Spatial genetic structure of a small rodent in a heterogeneous landscape

B. GAUFFRE,* A. ESTOUP,* V. BRETAGNOLLE† and J. F. COSSON*
*INRA, UMR CBGP (INRA/IRD/Cirad/Montpellier SupAgro), Campus international de Baillarguet, CS 30016, F-34988 Montferrier-sur-Lez cedex, France, †CNRS, Centre d’Etudes Biologiques de Chizé, 79360, Beauvoir sur Niort, France

Abstract

Gene flow in natural populations may be strongly influenced by landscape features. The integration of landscape characteristics in population genetic studies may thus improve our understanding of population functioning. In this study, we investigated the population genetic structure and gene flow pattern for the common vole, Microtus arvalis, in a heterogeneous landscape characterised by strong spatial and temporal variation. The studied area is an intensive agricultural zone of approximately 500 km² crossed by a motorway. We used individual-based Bayesian methods to define the number of population units and their spatial borders without prior delimitation of such units. Unexpectedly, we determined a single genetic unit that covered the entire area studied. In particular, the motorway considered as a likely barrier to dispersal was not associated with any spatial genetic discontinuity. Using computer simulations, we demonstrated that recent anthropogenic barriers to effective dispersal are difficult to detect through analysis of genetic variation for species with large effective population sizes. We observed a slight, but significant, pattern of isolation by distance over the whole study site. Spatial autocorrelation analyses detected genetic structuring on a local scale, most probably due to the social organisation of the study species. Overall, our analysis suggests intense small-scale dispersal associated with a large effective population size. High dispersal rates may be imposed by the strong spatio-temporal heterogeneity of habitat quality, which characterises intensive agroecosystems.

Keywords: clustering methods, computer simulation, gene flow, landscape genetics, microsatellites, spatial genetic structure

Received 4 June 2008; revisions received 24 August 2008; accepted 1 September 2008

Introduction

The environmental features characterising the mosaic of habitats that make up a landscape can either restrict or promote movements of individuals in natural populations and consequently the extent of genetic connectivity (Taylor et al. 1993). Establishing the genetic structure of populations and the pattern of gene flow in relation to landscape characteristics thus provides information on population functioning, which may be critically important for management and conservation decisions (Palsboll et al. 2007). For instance, when habitat is highly fragmented, genetic drift and mutation may lead to an overall patchy pattern of genetic variability (Rousset 2003). Small and isolated populations suffer from restricted dispersal and reduced genetic diversity (Frankham 2005). Over the past few years, many studies in the rapidly expanding field of landscape genetics (Manel et al. 2003; Storfer et al. 2007 for reviews), have been carried out to detect landscape and environmental features that may influence gene flow and hence shape the genetic structure of populations (Castric et al. 2001; Spear et al. 2005; Giordano et al. 2007; Dionne et al. 2008).

Because they limit biases due to the sampling of population units defined a priori (Manel et al. 2003), recently developed Bayesian clustering methods based on individual genotypes are particularly powerful in the identification of gene flow barriers — a major focus of landscape genetics research (Guillot et al. 2005a; Francois et al. 2006). Among physical barriers that may reduce genetic connectivity (such as rivers or mountains), anthropogenic features (e.g. roads) have been repeatedly shown to slow animal dispersal (Coulon...
et al. 2006; Gerlach & Musolf 2000; Riley et al. 2006). Additionally, because dispersal is generally restricted in space for many species, continuous spatial patterns of genetic differentiation with geographical distance have been theoretically predicted (Rousset 2000) and empirically verified in many natural populations (Broquet et al. 2006; Elmer et al. 2007). Extremely complex spatial patterns have been reported for some landscapes, resulting from the conjunction of both discrete barriers to dispersal and isolation by distance (IBD; e.g. Berthier et al. 2005).

The spatial organisation of populations is thus greatly influenced by the pattern of dispersal. On the other hand, the latter has been found to be considerably influenced by the environmental characteristics of the landscape. In particular, experimental studies suggest that spatial and temporal random environmental variations favour the evolution of increased dispersal propensity (Johnson & Gaines 1990; Clobert et al. 2001). Although numerous empirical studies have addressed patterns of gene flow in spatially heterogeneous landscapes (Diffendorfer 1998; Lin & Batzli 2001; Kreuzer & Huntly 2003; Morris & Diffendorfer 2004), only a few were concerned with landscapes that vary both in space and time (Denno et al. 1996; Tattersall et al. 2004). Agroecosystems represent an extreme and perhaps unique example of such spatio-temporal variability. Agricultural landscapes are by nature highly fragmented, providing to natural populations a patchwork of habitats of varying quality depending on species characteristics and life history. Additionally, land use and agricultural practices (i.e. harvesting, mowing and ploughing) introduce further temporally asynchronous alterations of habitat quality and suitability. Therefore, agroecosystems offer an ideal experimental-like setting to study the consequences of spatio-temporal variability on gene flow in natural populations. It remains difficult, however, to establish predictions on population spatial structure and pattern of gene flow in such highly variable ecosystems (see Diffendorfer 1998).

