ELSEVIER

Mammalian Biology
Zeitschrift für Säugetierkunde

www.elsevier.de/mambio

ORIGINAL INVESTIGATION

Revealing cryptic genetic structuring in an urban population of stray cats (*Felis silvestris catus*)

Sébastien Devillard*, Thibaut Jombart, Dominique Pontier

Laboratoire de Biométrie et Biologie Evolutive, UMR-CNRS 5558; Université de Lyon, Univ. Lyon 1, 69622, Villeurbanne Cedex, France

Received 14 November 2007; accepted 14 January 2008

Abstract

Urban stray cat populations have previously been shown to be organized in moderately to strongly genetically differentiated colonies. However, the spatial pattern of this genetic differentiation and the possibility that some cryptic genetic structure occurs have not been investigated. Firstly, we combine a pairwise *Fst* method, a fully Bayesian clustering method and a multivariate analysis to show that the spatial structure of 17 urban stray cat colonies (Nancy, France) does not match the underlying genetic structure. These colonies are structured in two clusters at the uppermost hierarchical level. Additionally, geographic proximity between colonies does not explain their genetic homogeneity. The spatial pattern in genetic differentiation seems more to be a mixture between low global structure and some highly localized structure, comprising isolated colonies. Variations in the local ecological and social constraints on immigration between colonies may explain this pattern. Secondly, we show that the overall current immigration rate is low (8%). Our results suggest that dispersal does not play a major role in the process of homogenization of allelic frequencies. Our study provides a case-study on the use of Bayesian clustering and assignment methods on a real, small data set with numerous fragmented populations; a population structure that is of high relevance in conservation biology.

Keywords: *Felis silvestris catus*; Stray cats; Dispersal; Bayesian method; Population genetic structure

Introduction

The relationships between dispersal, gene flow and population genetic structure have been the subject of many studies during the last three decades (Slatkin, 1985; Neigel, 1997; Bohonak, 1999). Spatial and temporal gene flow affect many important ecological and evolutionary aspects of populations (Chepko-Sade and Halpin, 1987; Hanski and Gilpin, 1997; Clobert et al., 2001).

In a spatially structured population, adjacent populations are expected to be more genetically similar than distant ones (isolation by distance, IBD, Wright, 1943, or stepping stone models, Kimura, 1953) because gene flow is more likely to occur at smaller spatial scales. But, environmental constraints due to complex features of the habitat might lead to cryptic spatial distribution of genetic variability. In a fragmented habitat with a strong heterogeneity in both habitat availability and quality, the linear geographic distance between two spatial entities makes less biological sense than any measure of functional connectivity (Baudry and Burel, 1998, see Coulon et al., 2004, for an example in roe deer *Capreolus*

*Corresponding author.
E-mail address: devillar@biomserv.univ-lyon1.fr (S. Devillard).

1616-5047/$ - see front matter © 2008 Deutsche Gesellschaft für Säugetierkunde. Published by Elsevier GmbH. All rights reserved.
doi:10.1016/j.mambio.2008.01.001

capreolus). Thus, in conditions where the complex structure of the habitat might confound any of our predictions (IBD or stepping stone models for example), we have to first assess the level of matching between the spatial and the genetic structures of the populations before making any inference about the patterns of dispersal and gene flow. This is especially crucial for social species where both spatial and social constraints interact to determine the patterns of emigration, dispersal costs, and effective immigration.

In the highly fragmented urban environment, social stray cats (*Felis silvestris catus*) live in a spatially structured population composed of different colonies separated from each other by distances largely below the dispersal abilities of the species (Liberg et al., 2000). Despite such short distances, urban colonies usually show moderate overall genetic differentiation when using *F*-statistics, based both on microsatellite markers and on coat color markers, suggesting a likely cryptic genetic structure and low gene flow (Say et al., 2003). This general finding is consistent with a long-term demographic monitoring of a large urban colony that has shown that immigration and emigration are low (Courchamp, 1996; Say, 2000; Devillard et al., 2003). This stray cat in an urban habitat system thus provides a good opportunity to assess congruity between the geographic distribution of colonies and the spatial distribution of genetic variability in a complex habitat. To our knowledge, however, no study has investigated this question in stray cats, and the associated immigration and gene flow patterns between colonies residing in contrasting sites within this habitat also remains unknown.

Several methodologies are available to deal with this question of congruity of geographic and genetic structure in the absence of information about functional connectivity. Firstly, conventional and indirect approaches to the study of gene flow and dispersal use *F*-statistics (Wright, 1951) to examine the relationship between the genetic and geographical distances between populations. The theoretical number of migrants per generation: $F_{ST} = 1/(4Nm+1)$ is then used as an index of the level of dispersal. Secondly, some recently developed methods for population genetic analysis can be viewed as a complementary approach to *F*-statistics (Neigel, 2002; Pearse and Crandall, 2004) when investigating population substructure and patterns of dispersal and gene flow among these populations. For example, the family of assignment methods allows individuals to be assigned to their populations of origin, to test whether the populations of origin differ from the populations from which the individuals were sampled and thus, to estimate recent immigration rates in each of the studied populations (Paetkau et al., 1995; Favre et al., 1997; Rannala and Mountain, 1997; Cornuet et al., 1999; Pritchard et al., 2000; Dawson and Belkhir, 2001; Corander et al., 2003; Wilson and Rannala, 2003; Pearse and Crandall, 2004; see Manel et al., 2005, for a review). These methods have been proven to estimate contemporary dispersal rates equivalent to those estimated by direct demographic methods such as multi-site mark-recapture methods (Berry et al., 2004). Of the assignment methods, the fully Bayesian clustering method (Pritchard et al., 2000) is the most widely used. Indeed, this method can be used both to detect cryptic genetic population structure different from that of the known geographic population structure, and to perform assignment tests using prior information about the geographic source of individuals sampled. However, both *Fst* and most Bayesian approaches are based on Hardy–Weinberg equilibrium expectations. One way to circumvent this restrictive hypothesis is the use of multivariate analyses (Menozzi et al., 1978) as a third approach. Such an approach aims to describe the spatial distribution of genetic differentiation without any assumption about the underlying genetic model.

Here, we used a complementary approach, employing these three methodologies, to study the spatial pattern of genetic differentiation and the pattern of immigration and gene flow in an urban population of stray cats (*Felis silvestris catus*) composed of two different types of colonies. Tabor (1983) has indicated that two colony types could co-exist in the urban environment: first, large and persistent multi-male, multi-female colonies with low-immigration rates which are likely to have their own genetic identity and second, small colonies with low-persistence time and composed of unrelated individuals from different origins. In our population, Xémar (1997) defined these two kinds of colonies as "closed" or "opened" (see Material and methods for more details) depending on the local environmental conditions.

Firstly, we tested if the geographically distinct colonies of Nancy (France) correspond to genetically distinct clusters by using: (i) a traditional pairwise *Fst* method with an analysis of the pattern of isolation by distance; (ii) the fully Bayesian clustering method of Pritchard et al. (2000); and (iii) a multivariate principal component analysis. For these two latter methods, we assessed whether genetic clustering matches the geographic distribution of colonies. We expected a cryptic genetic structure, i.e. a poor relationship between spatial and genetic structure, based on our life-history knowledge of the species. More specifically, we expected that the "closed" colonies would be genetically well-defined and thus, more prone to isolation in a cluster.

