Molecular Diversity After a Range Expansion in Heterogeneous Environments

Daniel Wegmann,* Mathias Currat*,† and Laurent Excoffier*,[1]

*Computational and Molecular Population Genetics Laboratory, Zoological Institute, University of Bern, 3012 Bern, Switzerland and †Laboratoire d'Anthropologie, Génétique et Peuplements, University of Geneva, 1227 Carouge, Switzerland

ABSTRACT

Recent range expansions have probably occurred in many species, as they often happen after speciation events, after ice ages, or after the introduction of invasive species. While it has been shown that range expansions lead to patterns of molecular diversity distinct from those of a pure demographic expansion, the fact that many species do live in heterogeneous environments has not been taken into account. We develop here a model of range expansion with a spatial heterogeneity of the environment, which is modeled as a gamma distribution of the carrying capacities of the demes. By allowing temporal variation of these carrying capacities, our model becomes a new metapopulation model linking ecological parameters to molecular diversity. We show by extensive simulations that environmental heterogeneity induces a loss of genetic diversity within demes and increases the degree of population differentiation. We find that metapopulations with low average densities are much more affected by environmental heterogeneity than metapopulations with high average densities, which are relatively insensitive to spatial and temporal variations of the environment. Spatial heterogeneity is shown to have a larger impact on genetic diversity than temporal heterogeneity. Overall, temporal heterogeneity and local extinctions are not found to leave any specific signature on molecular diversity that cannot be produced by spatial heterogeneity.

R ANGE expansions are a recurrent phenomenon in the history of many species. If one assumes that speciation occurs in a restricted portion of an ancestral species range, the fact that current species have a wide geographic distribution implies that they have spread from their initial location. This is, for instance, thought to be the case of modern humans (*e.g.*, Cavalli-Sforza and Feldman 2003). Climatic changes such as ice ages have also triggered episodes of range contractions and expansions in many plant and animal species (*e.g.*, Taberlet *et al.* 1998; Hewitt 2000; Shapiro *et al.* 2004; Tzedakis *et al.* 2004; Smith and Farrell 2005). Finally, invasive species that are introduced into new and favorable habitats often spread from their introductory site and then quickly expand to a large territory by a series of migrations and colonization events (*e.g.*, Silva *et al.* 2002; Smith *et al.* 2004).

Austerlitz *et al.* (1997) studied the pattern of molecular diversity after a colonization process in a one-dimensional stepping-stone model. They showed that the expansion leads to a series of bottlenecks inducing a gradient of differentiation from the initial deme unless migration rates between demes are very large. Thus, range expansions can somehow be viewed as a series of successive founder events (Austerlitz *et al.* 1997). Le Corre and Kremer (1998) analyzed the effects of these founder events on within-population heterozygosity and on genetic differentiation (as measured by F_{ST}) in two models of range expansions: a linear array of demes (one-dimensional stepping-stone model), where dispersal occurs only between neighboring demes, and an island model representing an extreme case of long-distance dispersal. While an increase in F_{ST} and a decrease in intrademe heterozygosity were found in both models, these effects were much more pronounced in the linear stepping-stone model. A recent study has shown that the pattern of molecular diversity observed after a range expansion is different from that expected after a pure demographic expansion and depends mainly on the age of the expansion as well as the number of migrants exchanged between neighboring demes (Ray *et al.* 2003). By simulating a range expansion in a two-dimensional stepping-stone world, Ray *et al.* (2003) showed that the expected distribution of the number of differences between pairs of genes (also called the mismatch distribution) is bimodal if the number of genes entering each deme per generation is not large (<50). This is in contrast to the unimodal distributions observed for larger migrations rates between neighboring demes, which are very similar to those obtained after a large demographic expansion in an unsubdivided population (Slatkin and Hudson 1991; Rogers and

HARPENDING 1992). These results were confirmed analytically for a simple instantaneous expansion in an infinite-island model (EXCOFFIER 2004). RAY *et al.* (2003) also found that summary statistics used in tests of selective neutrality (like Tajima's D and Fu's F_S) showed significant negative values only in demes exchanging a larger number of migrants. Contrastingly, demes receiving $< \sim 20$ genes per generation would not show any sign of departure from demographic equilibrium despite being part of a large expanding population.

All expansion models described so far assumed that deme sizes and migration rates had been constant over time, as expected in a homogeneous and stationary environment. However, many species do not meet these assumptions, since they are actually subdivided into locally breeding demes of different size showing metapopulation dynamics, with extinction and recolonization (*e.g.*, HANSKI and GILPIN 1997). In a pioneering work, SLATKIN (1977) extended the work of MARUYAMA (1970) and introduced a genetic metapopulation model with two different modes of recolonization: the propagule-pool model in which all colonists originate from a single, randomly chosen deme and the migrant-pool model, in which the colonists are drawn randomly from the whole population. The resulting pattern of genetic differentiation is complex and depends on the pattern of extinctions and recolonizations, since random extinctions can be considered as another form of genetic drift increasing differentiation between demes, while recolonizations can be seen as founder effects contributing to gene flow between demes, making them potentially more similar. PANNELL and CHARLESWORTH (1999) observed that colonization was more likely to follow the propagule-pool model in highly structured metapopulations, since colonists would tend to originate from neighboring colonies. WADE and McCAULEY (1988) reframed the central equations from SLATKIN's (1977) models in terms of F_{ST}, a measure of the genetic differentiation among demes. They showed that in SLATKIN's (1977) metapopulation model without mutation, population differentiation is indeed reduced when the number of colonists exceeds twice the number of migrants exchanged between demes. In all other cases, however, the process of extinction and recolonization increased population differentiation, and often very substantially. WHITLOCK and McCAULEY (1990) then generalized these recolonization models by introducing the probability of a common origin of colonizing gametes (Φ). In particular, they showed that population differentiation increases if the number of colonists is less than four times the number of migrants exchanged between demes when $\Phi = 0.5$. This is an interesting result, since several studies of natural populations (*e.g.*, WHITLOCK 1992a; INGVARSSON *et al.* 1997) estimated Φ to be ~ 0.5.

More recently, WAKELEY and ALIACAR (2001) and PANNELL (2003) studied the genealogical (coalescent)

structure of samples drawn from an island model with recurrent extinctions and recolonizations under the generalized recolonization process of WHITLOCK and McCAULEY (1990). In line with previous studies of simple finite-island models without extinctions (WAKELEY 1999), he showed that if the number of demes is large, there is a separation of timescales within the genealogical history of a sample, which consists of a brief "scattering" phase followed by a more ancient "collecting" phase that dominates the genealogical history. While the scattering phase is a stochastic and structured process shaping the pattern of variability within *vs.* between demes, the collecting phase is an unstructured coalescent process (KINGMAN 1982a,b) with an effective size determined by the migration and extinction–recolonization pattern of the metapopulation. While this approach leads to many insights about the effects of metapopulation dynamics on genetic diversity, it is not spatially explicit and assumes (like all previously presented metapopulation models) that all demes have the same size. In an early attempt to account for variability in population densities, WHITLOCK (1992b) studied the effect of variation in deme size, migration rate, and extinction rate on the probability of identity by descent. He concluded that this kind of temporal or spatial variation can significantly affect the amount of genetic variation among demes. WAKELEY (2001) also studied a finite-island model with variation in deme sizes and derived expressions for the effective populations size N_e. He showed that size variation among demes always reduces N_e because the decrease in coalescent times in small demes more than offsets its increase in large demes. WAKELEY and ALIACAR (2001) then analyzed this model with extinction and recolonization following the migrant-pool model of SLATKIN (1977) and showed that the expected number of segregating sites within deme was correlated with deme size and migration rate. Demes facing a higher extinction rate would also be less polymorphic unless the number of new incoming colonists was very large compared to the extinction rate.