We carried out a landscape genetics study to investigate the spatial pattern of genetic variation in the common vole Microtus arvalis — a small-sized rodent abundant in crop areas in Western Europe (Mitchell-Jones et al. 1999; Salamolard et al. 2000) — over a large, intensively cultivated study area (c. 500 km²). This model is important for conservation concerns because it represents a key resource in the trophic chain in agroecosystems (Lambin et al. 2006; Ims et al. 2008). Although the common vole was thought to be restricted to meadows, perennial crops and edge habitats (e.g. floral strips) in intensively cultivated areas, it has been found at high density in cereal fields or rapeseed (see Salamolard et al. 2000). Short-term effects of farming practices on common vole population dynamics and demographic parameters have been investigated (Jacob 2003; Jacob & Hempel 2003). Harvesting, mowing and mulching do not destroy nest sites (common vole build and inhabit subterraneous burrows), but potentially eliminate a substantial proportion of the local population by direct killing or increased risk of predation due to reduced vegetation height and cover. No dispersive movement appears to be associated with these activities (Jacob & Hempel 2003). Ploughing, which destroys common vole burrows, seems to be the only agricultural practice likely to eradicate common vole (Jacob 2003).

Dispersal distances are suspected to be short (a few hundred metres) in common vole (Spitz 1977b; Boyce & Boyce 1988a; Jacob 2003). However, the complexity of the processes involved in common vole population dynamics makes it difficult to predict the resulting pattern of genetic variation. In contrast to the above theoretical predictions, dispersal in common voles seems to be slowed in heterogeneous and highly disrupted landscapes (Delattre et al. 1992). Indeed, one study demonstrated small-scale genetic structuring among small and isolated populations (Schweizer et al. 2007). We would expect that less variable or more permanent habitats (e.g. grasslands) may provide refuge from where temporary habitats (e.g. annuals crops) are seasonally colonised. If refuge habitats are not connected by dispersal (directly or through reproduction in temporary habitats), strong genetic structuring is likely to prevail among isolated units (resulting from source-sink dynamics). However, balanced dispersed dynamics is more likely to result if temporary habitats allow breeding to occur and if human-associated perturbations cause individual movements. Therefore, the temporal instability of habitat quality may actually enhance effective dispersal and lead to low-level genetic differentiation between habitats, whether permanent or not. Landscape features, such as narrow and continuous zones of unfavourable habitats (hedged farmland, woodland, urbanised areas) or roads, are likely to result in further genetic discontinuities, as previously described for other rodent species (Gerlach & Musolf 2000; Orlowski & Nowak 2006).

The main aim of this study was to assess the pattern of gene flow in such spatially and temporally varying environments. We also analysed whether obvious habitat disruptions (a motorway in our case) produced significant spatial structuring of genetic variation. To this end, and in contrast to classical population genetics analyses relying on the a priori delineation and definition of population units, we used recent methodological and conceptual developments based on georeferenced individual genetic data.

Materials and methods

Study area

The study site is located in central-western France (Région Poitou-Charentes, 46·11°N, 0·28°W) and covers 500 km² of an agricultural landscape (Fig. 1). It is dominated by
intensive agricultural system including cereals, colza and spring-sown crops with temporal instability and varied vole abundance (Table 1). The study site is crossed by a motorway built in 1981 and has a network of unfavourable areas for common vole (urbanised zones, hedged farmland and woodland) as this species is known to be restricted to open habitats (Delattre et al. 1992). All these landscape features are likely to decrease dispersal rates.

**Tissue sampling**

Genetic sampling and estimation of common vole abundance were carried out during monthly trapping sessions from April to July 2006. A total of 90 trap lines in April and June, and 20 trap lines in May and July were set in different habitats, with the number of traps in each habitat proportionate to the landscape covered. Each trap line consisted of a 100-m transect with 51 live traps without bait spaced every 2 m, which was surveyed for 24 h. Common voles were also directly trapped in their colonies during the same period. Location of captured voles was recorded with a precision of 10 m. A total of 372 common voles were trapped in 193 different sampling locations for 155 different agricultural plots. The black line indicates the motorway path and the dashed area corresponds to hedged farmland.

**Genetic analyses**

Genomic DNA was extracted from tissue samples (toe or ear clipping) using the Puregene DNA Purification Kit (Gentra Systems) and following the manufacturer’s ‘DNA purification from solid tissue’ protocol. Genotyping was carried out using two polymerase chain reaction (PCR) multiplex sets comprising 14 microsatellites loci isolated from *M. arvalis* (Gauffre et al. 2007) and one (MSMM6) adapted from a locus isolated in *Microtus montebelli* (Ishibashi et al. 1999). For detailed information about the sets and PCR conditions, see Gauffre et al. (2007). We used an ABI PRISM 310 Genetic Analyser (Applied Biosystems) and

**Table 1** Characteristics and proportion of study surface area covered for the main habitats comprising the agricultural matrix. Intensity and timing of main agricultural practices are also given. Common vole abundance is expressed in number of captures per 100 traps.