Secondly, we used prior geographic information of colonies to assign each individual to one colony based on their multi-locus genotype using: (i) the fully Bayesian assignment method (Pritchard et al., 2000; Falush et al., 2003); and (ii) the frequency-based assignment method (Paetkau et al., 2004). This provided us with rough estimates of the immigration rates for

each colony, and facilitated the detection of patterns with respect to the sex and age of immigrants and any geographic pattern in immigration. Overall, we expected a low-immigration rate, as was previously shown for a large colony (Say, 2000; Devillard et al., 2003). More specifically, the "closed" colonies are expected to show the lowest immigration rates.

Materials and methods

Study species

Stray cats are small social felids (order Carnivora) that have colonized many different environments (from urban areas to sub-Antarctic islands, Pascal, 1980; Liberg et al., 2000). Stray cats show large flexibility in social organization, spatial structure, mating systems and dispersal patterns, depending on the environment (Liberg, 1980; Pontier and Natoli, 1996; Say et al., 1999; Liberg et al., 2000; Devillard et al., 2003) and more precisely, on the distribution and temporal stability of resources (food and shelters). In the urban environment, stray cat colonies settle in specific places (e.g. public gardens) where food is delivered by volunteers and shelters are available (Cahloon and Haspel, 1989). Very roughly described, a colony is composed of one or several social groups that are founded by a single female. Each group then grows through philopatry of both males and females, with some exchanges of males between groups of the same colony and exceptional dispersal of females outside the colony (Devillard et al., 2003). Persistent hostility occurs towards strange individuals of both sexes (Liberg et al., 2000) but exclusion of these individuals from the colony may depend on the colony size or demographic history (Tabor, 1983; S. Devillard, D. Pontier, L. Say, unpublished data). The mating system is promiscuous (Natoli and De Vito, 1991; Yamane, 1998; Say et al., 1999).

Colony sampling

Seventeen colonies, comprising a total of 237 stray cats, were sampled in the center of Nancy (Eastern France) during 2 years in 1995 and 1996 (Xémar, 1997). Colonies were located between 500 and 4000 m away from each other (Euclidean distance, Fig. 1) and in relatively isolated areas, such as hospital grounds and industrial or housing estates. Xémar (1997) defined two categories of colony with regard to their ecological properties: (1) "closed": in this study, colonies 2, 9, 10, 12, 13 and 16, which are surrounded by high-traffic roads and nearly isolated due to the lack of small streets or gardens that stray cats could use as corridors; (2) "opened": all the other colonies in this study, which are located in quieter urban environments with potential dispersal corridors. We used Xémar's classification for subsequent analyses. The number of stray cats sampled per colony varied from 9 to 23 (mean $\pm$ S.E. $= 13.94 \pm 4.2$). Density varied from 2.3 to 30 stray cats ha^{-1} (Xémar, 1997). In each colony, stray cats were captured using double-door baited traps. Stray cats were then anaesthetized with an intramuscular injection of ketamin chlorhydrate (Imalgène 1000 15 mg kg^{-1}, Rhône Mérieux, Lyon, France) and acepromazin (Vétranquil 5.5% 0.5 mg kg^{-1}, Sanofi, Paris, France). Sex and age (years) were determined (Table 1) and hair samples were collected before releasing the stray cats at their capture site. The overall proportion of males in the samples was 0.43, ranging from 0.23 in colony 13 to 0.73 in colony 6. The relative proportions of each age class (Table 1) were 0.27 (<1 year old), 0.46 (1–3 years old) and 0.27 (>3 years old). The ratio between the number of juvenile and adult stray cats was highly variable, ranging from 0 in colonies 3, 5, 6, 8 and 9 to 0.70 in colonies 1 and 11. This ratio strongly depends on the length of time the colony was monitored with respect to the reproductive cycle. No stray cat captured in a given colony was observed or recaptured in any other colony during the whole study period.

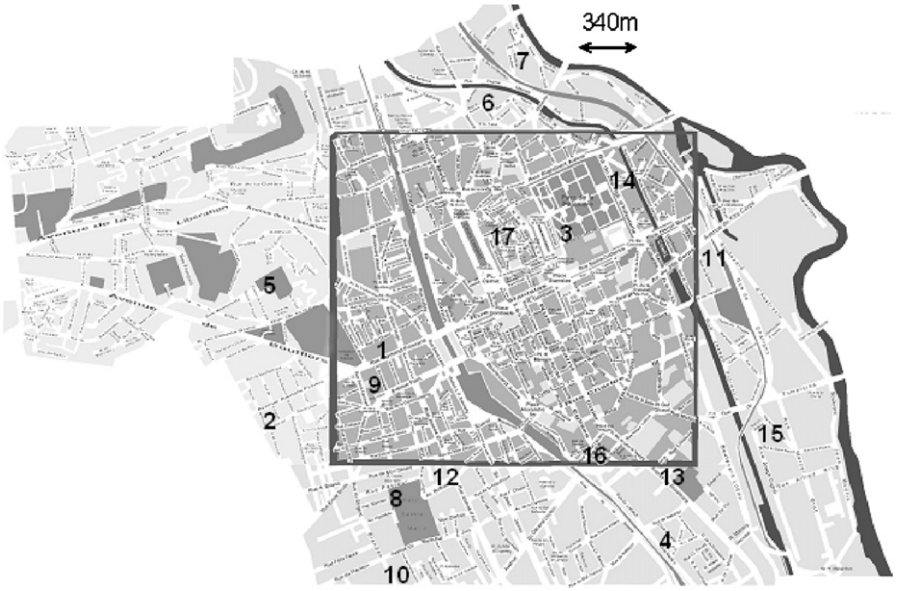

Fig. 1. Spatial distribution of the 17 colonies in Nancy. Colonies are numbered from 1 to 17.

Table 1. Number of sampled males and females per colony (numbered from 1 to 17) and distribution of the individuals per age class

Colony	Female	Male	<1 year old	[1–3] years old	>3 years old	Total
1	7	3	7	2	1	10
2	15	7	13	9	0	22
3	6	6	0	7	5	12
4	12	11	3	9	11	23
5	9	6	0	11	4	15
6	3	8	0	3	8	11
7	5	9	3	6	5	14
8	3	7	0	4	6	10
9	5	4	0	8	1	9
10	4	7	5	3	3	11
11	11	9	14	4	2	20
12	10	4	6	7	1	14
13	10	3	1	7	5	13
14	8	9	1	10	6	17
15	8	3	5	5	1	11
16	7	5	4	8	0	12
17	6	7	1	6	6	13
Total	129	108	63	109	65	237

Genetic marker analysis

We used a Chelex-based method to extract DNA, and amplification was carried out for nine unlinked microsatellite loci by polymerase chain reaction (PCR) using the fluorescent-labelled primers *fca*8, *fca*23, *fca*43, *fca*45, *fca*77, *fca*78, *fca*90, *fca*96 (Menotti-Raymond and O'Brien, 1995) and *fca*37 (Menotti-Raymond, pers. com.). Analysis of the PCR products was resolved on a denaturating polyacrylamide gel on a Pharmacia Automatic sequencer (ALFexpress). Data collection and analysis were carried out using Fragment Manager software supplied with the sequencer (see Say et al., 1999, for more details on the molecular protocol).

