Heredity (2014) 112, 291–299

www.nature.com/hdy

ORIGINAL ARTICLE

Genetic consequences of habitat fragmentation during a range expansion

S Mona[1,2,3], N Ray[4], M Arenas[1,2] and L Excoffier[1,2]

We investigate the effect of habitat fragmentation on the genetic diversity of a species experiencing a range expansion. These two evolutionary processes have not been studied yet, at the same time, owing to the difficulties of deriving analytic results for non-equilibrium models. Here we provide a description of their interaction by using extensive spatial and temporal coalescent simulations and we suggest guidelines for a proper genetic sampling to detect fragmentation. To model habitat fragmentation, we simulated a two-dimensional lattice of demes partitioned into groups (patches) by adding barriers to dispersal. After letting a population expand on this grid, we sampled lineages from the lattice at several scales and studied their coalescent history. We find that in order to detect fragmentation, one needs to extensively sample at a local level rather than at a landscape level. This is because the gene genealogy of a scattered sample is less sensitive to the presence of genetic barriers. Considering the effect of temporal changes of fragmentation intensities, we find that at least 10, but often >100, generations are needed to affect local genetic diversity and population structure. This result explains why recent habitat fragmentation does not always lead to detectable signatures in the genetic structure of populations. Finally, as expected, long-distance dispersal increases local genetic diversity and decreases levels of population differentiation, efficiently counteracting the effects of fragmentation.
Heredity (2014) **112,** 291–299; doi:10.1038/hdy.2013.105; published online 23 October 2013

Keywords: range expansion; fragmentation; surfing; long-distance dispersal; coalescent simulation

INTRODUCTION

Habitat fragmentation poses a serious threat to the genetic diversity of many species (Fahrig, 2003; Keller and Largiader, 2003; Jump and Penuelas, 2006). It lowers the connectivity between demes, thus reducing the intrapopulation genetic diversity and increasing population differentiation (Frankham, 1996; Young *et al.*, 1996; Johansson *et al.*, 2007; Dixo *et al.*, 2009). Fragmentation can arise frequently in real environments, for instance when a continuous landscape is divided into patches of demes separated by empty regions (Fahrig, 2003).

Fragmentation is one of the processes that can occur during the evolutionary history of a species. Indeed, species can be regarded as dynamic rather than static entities, which tend to modify their spatial distribution over time, often as a consequence of environmental changes (Davis and Shaw, 2001). On the other hand, range expansion is an important process that has been shown to considerably shape species genetic variability. Range expansions leave a distinct signature on the genetic structure of the populations (Austerlitz *et al.*, 1997; Ray *et al.*, 2003; Excoffier, 2004), as the colonization of new territories proceeds as a series of founder events followed by continuous migrations among neighboring demes (Austerlitz *et al.*, 1997; Ray *et al.*, 2003). This process leads to an increase in the total species population size, but the coalescent pattern can differ significantly from that left by a purely demographic expansion (Excoffier *et al.*, 2009).

Previous individual-based simulation studies have shown the effect of genetic barriers and the time required for their detection (Landguth *et al.*, 2010). However, habitat fragmentation and range expansion can occur simultaneously, both contributing to the genetic variability of a species. Their joint effect has not been considered so far, especially because of a lack of analytical results in the case of non-equilibrium models. Here we investigate range expansions occurring in a fragmented habitat by means of spatially explicit simulations. We consider the most general case of fragmentation, that is, when patterns of gene flow are not homogeneous due to barriers (natural or human induced), which partitions the habitat of a species into discrete patches. We did not consider habitat loss, as it is a more specific instance of the fragmentation process.

In this study, we particularly focused on the differences between a species experiencing a range expansion with or without fragmentation and how the mode of genetic sampling can affect our ability to detect environmental heterogeneity. Several parameters were considered: the carrying capacity of the demes, the connectivity within and among patches, the duration of the fragmentation period, the duration of the range expansion and the effect of long-distance dispersal. Our model addresses cases where species are distributed in patches that are connected by migration events. Patches are sectors of habitat delimited by areas unsuitable for species survival: an example could be a forest that has been progressively converted into agricultural crops and invaded by new species. Another example can be the

[1]Computational and Molecular Population Genetics Lab, Institute of Ecology and Evolution, University of Berne, Berne, Switzerland; [2]Swiss Institute of Bioinformatics, Lausanne, Switzerland; [3]Department of Systematics and Evolution, Muséum National d'Histoire Naturelle, Ecole Pratique des Hautes Etudes, UMR-CNRS 7205, Paris, France and [4]EnviroSPACE Laboratory, Institute for Environmental Sciences, University of Geneva, Carouge, Switzerland
Correspondence: Dr S Mona, Department of Systematics and Evolution, Muséum National d'Histoire Naturelle, Ecole Pratique des Hautes Etudes, UMR-CNRS 7205, 16 Rue Buffon, 75005 Paris, France.
E-mail: mona@mnhn.fr
Received 7 March 2013; revised 3 September 2013; accepted 6 September 2013; published online 23 October 2013

distribution of species into refugial areas, which can be connected, at some stages, by gene flow. Both plant and animal species can be affected by the distribution of habitats of these types (Hewitt, 2004; Aguilar *et al.*, 2008; Haag *et al.*, 2010). It is therefore important to provide a more rigorous framework to interpret the resulting genetic diversity. We also provide a genealogical interpretation of the consequence of habitat fragmentation after a range expansion. By exploring the shape of the gene genealogy at various sampling levels (using summary statistics as proxies) we suggest some practical guidelines to detect fragmentation. We analyze scenarios where the intensity of fragmentation was fixed throughout the history of a species and scenarios with temporal heterogeneity, mimicking recent shift in the pattern of gene flow connecting the demes (that is, particularly important to understand the effects of environmental changes due to human activities). Finally, we show how long-distance dispersal (LDD) can counteract the effects of fragmentation.