Contrary to genetic metapopulation models where an arbitrary probability of extinction is the key parameter distinguishing metapopulation from other subdivided population models (LEVINS 1969; SLATKIN 1977), ecological metapopulation models are often spatially explicit and more realistic, as they often use environmental information like patch size and connectivity to compute extinction probabilities (see, *e.g.*, HANSKI and OVASKAINEN 2003). An example of such an ecological model is the incidence function model developed by HANSKI (1994), which uses deme area, deme quality, and deme coordinates to predict the dynamics or the fraction of occupied demes in any metapopulation. However, while incorporating more biological information than genetic metapopulation models, ecological metapopulation models do not predict patterns of genetic

diversity but rather metapopulation stability or the number of occupied demes.

In an attempt to explicitly link environmental information to genetic data, Currat *et al.* (2004) have developed a computer program (SPLATCHE) to simulate genetic data, taking spatial environmental heterogeneity into account. Under the SPLATCHE framework, the demographic and migration history of a set of subpopulations arranged on a two-dimensional lattice is first simulated and then used to generate the genealogy of genes sampled at one or several locations under a coalescent backward approach (Currat *et al.* 2004). The separation of demographic and genetic simulations allows one to include more realism into the demographic model, while keeping the genetic simulations simple and fast. Spatial environmental heterogeneity is handled by assigning demes to particular eco-vegetation types, which are assumed to sustain different numbers of individuals and where migration may be more or less difficult. One can thus explicitly associate a given carrying capacity and a given migration friction to each ecotype, which will influence the demographic history of the demes and therefore their expected neutral genetic diversity. Even though the current SPLATCHE framework is relatively flexible and can even deal with some form of temporal changes of the environment (Ray *et al.* 2003), it assumes that a given type of environment is uniformly occupied, which is quite unrealistic. It may indeed be more realistic to assume that a particular environment is not uniformly occupied by a given species, which could therefore have a patchy distribution due to chance alone. Even though a given environment may be perceived as uniform, it will almost always present microheterogeneities in the distribution of resources (*e.g.*, water, sunlight, nutrients, preys) or various physical conditions (*e.g.*, humidity, temperature), which should translate into different population densities at different locations. For instance, even though most populations living in unfavorable environments (*e.g.*, deserts) have low densities, a few populations may show much larger densities due to microclimatic fluctuations or access to rare resources (*e.g.*, water).

In this article, we extend the SPLATCHE framework (Currat *et al.* 2004), and we introduce a convenient way to model spatial and temporal environmental heterogeneity with just two new parameters. We model spatial heterogeneity of the environment by assuming that the carrying capacities of the demes found in a given type of environment follow a gamma distribution, and we use the shape parameter (α) of this distribution to control the degree of spatial heterogeneity. Temporal heterogeneity is controlled by a second parameter, which is the autocorrelation of carrying capacities over time (ρ). We then study the effect of various types of environmental heterogeneity on different aspects of molecular diversity within and between demes by extensive simulations.

MATERIALS AND METHODS

Our simulation framework extends that introduced by Currat *et al.* (2004), who proposed to simulate the effects of heterogeneous environments on the molecular diversity after a range expansion in two steps: a first forward step simulates population densities and migration events, and a second backward step simulates molecular diversity on the basis of the demography simulated in the first step using a coalescent approach. While this approach is more fully described elsewhere (*e.g.*, Ray *et al.* 2003; Currat *et al.* 2004), we briefly outline here these two steps.

Demographic simulations: Following Ray *et al.* (2003), we performed simulations on a two-dimensional grid of 50×50 demes, representing subpopulations (demes) of haploid individuals interconnected by migration. The colonization of the grid arbitrarily starts at the center of the lattice (at position $<25;25>$). Each generation, a growth phase is followed by an emigration phase from every occupied deme to neighboring demes with an emigration rate per gene lineage m, which may vary over simulations, but which is assumed constant over time and space for a given simulation. The density of each deme is logistically regulated, with an intrinsic growth rate r (we used here a constant r value of 0.5) and a carrying capacity K depending on the environment assigned to the deme. In our simulations, only one type of environment is used for the whole metapopulation. Note, however, that the carrying capacity of the ith deme, K_i, may vary according to spatial and temporal heterogeneity, as described later in this section. The number of emigrants sent by the ith deme during the migration phase is thus $N_{it}m$, where N_{it} is the local deme size at time t. The $N_{it}m$ emigrants are then distributed equally among the four neighboring demes. The density of the demes is then regulated in the next growth phase. The way migration is modeled here differs from Slatkin's (1977) model and its followers (see, *e.g.*, Pannell and Charlesworth 2000), where migration is considered as the proportion of individuals that are replaced by new immigrants in each deme. In our model, the number of emigrants depends on local deme size, whereas the number of immigrants depends on the size of surrounding demes. In the case of a spatial variation in deme size, this would result in small demes receiving more migrants than they send to larger neighbors, this unbalance being then logistically regulated. In the SPLATCHE framework, migration also depends on a friction parameter that may differ according to the environment, but for the sake of simplicity we have chosen here to assume a uniform friction in the whole environment, such that migration patterns will just depend on deme densities and emigration probability m per lineage.

To get results comparable to those of a previous study of a range expansion in a uniform environment (Ray *et al.* 2003), we simulated the same expansion time of 4000 generations in the past. This time was then arbitrarily selected because these authors were interested in human evolution and 4000 generations roughly correspond to the emergence of modern humans ~120,000 years ago. However, results for expansion times of 1000 or 10,000 generations are reported in supplemental Tables S1–S4 at http://www.genetics.org/supplemental/. Note, however, that we did not consider very recent range expansions, because up to 1000 generations are necessary to colonize our simulated world, such that our results would not necessarily apply to the case of invasive species. We always performed 1000 demographic simulations per set of parameters.

Genetic simulations: The program SPLATCHE was modified to simulate the genealogy of sampled genes to take environmental heterogeneity into account as described below. The genealogy of a sample of 30 genes in a deme located at

position ⟨5;5⟩ was simulated to study intrademe DNA sequence diversity, and we allowed for multiple coalescent events per generation per deme. Mutations were assumed to follow a Poisson process with rate $\mu = 0.001$ for a sequence of 300 bp in length, such that four mutations were expected to have occurred since the expansion along any lineage, which roughly corresponds to the range of values observed for human mtDNA HV1 sequences (*e.g.*, Santos *et al.* 2005). Several summary statistics were computed on the simulated DNA sequences using the software package ARLEQUIN ver. 3 (Excoffier *et al.* 2005): (i) measures of genetic diversity within deme: average number of pairwise differences π and number of segregating sites S; (ii) measures of genetic diversity among demes: F_{ST} and average number of pairwise differences between demes π_{xy}; as well as (iii) statistics to detect departure from mutation–drift equilibrium: Tajima's D (Tajima 1989) and Fu's F_S (Fu 1997). To calculate F_{ST} and π_{xy} values, we sampled 30 additional genes at position ⟨25;45⟩ and we calculated F_{ST} between these two demes as θ_W (Weir and Cockerham 1984). Because the number of demes is high (2500), the average number of pairwise differences between demes π_{xy} is equivalent to the expected number of pairwise differences between two randomly chosen genes in the population. We always performed 10 genetic simulations per demographic simulation, such that a total of 10,000 genetic simulations were performed for each set of parameters.