<table>
<thead>
<tr>
<th>Habitat</th>
<th>Percentage of surface</th>
<th>Mowing (temporary disturbance)</th>
<th>Ploughing (destruction of colonies)</th>
<th>Spring common vole abundance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perennial Meadows</td>
<td>12</td>
<td>2 (spring and summer)</td>
<td>None</td>
<td>1.42</td>
</tr>
<tr>
<td>Alfalfa</td>
<td>3</td>
<td>3–4 (spring and summer)</td>
<td>None</td>
<td>1.59</td>
</tr>
<tr>
<td>Annual Ray grass</td>
<td>3</td>
<td>1–2 (spring)</td>
<td>1 (spring)</td>
<td>3.13</td>
</tr>
<tr>
<td>Cereal crops</td>
<td>40</td>
<td>1 (July)</td>
<td>1 (summer or winter)</td>
<td>1.58</td>
</tr>
<tr>
<td>Colza</td>
<td>12</td>
<td>1 (June)</td>
<td>1 (summer or winter)</td>
<td>0.99</td>
</tr>
<tr>
<td>Maize/sunflower</td>
<td>18</td>
<td>1 (September)</td>
<td>1 (spring)</td>
<td>0.24</td>
</tr>
<tr>
<td>Woodland/build-up</td>
<td>12</td>
<td>None</td>
<td>None</td>
<td>≈ 0</td>
</tr>
</tbody>
</table>

© 2008 The Authors
Journal compilation © 2008 Blackwell Publishing Ltd
PCR profiles were analysed using GeneScan and Genotyper software. For some loci, PCR amplification was difficult and an observed excess of homozygotes was detected regardless of the geographical scale of sampling. We used the program Micro-Checker (Van Oosterhout et al. 2004) to test for the most probable causes of this observation, including various genotyping errors (e.g. short allele dominance) and the presence of null alleles. We excluded two loci (Ma30 and Ma42) that failed to amplify for some individuals and for which the general excess of homozygotes was observed for most allele size classes indicating a presence of null alleles. Two other loci (Ma35 and Ma88) were also excluded due to poor quality PCR profiles of some alleles (confounded in the stutter of shorter alleles). Our final data set thus included genotypes obtained at 10 microsatellite loci (Table 2).

### Statistical analyses

We tested for deviation from Hardy–Weinberg equilibrium (HWE) using the exact test in GenePop 4.0 (Rousset 2008). Deviation from HWE was quantified by estimating and testing the significance of the $F_{\text{ST}}$ value using a permutation procedure (10 000 permutations) in Genetix 4.05.2 (Belkhir 2004). Expected ($H_E$; Nei 1987) and observed ($H_O$) heterozygosity levels were also calculated using GenePop 4.0. We tested linkage disequilibrium (LD) between loci using the G-test implemented in GenePop 4.0 followed by a false discovery rate (FDR) correction for multiple tests (Benjamini & Hochberg 1995) with a nominal level of 5%. We ran 10 independent tests using (for each test) the default conditions specified in the program. For each run, we recorded the number of pairs of loci showing significant LD as well as the identity of the concerned loci.

### Assessing genetic structure with clustering methods

The motorway and two zones of hedged farmland coupled with urbanised or forested areas were identified as linear barriers that had the potential to considerably reduce common vole dispersal. Therefore, we estimated the number of genetic clusters ($K$) and the spatial location of breaks in gene flow delineating such clusters within the studied area using Bayesian assignment approaches implemented in Structure 2.2 (Pritchard et al. 2000), Geneland 2.0.10 (Guillot et al. 2009b) and TESS 1.1 (Francois et al. 2006). The three Bayesian methods used allow statistical inferences to be made for individual multilocus genetic data. Subdivision of the data into clusters in the Bayesian methods is done by maximising HWE and minimising LD. Structure bases its inference on genetic data alone, whereas Geneland and TESS explicitly incorporate spatial organisation information for genotyped individuals. We followed the recommendations of the authors of these methods to address the problem of genetic structure, which can include populations with a low level of differentiation. For Structure, we assumed an admixture model with correlated allele frequencies (Falush et al. 2003). The number of populations ($K$) is a fixed parameter of this model. Five independent runs for $K = 1$ to 10 were carried out each with a Markov chain Monte Carlo (MCMC) of 500 000 iterations following a burn-in of 100 000 iterations. For each $K$, we calculated the mean posterior probability over its runs. We then used this mean value to estimate the posterior probability of each K using the formula given by Pritchard & Wen (2003). For Geneland, we used the Dirichlet model of allele frequencies and long MCMC runs to ensure convergence of the chain (Guillot et al. 2005a, b). Ten independent runs with 50 000 MCMC iterations (thinning = 50) were performed allowing $K$ to vary from 1 to 10. The maximum rate of the Poisson process was fixed at 500, the maximum number of nuclei in the Poisson-Voronoi tessellation at 300, and the potential error for spatial coordinates at 50 m. $K$ was inferred from the modal value of the run with the highest likelihood. TESS was run with the Dirichlet model of allelic frequencies and the interaction parameter (Ψ) fixed at 0.6, as described by Chen et al. (2007). We ran 10 replicates of 10 independent MCMC chains with 12 000 sweeps and a burn-in period of 2000 sweeps. $K$ was then inferred from the modal value of the replicate with the highest likelihood.