MATERIALS AND METHODS

We used a modified version of the software SPLATCHE (Currat *et al.*, 2004) to model habitat fragmentation. SPLATCHE simulates a range expansion of haploid individuals over a two-dimensional array of demes arranged on a lattice and exchanging migrants as in a two-dimensional stepping-stone model. Simulations are done in two steps, the forward (demographic) and the backward (coalescent) steps. The forward simulation starts from an ancestral deme, which sends migrants to its neighboring demes. Migrations to empty demes represent colonization events. Each deme has an intrinsic growth rate g and its density is logistically regulated by its carrying capacity (K). After the regulation step, migrants are sent to the four neighboring demes at rate m. The process is repeated generation by generation for each non-empty deme, resulting in a wave of advance of the whole population. The second phase of the algorithm then starts at the present generation, proceeding backward in time. The effective number of individuals present in a deme is used to compute the probability of a coalescent event, and the migration rates determine the probability of each sampled genes to emigrate, backward in time, to the surrounding demes. The coalescent process stops once all genes have coalesced. Previous versions of SPLATCHE used a single migration rate over the array of demes (hereafter, landscape). The total number of emigrants (Nm) sent by each deme was then equally allocated among its four neighbors. Here we relaxed these assumptions by implementing a map of migration rates, analogous to the existing maps of carrying capacities. Moreover, a deme can now send migrants at different rates to the four possible spatial directions (north, east, south and west), and the total number of emigrants is obtained as the sum of the four Nm_x values.

The size of the array of stepping stones was fixed to 50×50 demes and the origin of the expansion was set to position $<1;1>$ of the square lattice (North–West corner), with an initial effective population size of 100 haploid individuals. To model habitat fragmentation, the landscape was partitioned into groups of demes (hereafter, patches) by adding barriers to dispersal defining 100 patches of 5×5 demes (Figure 1). Barriers were generated by simulating anisotropic migration: demes were allowed to exchange migrants with the surroundings demes at rate m_1 if belonging to the same patch and at rate m_2 otherwise, with $m_1 \geqslant m_2$. When $m_1 = m_2$, the landscape is homogeneous as each deme exchanges on average Nm_1 genes with each of its surrounding demes. This limits dispersal between patches compared with dispersal within a patch. These barriers between patches could, for instance, correspond to the creation of roads in a previously free-ranging environment. Alternatively, this could be viewed as an increase in physical distance between patches, resembling a world where patches are connected by migrational corridors. The degree of fragmentation (that is,, the length of the corridors) is summarized by the ratio of the two migration rates $\rho = m_1/m_2$.

We performed simulations differing by the values of the demographic parameters, the intensity of fragmentation ρ and the age of the spatial expansion. We kept $g = 0.8$, as preliminary simulations and previous studies showed that it affects mainly the speed of the colonization and the probability of surfing (Klopfstein *et al.*, 2006), while having little effect on other aspects of genetic diversity. We considered three possible expansion times, namely $T = 1000$, 3000 or 15 000 generations, to cover a broad range of temporal scenarios. We therefore considered the following sets of simulations aimed at studying the role of different parameters on resulting genetic diversity:

Variable levels of fragmentation

We considered here three values of carrying capacity (held constant over the whole landscape; $K = 20$, 100 and 1000) and three expansion times ($T = 1000$, 3000 and 15 000 generations), but we also included $T = 2\,000$ in some scenarios. The number of immigrants at carrying capacity Km_1 was set to four values (namely 2, 4, 8 and 100) and m_2 was varied accordingly so as to examine the ρ-values of 1 (homogeneous landscape), 5, 15 and 50.

Adding or removing dispersal barriers

We fixed $T = 15\,000$, $K = 1000$, $m_1 = 0.025$ and $\rho = 50$, but changed the migration map at a specified number of generations before present. Range expansions were simulated in: (i) homogeneous landscape where dispersal barriers were added $T_f = 10$, 100 or 1000 generations before present; and (ii) fragmented landscape where dispersal barriers were removed $T_r = 10$, 100 or 1000 generations before present.

Long-distance dispersal

Following Ray and Excoffier (2010), the distance of LDD events was drawn from a γ-dispersal kernel with shape parameter $\alpha = 0.0419$ and scale parameter $\theta = 488.5$ (these values lead to a mean LDD distance $\alpha\theta$ of around 20 demes), as estimated from human demographic data (Novembre *et al.*, 2005). We studied the effect of four different proportions of LDD events, namely 0, 0.001, 0.01 and 0.05, for several fragmentation intensities. We fixed $K = 1000$, $m_1 = 0.025$ and the expansion time at $T = 3000$ or $T = 15\,000$ generations. We varied ρ as in the 'Variable levels of fragmentation' section; that is, we tested values of 1 (homogeneous landscape), 5, 15 and 50.

In order to study the effect of habitat fragmentation at different spatial scales, we arbitrarily sampled genes from 32 demes distributed over the whole array, as shown in Figure 1. We defined four levels (scales) of sampling and computed summary statistics for each of them:

Deme level. Summary statistics are computed on a sample of size $n = 100$ genes in a single deme.

Patch level. Summary statistics are computed by pooling genes from five demes collected within a patch ($n = 20$ per deme).

Region level. Summary statistics are computed by pooling samples of four demes collected from four adjacent patches ($n = 25$ per deme).

Landscape level. Summary statistics are computed by pooling samples from all 32 demes, sampling either $n = 20$ (intense sampling) or $n = 3$ (minimal sampling) gene copies per locus per deme.

We generated 1000 coalescent trees for each combination of parameters. We simulated two types of genetic markers: (i) a DNA sequence of 300 bp with mutation rate of 0.001 per generation for the whole sequence; and (ii) 50 unlinked short tandem repeats (STRs) with a mutation rate of 0.0005 per generation per locus under a pure stepwise mutation model.

Summary statistics were computed with Arlequin 3.5 (Excoffier and Lischer, 2010). For sequence data, we computed the number of haplotypes (H), Tajima's D and F_{ST} (Excoffier *et al.*, 1992). The significance of Tajima's D was determined from an empirical distribution obtained from 1,000 coalescent simulations in a stationary population. For STRs, we computed the average number of alleles over loci (H) and R_{ST} (Michalakis and Excoffier, 1996).

RESULTS

Coalescence pattern in a range expansion with or without fragmentation

Habitat fragmentation influenced the duration of the colonization process. For instance, with $K = 1000$, $m_1 = 0.025$ and $\rho = 1$, the whole array was colonized after 280 generations, whereas with the same