Genetic structure using pairwise *Fst*

Exact tests were performed for allele frequency differences for all pairs of colonies. Genetic distances between populations were quantified using the pairwise distance statistic *Fst*. Pairwise *Fst* values were calculated using FSTAT 2.9.3 (Goudet, 2001) and the significance of pairwise distances was tested using a randomization test (5000 random permutations). We also took into account the Allendorf–Phelps effect (Waples, 1998) by subtracting $1/(2S)$ from the *Fst* values, S being the mean number of sampled individuals of each pair of colonies. The Allendorf–Phelps effect is expected to artificially inflate *Fst* estimates due to intra-locus sampling error when sample sizes are small. We tested for an isolation by distance pattern of genetic differentiation between pairs of colonies using the Mantel test (9999 permutations, Mantel, 1967) on the relationship $Fst/(1-Fst) = f[\ln(d)]$ where $Fst/(1-Fst)$ is the linearized *Fst* and d is the geographic distance between each pair of colonies (Rousset, 1997). This test was performed with functions from the *ade*4 library (Chessel et al., 2004) of the *R* software (Ihaka and Gentleman, 1996).

Genetic structure using the fully Bayesian model STRUCTURE

The fully Bayesian clustering model takes a sample of multilocus genotypes and by limiting deviations from Hardy–Weinberg and linkage equilibrium assumptions within populations identifies the number of populations K that best fits the data. Analysis was carried out using *STRUCTURE* 2.1 software (Pritchard et al., 2000). As recommended by Pritchard and Wen (2003), we used the admixture model and the option of correlated allele frequencies between populations, which is considered optimal when subtle cryptic population structure is expected (Falush et al., 2003). We followed Pritchard and Wen (2003)'s recommendation to estimate λ, the Dirichlet parameter, which describes how allele frequencies are correlated between populations. We inferred λ using the uncorrelated allele frequencies model with $K = 1$. We then fixed λ at the inferred value in the correlated allele frequencies model ($\lambda = 0.7206$). From a pilot study, we found that 1,000,000 iterations for the MCMC with burn-in of 100,000 were sufficient for the summary statistics to converge (Pritchard and Wen, 2003). The values of possible Ks we tested varied from 1 to 20 (the number of geographically distinct colonies was 17). For each K value, 10 runs of structure were carried out to take into account the variation of the likelihood value $L(K)$ for each K.

Evanno et al. (2005) recently proposed a new criteria ΔK, rather than the commonly used $L(K)$, to select the most likely number of populations K. Indeed, the distribution of the likelihood of the data with increasing values of K often plateaus or slightly increases after the "real" number of

populations is reached. Thus, there may not be a clear optimal value for $L(K)$. ΔK is based on the rate of change in the likelihood of the data between successive K values. ΔK is calculated as follows (see Evanno et al., 2005, for more details):

- for each K value, $mL(K)$ and $sL(K)$ are estimated as the mean and the standard deviation of $L(K)$ over 10 runs
- for $K = 2$ to 19, $L''(K)$ is calculated as $L''(K) = |mL(K+1) - 2mL(K) + mL(K-1)|$, and
- ΔK is then calculated as $\Delta K = L''(K)/sL(K)$

The modal value of the distribution ΔK was found to be located at the real K number of populations in the simulation study of Evanno et al. (2005). As recommended by these authors, we used the height of this modal value as the signal for the uppermost hierarchical level of genetic structure in the data set. If each colony corresponds to a single cluster, the highest likelihood of the data should be obtained for $K = 17$. The pre-defined populations were partitioned into K' different isolated clusters based on their proportion p of membership for each of the K' clusters. We arbitrarily differentiated two types of colonies: (1) colonies which have a p higher than 75% for a given cluster, i.e. strongly attached with one cluster and (2) colonies which have no p higher than 75%, i.e. shared membership between numerous clusters.

Genetic structure using multivariate analyses

Principal component analysis (PCA, Menozzi et al., 1978; Bertranpetit and Cavalli-Sforza, 1991; Cavalli-Sforza et al., 1993) was also used to investigate the spatial pattern of the genetic variability among colonies. As spatial structure is not always consistent across loci (Smouse and Peakall, 1999), we performed a centered PCA for each marker separately.

Within a multidimensional framework, the allelic frequencies of each marker define Euclidian distances between the colonies in a multivariate space. The PCA seeks synthetic variables which best summarize these distances by optimizing a variance criterion. The new synthetic variables are the scores of the colonies plotted onto the principal axes of the multivariate space. We retained only the scores of the first principal axis, which are by definition the most scattered, and thus highlight the major genetic differences between colonies. Following Menozzi et al. (1978), we plotted the PCA scores onto geographic maps of the sampled area, in order to appreciate their spatial structure.

We used a neighborhood approach to test the significance of the spatial structures. First, the spatial relationships between colonies were established, consisting of a binary state between every pair of colonies: 'neighbors' (i.e. connected), or 'not neighbors' (i.e. not connected). Because information about connectivity between colonies was lacking, we used the Gabriel criterion (Gabriel and Sokal, 1969) to draw the neighboring graph: two colonies A and B were connected if, and only if, there was no other colony inside the circle whose diameter is AB. Once determined, this spatial connectivity was used to compute Moran's I (Moran, 1948, 1950; Cliff and Ord, 1981), which facilitates the assessment of the nature and strength of the spatial pattern of the colonies. This statistic measures the spatial autocorrelation of a given variable. The expected value of I when no structure arises is $-(n-1)^{-1} = -0.0625$, n being, here, the number of colonies. This index takes positive or negative values when neighbors tend to have similar or dissimilar values, respectively. For each locus, the Moran's I of the first PCA scores was computed and tested by a Monte Carlo procedure. The analyses were performed with functions from the *ade*4 library (Chessel et al., 2004) of the *R* software (Ihaka and Gentleman, 1996).

Individual assignment incorporating prior geographical colony structure

In *STRUCTURE*, we assumed that our sample consisted of 17 colonies and that prior dispersal between colonies was rare $v = 0.05$ (as recommended by Pritchard and Wen (2003), and suggested by previous work on stray cats), where v is the probability that an individual is an immigrant from a random colony. We used prior geographical information about colony structure, assuming that each individual originated from the colony where it was sampled, but allowing a small probability that it was an immigrant. We ran *STRUCTURE* with 1,000,000 steps, after a burn-in period of 100,000 steps, for $K = 17$ colonies using geographic information to obtain the assignment probability of each individual to colonies. Each individual was assigned arbitrarily to the colony for which it had a P_{ac} (probability of being from an assumed colony) higher than 80%. Individuals were classified as philopatric if they were assigned in the population from which they were sampled, and immigrant if not. Furthermore, we distinguished two types of immigrants: (i) probable immigrants if $P_{ac} < 80\%$, but higher than the probability of being from any other colony; (ii) true immigrants if $P_{ac} < 80\%$ and a probability of being from another colony higher than 80%.

The probability for each individual to be assigned to one or more of the defined colonies was also estimated using the frequency-based test Monte Carlo re-sampling method (Paetkau et al., 2004) implemented in *Geneclass*2, applying 10,000 re-samples and a type I error of $\alpha = 0.01$ (Piry et al., 2004). For both methods, we calculated the overall immigration rate, *IR*, and the *IR* per colony, respectively, as the proportion of immigrants in the overall data set and within each colony.