### Detection of recent linear barrier to gene flow

We tested for genetic differentiation among individuals across the motorway. We used GenePop 4.0 to estimate pairwise $F_{\text{ST}}$ values and assess their significance with permutations in FSTAT 2.9.3.2 (Goudet 2001). Calculating $F_{\text{ST}}$ on a priori-defined populations may lead to erroneous conclusion (Manel et al. 2003). We thus processed the latter analysis on five arbitrary partitions of the study area into two continuous spatial units independently of any landscape element.

---

**Table 2** Number of alleles ($A$), expected heterozygosity ($H_E$) (Nei 1987), and observed heterozygosity ($H_O$), and $P$ values for Hardy–Weinberg equilibrium (HWE) exact tests, computed over the entire sampled area. Two loci, Ma36 and Ma68, significantly deviate from HWE.

<table>
<thead>
<tr>
<th>Locus</th>
<th>$A$</th>
<th>$H_E$</th>
<th>$H_O$</th>
<th>HWE $P$ value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ma09</td>
<td>18</td>
<td>0.88</td>
<td>0.88</td>
<td>0.066</td>
</tr>
<tr>
<td>Ma25</td>
<td>44</td>
<td>0.96</td>
<td>0.94</td>
<td>0.287</td>
</tr>
<tr>
<td>Ma29</td>
<td>17</td>
<td>0.91</td>
<td>0.88</td>
<td>0.250</td>
</tr>
<tr>
<td>Ma36</td>
<td>38</td>
<td>0.94</td>
<td>0.93</td>
<td>0.028</td>
</tr>
<tr>
<td>Ma54</td>
<td>25</td>
<td>0.88</td>
<td>0.86</td>
<td>0.056</td>
</tr>
<tr>
<td>Ma66</td>
<td>24</td>
<td>0.91</td>
<td>0.87</td>
<td>0.501</td>
</tr>
<tr>
<td>Ma68</td>
<td>17</td>
<td>0.87</td>
<td>0.78</td>
<td>0.006</td>
</tr>
<tr>
<td>Ma75</td>
<td>20</td>
<td>0.89</td>
<td>0.86</td>
<td>0.584</td>
</tr>
<tr>
<td>Ma81</td>
<td>20</td>
<td>0.63</td>
<td>0.63</td>
<td>0.224</td>
</tr>
<tr>
<td>MSMM6</td>
<td>18</td>
<td>0.88</td>
<td>0.86</td>
<td>0.207</td>
</tr>
<tr>
<td>All loci</td>
<td>24</td>
<td>0.88</td>
<td>0.85</td>
<td>0.007</td>
</tr>
</tbody>
</table>
Finally, we also adopted an approach without a priori spatial information using the above Bayesian clustering methods with the number of populations fixed at two. The inferred spatial delineation between the two populations was then compared to that of the motorway.

We then used computer simulations with our own program based on coalescence theory to simulate genetic data under a simple model of one ancestral population splitting into two completely isolated populations. Demographic parameters included the split time corresponding to the age of the motorway expressed as number of common vole generations (G) and the effective common vole population size (Np), which were assumed to be identical for each population. Simulations were run for same sample size as our real data set (i.e. one population with 98 individuals and one with 274) and for 10 microsatellite loci, assuming a generalised stepwise mutational model with a maximum number of continuous allelic states set to 60 and the parameter of the geometric distribution of allelic mutation size fixed at 0.22 (Estoup et al. 2002). The mutation rate was fixed at 5 × 10⁻⁴, as widely assumed for microsatellite loci (Goldstein & Schlötterer 1999). Pairwise FST and HET were calculated for each simulated data set. The level of uncertainty for the value of G was assessed using the rationale that the motorway was built 25 years before our sampling and the number of common vole generations per year is between 1.5 and 3.5 (JP Quéré, personal communication). Thus, in our simulations, we considered values of G between 37 and 88. HET was used instead of Np, because the range was unknown and the parameters are related (Hartl & Clark 1997). In the real data set, HET was found to be 0.88 (SD = 0.08). We ran 50,000 simulations with G randomly drawn in a uniform distribution bounded between 37 and 88, and Np drawn in a uniform distribution bounded between 100 and 20,000. We then selected those simulated data sets for which HET was between 0.80 and 0.96 (i.e. 30,000 of the 50,000 simulated data sets), and computed an FST distribution from these sub-data sets. Ten of these sub-data sets were randomly chosen and analysed with Structure, Geneland and TESS using a similar approach as above and using the same individual geographical location as in our real sample. We also tested the differentiation between the two populations for each of the 10 data sets using GenePop 4.0.