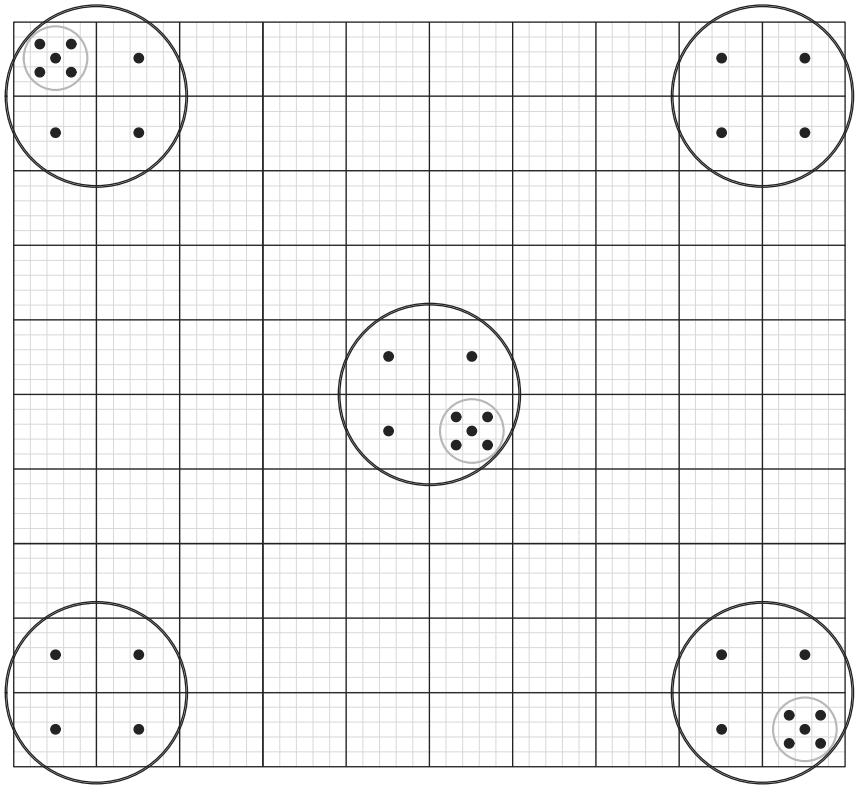

Figure 1 The simulated array of 50×50 demes. Dispersal barriers are shown in black and define 100 patches of 5×5 demes. Black dots: sampled demes. Gray circles: patch sampling level. Black circles: region sampling level.

parameters but $\rho = 50$ the process took around 110 generations longer. We qualitatively compared the spatial distribution of coalescent events between homogeneous and fragmented landscape ($\rho = 50$; Figure 2). Coalescent events are much more randomly distributed in homogeneous landscapes than in fragmented landscapes, where they tend to be concentrated within the sampled patches (Figure 2). This is because barriers to dispersal maintain gene lineages for longer time in the same demes increasing their probability to coalesce. In the following, we present results at the region, patch and deme level obtained for the center of the landscape. Similar patterns were also observed at the periphery, the only difference being an overall reduction in diversity, consistent with the expectations of metapopulation models (Austerlitz *et al.*, 1997; Wilkins and Wakeley, 2002; Wilkins, 2004).

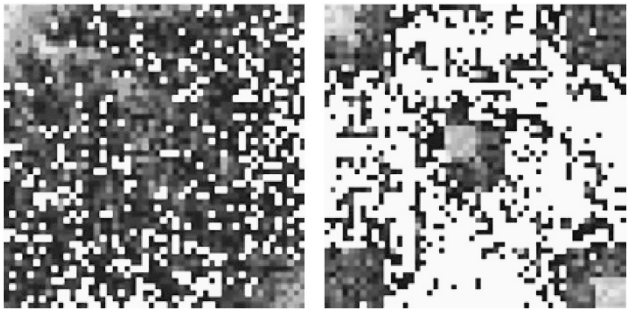

Figure 2 Distribution of coalescent events in a homogeneous landscape (left panel) and a fragmented landscape with $\rho = 50$ (right panel). The number of coalescent events is proportional to the shade of gray (that is, more coalescent events in light gray demes). White demes indicate an absence of coalescent events. Demographic parameters are $K = 1000$, $m_1 = 0.025$ and $m_2 = 0.0005$. Initial expansion occurred 15 000 generations ago. Origin of the expansion and sampled demes located as in Figure 1.

Effects of range expansion and habitat fragmentation on genetic diversity

In Figure 3, we present the decrease in the number of haplotypes for DNA sequence data, but similar patterns were observed for STRs also (Supplementary Figure S1). As expected, the stronger the level of fragmentation, the lower the number of observed haplotypes, with such reduction only slightly dependent on T (Figure 3). Moreover, as can be seen in Figure 3, an equilibrium value in the number of haplotypes and its decrease due to fragmentation is reached already for T around 3000 generations.

The loss of diversity due to habitat fragmentation is not homogeneous across the sampling scales, being more pronounced at the deme and patch level than at the landscape level (Figure 3). For instance, when $T = 15\,000$ and $\rho = 50$, we observe a 26% decrease in the number of haplotypes within a patch, whereas the decrease is only 5% and 3% at the region and at the landscape minimal levels, respectively (Figure 3). The relative reduction in the number of haplotypes is smaller for STRs than for sequence data, which is likely due to the mutational process and not due to the number of simulated markers. Indeed, we repeated the analysis by averaging the number of haplotypes over 50 unlinked DNA loci and obtained exactly the same picture as Figure 3 (results not shown).

Detecting range expansions

The intensity of fragmentation affects both the distribution and the significance of Tajima's D (Table 1, Supplementary Tables S1–S3).