Spatially heterogeneous environments: We assume that differences in environmental conditions affect the carrying capacity (K) of the demes. While a heterogeneity in carrying capacities over space could certainly result from adaptive or selective processes, we concentrate on neutral diversity at an unlinked marker, the diversity of which depends only on population demography. Environmental heterogeneity was thus modeled by assuming that the carrying capacity of each deme was a random variable drawn from a gamma distribution with mean $\overline{K}$, shape parameter α, and variance $\overline{K}^2/\alpha$. The gamma distribution is commonly used to reflect heterogeneous processes like spatial heterogeneity of mutation rates along chromosomes (*e.g.*, Nei *et al.* 1976; Wakeley 1993; Aris-Brosou and Excoffier 1996). In the context of environmental heterogeneity, small α-values correspond to environments that are spatially more heterogeneous than those simulated with larger α-values. We considered here four different degrees of spatial heterogeneity: large ($\alpha = 0.5$), intermediate ($\alpha = 1$), small ($\alpha = 5$), and none ($\alpha = \infty$, implying that $K_i = \overline{K}, \forall i$). The distributions of the resulting carrying capacities for $\overline{K} = 100$ are reported in Figure 1, as well as random examples of their corresponding spatial distribution. As can be seen in Figure 1, α-values <1 lead to L-shaped distributions of carrying capacities, indicating that most of the demes have very small carrying capacities, while a few may have large densities, which could be considered as density hotspots. Such a distribution could correspond to population densities observed in very inhospitable habitats such as deserts, in unsaturated environments, or in mixed environments such as cultivated landscapes, with a very patchy distribution of resources and individuals. α-values >1 lead to bell-shaped distributions of carrying capacities and would correspond to mildly heterogeneous and saturated environments, where resources are more evenly distributed. Our simulations results should therefore cover a wide range of situations for species living in very different environments. Globally, the proportion of empty demes (indicated by open squares in Figure 1B) is largest for highly heterogeneous environments. While the simulated samples are always drawn from the same deme across simulations, the initial size (going backward in time) of this deme varies over simulations. Cases where the initial deme size was zero have been discarded. Note also that we did not
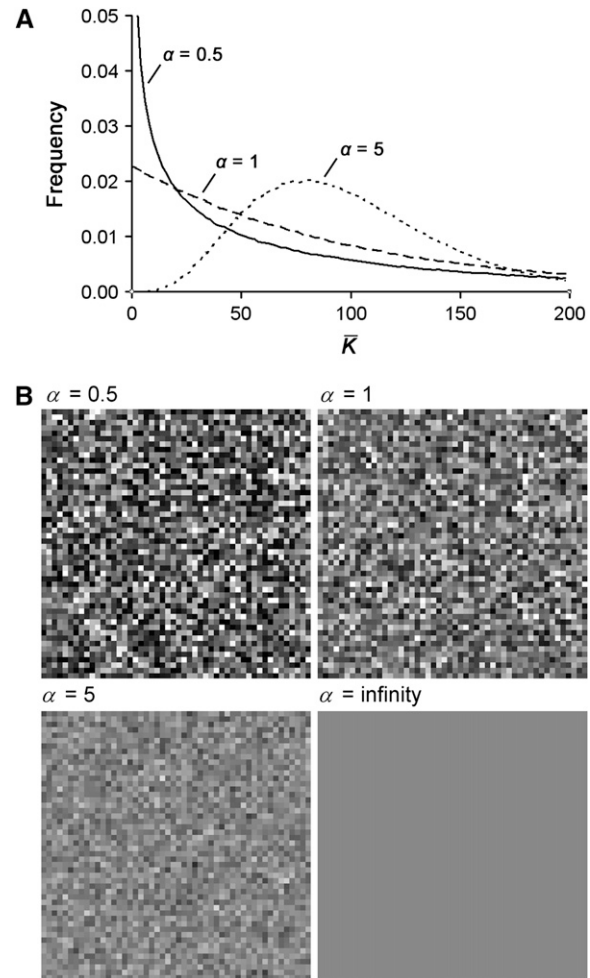

Figure 1.—Distribution of carrying capacities (K) for different levels of spatial heterogeneity. (A) Gamma distributions of carrying capacities for different values of the shape parameter α. (B) Examples of random spatial distribution of K for different values of α. From the top left to the bottom right α was set to 0.5, 1, 5, and infinity, respectively. $\overline{K} = 100$.

assume any spatial autocorrelation between the carrying capacities of neighboring demes.

Temporally heterogeneous environments: Temporal heterogeneity was incorporated into the demographic simulations by changing the carrying capacity of each deme every generation, while keeping the mean carrying capacity $\overline{K}$ of the demes constant over time. We propose the autocorrelation coefficient (ρ) of the carrying capacities between generations as a measure of the degree of temporal heterogeneity. Small ρ-values therefore indicate a large temporal variability, whereas large ρ-values indicate small temporal changes in carrying capacities for a given deme. Note that temporally stable environments can be simulated by setting ρ to 1. For fixed α- and ρ-values, the absolute magnitude of the temporal change in carrying capacity of a deme will depend on its current value. It implies that the probability for a deme to go extinct will depend on its current size and will thus be larger in currently small than in currently large demes, which seems realistic. This situation differs from a previous genetic metapopulation model (*e.g.*, Slatkin 1977), where the extinction probability is assumed constant over time and identical for all demes. Our model is therefore closest to a conventional genetic metapopulation model when we assume very low ρ-values ($\rho \rightarrow 0$),

<div align="center">

TABLE 1

Pattern of molecular diversity in spatially heterogeneous environments

</div>

m	$\overline{K}$	$\overline{K}m$	α	π^a	CV (π)	S^b	CV (S)	F_{ST}	CV (F_{ST})	$\pi_{xy}{}^c$	CV (π_{xy})
0.05	50	2.5	∞	4.92	0.34	18.79	0.31	0.353	0.294	8.04	0.22
			5	4.40	0.40	16.21	0.35	0.415	0.292	8.00	0.24
			1	2.95	0.60	10.55	0.50	0.602	0.256	7.94	0.28
			0.5	2.24	0.72	8.13	0.59	0.695	0.226	7.65	0.31
	100	5	∞	6.05	0.25	27.49	0.25	0.220	0.310	8.03	0.18
			5	5.71	0.27	24.57	0.27	0.259	0.316	8.04	0.19
			1	4.55	0.37	17.83	0.35	0.402	0.320	8.00	0.23
			0.5	3.65	0.48	14.18	0.43	0.515	0.299	7.88	0.26
	200	10	∞	6.84	0.19	39.12	0.21	0.131	0.322	8.03	0.14
			5	6.63	0.21	35.82	0.22	0.155	0.334	8.04	0.15
			1	5.88	0.26	27.72	0.28	0.246	0.354	8.02	0.18
			0.5	5.12	0.33	22.84	0.35	0.342	0.356	8.02	0.21
	1000	50	∞	7.72	0.12	72.40	0.13	0.034	0.504	8.04	0.09
			5	7.65	0.13	68.93	0.14	0.042	0.501	8.05	0.10
			1	7.41	0.15	59.19	0.17	0.071	0.498	8.05	0.11
			0.5	7.10	0.17	51.79	0.21	0.108	0.489	8.06	0.13
0.2	50	10	∞	6.66	0.21	34.53	0.22	0.139	0.321	8.02	0.15
			5	6.38	0.23	30.74	0.24	0.169	0.326	8.01	0.16
			1	5.29	0.32	21.46	0.31	0.291	0.339	8.03	0.20
			0.5	4.14	0.45	15.62	0.40	0.435	0.325	7.97	0.24
	100	20	∞	7.25	0.17	47.87	0.18	0.079	0.318	8.03	0.12
			5	7.09	0.18	43.76	0.20	0.096	0.329	8.03	0.13
			1	6.39	0.23	32.01	0.25	0.174	0.351	8.03	0.16
			0.5	5.47	0.30	23.91	0.32	0.276	0.368	8.03	0.19
	200	40	∞	7.61	0.14	61.93	0.15	0.045	0.344	8.04	0.10
			5	7.50	0.15	58.06	0.16	0.053	0.349	8.02	0.11
			1	7.08	0.18	45.30	0.20	0.098	0.361	8.03	0.13
			0.5	6.45	0.23	34.95	0.27	0.168	0.391	8.04	0.16
	1000	200	∞	7.94	0.11	87.33	0.11	0.010	0.573	8.04	0.08
			5	7.90	0.11	85.21	0.11	0.013	0.537	8.03	0.08
			1	7.80	0.12	76.81	0.13	0.025	0.511	8.04	0.09
			0.5	7.59	0.13	67.09	0.16	0.045	0.537	8.03	0.10

In all cases the expansion started 4000 generations ago. The statistics summarizing genetic diversity within demes were obtained by simulating a sample of 30 DNA sequences of 300 bp in a single deme at lattice position $<5;5>$. For the calculation of the F_{ST}, we sampled genetic diversity in another deme at position $<25;45>$ on the lattice. CV, the coefficient of variation defined as the standard deviation divided by the mean.