Assessing the scale of genetic differentiation. We determined IBD by testing the correlation between pairwise genetic distances (q, Rousset 2000) and the logarithm of geographical distances using GenePop 4.0. The linear relationship of these variables is expected to hold best at distances greater than the mean parent–offspring dispersal distance (σ) and to progressively deviate from linearity at distances larger than 10 to 15σ (Rousset 2000; Leblois et al. 2003). Thus, we excluded pairwise distances between individuals separated by less than 250 m, as the latter was roughly estimated from literature (Boyce & Boyce 1988b; Spitz 1977b). To visualise the relationship between pairwise genetic and geographical distances for this large data set, we fit a spline correlogram (Chambers & Hastie 1992) with a 95% bootstrap confidence envelope, using r (r Development Core Team 2007), evaluated for 50 points with a target of 7 degree of freedom. To calculate an indirect estimate of σ from our genetic data, we used the inverse of the slope to determine the neighbourhood size (4Dπσ²), where D is the effective density in the population (Rousset 2000). Since the ideal sampling area is 10σ × 10σ for this calculation, we used the slope value given by regression between 250 m and 2500 m. D was independently estimated to be 1000 individuals/km². The latter value was inferred from trapping data (Table 1) by calculating the overall density of adults using Spitz’s (1977a) conversion indices with weighted adult common vole densities in each habitat as a proportion of the studied area covered.

Finally, the spatial pattern of genetic variation was investigated using spatial autocorrelation analyses which assess the genetic similarity between pairs of individuals at different distance classes, thus providing results on the scale at which spatial patterns occur. We adopted the approach developed by Smouse & Peakall (1999) using GenAlEx (Peakall & Smouse 2006) and ran two different analyses using 500-m and 1-km distance class sizes. Each distance class (except distance classes above 25 km) included more than 350 pairwise genotypic data. Significant deviation of spatial autocorrelation patterns from the random distribution of genotypes was tested by plotting the 95% null hypothesis confidence intervals estimated by 1000 random permutations of individual genotypes among the geographical locations. Although distance and rate of dispersal influence the shape of the correlogram, it is difficult to directly translate spatial autocorrelation diagrams into dispersal patterns (Aars et al. 2006). However, the correlogram should flatten out at the scale where dispersal is not connecting ‘subpopulations’ (Aars et al. 2006).

Results

Genotypic data revealed high levels of genetic diversity at all 10 microsatellite loci (Table 2). The number of alleles per locus ranged between 17 and 44. Expected heterozygosity for each locus ranged from 0.63 to 0.96, with an average value of 0.88 for all loci. Observed heterozygosity ranged from 0.63 to 0.94, with an average value of 0.85 for all loci. There was a significant overall heterozygosity deficit in the total sample (P = 0.007) with a low (0.032) but significant global FST value (P < 0.001) that could be related with some spatial structure in our data. We found that the number of pairs of loci showing significant LD was unstable among the 10 independent runs of GenePop, varying from zero to four pairs of loci depending on the run. Moreover, the loci
involved in the significant tests were most of the time different in the 10 runs. We hence did not find any evidence for strong and stable LD in our data set. Again, this LD pattern most probably results from genetic structure within the study area rather than from physical linkage between loci.

Genetic clustering

Results of the clustering analyses are summarised in Table 3. The estimated logarithm of likelihood for data analysed with Structure was highest for \( K = 1 \) and the posterior probability of having only one population was 1. For \( K > 1 \), none of the 372 individuals could be assigned to a given cluster as each individual had a similar probability of belonging to each cluster. Geneland inferred the presence of two populations; however, one, turned out to be a ‘ghost’ population corresponding to a population to which none of the sampled individuals had been assigned. TESS inferred four clusters from the highest likelihood run; however, over 85% of individuals fell into a single cluster and the individuals assigned to the other three clusters were randomly distributed across the study site, without spatial coherency.

Barrier effect on gene flow

The level of genetic differentiation between individuals sampled across the motorway was low, but significant (\( F_{ST} = 0.0018, \ P < 0.01 \)). However, the five arbitrary subdivisions of the data set into two spatially coherent samples also gave significant \( F_{ST} \) values of the same order of magnitude (\( F_{ST} \) range = 0.002 to 0.0045). Therefore, we suspected that the genetic differentiation was not directly associated with the motorway. Geneland assigned all individuals to one population (with a probability of membership > 0.99) and TESS assigned 339 individuals to the first population and 33 to the second, with the individuals assigned to this second being randomly distributed across the study site.