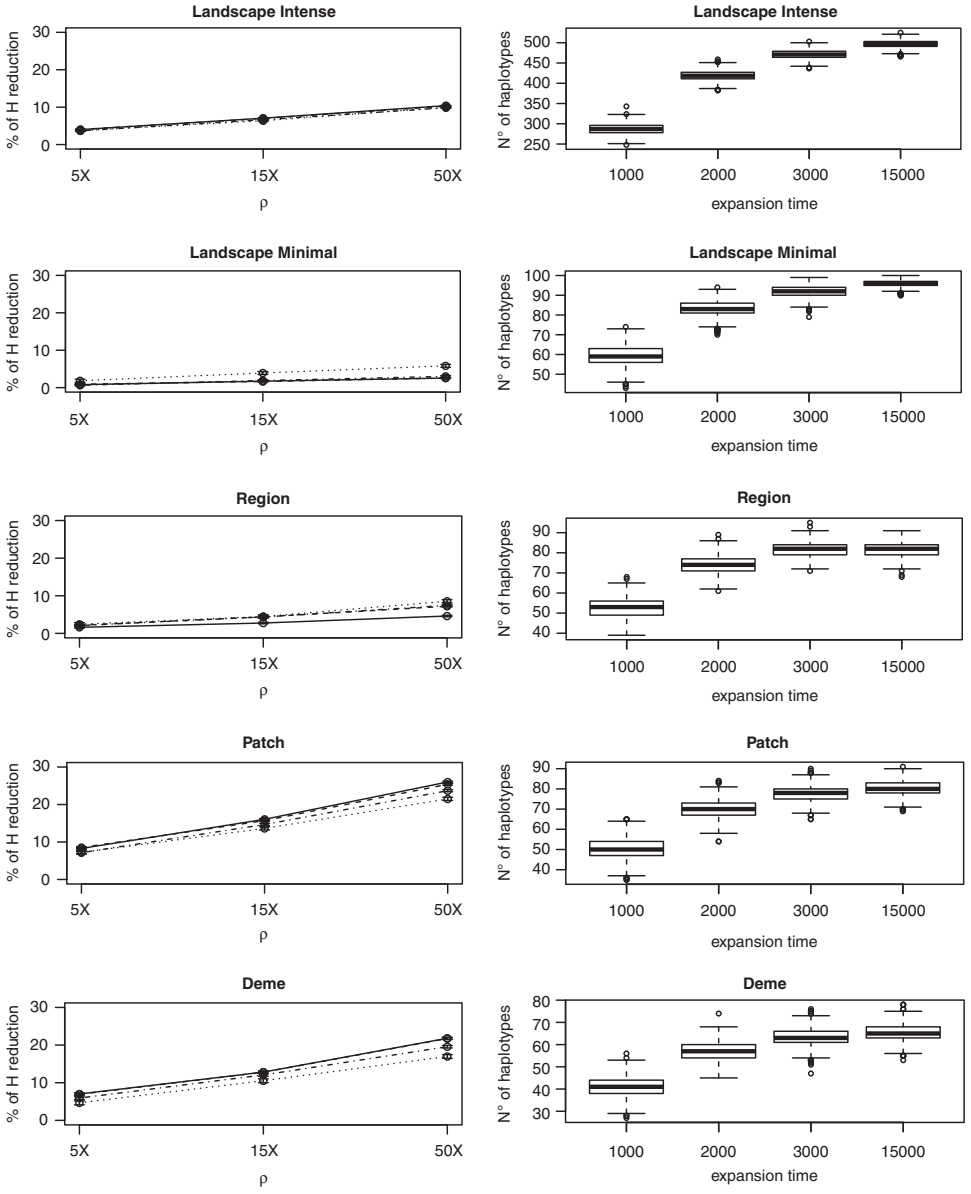

Figure 3 Left panel: decrease in number of DNA sequence haplotypes observed for various fragmentation levels relative to a homogeneous landscape for different sampling levels. *y Axis*: percentage of haplotype reduction. *x Axis*: fragmentation intensity. Demographic parameters are $K=1000$ and $m_1=0.025$; ρ of 5, 15, 50 correspond to m_2 values of 0.005, 0.0016 and 0.0005, respectively. Solid line: $T=15\,000$. Dashed line: $T=3000$. Dot-dashed line: $T=2000$. Dotted line: $T=1000$. Right panel: distribution of the number of haplotypes in a homogeneous landscape. *y Axis*: number of haplotypes. *x Axis*: expansion time.

Fragmentation intensity is positively correlated with Tajima's D and negatively correlated with its significance (that is, it reduces the power to detect a range expansion). Both trends are stronger for older expansions. Indeed, when $T=15\,000$ and $K=1000$, the percentage of negative significant values of Tajima's D dropped from 94% to 2.7% at the deme level, with increasing ρ (Table 1). Conversely, all Tajima's D statistics remain significantly negative with increasing ρ for $T=1000$ (Table 1). Although patch and deme levels are strongly affected by fragmentation, particularly for old expansion times, regions and minimal samples (3 genes from each of the 32 demes) are not (Table 1, Supplementary Tables S1–S3). An intense sampling (20 genes from each of the 32 demes) after an old range expansion shows no significant test statistics irrespective of ρ, at odds with the minimal sampling where we found 100% of tests to be significant (Table 1). Values of D can be quite similar between minimal and intense samplings but the percentage of significant tests can drastically change (Table 1, Supplementary Tables S1 and S2). As an example, for $T=15\,000$, $K=100$ and $m_1=0.05$, average Tajima's D values are negative in both cases, but never significant with an intense sampling and almost always significant with a minimal sampling (Supplementary Table S2).

Effects of range expansions and habitat fragmentation on population differentiation

In Figure 4, we report the F_{ST} distribution computed on DNA sequences collected among the 32 sampled demes for $K=1000$, and a

Table 1 Tajima's *D* distribution for various fragmentation intensities

		T = *15 000*			T = *1000*		
Sampling	ρ*	*Mean*	%P†	%N†	*Mean*	%P†	%N†
Intense	1	−1.05	0.0	0.0	−2.75	0.0	100.0
	5	−1.05	0.0	0.0	−2.75	0.0	100.0
	15	−1.05	0.0	0.0	−2.75	0.0	100.0
	50	−1.06	0.2	0.0	−2.74	0.0	100.0
Minimal	1	−1.74	0.0	99.9	−2.79	0.0	100.0
	5	−1.74	0.0	99.8	−2.77	0.0	100.0
	15	−1.73	0.0	99.7	−2.76	0.0	100.0
	50	−1.72	0.0	99.5	−2.74	0.0	100.0
Region	1	−1.72	0.0	99.7	−2.72	0.0	100.0
	5	−1.71	0.0	98.1	−2.71	0.0	100.0
	15	−1.69	0.0	97.6	−2.68	0.0	100.0
	50	−1.61	0.0	76.4	−2.68	0.0	100.0
Patch	1	−1.71	0.0	98.6	−2.69	0.0	100.0
	5	−1.68	0.0	95.5	−2.63	0.0	100.0
	15	−1.61	0.0	75.4	−2.59	0.0	100.0
	50	−1.36	0.0	7.3	−2.53	0.0	100.0
Deme	1	−1.67	0.0	94.0	−2.56	0.0	100.0
	5	−1.63	0.0	80.1	−2.52	0.0	100.0
	15	−1.55	0.0	50.2	−2.46	0.0	100.0
	50	−1.28	0.0	2.7	−2.42	0.0	100.0

Parameters of the range expansion: $K = 1000$, and $m_1 = 0.025$; ρ of 5, 15 and 50 corresponds to m_2 values of 0.005, 0.0016 and 0.0005, respectively. Expansion times: $T = 15\,000$ and $T = 1000$.
*: fragmentation intensity (see text for details).
†: percentage of simulated datasets showing significant positive (%P) or negative (%N) values of *D*.

