELL, J. R., AND S. C. H. BARRETT. 1998. Baker's law revisited: reproductive assurance in a metapopulation. *Evolution* 52:657–668.
- SCHOEN, D. J., AND A. H. D. BROWN. 1991. Intraspecific variation in population gene diversity and effective population size correlates with the mating system in plants. *Proc. Natl. Acad. Sci. USA* 88:4494–4497.
- SLATKIN, M. 1977. Gene flow and genetic drift in a species subject to frequent local extinction. *Theor. Pop. Biol.* 12:253–262.

- . 1985. Gene flow in natural populations. *Annu. Rev. Ecol. Syst.* 16:393–430.
- . 1987. The average number of sites separating DNA sequences drawn from a subdivided population. *Theor. Pop. Biol.* 32:42–49.
- SLATKIN, M., AND N. H. BARTON. 1989. A comparison of three indirect methods for estimating average levels of gene flow. *Evolution* 43:1349–1368.
- STROBECK, C. 1987. Average number of nucleotide differences in a sample from a single subpopulation: a test for population subdivision. *Genetics* 117:149–153.
- TAJIMA, F. 1983. Evolutionary relationship of DNA sequences in finite populations. *Genetics* 105:437–460.
- VALVERDE, T., AND J. SILVERTOWN. 1997. A metapopulation model for *Primula vulgaris*, a temperate forest understorey herb. *J. Ecol.* 85:193–210.
- WADE, M. J., AND D. E. MCCAULEY. 1988. Extinction and recolonization: their effects on the genetic differentiation of local populations. *Evolution* 42:995–1005.
- WHITLOCK, M. C. 1992a. Temporal fluctuations in demographic parameters and the genetic variance among populations. *Evolution* 46:608–615.
- WHITLOCK, M. C. 1992b. Nonequilibrium population structure in forked fungus beetles: extinction, colonization, and the genetic variance among populations. *Am. Nat.* 139:952–970.
- WHITLOCK, M. C., AND N. H. BARTON. 1997. The effective size of a subdivided population. *Genetics* 146:427–441.
- WHITLOCK, M. C., AND D. E. MCCAULEY. 1990. Some population genetic consequences of colony formation and extinction: genetic correlations within founding groups. *Evolution* 44:1717–1724.
- WRIGHT, S. 1940. Breeding structure of populations in relation to speciation. *Am. Nat.* 74:232–248.
- . 1951. The genetical structure of populations. *Ann. Eugen.* 15:323–354.

Corresponding Editor: T. Markow

APPENDIX

Here, we define the parameters a , b , A , B_1 , and B_2 , as introduced by Slatkin (1977) and as used in the recurrence equations (1) and (2) in the text. The parameter

$$a = (1 - m)^2 + \frac{m^2}{(1 - e)n} + \frac{2m(1 - m)}{(1 - e)n} \quad (\text{A1})$$

represents the sum of probabilities that (1) neither of the two genes sampled are immigrants; (2) both are immigrants from the same source population; and (3) one is not an immigrant and the other is an immigrant that has returned to its source population, which we are sampling. Slatkin (1977) attributed this last term to Latter (1973), who suggested it as an improvement to Maruyama's (1970) formulation in which it was omitted. However, we may choose to

define immigrants as only those individuals that have originated elsewhere in the metapopulation, which might be biologically more meaningful. In this case, we omit the third term in equation (A1) altering the second term such that

$$a = (1 - m)^2 + \frac{m^2}{(1 - e)n - 1}. \quad (\text{A2})$$

Making Slatkin's formulation somewhat more explicit, the parameter

$$b = \frac{m^2}{(1 - e)n} + \frac{2m(1 - m)}{(1 - e)n} \quad (\text{A3})$$

represents the sum of probabilities that both genes sampled from distinct populations are immigrants from the same source population and the gene sampled in the first population is not an immigrant, whereas that sampled in the second population is an immigrant that originated from the first population. If immigrants are viewed as individuals originating only from other populations, as suggested above, then we substitute $[(1 - e)n - 1]$ for $(1 - e)n$ in equation (A3). We assume that $m \ll 1$ and so approximate equations (A1) and (A3) by assuming $a \approx 1 - 2m$ and $b \approx 2m/[(1 - e)n]$.

In equations (1) and (2) in the text, A is the probability that the two genes chosen from the same newly founded population are identical; B_1 is the probability that two identical genes are chosen from different populations, one from a newly founded population and the other from an existing population; and B_2 is the probability that two identical genes are sampled from different newly founded populations. For the propagule-pool model, Slatkin (1977) gives

$$A = \frac{1}{2N} + \left(1 - \frac{1}{2N}\right) \left(\frac{1}{2k} + \left(1 - \frac{1}{2k}\right)(1 - h_0) \right). \quad (\text{A4})$$

Here, we account for both the sampling of k individuals during colonization and the sampling involved when these k individuals mate randomly to form a population of size N .

For the migrant-pool model, in which colonists are drawn randomly from the metapopulation as a whole, we have

$$A = \frac{1}{2Nn(1 - e)} + \left(1 - \frac{1}{2Nn(1 - e)}\right) \left(\frac{1}{2k} + \left(1 - \frac{1}{2k}\right)(1 - \tilde{h}) \right), \quad (\text{A5})$$

where

$$1 - \tilde{h} = \frac{1}{n(1 - e)}(1 - h_0) + \left(1 - \frac{1}{n(1 - e)}\right)(1 - h_1) \quad (\text{A6})$$

is the probability that two genes chosen from the migrant pool are identical (Slatkin 1977). To calculate B_1 and B_2 , we must take into account the possibility that colonists originate from the same or from different populations. Slatkin (1977) gives us, for both the propagule- and migrant-pool models,

$$B_1 = B_2 = \frac{1}{(1 - e)n}(1 - h_0) + \left(1 - \frac{1}{(1 - e)n}\right)(1 - h_1) = 1 - \tilde{h}. \quad (\text{A7})$$