^{a} Mean number of differences between all pairs of sequences in a sample.
^{b} Number of segregating sites.
^{c} Mean number of differences between all pairs of sequences between two samples.

where the probability of extinction is not correlated to deme size and is equal to the expected proportion of empty demes for given α- and $\overline{K}$-values. In temporally stable environments ($\rho \approx 1$), only demes with currently small carrying capacities face the risk of extinction and empty demes will stay empty longer, since they cannot be recolonized as long as their carrying capacity is zero.

RESULTS

The results of our simulations are divided into three sections. We first present results obtained in the absence of environmental heterogeneity and outline here several aspects of DNA sequence diversity within and between samples. We then present comparable results under a model with spatial heterogeneity and finally results obtained with both spatial and temporal environmental heterogeneity.

Constant environment: The pattern of genetic diversity within and between demes is reported in Table 1 for various amounts of spatial heterogeneity, various average carrying capacities ($\overline{K}$), and two different rates of migration ($m = 0.05$ and $m = 0.2$). When considering

homogeneous environments ($\alpha = \infty$), we find results equivalent to those described in RAY *et al.* (2003). Measures of intrademe diversity, like the average number of pairwise differences (π) and the number of polymorphic sites (S), increase with the number of migrants exchanged between neighboring demes ($\overline{K}m$), while their coefficient of variation decreases. On the other hand, demes become less differentiated (as measured by F_{ST}) with increasing $\overline{K}m$ values. Note that in all cases the overall genetic diversity as measured by π_{xy} is not much affected by the different demographic parameters (but its coefficient of variation is slightly larger for small $\overline{K}m$ values), suggesting that the increase of F_{ST} with a lower number of migrants exchanged between demes is primarily due to a reduction in genetic diversity within deme. Note, however, that we have measured F_{ST} by considering two demes geographically quite far from each other. The coalescence time of two genes drawn from these two demes will always be close to the expansion time, irrespective of the heterogeneity of the environment. It is possible that the coalescence time of two genes drawn from two geographically close demes would be more affected by environmental heterogeneity and therefore that F_{ST} would also depend on the level of genetic diversity between demes in that case.

In Table 2 we present the performance of two statistics used to detect departure from selective neutrality and population equilibrium: Tajima's D and Fu's F_S. In homogeneous environments ($\alpha = \infty$), both statistics change from very significant negative values when $\overline{K}m$ is large (>20) to mostly nonsignificant negative values when $\overline{K}m \leq 20$, even reaching positive values when $\overline{K}m < 10$. These results are again consistent with those from RAY *et al.* (2003).

Spatial heterogeneity: Results on the amount of genetic diversity within and between populations after an expansion in a heterogeneous environment are presented in Table 1. We find that an increase in the degree of environmental heterogeneity has globally the same effect for different sets of $\overline{K}$- and m-values: with decreasing α-values, genetic diversity within sample is reduced (π and S decrease), while it is more variable between replicates (the coefficients of variation of π and S increase with α), and F_{ST} is increased while its coefficient of variation remains relatively unaffected by α. However, these effects appear much less marked for large than for small $\overline{K}m$ values. Interestingly, F_{ST} levels seem more affected by average levels of gene flow (as measured by $\overline{K}m$) than by environmental heterogeneity: for instance, F_{ST} changes from 0.01 for $\overline{K}m = 200$ to 0.35 for $\overline{K}m = 2.5$ in a constant environment, while F_{ST} is increased only by a factor of 2–4 for a fixed $\overline{K}m$ value between a homogeneous ($\alpha = \infty$) and a highly heterogeneous environment ($\alpha = 0.5$) (Table 1). Note that π_{xy} is also marginally affected by the degree of environmental heterogeneity. While the age of the expansion will affect the extent and pattern of molecular diversity,

the effect of environmental heterogeneity is qualitatively similar for different expansion times and is not discussed hereafter (see supplemental Tables S1–S4 at http://www.genetics.org/supplemental/, reporting results analogous to those in Tables 1 and 2 for expansion times of 1000 or 10,000 generations).

In Figure 2, we report F_{ST} values as a function of the average carrying capacity of the environment for different degrees of spatial heterogeneity, but for a constant $\overline{K}m$ value of 10. As already shown in Table 1, the level of genetic structure is lowest for the homogeneous environment and increases with environmental heterogeneity. However, Figure 2 shows that F_{ST} is also very sensitive to the average size of the demes in heterogeneous environments, which is much less the case in homogeneous environments. The level of genetic structure actually increases with lower $\overline{K}$-values, and thus F_{ST} does not simply depend on the product $\overline{K}m$, as shown previously in a uniform and spatially unstructured environment (EXCOFFIER 2004).

The variability of Tajima's D- and Fu's F_S-statistics computed under scenarios with environmental heterogeneity is reported in Table 2. Both statistics are shifted toward less negative or more positive values with increasing levels of environmental heterogeneity, and they tend to become less often significant. For instance, for $\overline{K}m = 40$ the power of Tajima's D-statistic to detect past expansion is reduced from 96% in homogeneous environments down to only 13% in heterogeneous environments with $\alpha = 0.5$.

Temporal heterogeneity: In Table 3, we report the effect of different levels of temporal heterogeneity (ρ) on genetic diversity within samples for the case of $\overline{K}m = 40$, because it is for this case that Fu's F_S- and Tajima's D-statistics show the largest sensitivity to spatial heterogeneity (Table 2). We can see that temporal heterogeneity does not affect much molecular diversity when spatial heterogeneity is weak ($\alpha = 5$) or intermediate ($\alpha = 1$). It has, however, a marked effect when environmental heterogeneity is strong ($\alpha = 0.5$), where π and S drop sharply in highly changing environments ($\rho < 0.4$). It seems that S is more affected by temporal heterogeneity than π, since there is only a 14% decrease in π (6.5 to 5.6) when the autocorrelation coefficient (ρ) changes from 1 to 0, while S decreases by 38% (35 to 22 segregating sites) when ρ goes to 0. Interestingly, the coefficients of variation of S and π seem almost unaffected by temporal heterogeneity for all levels of spatial heterogeneity considered here. Qualitatively similar results are shown for $\overline{K}m = 10$ in supplemental Table S5 at http://www.genetics.org/supplemental/.