We ran computer simulations under the extreme conditions of no effective dispersal across a recent anthropogenic barrier to assess trends in the genetic data and methods to identify recent barriers to gene flow such as the motorway. When using parameter values for \( H_E \) and \( G \) close to those of our real data set, the mean \( F_{ST} \) value of the simulations was 0.0031, the mode was 0.0019 and the 2.5% and 97.5% quantiles were 0.0006 and 0.0087, respectively (Fig. 2). The modal value of the simulation set (0.0019) was thus very close to the observed value (0.0018) in our real data set. Both STRUCTURE and Geneland inferred a single cluster of individuals for each of the 10 randomly selected simulated data sets, for which \( F_{ST} \) values ranged from 0.0005 to 0.005. TESS inferred higher values of \( K \) (between 4 and 8), although similar to the estimation \( K \) in the real data set, with most individuals assigned to a single population and other individuals randomly distributed over the entire study area. We found significant genetic differentiation in all but one of these simulated data sets, using approximated Fisher’s exact tests between the two sides of the motorway.

Table 3 Results of individual-based Bayesian clustering analyses. For Geneland and TESS, the 10 runs were ranked by decreasing value of likelihood (in the log scale) and the \( K \) value estimated for each run is presented. In Structure, \( K \) is fixed from 1 to 10 and runs are ranked by decreasing value of likelihood.

<table>
<thead>
<tr>
<th>Geneland</th>
<th>TESS</th>
<th>Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likelihood</td>
<td>( K )</td>
<td>Likelihood</td>
</tr>
<tr>
<td>-18034.7</td>
<td>2</td>
<td>-18049.9</td>
</tr>
<tr>
<td>-18063.3</td>
<td>2</td>
<td>-18056.8</td>
</tr>
<tr>
<td>-18080.0</td>
<td>2</td>
<td>-18077.7</td>
</tr>
<tr>
<td>-18111.1</td>
<td>2</td>
<td>-18080.8</td>
</tr>
<tr>
<td>-18130.1</td>
<td>2</td>
<td>-18089.2</td>
</tr>
<tr>
<td>-18133.6</td>
<td>2</td>
<td>-18092.4</td>
</tr>
<tr>
<td>-18137.7</td>
<td>2</td>
<td>-18102.5</td>
</tr>
<tr>
<td>-18140.6</td>
<td>2</td>
<td>-18104.1</td>
</tr>
<tr>
<td>-18143.2</td>
<td>2</td>
<td>-18104.6</td>
</tr>
<tr>
<td>-18159.3</td>
<td>2</td>
<td>-18108.9</td>
</tr>
</tbody>
</table>

Fig. 2 Density distribution of \( F_{ST} \) values for simulated data sets under a simple model of one ancestral population splitting into two completely isolated populations. Data sets have been simulated using a parameter setting reflecting the actual case of a motorway splitting the studied common vole population (splitting time between 37 and 88 generations and mean heterozygosity between 0.80 and 0.96; see text for details). Grey areas represent the 2.5% and 97.5% quantiles. Dotted line indicates the \( F_{ST} \) value observed in the real common vole data set (\( F_{ST} = 0.0018 \)).
Our simulation study highlights the difficulties involved in identifying recent barriers to gene flow for species with a high level of heterozygosity, and thus, probably for those with a large effective population size. We cannot draw firm conclusions about the exact impact of the motorway on common vole gene flow in this study area.

Spatial scale of genetic differentiation

The observed genetic differentiation across the motorway was similar to all arbitrary subdivisions of the studied landscape, suggesting that factors other than the motorway may underlie the genetic structure. A significant pattern of IBD was observed among individuals in the study area ($P = 0.005$). The spline correlogram (Fig. 3) illustrates that genetic differentiation among individuals was substantial (mean $r = 0.033$), and increased with increasing geographical distance. Limiting the IBD analysis to minimum and maximum distances of 0.25 and 2.5 km (1σ and 10σ, as recommended by Rousset 2000) resulted in a significant positive slope. As expected from theoretical studies, the slope computed with distances between 1σ and 10σ was two times higher than the slope estimated over the entire study area (i.e. 0.01 and 0.0048, respectively). Using this value of the slope and assuming a density of 1000 adult common voles per km² in the area of study, we estimated an indirect dispersal distance of 88 m (95% CI: 76–110 m).