Results

Genetic diversity and heterozygosity

The number of alleles per locus across all colonies ranged from 8 (*fca*78) to 18 (*fca*37). Mean observed heterozygosity per locus ranged from 0.46 ± 0.12 (S.E.) to 0.71 ± 0.15 (S.E.), and ranged over all loci from 0.55 ± 0.11 (S.E.) to 0.75 ± 0.17 (S.E.) depending on the colony. Average genetic differentiation between colonies is moderate and highly significant (microsatellite-based $F_{ST} = 0.085$, $P < 0.001$, Wright, 1978), and the multilocus F_{IS} values across all the colonies showed a significant departure from Hardy–Weinberg expectations (microsatellite-based $F_{IS} = 0.12$, $P < 0.001$). All

these results were highly consistent with those previously obtained on a smaller number of colonies ($n = 9$, Say et al., 2003).

Spatial pattern of genetic differentiation

Genetic differentiation using pairwise Fst

All except 22 pairwise values indicated significant differentiation between pairs of colonies ($P < 3.0 \times 10^{-4}$ adjusted P-value for multiple tests, Table 2). The expected contribution to the *Fst* from intra-locus sampling error (Waples, 1998) ranged from 0.022 to 0.052 (mean $\pm$ SD $= 0.037 \pm 0.007$) suggesting that the genetic differentiation between the colonies is probably lower than the degree of differentiation expected from the uncorrected *Fst* values. The geographical distance between the colonies did not explain their genetic differentiation (Mantel correlation $r_M = -0.07$, $P = 0.77$, 9999 permutations, and Mantel correlation $r_M = -0.01$, $P = 0.56$, 9999 permutations, respectively for the raw *Fst* and the corrected *Fst*).

Number of genetic clusters using the fully Bayesian method

As is often found in studies using *STRUCTURE*, we observed no clear maximum for the likelihood $L(K)$ of our data set for successive values of K. Indeed, $L(K)$ increased from $K = 2$ to 7, and then reached a plateau for larger Ks (Fig. 2). In addition, variance between different *STRUCTURE* runs also increased with increasing K. The ΔK criteria (Evanno et al., 2005) reached its modal value for $K = 2$, suggesting that the uppermost level of hierarchical genetic structure has two distinct clusters $C1$ and $C2$ (Fig. 2). Each of the 17 colonies was then assigned to one cluster based on their proportion of membership to both clusters (threshold = 75%). Clusters $C1$ and $C2$ consisted of 9 and 8 colonies, respectively. Eleven of the 17 colonies were strongly assigned to one cluster (4/9 for $C1$, 7/8 for $C2$). All colonies had a proportion of membership in one cluster of at least 60%.

Sub-clustering

For each of the $K = 2$ clusters $C1$ and $C2$, we conducted the same analysis as above to clarify the

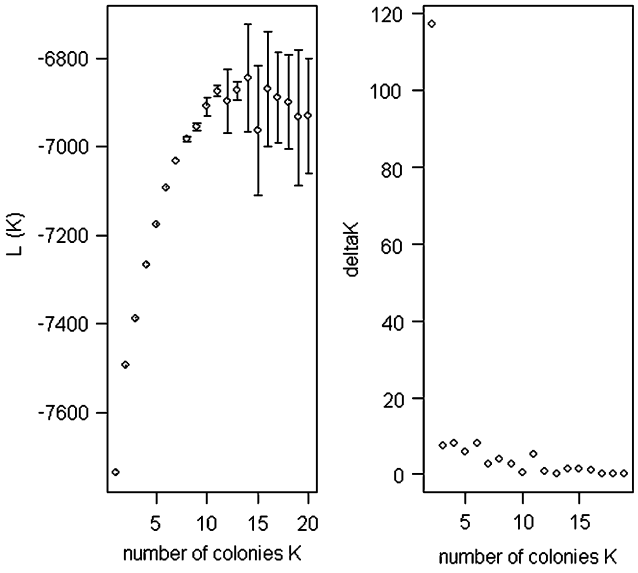

Fig. 2. Mean ($\pm$SD) of $L(K)$ over 10 *Structure* runs for successive K values on the overall data set (left). ΔK as calculated by Evanno et al. (2005): the modal value (here for $K = 2$) shows the uppermost level of genetic structure (right).

Table 2. Genetic differentiation matrix of pairwise *Fst* (Weir and Cockerham, 1984) between the 17 colonies of Nancy

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
2	0.129*	#														
3	0.092*	0.138*	#													
4	0.075*	0.108*	0.022	#												
5	0.072*	0.144*	0.066*	0.046*	#											
6	0.093*	0.099*	0.025	0.040*	0.075	#										
7	0.077*	0.096*	0.105*	0.083*	0.120*	0.082*	#									
8	0.051*	0.076*	0.068*	0.038	0.091*	0.051	0.033	#								
9	0.110*	0.131*	0.108*	0.073*	0.139*	0.094*	0.139	0.112	#							
10	0.062*	0.076*	0.103*	0.101*	0.105*	0.076*	0.100*	0.054*	0.114*	#						
11	0.042	0.099*	0.025	0.025*	0.060*	0.016	0.065*	0.051*	0.072*	0.050	#					
12	0.073*	0.110*	0.083*	0.059*	0.085*	0.093*	0.083*	0.065	0.069	0.101*	0.063*	#				
13	0.107*	0.132*	0.125*	0.111*	0.123*	0.101*	0.142*	0.107	0.157*	0.129*	0.102*	0.128*	#			
14	0.080*	0.107*	0.048*	0.041*	0.085*	0.054*	0.103*	0.056	0.071*	0.096*	0.043*	0.068*	0.103*	#		
15	0.069*	0.094*	0.074*	0.057*	0.108*	0.084*	0.107*	0.051	0.076*	0.053*	0.061*	0.098*	0.113*	0.061*	#	
16	0.116*	0.113*	0.104*	0.081*	0.096*	0.054*	0.127*	0.097	0.101*	0.105*	0.065*	0.121*	0.129*	0.077*	0.095*	#
17	0.066	0.142*	0.101*	0.081*	0.078*	0.124*	0.099	0.116	0.123*	0.110	0.066	0.062	0.129*	0.086	0.098	0.151

*Significant *Fst* values at $P < 3.0 \times 10^{-4}$ (adjusted P-value for multiple tests).

genetic structure of the 17 colonies at a lower hierarchical level (as advocated by Rosenberg et al., 2001). We ran *STRUCTURE* 10 times for each value of K (K from 1 to 9 for $C1$, and for 1 to 8 for $C2$; Burn-in period = 100,000, MCMC = 1,000,000). The ΔK criterion distinguished 3 and 5 genetic sub-clusters, respectively, for $C1$ and $C2$ (Fig. 3). We assigned a given colony to the cluster in which its probability of membership was the highest (Fig. 4). However, to decide if a colony is strongly attached to its cluster, we applied the arbitrary rule of "two times the probability of belonging by chance to a cluster", i.e. a threshold equal to 66% for $C1$ and equal to 40% for C2; the probabilities for a colony to belong by chance to one of the sub-clusters being respectively $1/3 = 33\%$ and $1/5 = 20\%$ for C1 and C2. Three out of the nine colonies belonging to $C1$ were strongly attached to their cluster (Fig. 4) and none out of the remaining six that shared membership between clusters showed a P-value lower than 48.1% (for colony 15). Only two out of the eight colonies in $C2$ were strongly attached to their cluster and one (colony 11) of the remaining six that shared membership between clusters had a value of P very close to 20%. As expected, among the five colonies strongly