In highly fluctuating environments (*i.e.*, when $\rho < 0.4$ and $\alpha = 0.5$ in Table 3), genetic diversity within demes is considerably reduced, while the overall genetic diversity as measured by π_{xy} is only slightly decreased, resulting in a sharp increase in F_{ST}. We therefore see that increasing temporal heterogeneity has a similar effect as increasing

TABLE 2

Summary statistics used to detect departure from population equilibrium computed within demes drawn from a spatially heterogeneous environment

m	$\overline{K}$	$\overline{K}m$	α	$F_S{}^a$	$\Pr(F_S) < 0.05{}^b$	D^c	$\Pr(D) < 0.05{}^d$
0.05	50	2.5	∞	3.02	0.00	0.18	0.02
			5	3.14	0.00	0.29	0.03
			1	2.74	0.01	0.32	0.07
			0.5	2.03	0.03	0.22	0.08
	100	5	∞	1.38	0.00	−0.40	0.02
			5	1.65	0.00	−0.22	0.02
			1	2.16	0.00	0.09	0.03
			0.5	1.86	0.01	0.10	0.05
	200	10	∞	−0.81	0.02	−1.09	0.18
			5	−0.44	0.02	−0.93	0.11
			1	0.36	0.01	−0.50	0.04
			0.5	0.56	0.02	−0.30	0.04
	1000	50	∞	−11.06	0.97	−2.18	1.00
			5	−10.01	0.95	−2.10	1.00
			1	−7.65	0.79	−1.87	0.92
			0.5	−6.46	0.64	−1.65	0.72
0.2	50	10	∞	−0.02	0.01	−0.81	0.07
			5	0.52	0.00	−0.59	0.03
			1	1.86	0.00	−0.03	0.02
			0.5	2.22	0.01	0.20	0.04
	100	20	∞	−2.73	0.17	−1.46	0.51
			5	−2.04	0.10	−1.29	0.33
			1	−0.22	0.02	−0.70	0.06
			0.5	0.84	0.01	−0.26	0.03
	200	40	∞	−6.53	0.74	−1.91	0.96
			5	−5.66	0.62	−1.81	0.89
			1	−3.11	0.24	−1.37	0.43
			0.5	−1.50	0.11	−0.90	0.13
	1000	200	∞	−18.62	1.00	−2.43	1.00
			5	−17.70	1.00	−2.40	1.00
			1	−14.38	1.00	−2.26	1.00
			0.5	−11.39	0.95	−2.06	0.98

In all cases the expansion started 4000 generations ago.
[a] Fu's F_S-statistic (FU 1997).
[b] Probability that Fu's F_S-statistic is significant at the 5% level, as estimated from 10,000 simulations.
[c] Tajima's D-statistic (TAJIMA 1989).
[d] Probability that Tajima's D-statistic is significant at the 5% level, as estimated from 10,000 simulations.

spatial heterogeneity. Note that while it is quite difficult to strictly compare our model to previous metapopulation models, the observed magnitude of the increase in F_{ST} due to temporal changes in spatially highly heterogeneous environments is similar to the predictions of WADE and MCCAULEY (1988). For instance, in the case of $\overline{K} = 200$, $\rho = 0$, and $\alpha = 0.5$, we observe that 5.6% of the demes go extinct per generation on average and this extinction rate is expected to result in a 50% increase in F_{ST} under WADE and MCCAULEY's (1988) migrant-pool colonization model as compared to the no-extinction case, whereas our simulations show a 46% increase (from 0.168 for $\rho = 1$ to 0.248 for $\rho = 0$). So, despite the spatially explicit context of our simulations and the assumed past expansion, our results are in relatively good agreement with former predictions. However, our simulations reveal a previously unnoted behavior for spatially very heterogeneous but temporally slightly changing environments. Indeed, as compared to temporally stable environments, we find for $\rho = 0.8$ and $\alpha = 0.5$ (see
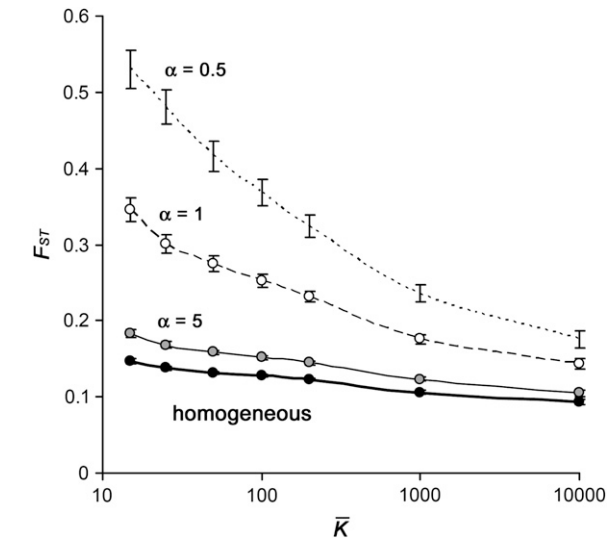
FIGURE 2.—F_{ST} values between two distant demes (see text) drawn from environments with various degrees of heterogeneity and different average carrying capacities $\overline{K}$. Note that the average number of genes exchanged between neighboring demes ($\overline{K}m$) is kept constant and equal to 10 in all cases.

Table 3) that genetic diversity within deme (π) increases and that F_{ST} decreases, before increasing again for lower ρ-values.

In Table 4, we report the influence of temporal heterogeneity on the performance of Fu's F_S- and Tajima's D-statistics. When environmental heterogeneity is strong or intermediate, both statistics are shown to change from negative values in the case of a temporally constant environment to less negative or even positive

values when $\rho = 0$. In the case of spatially highly heterogeneous environments ($\alpha = 0.5$) the power of these statistics to detect departure from population equilibrium is considerably lowered in the presence of temporal heterogeneity, since the probability to get significant statistics is considerably less than the significance level and even drops to zero for Fu's F_S when $\rho \le 0.4$. For a smaller $\overline{K}m$ value of 10, these statistics will have even lower power to evidence past range expansion (see supplemental Table S6 at http://www.genetics.org/supplemental/).

Effect of environmental heterogeneity on mismatch distributions: We report empirical distributions of the number of differences between pairs of genes (mismatch distributions) in Figure 3 for a subset of the cases described in Table 1. As reported previously (RAY *et al.* 2003), the expected mismatch distributions are unimodal for ($\overline{K}m > 50$) and this mode is indicative of the age of the spatial expansion (SLATKIN and HUDSON 1991; ROGERS and HARPENDING 1992; EXCOFFIER 2004). When the environment is homogeneous, expected mismatch distributions are clearly bimodal for $\overline{K}m$ values <50, with the first mode at zero due to recent coalescent events (Figure 3). We find overall that spatial heterogeneity affects mainly mismatch distributions if $\overline{K}m$ is also <50 (Figure 3). In that case, the first mode of the mismatch distribution becomes more important with increasing levels of spatial heterogeneity, which is again due to an increase in the proportion of recent coalescent events. We also note a shift of the second mode toward lower values with increasing environmental heterogeneity, suggesting that coalescent events occur just

TABLE 3

Pattern of molecular diversity in temporally heterogeneous environments

α	ρ	π^a	CV (π)	S^b	CV (S)	F_{ST}	CV (F_{ST})	$\pi_{xy}{}^c$	CV (π_{xy})
0.5	1	6.45	0.23	34.95	0.27	0.168	0.391	8.04	0.16
	0.8	6.60	0.21	33.93	0.23	0.140	0.329	7.95	0.15
	0.6	6.47	0.22	31.83	0.24	0.154	0.334	7.97	0.15
	0.4	6.27	0.24	29.12	0.25	0.174	0.332	7.96	0.16
	0.2	5.95	0.26	25.59	0.27	0.206	0.335	7.94	0.17
	0	5.57	0.29	22.04	0.29	0.246	0.328	7.93	0.19
1	1	7.08	0.18	45.30	0.20	0.098	0.361	8.03	0.13
	0.8	7.20	0.17	46.60	0.19	0.083	0.319	8.01	0.12
	0.6	7.18	0.17	45.87	0.19	0.086	0.325	8.02	0.12
	0.4	7.14	0.17	44.78	0.19	0.089	0.323	8.00	0.12
	0.2	7.07	0.18	43.02	0.20	0.095	0.325	8.00	0.13
	0	7.02	0.19	41.78	0.20	0.100	0.317	8.00	0.13
5	1	7.50	0.15	58.06	0.16	0.053	0.349	8.02	0.11
	0.8	7.53	0.15	58.82	0.16	0.051	0.329	8.02	0.10
	0.6	7.53	0.15	58.89	0.16	0.050	0.329	8.01	0.10
	0.4	7.53	0.15	58.54	0.16	0.051	0.336	8.02	0.11
	0.2	7.54	0.15	58.62	0.16	0.051	0.327	8.03	0.10
	0	7.53	0.15	58.46	0.16	0.052	0.325	8.03	0.11

Simulation conditions are as in Table 1, but here $\overline{K} = 200$ and $m = 0.2$ in all cases.