The spatial genetic autocorrelation analysis displayed positive genetic correlation values ($r$) for the first distance classes (zero and 1 km), and negative and significant $r$-values only for certain classes above 12 km (Fig. 4). Genetic similarities between individuals decreased between distance classes 0 km and 2 km and then flatten out. For distances greater than 2 km, we observed a tendency for a slight decrease in similarity with distance, consistent with the IBD pattern previously described. Setting distance class sizes to 500 m did not change overall relationships but demonstrated patterns of genetic similarity on a smaller scale (up to 1 km). In contrast to the pattern usually observed for rodent populations under IBD (Berthier et al. 2005; Aars et al. 2006), we did not observe a clear shift toward negative $r$-values at larger distance classes but rather a slight decreasing trend of $r$-values with increasing distance classes. We suggest that this pattern could be the consequence of large effective population size.

Discussion

We demonstrate that, in a large 500 km² area with spatial heterogeneity and temporal instability, gene flow in a small rodent population is not sufficiently limited to result in a
strong genetic structure. Rather, we observed a slight, but significant, IBD pattern shaping the roughly homogeneous genetic structure of the studied common vole population. A commonly proposed model for the spatial structure of common vole populations in heterogeneous landscapes such as agricultural mosaics is a source-sink dynamics between permanent and temporary habitats (Butet & Leroux 2001). Our findings do not contradict this model but suggest that if such dynamics exists, it occurs in the context of large effective population sizes and high movement rates among the different types of habitats constituting the agricultural matrix.

The combination of high dispersal rates with high densities may explain the observed spatial pattern of genetic variation, despite the presence of very local substructures (mainly due to the social organisation of the species) and the relatively short individual dispersal distances. Our results for dispersal are consistent with empirical measures of movement distances (Spitz 1977b; Jacob 2003) and predictions of the limited dispersal ability of these small rodents (van de Zande et al. 2000; Heckel et al. 2005). For example, direct observations suggest that short-range natal dispersal of females is frequent and occurs over distances of 50 to 160 m (Boyce & Boyce 1988a). We obtained a short distance estimate of 88 m for the mean parent–offspring dispersal distance (σ), derived from the inverse of the IBD regression slope. Although this method relies on numerous assumptions (Rousset 2000; Leblois et al. 2003; Vekemans & Hardy 2004), it provides valuable information about dispersal, especially when using highly polymorphic microsatellite markers (Leblois et al. 2003). The slope of the correlogram obtained from spatial autocorrelation analysis flattens out after the 2-km distance class (1 km when defining 500-m distance classes), confirming small-scale dispersal. An informative result in terms of effective population size (N_e) was the high level of heterozygosity in the population (mean H_0 = 0.85). This value was higher than those observed in previous microsatellite-based genetic studies on rodents (Ehrich et al. 2001; Berthier et al. 2005; Aars et al. 2006; Brouat et al. 2007; Kraaijeveld-Smit et al. 2007). Heterozygosity is lost through genetic drift at the rate of one-half N_e per generation; thus, the level of heterozygosity in a population at a given time is a good indicator of its past and present demography (Hartl & Clark 1997). The high level of heterozygosity observed in our study therefore suggests that the population is characterised by a large N_e and hence large densities, which we estimated from trap data at 1000 adults per km².

Previous studies on common vole biology have suggested that genetic diversity and gene flow should be low in heterogeneous agricultural landscapes (Delattre et al. 1992; Topping et al. 2003). So far, these predictions have been at least partly supported by a recent study in Switzerland (Schweizer et al. 2007), which concluded that Microtus arvalis populations were genetically structured at a fine spatial scale (< 2 km). As reported for other socially organised species (Scribner & Chesser 1993), we suggest that this genetic structuring at a small geographical scale is due to the social structure (i.e. related individuals clustered in colonies) rather than to a low level of effective small-scale dispersal between isolated subpopulations. In line with this, the strong positive autocorrelation observed in our data set only for the first distance class, which includes individuals originating from the same colony, supports this interpretation. The lower level of heterozygosity observed in the study of Schweizer et al. (2007) than in our study, mean H_0 = 0.77 vs. 0.85 reflects a smaller effective population size. The latter account, at least partly, for the occurrence of greater genetic differentiation between close sampling sites in Schweizer’s study.

Interpopulation proximity and connectivity are crucial for population persistence in heterogeneous landscapes (Gulve 1994). A patchy distribution of suitable habitats is expected to create a patchy genetic structure, especially for species with dispersal limited in space. It was therefore surprising to find such large-scale genetic homogeneity given the small-scale dispersal abilities of this small rodent and the patchy distribution of perennial habitats in the landscape (Fig. 1). Common voles were trapped in every habitat of our studied area. In particular, spring densities were found to be high in cereals crops (covering 40% of the study site), despite the fact that ploughing eradicates common voles annually (Jacob 2003). Common voles are characterised by a short life cycle and high reproductive capacities. They produce several generations each year, as young common voles are able to reproduce at only 3 weeks of age (Toussaint 1990) and breeding can occur throughout the year when climatic and nutritional conditions allow sexual activity in winter (Le Louarn & Quére 2003). Given this rapid reproductive cycle and short, but frequent dispersal events, cereal crops appear to be suitable habitats with sufficient time for gene flow to occur. This highlights the potential role of temporary suitable habitats for facilitating gene flow and mitigating the apparent spatial fragmentation (Blau & Wichmann 2007). Borders and floral strips also play a major role in the population dynamics of M. arvalis (Briner et al. 2005) and can enhance the overall landscape connectivity. Additionally, the disturbance caused by mowing (substantial increase in predation risk), which occurs in every habitat (perennials and annuals) with different timing and intensity, is likely to cause some movement to adjacent fields where the cover is higher (Jacob & Brown 2000). This would thus artificially enhance dispersal in common vole populations.