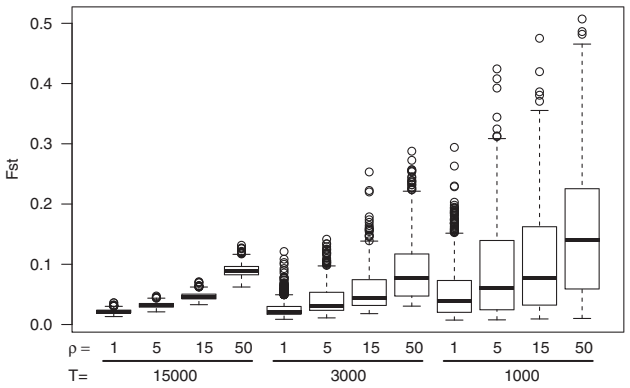

Figure 4 F_{ST} distribution for increasing levels of fragmentation (ρ) and three expansion times (T) for DNA sequence data with $K = 1000$ and $m_1 = 0.025$; ρ-values are as in Figure 3. The distributions were computed from 1000 coalescent simulations in each case.

combination of T and ρ (Supplementary Figure S2 shows the distribution of R_{ST} computed over 50 unlinked STRs). F_{ST} increases with fragmentation level. Both the expansion time and the level of fragmentation have an impact on the mean and the variance of the F_{ST} distribution. Indeed, for a given level of fragmentation, the s.d. of the F_{ST} distribution is 10-fold smaller in old expansions ($T = 15\,000$) than in recent ones ($T = 1000$), whereas mean F_{ST} values are less affected, with only a 2-fold decrease in old expansions. Conversely, ρ affects the mean and s.d. to a similar extent with an increase of up to fourfold for both statistics when $\rho = 50$ as compared with a homogeneous landscape (Figure 4). We found similar effects of T and ρ when computing a multilocus F_{ST} averaged over 50 loci (Supplementary Figure S3). As expected, we find a reduced variance of the F_{ST} distribution compared with the single locus case for any combination of the two parameters (see Figure 4 and Supplementary Figure S3).

When testing other combinations of K and ρ ($K = 100$ and $\rho = 10$; $K = 20$ and $\rho = 4$), we find a similar pattern: T has larger influence on the variance of the F_{ST} distribution than on the mean, whereas ρ has similar influence on both moments of the distribution (results not shown). Therefore, a higher variance of F_{ST} is expected in fragmented compared with homogeneous landscapes for any K-values. Similarly, more recent expansions are characterized by larger variance in F_{ST} than in older expansions.

Influence of temporal change in fragmentation intensity

In Figure 5, we present the effect of recent changes in the level of fragmentation in the context of an ancient range expansion ($T = 15\,000$). After the introduction of a strong fragmentation ($\rho = 50$) in a homogeneous landscape, levels of genetic diversity within demes and within regions take a relatively long time to decline, being barely affected after 100 generations (Figure 5). However, after 1000 generations of habitat fragmentation, we observed a reduction in genetic diversity comparable to that in a continuously fragmented landscape. The effect of habitat fragmentation is more visible at the patch level: genetic diversity within a patch is already substantially reduced by 100 generations after the onset of barriers to dispersal. By contrast, even 1000 generations after the introduction of genetic barriers, F_{ST} only reaches half of what is expected in a continuously fragmented habitat (Figure 5), suggesting that an important component of population genetic structure results from processes ocurring during the expansion in a fragmented landscape.

When we remove dispersal barriers, we note that 100 generations are not sufficient to completely restore the level of genetic diversity expected after a range expansion in a homogeneous environment (Figure 5). Patterns of differentiation between populations take an even longer time to be restored. Indeed, after 1000 generations F_{ST} is still on average 1.6 times higher than that expected in a constantly homogeneous landscape (Figure 5).

Genetic diversity responds more quickly to change in levels of habitat fragmentation than F_{ST}, and this holds true for all the sampling levels considered here. One thousand generations are sufficient to restore genetic diversity to levels expected in a homogeneous world, and 1000 generations of evolution in a fragmented landscape erases diversity previously accumulated in a homogeneous world (Figure 5). We observe a similar pattern when fragmentation intensity is lower ($\rho = 15$, Supplementary Figure S4).

Influence of LDD

LDD has a different effect on genetic diversity at the landscape and at the patch levels (Figure 6 and Supplementary Figure S5). High LDD proportions reduce the extent of population differentiation (measured by F_{ST}) and increase the genetic diversity observed at the landscape level (Figure 6, Supplementary Figures S5–S7, left panels). Conversely, LDD increases genetic diversity having little effect on F_{ST} at the patch level. The number of haplotypes within a patch increases proportionally more than at the landscape level (Supplementary Figures S6 and S7) for higher LDD proportion, consistent with the observations of Ray and Excoffier, (2010) in a homogenous landscape. The strongest effects of LDD occur for high levels of fragmentation (Figure 6, Supplementary Figures S5–S7). For instance, 5% of LDD with $\rho = 50$ resulted in a reduction of F_{ST} of 70% at the landscape scale (Supplementary Figures S6 and S7).