TABLE 4

Summary statistics used to detect departure from population equilibrium computed within demes drawn from a spatially and temporally heterogeneous environment

α	ρ	F_S	$\Pr(F_S) < 0.05$	D	$\Pr(D) < 0.05$
0.5	1	−1.50	0.11	−0.90	0.13
	0.8	−0.05	0.01	−0.78	0.06
	0.6	0.38	0.01	−0.65	0.04
	0.4	0.89	0.00	−0.47	0.02
	0.2	1.58	0.00	−0.22	0.01
	0	2.37	0.00	0.06	0.01
1	1	−3.11	0.24	−1.37	0.43
	0.8	−2.60	0.15	−1.41	0.45
	0.6	−2.37	0.13	−1.38	0.41
	0.4	−2.15	0.11	−1.33	0.37
	0.2	−1.80	0.08	−1.25	0.29
	0	−1.51	0.06	−1.20	0.25
5	1	−5.66	0.62	−1.81	0.89
	0.8	−5.63	0.61	−1.83	0.91
	0.6	−5.65	0.62	−1.83	0.91
	0.4	−5.54	0.60	−1.82	0.91
	0.2	−5.52	0.60	−1.82	0.90
	0	−5.50	0.59	−1.82	0.90

Simulation conditions are as in Table 2.

after the onset of the expansion and thus probably during the range expansion and not mainly when genes are found in the ancestral deme as in homogeneous environments (RAY *et al.* 2003; CURRAT and EXCOFFIER 2005). When $\overline{K}m > 50$, the mismatch distribution is only weakly affected by any degree of spatial heterogeneity.

In Figure 4, we show the mismatch distributions expected after a range expansion in spatially heterogeneous environments for different degrees of temporal heterogeneity ρ. Temporal and spatial heterogeneity have also similar effects on the mismatch distribution,

as pairs of genes have more recent ancestors in temporally more fluctuating environments. This pattern is, however, found only in spatially highly heterogeneous environments (Figure 4, $\alpha = 0.5$), since we find no difference between the mismatch distributions taken from samples in spatially more homogeneous environments ($\alpha = 5$, Figure 4). Finally, we note that temporal heterogeneity also adds to the variance of the mismatch distribution when environmental heterogeneity is not weak (if $\alpha \leq 1$, see Figure 5).

DISCUSSION

Effect of environmental heterogeneity on genetic diversity: We have shown that levels of genetic diversity both within and between demes are affected by environmental heterogeneity after a relatively old range expansion. While genetic diversity is smaller in samples taken in more heterogeneous environments, the overall degree of genetic structure is usually stronger. These results can be understood by considering two factors that are affected by environmental heterogeneity. First, the total number of demes where gene lineages can actually migrate will decrease with larger environmental heterogeneity, implying that pairs of genes will have a larger chance to occur in the same deme and thus to coalesce with smaller α-values. Second, the rate of a coalescent event within deme is increased in environments with larger environmental heterogeneity, because the increase in coalescent probability in small demes is more than compensated by its decrease in large demes (WAKELEY and ALIACAR 2001). Therefore, increasing environmental heterogeneity has the same effect as a decrease in the Km value in a homogeneous environment. Note, however, that there is one exception to this general rule, which occurs in spatially highly heterogeneous environments, where low levels of temporal
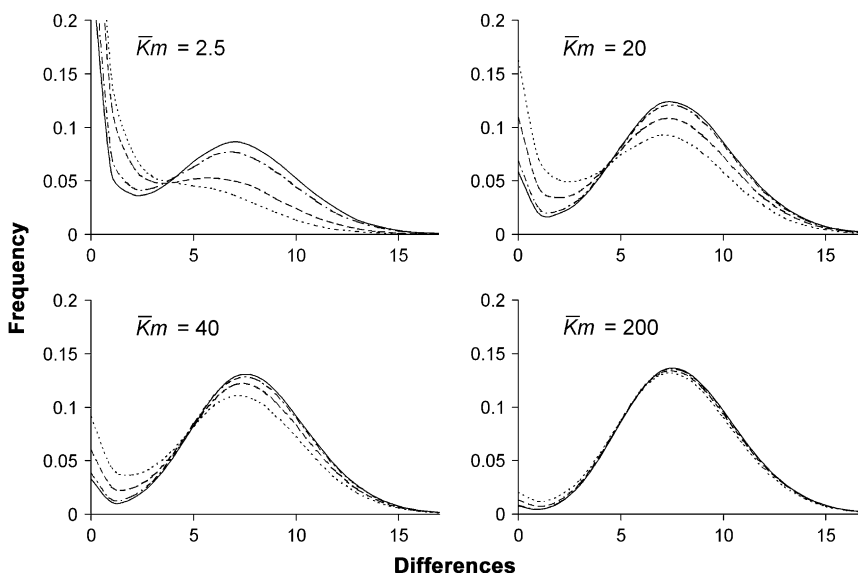

FIGURE 3.—Expected mismatch distributions for different degrees of spatial heterogeneity α and for different $\overline{K}m$ values. The four different lines correspond to $\alpha = $ infinity (—), $\alpha = 5$ (– · –), $\alpha = 1$ (– – –), and $\alpha = 0.5$ (- - -).
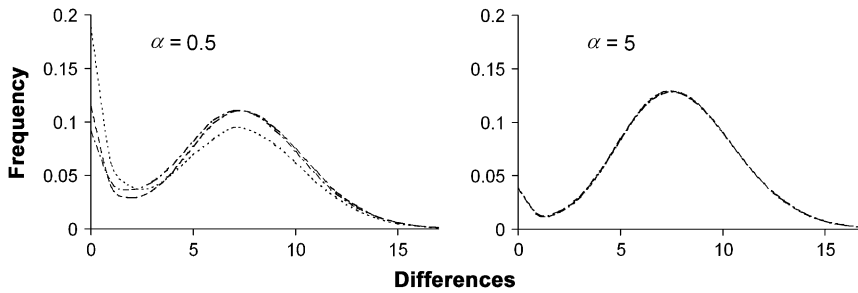
FIGURE 4.—Expected mismatch distributions for different degrees of temporal heterogeneity ρ. The three different lines correspond to $\rho = 1$ (– · –), $\rho = 0.6$ (– – –), and $\rho = 0$ (- - -).

heterogeneity lead to a slight decrease in F_{ST}. In that case, it seems that slight temporal deme size fluctuations can prevent some recent coalescent events from happening as small demes can have their size increased with temporal heterogeneity. Therefore, slight temporal heterogeneity leads to a virtual homogenization of a highly fragmented and heterogeneous landscape.

Contrary to what we found in homogeneous environments (RAY *et al.* 2003; EXCOFFIER 2004), the pattern of diversity is not just sensitive to the average number of genes exchanged between neighboring demes ($\overline{K}m$), but it is highly affected by the degree of environmental heterogeneity (Table 1, Figure 2). F_{ST} also depends strongly on $\overline{K}$ and m separately in highly heterogeneous environments and very little in a homogeneous environment. For instance, when $\alpha \leq 1$, F_{ST} is much larger for small ($\overline{K} = 20$) than for large ($\overline{K} = 10,000$) average deme sizes (Figure 5). In more homogeneous environments, the separate effects of $\overline{K}$ and m tend to be very small on F_{ST} (Figure 5), but it also affects patterns of diversity within demes (see Table 1 for two distinct entries with $\overline{K}m = 10$ and $\alpha = \infty$). Note that the relationship between genetic diversity and absolute deme size was previously unnoted in homogeneous environments (RAY *et al.* 2003), since deme sizes >100 were always used in previous simulation studies and this relationship is barely visible for large deme size. Overall, our results suggest that F_{ST} can reach high values in a structured population where demes exchange a large number of migrants if environmental heterogeneity contributes to a large variance in deme size. This effect is reinforced by additional temporal heterogeneity, such that the relationship between F_{ST} and migration rate in realistic

and heterogeneous environments is far more complex than its expectation under the infinite-island model (WADE and MCCAULEY 1988; WHITLOCK and MCCAULEY 1999).