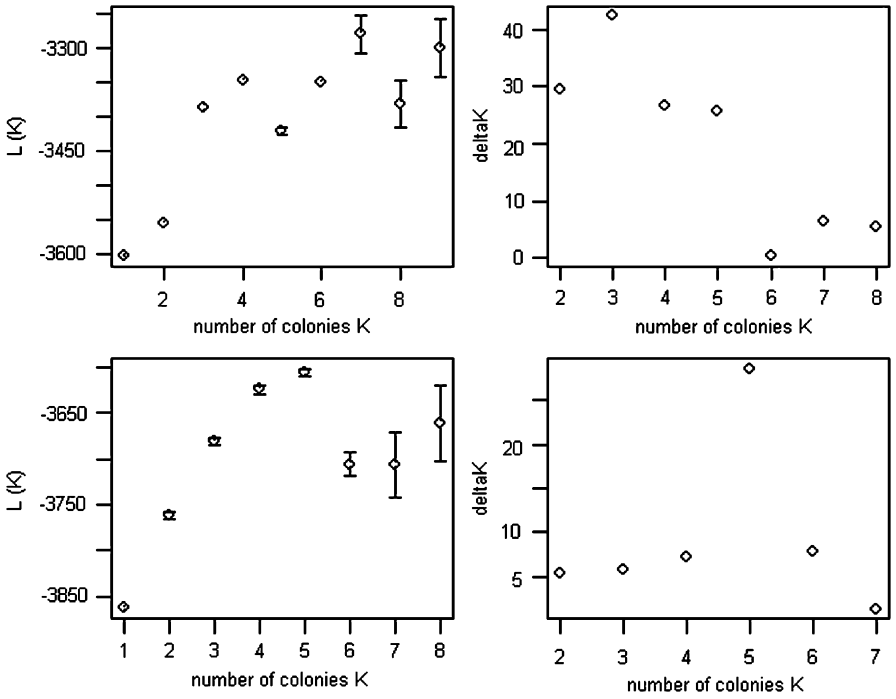

Fig. 3. Mean (±SD) of $L(K)$ over 10 *STRUCTURE* runs for successive K values on the restricted data set for the sub-clusters $C1$ (top) and $C2$ (bottom). ΔK shows a modal value for $K = 3$ and 5, respectively, for the sub-clusters $C1$ and $C2$.

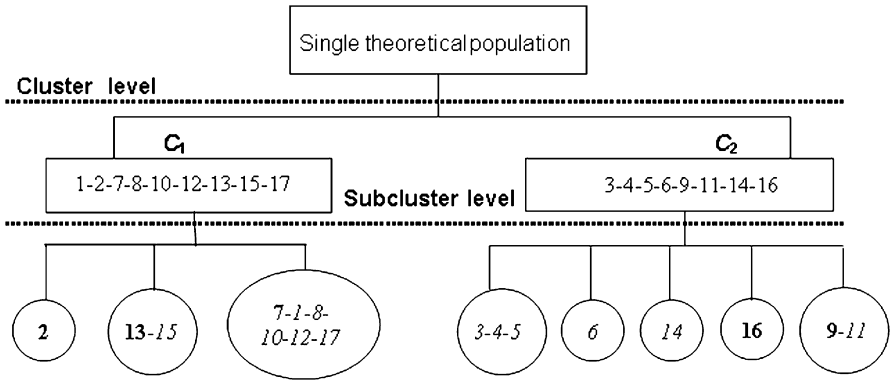

Fig. 4. Clustering and sub-clustering solution from a single theoretical population (Pritchard et al., 2000): the uppermost hierarchical level of genetic structure is two clusters ($C1$ and $C2$), and, at the sub-cluster level, three and five sub-clusters were found for $C1$ and $C2$, respectively. See Materials and methods section for the definitions of the arbitrarily assigned thresholds. Colonies in bold are those strongly attached to their sub-cluster.

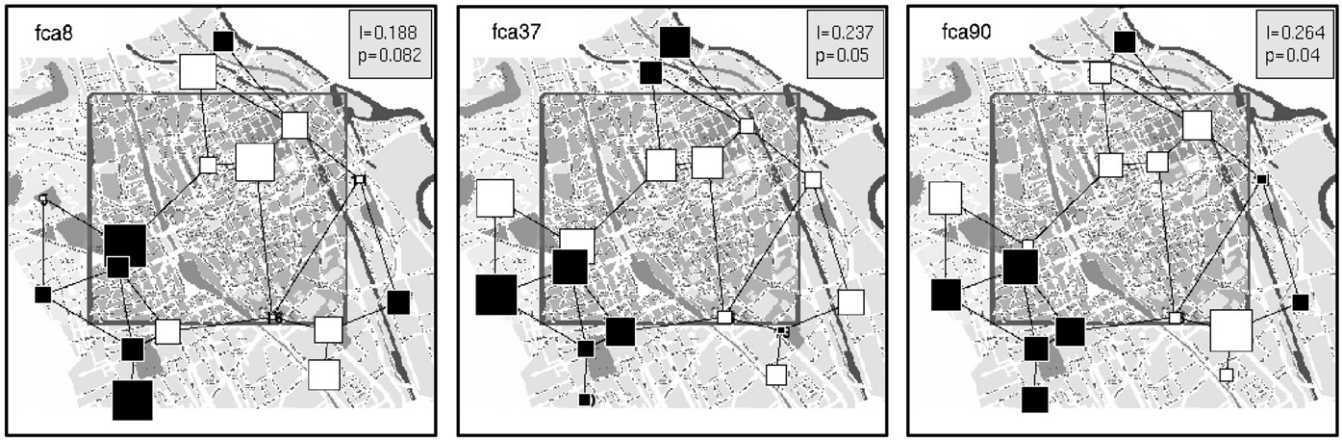

Fig. 5. First axis PCA scores of the 17 colonies for the three informative markers. The marker names are provided on the top-left side of each plot. Large black squares correspond to high-positive scores, whereas large white squares correspond to high-negative scores. Small squares indicate colonies that are not highly differentiated onto the first principal axis. The neighbors defined by the Gabriel criterion are linked by a line. The Moran's *I* values computed for the scores and their associated *P*-values ($p(X \geqslant I)$ after 9999 randomizations) are also provided on each plot.

Table 3. Immigration rate per colony (*IRc*) estimated with *STRUCTURE* ((1), true and probable immigrants) for the *a priori* probability of being an immigrant $v = 0.05$ and *Geneclass2* ((2), immigrant) with a type I error $\alpha = 0.01$

Colony	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Sample size	10	22	12	23	15	11	14	10	9	11	20	14	13	17	11	12	13
True immigrant (1)				1				1		1	1	1		1			
Probable immigrant (1)		1		2		1					2	2		2			1
Immigrant (2)	1	1		1	1	1	2	1	1	1	1	2	1	3	1	1	
IRc (%):																	
STRUCTURE	0	5	0	13	0	9	0	10	0	9	15	21	2	18	0	0	8
Geneclass2	10	5	0	4	7	9	14	10	11	9	5	14	8	18	9	8	0

See Materials and methods section for the definitions of the arbitrarily assigned thresholds in *STRUCTURE*.

attached to their sub-cluster, four were *a priori* defined as "closed".