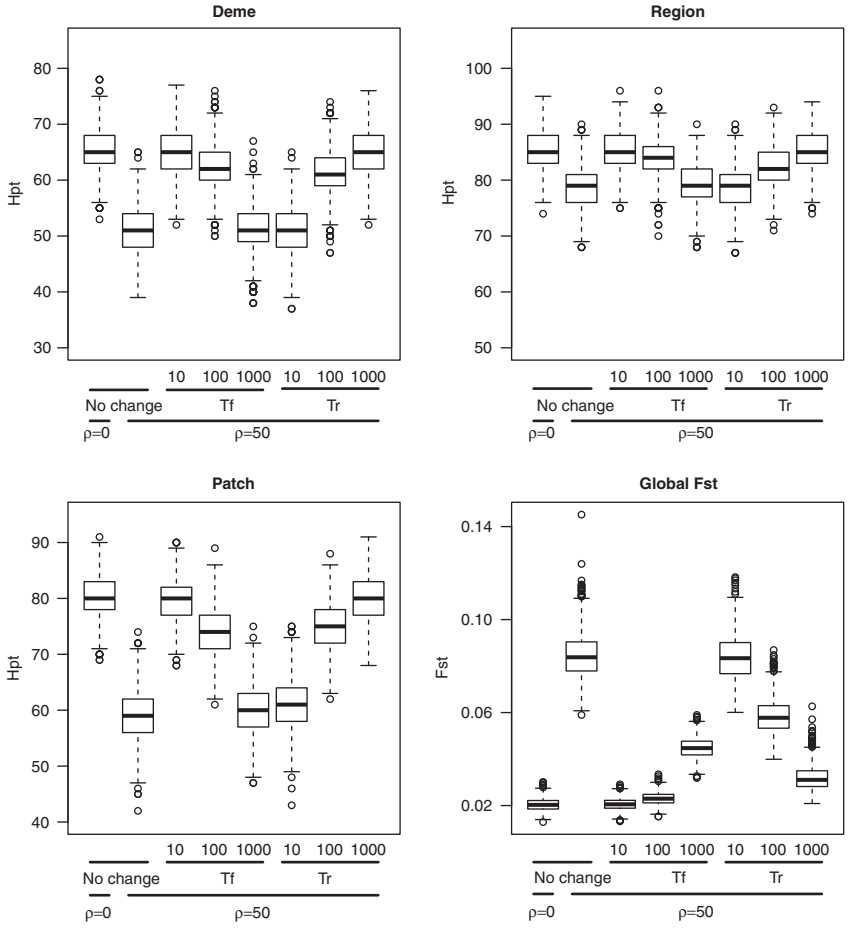

Figure 5 Distribution of the number of DNA sequence haplotypes (*Hpt*) at various sampling levels and F_{ST} when habitat fragmentation has occurred recently. Expansion occurred 15 000 generations ago, $K = 1000$ and $m_1 = 0.025$. *Tf (Tr)*: expansion in a homogeneous landscape with introduction (removal) of fragmentation 10, 100 and 1000 generations before present. No change: fragmentation constant through time (either $\rho = 1$ or $\rho = 50$); $\rho = 50$ corresponds to $m_2 = 0.0005$. The distributions were estimated from 1000 coalescent simulations in each case.

DISCUSSION

Previous individual-based simulation studies have described the effects of the emergence (or disappearance) of a single genetic barrier dividing a spatially continuous population (Landguth *et al.*, 2010). Here we analyzed a more complex setting, where habitat fragmentation occurs in a non-equilibrium metapopulation divided into discrete demes and experiencing a range expansion. For the sake of simplicity, fragmentation was modeled as an instantaneous process, but a more gradual change is unlikely to affect the general pattern we find, as fragmentation requires several generations to affect genetic diversity (see below). Our aim is to understand the interaction between fragmentation and range expansion, and to provide some guidelines on the sampling procedure to follow, in order to detect the genetic signature left by the two processes. Fragmentation is a dynamic characteristic of a habitat and its intensity may vary through time (Fahrig, 2003). We therefore explore the consequence of such changes on the genetic diversity of a species and we finally discuss the importance of LDD to counteract the onset of dispersal barriers. The relevance of our results is important to understand the loss of genetic diversity, which is frequently observed in nature during habitat fragmentation.

In order to interpret our findings, we provide a description of the effect of habitat fragmentation on the gene genealogy. In his seminal work, Wakeley (1999) separated the coalescent process of genes sampled in a structured population into two phases, the scattering and the collecting phase. Going backward in time, the scattering phase starts at the present generation and ends when all sampled genes have either coalesced or migrated to different demes. At this point, the collecting phase begins. It is characterized by a large number of migration events and a few coalescent events that can happen only when two lineages end in the same deme (see Figure 1 from Wakeley (1999) for a graphical interpretation). The relative length of the two phases depends on the demographic parameters characterizing the metapopulation (the number of genes entering in a deme per generation, Nm) and determines the shape of the gene genealogy (Ray *et al.*, 2003). A similar separation of time scales was shown to hold in several complex and realistic models taking into account geographic features (Wakeley and Aliacar, 2001), in extinction/recolonization models (Wakeley and Aliacar, 2001; Wakeley, 2004), in continuous populations (Wilkins and Wakeley, 2002; Wilkins, 2004), as well as in range expansions (Ray *et al.*, 2003; Wegmann *et al.*, 2006). When the number of immigrant genes per deme is large enough (say when $Nm > 20$), the scattering phase will be very short and will have little influence on the distribution of coalescent events. In this case, lineages sampled within a deme will display a gene genealogy similar to that expected after a simple

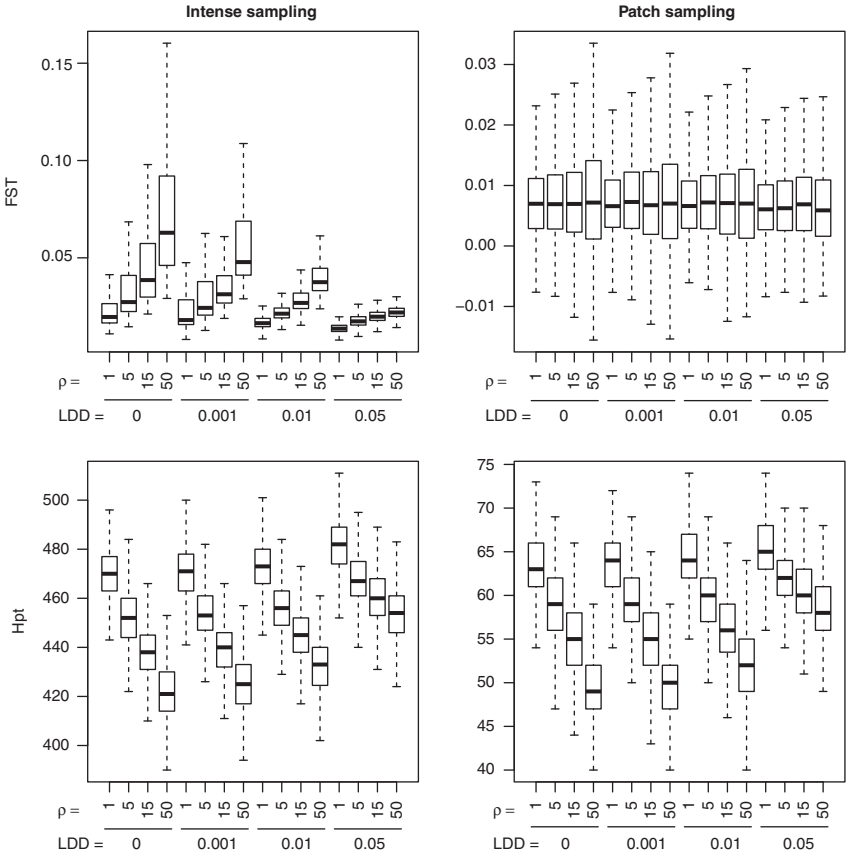

Figure 6 Effect of LDD on the distribution of F_{ST} (first row) and number of DNA sequence haplotypes (*Hpt*, second row) 3000 generations after the onset of the expansion. Estimates are performed for various levels of fragmentation (ρ) and LDD proportions at landscape (intense) and patch levels; ρ-values are as reported in Figure 3.