While our implementation of immigration rate depending on the size of the surrounding demes is certainly more realistic than previous approaches, it has the disadvantage that any variation in deme sizes will affect immigration rates, making it impossible to separate the effect of these two forces on genetic diversity. Under a finite-island model, WAKELEY (2001) found that in contrast to a variation in deme sizes that always reduces the effective size of the population, variation in migration rates would decrease the effective size if $4\overline{K}m > 1$ and increase it otherwise. Since most of our simulations correspond to cases where $4\overline{K}m \gg 1$ we expect that variation in immigration rates will reduce the effective size of our metapopulation and thus enhance the effect due to the variation in deme sizes.

The role of extinctions: In our model, extinctions simply follow from temporal fluctuations in deme densities. Extinction rates depend on the global degree of temporal and spatial environmental heterogeneity, as well as the current densities of particular demes. The probability of going extinct is thus larger for small than for large demes, unless there is no correlation in deme sizes between generations. This implementation differs from previous genetic metapopulation models (*e.g.*, SLATKIN 1977; MARUYAMA and KIMURA 1980; WAKELEY and ALIACAR 2001), where the probability of extinction was assumed constant over space and time. However, our results suggest that extinction is not a key parameter to predict a pattern of molecular diversity, since we
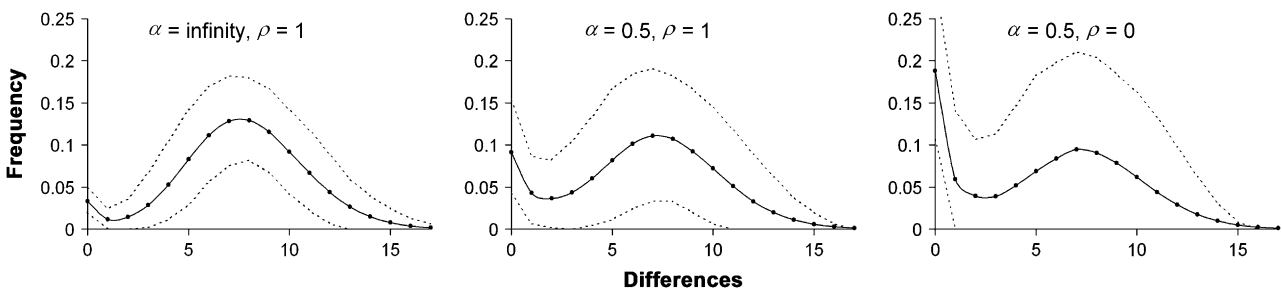

FIGURE 5.—Expected mismatch distributions after a spatial range expansion for different amounts of spatial and temporal heterogeneity. The dashed lines represent the 10 and 90% quantiles of the mismatch distribution. In all cases $\overline{K}m = 40$.

find that temporal and spatial heterogeneities have virtually the same effect (see Tables 1 and 3). Any local reduction in genetic diversity due to temporal heterogeneity added to spatial heterogeneity could thus be obtained by a more pronounced spatial heterogeneity, without any extinction. This suggests that extinction rates should be extremely difficult to estimate in practice if there is some degree of environmental heterogeneity.

The results we obtain under our metapopulation model after a range expansion differ from previous ones by two other aspects. First, a study of an island model (PANNELL and CHARLESWORTH 1999) suggested that population turnover would significantly affect genetic diversity only if the extinction rate was of the same order of magnitude or greater than the migration rate. We actually find that genetic diversity is clearly affected even for extinction rates much smaller than the migration rate. If we estimate the extinction rate as the proportion of empty demes when $\rho = 0$, then for the case of $\overline{K} = 200$ and $m = 0.2$, extinction rates (ε) ≈ 0.056 and 0.005 for $\alpha = 0.5$ and 1, respectively. From Table 3, it is clear that all aspects of genetic diversity within and between demes are affected by temporal heterogeneity for $\alpha = 0.5$ ($\varepsilon / m = 0.28$), and π and S are found slightly lower than in the absence of extinctions for $\alpha = 1$ ($\varepsilon / m = 0.025$). Second, the increased level of genetic structure measured by F_{ST} observed in previous metapopulation models (*e.g.*, WADE and MCCAULEY 1988; WHITLOCK and MCCAULEY 1990) can be attributed to the effect of extinction and recolonization on the total amount of diversity maintained in the metapopulation (PANNELL and CHARLESWORTH 1999). However, in our case the total amount of diversity (as measured by π_{xy} in Tables 1 and 3) is primarily affected by the expansion time and is relatively insensitive to spatial or temporal heterogeneity, and therefore the increase in F_{ST} is primarily due to a decrease in genetic diversity within deme with increasing environmental heterogeneity.

Detecting range expansions: Several statistics and approaches have been proposed to detect past range expansions from patterns of molecular diversity (HARPENDING *et al.* 1998; KIMMEL *et al.* 1998; RAY *et al.* 2003; EXCOFFIER 2004; HAMILTON *et al.* 2005). The typical signature of such expansions consists of long external branches of gene genealogies (SLATKIN and HUDSON 1991; ROGERS and HARPENDING 1992) leading, for instance, to large negative Fu's F_S- and Tajima's D-statistics. In homogeneous environments, these signatures of population range expansion are observed provided that the number of migrants (Km) exchanged between neighboring demes is $> \sim 20$ (RAY *et al.* 2003), while in heterogeneous environments, we find that these signatures are visible only for Km values >50. In highly heterogeneous environments ($\alpha = 0.5$), Fu's F_S- and Tajima's D-statistics can even become positive for $\overline{K}m \leq 20$ (see Table 1), which is usually indicative of an excess of recent coalescent events and therefore of recent bottlenecks. This

burst of recent coalescent events, which certainly occurs in demes of small sizes, is confirmed when considering mismatch distributions, as the fractions of pairs of genes showing no differences increase for small α-values when $\overline{K}m < 50$ (Figure 2). Spatial and demographic bottlenecks thus make it more difficult to evidence past range expansions from molecular data in heterogeneous than in homogeneous environments. The use of these statistics and F_{ST} to detect selection, such as in genome scans (see, *e.g.*, STORZ *et al.* 2004; TESHIMA *et al.* 2006), may also be problematic. Test based on Fu's F_S or Tajima's D-statistics would be expected to show many false negative results in the case of environmental heterogeneity, if it is not explicitly taken into account, while tests based on F_{ST} (BEAUMONT 1999; BEAUMONT and BALDING 2004) should show an excess of false positives. A future challenge would certainly be to estimate the extent of environmental heterogeneity from molecular data.

LITERATURE CITED

ARIS-BROSOU, S., and L. EXCOFFIER, 1996 The impact of population expansion and mutation rate heterogeneity on DNA sequence polymorphism. Mol. Biol. Evol. **13:** 494–504.

AUSTERLITZ, F., B. JUNG-MULLER, B. GODELLE and P.-H. GOUYON, 1997 Evolution of coalescence times, genetic diversity and structure during colonization. Theor. Popul. Biol. **51:** 148–164.

BEAUMONT, M. A., 1999 Detecting population expansion and decline using microsatellites. Genetics **153:** 2013–2029.

BEAUMONT, M. A., and D. J. BALDING, 2004 Identifying adaptive genetic divergence among populations from genome scans. Mol. Ecol. **13:** 969–980.