Spatial pattern of genetic differentiation using multivariate analysis

The scores of the colonies onto the first principal axis of the PCA are plotted onto the sampled area for each locus (Fig. 5). The indicated *P*-value corresponds to the probability that a value as high as or higher than the initial *I* occurs with randomly chosen neighbors. The marker *fca*90 displayed a significant spatial structure ($I = 0.264$; $P = 0.04$) consisting of a group of colonies in the southwest corner of Nancy. This trend was also observed in markers *fca*8 ($I = 0.19$; $P = 0.09$) and *fca*37 ($I = 0.24$; $P = 0.05$).

Geographical proximity and clustering using STRUCTURE and multivariate analysis

Using both approaches, no clear geographical pattern appeared in the constitution of genetic clusters: some spatially proximal populations were not in the same cluster (e.g., colonies 1 and 9, 3 and 17 or, 6 and 7) whereas colonies that are geographically distant from each other were sometimes observed in the same sub-cluster (e.g., colonies 3 and 4). Nonetheless, we can identify a unique uppermost hierarchical level of spatial patterning in genetic differentiation (consistent across the two approaches) in a southwest group of colonies (at least colonies 2, 8, 10 and 12, Figs. 1, 4 and 5).

Individual assignment using prior geographical colony information in the fully Bayesian model

Overall, with $v = 0{\cdot}05$, the mean Immigration Rate (*IR*) was 7.2% (range 0–21.4%, Table 3). Eight of the 17 colonies showed an *IRc* equal to zero but no support exists for lower *IR* in "closed" colonies (Table 3). Only six true immigrants and 11 probable immigrants were detected. Only one out of the six true immigrants came from a colony close to the colony in which it was sampled (sampled in colony 8, originated from colony

10, Fig. 1). Four out of the six true immigrants were males (two less than 1-year old, one 2-year old and one 5-year old) and two were females (2 and 3 years old). Of the probable immigrants, four were males (3, 5, 6 and 7 years old) and seven were females (three less than 1-year old, one each of 2, 3, 5, and 6 years old). Among true immigrants, the sex ratio was highly male-biased whereas this tendency disappeared when true and probable immigrants were pooled (nine males vs. eight females). Among the 17 immigrants (true and probable), the age distribution was significantly different from the age distribution in the overall sample ($X^2 = 14.13$, $P = 0.04$ on 10,000 Monte Carlo permutations, Agresti, 1996). This was mainly due to an excess of old immigrants (more than 5 years old), and to a lesser extent, to an excess of young immigrants (less than 1-year old).

Individual assignment using the frequency-based method

Overall, 19 individuals (eight females and 11 males) were detected as immigrant ($\alpha = 0.01$) generating a mean *IR* (8%, range 0–17.6%, Table 3) similar to the one estimated by *STRUCTURE*. The age structure of these 19 immigrants was evenly distributed across the age classes: three individuals of less than 1-year old; four between 1 and 3 years old, and three of more than 3 years old. This age structure was not different from the age distribution in the overall sample ($X^2 = 9.16$, $P = 0.24$ on 10,000 Monte Carlo permutations, Agresti, 1996).

The six true immigrants in *STRUCTURE* were also classified as immigrants in G*eneclass*2. Only four out of the 11 probable immigrants were also detected as immigrants in this second analysis. Thus nine individuals were immigrants according to analysis using *Geneclass*2 but not for analysis with *STRUCTURE*. Of the 10 individuals classified as immigrants in both methods, only two of them had a different colony of origin: both were sampled in colony 14, but were assumed to have migrated from colonies 5 and 13 for *STRUCTURE* and G*eneclass*2, respectively. Of the nine individuals only classified as immigrants by *Geneclass*2, *STRUCTURE* estimated their probabilities of being from the sampled colony as between 0.825 and 0.999. Of the seven individuals only classified as probable immigrants by *STRUCTURE*, *Geneclass*2 estimated their probabilities of being a first generation migrant as between 0.012 and 0.054. Note that all but two individuals only classified as immigrants by *Geneclass*2 are between 1 and 3 years old, whereas all but one individual classified as probable immigrants by *STRUCTURE* are less than 1-year old or more than 3 years old.

Discussion

Low-immigration rate

Unsurprisingly, in respect of what is known about the urban habitat (Liberg et al., 2000; Devillard et al., 2003), the overall immigration rate was very low (equal to 7–8% with both the Bayesian and frequency-based methods) ranging from 0% to 21% depending on the colony and the method. However, our prediction for lower immigration rates in "closed" colonies is not supported (Table 3). In addition, no sex-bias in immigration rates has been shown by either method. The immigrants were more evenly distributed among colonies with the frequency-based method. Unsurprisingly, increasing ν and α up to 0.10 and 0.05 led to an increase in the number of immigrants from 17 and 19 to 34 and 37 respectively for *STRUCTURE* and *Geneclass*2 (result not shown). Concordance between the methods was good for the identification of the six "true" immigrants. This was less so when considering the "probable" immigrants, both regarding their identification and their colony of origin (see results). Interestingly, the difference between the two methods led to contrasting interpretations of age-specific dispersal patterns. Using *STRUCTURE*, both very young and old individuals were detected as immigrants. Such an immigration pattern is surprising given the usual dispersal pattern in this species: males or females between 1 and 3 years old are usually the dispersers whatever the habitat (Liberg, 1980, in rural habitat, and Devillard et al., 2003, in urban habitat). Thus, this age distribution of immigrants strongly confounded our "life-history based" expectations. However, using *Geneclass*2, no such pattern appeared, both because the individuals detected as immigrants solely by *Geneclass*2 were mainly aged from 1 to 3 years old (which is more consistent with our "life-history based expectations") and because individuals detected as probable immigrants solely by *STRUCTURE* were mainly old stray cats. All these individuals could be misassigned for several biological and methodological reasons.

Firstly, old "probable immigrants" might be past immigrants that contribute little to the genetic profile of a colony and the "probable immigrants" of less than 1 year old might be the offspring of previous immigrants or the offspring of females having reproduced in another colony (i.e. temporary breeding excursions). Secondly, the mismatch between the classifications of *STRUCTURE* and *Geneclass*2 demonstrates that *STRUCTURE* is likely to classify as philopatric some immigrants originating from unsampled colonies (Berry et al., 2004) and that *Geneclass*2 is likely to classify as immigrant some philopatric individuals when sample size is small due to an inflated type I error (Paetkau et al., 2004). As we believe our sample to be nearly exhaustive, we think

that this latter scenario is more likely. Indeed, most of the stray cats only classified as immigrants by *Geneclass*2 have a very high probability of being from the colony they were sampled in (all $P > 0.85$ estimated with *STRUCTURE*), whereas the stray cats only classified as immigrants by *STRUCTURE* were all classified as migrants when we increased α up to 5% (results not shown).

It was thus difficult to confidently detect any pattern of immigration with respect to age and, without any doubt, the small sample sizes in our analysis strongly contributed to that effect. Our data set is at the lower bound to what is generally considered acceptable with respect to the number of individuals per population, the number of loci and the level of genetic differentiation between those populations in studies assessing the performance of these methods (see Paetkau et al., 2004, for a simulation study on a small data set). Unfortunately, conservation biologists will typically face this kind of situation more often in the future. Paetkau et al. (2004) underlined that it is especially valuable to assess the power and practical utility of such assignment methods when studying small and isolated populations. Our study provides a good example of such an empirical data set, allowing us to highlight the main difficulties one can encounter in analyzing results such as presented here.