demographic expansion in an unstructured population (Ray *et al.*, 2003). Conversely, when the number of immigrant genes per deme is lower, the scattering phase will be longer, so that many coalescent events are likely to occur during this phase as well. In that case, the gene genealogy will exhibit a mixture of short- and long-terminal branches, potentially hiding the signature of the expansion (Ray *et al.*, 2003). Our first and most important result is that coalescent events are differently distributed in homogeneous and fragmented landscapes (Figure 2). Fragmentation reduces migration between patches (Fahrig, 2003; Dixo *et al.*, 2009), which means that more coalescent events will occur within a patch (genes will not be able to 'escape' from the patch) before the collecting phase. In other words, the length of scattering phase relative to the collecting phase will be longer for higher fragmentation intensities. The gene genealogy of lineages sampled within a patch or a deme will have shorter terminal branches, an overall reduced total length and hence fewer mutations. This translates into a smaller number of haplotypes and less negative Tajima's *D* (Table 1, Supplementary Tables S1–S3, Figure 3). Such effect is much less pronounced at the region and particularly at the landscape level with minimal sampling; this is because the scattering phase is shorter than in local samples for these sampling schemes, which minimizes the influence of fragmentation. Conversely, at the landscape level with intense sampling we still have a longer scattering phase, owing to the fact that many lineages come from the same patch. For this reason, the intense sampling is a mixture of the minimal and local sampling, which determines that even though neutrality statistics (that is, Tajima's *D*) are negative, they are not significant for old range expansions. This genealogical pattern explains why habitat fragmentation affect more local than global (at the species level) measures of genetic diversity. These results have practical implications for conservation genetics: to detect a reduction in genetic diversity due to habitat fragmentation, it is necessary to sample many individuals from the same deme or patch, rather than to perform extensive geographic sampling with one or few individuals per site, but a mixed strategy could also be envisioned. Similarly, pooling lineages belonging to different patches may make it difficult to detect a loss of diversity having occurred in a fragmented species. Indeed, most of the coalescent events will occur in the vicinity of the expansion origin (Excoffier *et al.*, 2009), largely reducing the impact of dispersal barriers on the ratio of external to internal branch lengths. It is important to stress that the same line of reasoning also applies to homogeneous environments, as pooling neighboring demes separated by a dispersal barrier is analogous to pooling demes that are geographically distant. Therefore, in a stepping-stone model the genealogy of pooled demes that are geographically distant will look like the genealogy of a minimal sample (in agreement with the results of Stadler *et al.* (2009)). Contrastingly, the signature left by a range expansion would be better captured by a minimal sampling. It implies that the sampling design should depend on the biological questions to be addressed.

It has been shown that species with large dispersal abilities are less sensitive to genetic barriers (Landguth *et al.*, 2010) and present lower levels of population structure (Austerlitz and Garnier-Gere, 2003). Using a coalescent argument, it is clear why LDD can efficiently

decrease the effect of fragmentation. LDD decreases the length of the scattering phase, as lineages can more easily escape out of the patch crossing the dispersal barriers. At the same time, the effect of LDD is weaker at the landscape level because, as discussed above, the importance of the scattering phase is reduced by the sampling scheme (Supplementary Figures S6 and S7).

The reduced migration between patches due to fragmentation tends to increase the genetic differentiation (Frankham, 1996). However, the effect of fragmentation during a dynamic process such as a range expansion is more complex. Its influence on the F_{ST} distribution may be explained by the surfing phenomenon (Klopfstein *et al.*, 2006). The surfing behavior of newly arising mutations can produce sectors of genetic variation perpendicular to the leading front of the expansion (Hallatschek *et al.*, 2007; Excoffier and Ray, 2008). The probability of surfing is generally low (Klopfstein *et al.*, 2006); therefore, surfing will have a greater effect on the variance of the F_{ST} distribution rather than on its mean. Migration will progressively erase sectors, homogenizing the metapopulation. However, in scenarios of recent expansion and/or in fragmented landscapes, migration will have less time to occur than in older expansions and/or in homogeneous landscapes. Hence, sectors of diversity will persist for a longer period, leading to the larger observed variance of F_{ST}. Conversely, LDD reduces significantly the variance of F_{ST}: lineages will move faster through the landscape preventing the emergence of sectors even in highly fragmented habitat (Figure 6). LDD does not influence the F_{ST} distribution within a patch (Figure 6), where there are no barriers to dispersal and the sampled demes will likely be part of the same sectors (being close to each other). Our findings are consistent with the simulation study of Wegmann *et al.* (2006) and the theoretical results of Barton (2008), where an increase in the variance of F_{ST} was observed in the presence of spatial heterogeneity or dispersal barriers.

Habitat fragmentation is generally thought to lead to a reduction of intrapopulation genetic diversity and to increase population differentiation (Frankham, 1996; Johansson *et al.*, 2007; Dixo *et al.*, 2009), but many species exposed to fragmentation do not show this expected pattern (Petit *et al.*, 2002; Sumner *et al.*, 2004; Banks *et al.*, 2005). Our results show that genetic diversity need many generations of fragmentation to be severely affected (Figure 5). This can be understood from a coalescent perspective. The length of the scattering phase depends mostly on K, whereas fragmentation alters the relative length of the scattering and collecting phases. However, if fragmentation occurred recently (in generations) compared with the expected length of the scattering phase given, K, the relative length of the scattering compared with the collecting phase will not be affected. For this reason, in species with long generation time (or with high K) human-induced fragmentation may have occurred too recently to be detected. Species with lower K or shorter generation time should conversely exhibit the effect of fragmentation earlier. Clearly, a minimal sampling will show similar level of genetic diversity independently of fragmentation intensities and of the onset of dispersal barriers because the scattering phase will be, in any case, of little importance. This also implies that a failure to detect a reduction in diversity in recently fragmented species may be simply an artifact of the sampling design.

In summary, habitat fragmentation changes the shape of the gene genealogy by altering the relative length of the scattering and collecting phase. Processes such as LDD that shorten the scattering phase will therefore counteract the deleterious effect of fragmentation. Similarly, the sampling scale is extremely important: it is crucial to perform the correct sampling to answer specific questions. If the aim of a study is to detect fragmentation, then it would be best to perform a local sampling. Indeed, fragmentation cannot be detected with a minimal sampling, as its scattering phase is too short to be affected by dispersal barriers. For the same reason (that is, its short scattering phase), a minimal sampling will best describe the long-term demographic history of a species, as already noticed by Wakeley (1999). Finally, conservation genetics programs aiming at restoring genetic diversity loss resulting from landscape fragmentation need to be maintained over dozens or hundreds of generations of the threatened species to have a significant effect, especially if they occur in species with large K.