We did not identify any of the linear landscape features, such as the motorway or the unfavourably hedged farmlands coupled with permanent areas (forested and urbanised), as obvious barriers to gene flow, using individual-based
Bayesian clustering methods. However, our analysis demonstrates a potential bias when searching for breaks in gene flow in a landscape. The calculation of $F_{ST}$ between a priori-defined populations for a landscape element such as a motorway can lead to a pre-emptive conclusion that it is an effective barrier to gene flow. In our study, we observed that other arbitrary partitions of our data set resulted in similar and significant $F_{ST}$ values. Therefore, the significant differentiation across the motorway in our study was more likely to be due to the overall IBD pattern, although some recent studies have demonstrated that patterns of genetic divergence in terrestrial animals correlate with various elements present in the landscape, such as roads and motorways (Gerlach & Musolf 2000; Keller et al. 2004; Coulon et al. 2006). Roads may be quasi-impenetrable barriers even for large and highly mobile mammals like bobcats, coyotes (Riley et al. 2006) and ungulates (Coulon et al. 2006). Reasoning for common vole genetic structure to be not affected by the motorway is thus challenging. One possible reason is that dispersal may indeed be unaffected. There are two studies documenting road crosses by common voles. Orlowski & Nowak (2006) surveyed a road network (15 roads) with different traffic volumes (350–10 500 cars per 24 h) located in an agricultural landscape of southwestern Poland. The most frequently killed group of mammals was small rodents (40%), with the common vole accounting for the largest proportion among them (26%). The major factor affecting mortality rate was the daily vehicle traffic volume. Van Wieren & Worm (2001) studied how mammals used a wildlife overpass across a motorway in central Netherlands using track counts. They concluded that wildlife overpasses can be an effective means of connecting habitats for small mammals. However, given that there was no wildlife overpass in our study site and the high mortality risk associated with crossing a motorway with substantial traffic, it is unlikely that dispersal is not affected by such a barrier.

Computer simulation provided another explanation for our findings. Given the large population size of the studied common vole population, the motorway may simply be too recent to allow genetic differentiation to become sufficiently high to be detected through $F$-statistics or clustering methods. Additional simulations show that in the extreme case of complete isolation of two populations characterised by heterozygosity levels (i.e. effective sizes) similar to those observed in our study, more than 200 generations (i.e. 80 years in the case of $M. arvalis$) would be required to reach $F_{ST}$ values > 0.01. The latter value is often mentioned as a threshold to confidently detect an effective barrier to gene flow using Bayesian clustering methods (e.g. Chen et al. 2007).

In conclusion, our study shows that only geographical distance tends to limit common vole gene flow in this agricultural landscape. Gene flow might not be affected by habitat fragmentation because common vole population dynamics is faster than the perturbation dynamics of the landscape inhabited, hence mitigating the temporary suitability of most habitats (i.e. crops). Spatio-temporal variation in habitat quality resulting from land use favours increased dispersal propensity, and thus, should be a crucial determinant of such spatially extended populations. The exact effect of each habitat type on genetic variation over space and time and of agricultural practices on dispersal remains to be studied using specific sampling strategies. Social structure should also be explicitly accounted for in further sampling to test its likely effect on the structuring of genetic variation.

Acknowledgements

We thank Eric Petit, Raphael Leblois and three anonymous referees for constructive comments on the manuscript and Thomas Cornulier for comments on an earlier draft and help in statistical treatments. We thank Steve Augiron, Alexandre Villers, Mirabelle Gouat, Maxime Galan and Camille Fontaine for help in trapping voles. This work was supported by INRA (ECOGER national program) and, B. Gaffre received a grant from ACI ECCO.

References


© 2008 The Authors 
Journal compilation © 2008 Blackwell Publishing Ltd
COMMON VOLE GENE FLOW IN AGRICULTURAL LANDSCAPE


Bertrand Gauiffe is a PhD student interested in population biology and more particularly in the inference of dispersal (and factors influencing this behaviour) from genetic data. Arnaud Estoup’s current research mainly focuses on evolutionary biology of invasive species. Vincent Bretagnolle research focuses on the conservation, ecology and behaviour of wildlife with particular emphasis on birds. J.F. Cosson is mainly interested in population genetics of fluctuating populations in relation with landscape ecology.