CAVALLI-SFORZA, L. L., and M. W. FELDMAN, 2003 The application of molecular genetic approaches to the study of human evolution. Nat. Genet. **33:** 266–275.

CURRAT, M., and L. EXCOFFIER, 2005 The effect of the Neolithic expansion on European molecular diversity. Proc. Biol. Sci. **272:** 679–688.

CURRAT, M., N. RAY and L. EXCOFFIER, 2004 SPLATCHE: a program to simulate genetic diversity taking into account environmental heterogeneity. Mol. Ecol. Notes **4:** 139–142.

EXCOFFIER, L., 2004 Patterns of DNA sequence diversity and genetic structure after a range expansion: lessons from the infinite-island model. Mol. Ecol. **13:** 853–864.

EXCOFFIER, L., G. LAVAL and S. SCHNEIDER, 2005 Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evol. Bioinform. Online **1:** 47–50.

FU, Y.-X., 1997 Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics **147:** 915–925.

HAMILTON, G., M. CURRAT, N. RAY, G. HECKEL, M. BEAUMONT *et al.*, 2005 Bayesian estimation of recent migration rates after a spatial expansion. Genetics **170:** 409–417.

HANSKI, I., 1994 A practical model of metapopulation dynamics. J. Anim. Ecol. **63:** 151–162.

HANSKI, I., and M. GILPIN, 1997 *Metapopulation Biology: Ecology, Genetics and Evolution.* Academic Press, San Diego.

HANSKI, K., and O. OVASKAINEN, 2003 Metapopulation theory for fragmented landscapes. Theor. Popul. Biol. **64:** 119–127.

HARPENDING, H. C., M. A. BATZER, M. GURVEN, L. B. JORDE, A. R. ROGERS *et al.*, 1998 Genetic traces of ancient demography. Proc. Natl. Acad. Sci. USA **95:** 1961–1967.

HEWITT, G., 2000   The genetic legacy of the Quaternary ice ages. Nature **405:** 907–913.

INGVARSSON, P. K., K. OLSSON and L. ERICSON, 1997   Extinction-recolonization dynamics in the mycophagous beetle Phalacrus substriatus. Evolution **51:** 187–195.

KIMMEL, M., R. CHAKRABORTY, J. P. KING, M. J. BAMSHAD, W. S. WATKINS *et al.*, 1998   Signatures of population expansion in microsatellite repeat data. Genetics **148:** 1921–1930.

KINGMAN, J. F. C., 1982a   The coalescent. Stoch. Proc. Appl. **13:** 235–248.

KINGMAN, J. F. C., 1982b   On the genealogy of large populations. Adv. Appl. Probab. **19A:** 27–43.

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.

LEVINS, R., 1969   Some demographic and genetic consequences of environmental heterogeneity for biological control. Bull. Entomol. Soc. Am. **15:** 237–240.

MARUYAMA, T., 1970   Effective number of alleles in a subdivided population. Theor. Popul. 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.

NEI, M., R. CHAKRABORTY and P. A. FUERST, 1976   Infinite allele model with varying mutation rate. Proc. Natl. Acad. Sci. USA **73:** 4164–4168.

PANNELL, J. R., 2003   Coalescence in a metapopulation with recurrent local extinction and recolonization. Evolution **57:** 949–961.

PANNELL, J. R., and B. CHARLESWORTH, 1999   Neutral genetic diversity in a metapopulation with recurrent local extinction and recolonization. Evolution **53:** 664–676.

PANNELL, J. R., and B. CHARLESWORTH, 2000   Effects of metapopulation processes on measures of genetic diversity. Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci. **355:** 1851–1864.

RAY, N., M. CURRAT and L. EXCOFFIER, 2003   Intra-deme molecular diversity in spatially expanding populations. Mol. Biol. Evol. **20:** 76–86.

ROGERS, A. R., and H. HARPENDING, 1992   Population growth makes waves in the distribution of pairwise genetic differences. Mol. Biol. Evol. **9:** 552–569.

SANTOS, C., R. MONTIEL, B. SIERRA, C. BETTENCOURT, E. FERNANDEZ *et al.*, 2005   Understanding differences between phylogenetic and pedigree-derived mtDNA mutation rate: a model using families from the Azores Islands (Portugal). Mol. Biol. Evol. **22:** 1490–1505.

SHAPIRO, B., A. J. DRUMMOND, A. RAMBAUT, M. C. WILSON, P. E. MATHEUS *et al.*, 2004   Rise and fall of the Beringian Steppe bison. Science **306:** 1561–1565.

SILVA, T., L. M. REINO and R. BORRALHO, 2002   A model for range expansion of an introduced species: the common waxbill Estrilda astrild in Portugal. Divers. Distrib. **8:** 319–326.

SLATKIN, M., 1977   Gene flow and genetic drift in a species subject to frequent local extinctions. Theor. Popul. Biol. **12:** 253–262.

SLATKIN, M., and R. R. HUDSON, 1991   Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. Genetics **129:** 555–562.

SMITH, C. I., and B. D. FARRELL, 2005   Range expansions in the flightless longhorn cactus beetles, Moneilema gigas and Moneilema armatum, in response to Pleistocene climate changes. Mol. Ecol. **14:** 1025–1044.

SMITH, M. M., H. T. SMITH and R. M. ENGEMAN, 2004   Extensive contiguous north-south range expansion of the original population of an invasive lizard in Florida. Int. Biodeterior. Biodegrad. **54:** 261–264.

STORZ, J. F., B. A. PAYSEUR and M. W. NACHMAN, 2004   Genome scans of DNA variability in humans reveal evidence for selective sweeps outside of Africa. Mol. Biol. Evol. **21:** 1800–1811.

TABERLET, P., L. FUMAGALLI, A. G. WUST-SAUCY and J. F. COSSON, 1998   Comparative phylogeography and postglacial colonization routes in Europe. Mol. Ecol. **7:** 453–464.

TAJIMA, F., 1989   Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics **123:** 585–595.

TESHIMA, K. M., G. COOP and M. PRZEWORSKI, 2006   How reliable are empirical genomic scans for selective sweeps? Genome Res. **16:** 702–712.

TZEDAKIS, P. C., K. H. ROUCOUX, L. DE ABREU and N. J. SHACKLETON, 2004   The duration of forest stages in southern Europe and interglacial climate variability. Science **306:** 2231–2235.

WADE, M. J., and D. E. MCCAULEY, 1988   Extinction and recolonization—their effects on the genetic differentiation of local-populations. Evolution **42:** 995–1005.

WAKELEY, J., 1993   Substitution rate variation among sites in hypervariable region I of human mitochondrial DNA. J. Mol. Evol. **37:** 613–623.

WAKELEY, J., 1999   Nonequilibrium migration in human history. Genetics **153:** 1863–1871.

WAKELEY, J., 2001   The coalescent in an island model of population subdivision with variation among demes. Theor. Popul. Biol. **59:** 133–144.

WAKELEY, J., and N. ALIACAR, 2001   Gene genealogies in a metapopulation. Genetics **159:** 893–905.

WEIR, B. S., and C. C. COCKERHAM, 1984   Estimating F-statistics for the analysis of population structure. Evolution **38:** 1358–1370.

WHITLOCK, M. C., 1992a   Nonequilibrium population-structure in forked fungus beetles—extinction, colonization, and the genetic variance among populations. Am. Nat. **139:** 952–970.

WHITLOCK, M. C., 1992b   Temporal fluctuations in demographic parameters and the genetic variance among populations. Evolution **46:** 608–615.

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.

WHITLOCK, M. C., and D. E. MCCAULEY, 1999   Indirect measures of gene flow and migration: FST not equal to $1/(4Nm + 1)$. Heredity **82**(2): 117–125.

Communicating editor: J. WAKELEY