Genetic structure does not fit the geographic structure

Our results have shown that urban colonies of stray cats do not all genetically differ from each other, whatever methodological approach is used: fully Bayesian clustering method or multivariate analysis. The 17 colonies spatially distributed over the entire study area seemed to be grouped into two genetic clusters at the uppermost hierarchical level, and the only area matching both geographical and genetic proximities seemed to be in the south-west of the sampled area. At the subcluster level in the fully Bayesian approach, eight subclusters were identified with no additional spatial pattern. As highlighted by Pritchard et al. (2000), we believe that the genetic clusters do not always have biological meaning and, for example, the eight colonies in cluster *C*1 probably do not constitute a single colony from a demographic point of view, although these colonies are genetically similar. Alternatively, the low sample size in each colony (14 stray cats per colony on average), the unbalanced sampling design, and the limited number of loci (nine) and alleles per locus with respect to previous studies using *STRUCTURE* software (Randi et al., 2001, Rosenberg et al., 2001; Manel et al., 2002; Rosenberg et al., 2002, 2005) may be problematic for the Bayesian clustering method to separate the colonies, leading to an "artificial" genetic homogeneity among several colonies (Rosenberg et al., 2005). Nonetheless, we did isolate some colonies that do not have the largest sample size (see below) suggesting that our data set was sufficiently informative to identify population substructure. In addition, our multivariate analysis showed, for 3 out of 9 loci, a spatial pattern in genetic differentiation. This spatial pattern, free of any assumptions in terms of the underlying genetic model, is relatively consistent with the fully Bayesian clustering solution in the determination of the uppermost level of hierarchical structure. Such consistency has two implications: (i) *STRUCTURE* is able to detect the uppermost level of hierarchical structure despite a low number of loci, small sample sizes and arbitrarily-assigned thresholds for population assignment; (ii) a free-hypothesis multivariate method gives similar results to that of *STRUCTURE*, while also allowing us to test the significance of the detected spatial pattern. Finally, and as already shown (Smouse and Peakall, 1999), our multivariate analysis showed that all the individual loci were not informative in terms of detecting the spatial pattern of genetic differentiation. Increasing the number of loci is often advocated to enhance the power of genetic analysis. Our results suggest that by increasing the probability of genotyping an informative locus, rather than increasing the number of loci *per se,* will improve efficiency and statistical power. This observation is consistent with previous findings showing that the information provided about populations is not necessarily consistent across loci (Moazami-Goudarzi and Laloë, 2002; Laloë et al., 2007). As multivariate analyses seek common information in a set of descriptors, analyzing a set of weakly congruent markers by a single multivariate analysis would likely not provide an informative typology. Laloë et al. (2007) have shown how the existence of a consensus typology of a population from different markers can be assessed using a multiple co-inertia analysis. In our case, this approach confirmed that the studied loci provided no common information about the cat colonies (results not shown). Thus, it is not surprising to observe a significant spatial pattern in only some of the loci, which might be linked to adjacent selected and spatially structured loci. This raises the open question in relation to the subjective choice of informative loci over uninformative ones.

In a broader context, such low genetic differentiation between several colonies of the population is surprising given the moderate overall *Fst* value across these colonies (Say et al., 2003). The first theoretical explanation for this relative lack of genetic structure might be gene flow, via dispersal processes, which homogenize allele frequencies across these colonies. However, we have shown above that the low overall immigration rate is unlikely to explain the low level of genetic structure at the scale of the sampling area. On the one hand, the

geographic distances separating any pair of colonies can be bridged by a stray cat (Liberg et al., 2000). On the other hand, since functional connectivity in the urban habitat is undoubtedly low, distance alone has no real biological meaning in this habitat. Thus, gene flow might decrease genetic differentiation at a local scale (e.g. in the south-west area) but we do not believe that it is sufficient to homogenize allele frequencies across colonies distant from one another by more than 2 or 3 km in the urban habitat (e.g. colonies 4 and 5, 7 and 10). Dispersal costs could be too high to prevent the immigration of a sufficient number of individuals (Devillard et al., 2003) to homogenize allele frequencies (see Riley et al., 2006, for an example on the effect of a highway barrier on two carnivore species).

Not all colonies are equal

Two significant findings from our fully Bayesian analysis are: (i) four out of five colonies found as strongly linked to their cluster were *a priori* defined as ''closed'' colonies; and (ii) four out of six colonies that were found alone in a cluster or with only one poorly attached colony (colonies 11 and 15, Fig. 4) were *a priori* defined as ''closed'' colonies. These colonies are located in hospital parks or private gardens which are well-delimited by busy roads or high walls. Such ''closed'' colonies have previously been shown to be able to prevent immigration by social and environmental barriers, leading to a high level of philopatry and inbreeding (Liberg et al., 2000; Devillard et al., 2003) and are thus likely to represent family units, facilitating their clustering. In contrast, two colonies (number 6 and 14) that were *a priori* defined as ''opened'' colonies were also alone in their sub-cluster. Overall, our results matched our *a priori* expectation that ''closed'' colonies have their own genetic identity. In this study, we did not find any relationship between the size, the density, the age and sex-composition and the type of the colonies (results not shown). Thus, we think that the variability of local environmental constraints alone can differentiate the two colony types and explain the observed pattern. Previously, stray cat colonies would have constituted a large homogenous genetic pool before the significant increase in human activities of the past 50 years, which has significantly fragmented the urban landscape. Increasing habitat fragmentation has led to a decrease in long-distance gene flow. Consequently, colonies are spatially isolated from one another, but only those showing particular local socio-ecological characteristics (especially in ''closed'' colonies where resource availability is sufficient to allow a high number of resident stray cats to live and to defend access to those resources against intruders) have the potential to be highly genetically differentiated. Concordantly, there is a mixture of both a low overall global genetic structure coming from an ancestral unit and highly localized structure at the colony scale, which may be explained by habitat fragmentation and social constraints.

This finding can be placed in a conservation biology framework. Indeed, if the local structures (i.e. isolated populations) arise due to social barriers to immigration, it might, for example, prevent the success of population reinforcement experiments in an endangered species or the spread of a disease to control or eradicate a pest species (Hamilton et al., 2006).

Finally, our study provides a case-study on the successful use of Bayesian clustering and assignment methods on a real, small data set with numerous fragmented populations that is high pertinent in conservation biology (see also Kotze et al., 2008, for an example in forensic sciences). Despite a limited sample size, the clustering solution matched the ''opened/closed'' classification, suggesting that conservation biologists might infer information on the demographic and social processes from such a clustering analysis in endangered and fragmented populations where such information is lacking.

Acknowledgements

We are very grateful to D. Chessel, J. O'Brien and two anonymous referees for their constructive comments on a previous version of the paper. We thank V. Xémar for collecting the data, L. Say for genotyping the stray cats.

References

Agresti, A., 1996. An Introduction to Categorical Data Analysis. Wiley Series in Probability and Statistics, New York.

Baudry, J., Burel, F., 1998. In: Dover, J.W., Bunce, R.G.H. (Eds.), Dispersal, Movement, Connectivity and Land use Processes. Iale, Preston, UK, pp. 323–339.