DATA ARCHIVING

There were no data to deposit.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We are grateful to three anonymous reviewers for their comments and suggestions. This work was supported by a SCALES FP7 EU research project number 226852 to LE.

Aguilar R, Quesada M, Ashworth L, Herrerias-Diego Y, Lobo J (2008). Genetic consequences of habitat fragmentation in plant populations: susceptible signals in plant traits and methodological approaches. *Mol Ecol* **17**: 5177–5188.

Austerlitz F, Garnier-Gere PH (2003). Modelling the impact of colonisation on genetic diversity and differentiation of forest trees: interaction of life cycle, pollen flow and seed long-distance dispersal. *Heredity (Edinb)* **90**: 282–290.

Austerlitz F, Jung-Muller B, Godelle B, Gouyon PH (1997). Evolution of coalescence times, genetic diversity and structure during colonization. *Theor Popul Biol* **51**: 148–164.

Banks SC, Lindenmayer DB, Ward SJ, Taylor AC (2005). The effects of habitat fragmentation via forestry plantation establishment on spatial genotypic structure in the small marsupial carnivore, Antechinus agilis. *Mol Ecol* **14**: 1667–1680.

Barton NH (2008). The effect of a barrier to gene flow on patterns of geographic variation. *Genet Res* **90**: 139–149.

Currat M, Ray N, Excoffier L (2004). SPLATCHE: a program to simulate genetic diversity taking into account environmental heterogeneity. *Mol Ecol Notes* **4**: 139–142.

Davis MB, Shaw RG (2001). Range shifts and adaptive responses to Quaternary climate change. *Science* **292**: 673–679.

Dixo M, Metzger JP, Morgante JS, Zamudio KR (2009). Habitat fragmentation reduces genetic diversity and connectivity among toad populations in the Brazilian Atlantic Coastal Forest. *Biol Conserv* **142**: 1560–1569.

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, Foll M, Petit RJ (2009). Genetic consequences of range expansions. *Ann Rev Ecol Evol Syst* **40**: 481–501.

Excoffier L, Lischer HEL (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour* **10**: 564–567.

Excoffier L, Ray N (2008). Surfing during population expansions promotes genetic revolutions and structuration. *Trends Ecol Evol* **23**: 347–351.

Excoffier L, Smouse PE, Quattro JM (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* **131**: 479–491.

Fahrig L (2003). Effects of habitat fragmentation on biodiversity. *Ann Rev Ecol Evol Syst* **34**: 487–515.

Frankham R (1996). Relationship of genetic variation to population size in wildlife. *Conserv Biol* **10**: 1500–1508.

Haag T, Santos AS, Sana DA, Morato RG, Cullen L Jr., Crawshaw PG Jr. *et al.* (2010). The effect of habitat fragmentation on the genetic structure of a top predator: loss of diversity and high differentiation among remnant populations of Atlantic Forest jaguars (Panthera onca). *Mol Ecol* **19**: 4906–4921.

Hallatschek O, Hersen P, Ramanathan S, Nelson DR (2007). Genetic drift at expanding frontiers promotes gene segregation. *Proc Natl Acad Sci USA* **104**: 19926–19930.

Hewitt GM (2004). Genetic consequences of climatic oscillations in the Quaternary. *Philos Trans R Soc Lond B Biol Sci* **359**: 183–195.

Johansson M, Primmer CR, Merila J (2007). Does habitat fragmentation reduce fitness and adaptability? A case study of the common frog (Rana temporaria). *Mol Ecol* **16**: 2693–2700.

Jump AS, Penuelas J (2006). Genetic effects of chronic habitat fragmentation in a wind-pollinated tree. *Proc Natl Acad Sci USA* **103**: 8096–8100.

Keller I, Largiader CR (2003). Recent habitat fragmentation caused by major roads leads to reduction of gene flow and loss of genetic variability in ground beetles. *Proc Biol Sci* **270**: 417–423.

Klopfstein S, Currat M, Excoffier L (2006). The fate of mutations surfing on the wave of a range expansion. *Mol Biol Evol* **23**: 482–490.

Landguth EL, Cushman SA, Schwartz MK, McKelvey KS, Murphy M, Luikart G (2010). Quantifying the lag time to detect barriers in landscape genetics. *Mol Ecol* **19**: 4179–4191.

Michalakis Y, Excoffier L (1996). A generic estimation of population subdivision using distances between alleles with special reference for microsatellite loci. *Genetics* **142**: 1061–1064.

Novembre J, Galvani AP, Slatkin M (2005). The geographic spread of the CCR5 Delta32 HIV-resistance allele. *PLoS Biol* **3**: e339.

Petit RJ, Latouche-Halle C, Pemonge MH, Kremer A (2002). Chloroplast DNA variation of oaks in France and the influence of forest fragmentation on genetic diversity. *Forest Ecol Manag* **156**: 115–129.

Ray N, Currat M, Excoffier L (2003). Intra-deme molecular diversity in spatially expanding populations. *Mol Biol Evol* **20**: 76–86.

Ray N, Excoffier L (2010). A first step towards inferring levels of long-distance dispersal during past expansions. *Mol Ecol Resour* **10**: 902–914.

Stadler T, Haubold B, Merino C, Stephan W, Pfaffelhuber P (2009). The impact of sampling schemes on the site frequency spectrum in nonequilibrium subdivided populations. *Genetics* **182**: 205–216.

Sumner J, Jessop T, Paetkau D, Moritz C (2004). Limited effect of anthropogenic habitat fragmentation on molecular diversity in a rain forest skink, Gnypetoscincus queenslandiae. *Mol Ecol* **13**: 259–269.

Wakeley J (1999). Nonequilibrium migration in human history. *Genetics* **153**: 1863–1871.

Wakeley J (2004). Metapopulation models for historical inference. *Mol Ecol* **13**: 865–875.

Wakeley J, Aliacar N (2001). Gene genealogies in a metapopulation. *Genetics* **159**: 893–905.

Wegmann D, Currat M, Excoffier L (2006). Molecular diversity after a range expansion in heterogeneous environments. *Genetics* **174**: 2009–2020.

Wilkins JF (2004). A separation-of-timescales approach to the coalescent in a continuous population. *Genetics* **168**: 2227–2244.

Wilkins JF, Wakeley J (2002). The coalescent in a continuous, finite, linear population. *Genetics* **161**: 873–888.

Young A, Boyle T, Brown T (1996). The population genetic consequences of habitat fragmentation for plants. *Trends Ecol Evol* **11**: 413–418.

Supplementary Information accompanies this paper on Heredity website (http://www.nature.com/hdy)