Berry, O., Tocher, M.D., Sarre, S.D., 2004. Can assignment tests measure dispersal? Mol. Ecol. 13, 551–561.

Bertranpetit, J., Cavalli-Sforza, L.L., 1991. A genetic reconstruction of the history of the population of the Iberian Peninsula. Ann. Hum. Genet. 55, 51–67.

Bohonak, A.J., 1999. Dispersal, gene flow and population structure. Quat. Rev. Biol. 74, 21–45.

Cahloon, R.E., Haspel, C., 1989. Urban cat populations comparisons by season, subhabitat and supplemental feeding. J. Anim. Ecol. 58, 321–328.

Cavalli-Sforza, L.L., Menozzi, P., Piazza, A., 1993. Demic expansions and human evolution. Science 259, 639–646.

Chepko-Sade, B.D., Halpin, Z.T., 1987. Mammalian Dispersal Patterns: The Effects of Social Structure on Population Genetics. University of Chicago Press, Chicago, IL.

Chessel, D., Dufour, A.-B., Thioulouse, J., 2004. The ade4 package-I—One-table methods. R News 4, 5–10.

Cliff, A.D., Ord, J.K., 1981. Spatial Processes. Model & Applications. Pion, London.

Clobert, J., Danchin, E., Dhondt, A.A., Nichols, J.D., 2001. Dispersal: Individual, Population and Community. Oxford University Press, Oxford.

Corander, J., Waldman, P., Sillanpää, J., 2003. Bayesian analysis of genetic differentiation between populations. Genetics 163, 367–374.

Cornuet, J.M., Piry, S., Luikart, G., Estoup, A., Solignac, M., 1999. New methods employing multilocus genotypes to select or exclude populations as origins of individuals. Genetics 153, 1989–2000.

Coulon, A., Cosson, J.F., Angibault, J.M., Cargnelutti, B., Galan, M., Morellet, N., Petit, E., Aulagnier, S., Hewison, A.J.M., 2004. Landscape connectivity influences gene flow in a roe deer population inhabiting a fragmented landscape: an individual-based approach. Mol. Ecol. 13, 2841–2850.

Courchamp, F., 1996. Etude de l'épidémiologie du virus de l'Immunodéficience Féline dans les populations de chats domestiques (*Felis catus*). Ph.D. Thesis, University of Lyon, 263pp.

Dawson, K.J., Belkhir, K., 2001. A Bayesian approach to the identification of panmictic populations and the assignment of individuals. Genet. Res. 78, 59–77.

Devillard, S., Say, L., Pontier, D., 2003. Dispersal pattern of domestic cats (*Felis catus*) in a promiscuous urban population: do females disperse or die? J. Anim. Ecol. 72, 203–211.

Evanno, G., Regnaut, S., Goudet, J., 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol. Ecol. 14, 2611–2620.

Falush, D., Stephens, M., Pritchard, J.K., 2003. Inference of population structure from multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164, 1567–1587.

Favre, L., Balloux, F., Goudet, J., Perrin, N., 1997. Female-biased dispersal in the monogamous mammal *Crocidura russula*: Evidence from field data and microsatellite patterns. Proc. R. Soc. Biol. Ser. B 264, 127–132.

Gabriel, K.R., Sokal, R.R., 1969. A new statistical approach to geographic variation analysis. Syst. Zool. 18, 259–278.

Goudet, J., 2001. FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). Available at ⟨http://www2.unil.ch/popgen/softwares/fstat.htm⟩.

Hamilton, G.S., Mather, P.B., Wilson, J.C., 2006. Habitat heterogeneity influences connectivity in a spatially structured pest population. J. Anim. Ecol. 43, 219–226.

Hanski, I., Gilpin, M.E., 1997. Metapopulation Biology: Ecology, Genetics, and Evolution. Academic Press, San Diego, CA.

Ihaka, R., Gentleman, R., 1996. R: a language for data analysis and graphics. J. Comp. Graph. Stat. 5, 299–314.

Kimura, M., 1953. 'Stepping-stone' model of population. Annual Report of National Institute of Genetics, vol. 3, pp. 62–63.

Kotze, A., Ehlers, K., Cilliers, D.C., Grobler, J.P., 2008. The power of resolution of microsatellite markers and assignment tests to determine the geographic origin of cheetah (*Acinomyx jubatus*) in Southern Africa. Mamm. Biol., doi:10.1016/j.mambio.2007.10.011.

Laloë, D., Jombart, T., Dufour, A.-B., Moazami-Goudarzi, K., 2007. Consensus genetic structuring and typological value of markers using multiple co-inertia analysis. Gen. Sel. Evol. 39, 545–567.

Liberg, O., 1980. Spacing pattern in a population of rural free roaming domestic cats. Oikos 35, 336–349.

Liberg, O., Sandell, M., Pontier, D., Natoli, E., 2000. In: Turner, D.C., Bateson, P. (Eds.), Density, Spatial Organisation, and Reproductive Tactics in the Domestic Cats and Other Felids. Cambridge University Press, Cambridge, pp. 119–147.

Manel, S., Berthier, P., Luikart, G., 2002. Detecting wildlife poaching: identifying the origin of individuals with Bayesian assignment tests and multilocus genotypes. Conserv. Biol. 16, 650–659.

Manel, S., Gaggiotti, O., Waples, R., 2005. Assignment methods: matching biological questions with appropriate techniques. Trends Ecol. Evol. 20, 136–142.

Mantel, N., 1967. The detection of disease clustering and a generalized regression approach. Cancer Res. 27, 209–220.

Menotti-Raymond, M.A., O'Brien, S.J., 1995. Evolutionary conservation of ten microsatellite loci in four species of Felidae. J. Hered. 86, 319–321.

Menozzi, P., Piazza, A., Cavalli-Sforza, L.L., 1978. Synthetic maps of human gene frequencies in Europeans. Science 201, 786–792.

Moazami-Goudarzi, K., Laloë, D., 2002. Is a multivariate consensus representation of genetic relationships among populations always meaningful? Genetics 162, 473–484.

Moran, P.A.P., 1948. The interpretation of statistical maps. J. Royal Stat. Soc. 10, 243–251.

Moran, P.A.P., 1950. Notes on continuous stochastic phenomena. Biometrika 37, 17–23.

Natoli, E., De Vito, E., 1991. Agonistic behaviour, dominance rank and copulatory success in a large multi-male feral cat, *Felis catus* L, colony in central Rome. Anim. Behav. 42, 227–241.

Neigel, J.E., 1997. A comparison of alternative strategies for estimating gene flow from genetic markers. Annu. Rev. Ecol. Syst. 28, 105–128.

Neigel, J.E., 2002. Is Fst obsolete? Cons. Gen. 3, 167–173.

Paetkau, D., Calvert, W., Stirling, I., Strobeck, C., 1995. Microsatellite analysis of population structure in Canadian polar bears. Mol. Ecol. 4, 347–354.

Paetkau, D., Slade, R., Burden, M., Estoup, A., 2004. Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. Mol. Ecol. 13, 55–65.

Pascal, M., 1980. Structure et dynamique de la population de chats harets de l'archipel des Kerguelen. Mammalia 44, 161–182.

Pearse, D.E., Crandall, K.A., 2004. Beyond Fst: analysis of population genetic data for conservation. Cons. Gen. 5, 585–602.

Piry, S., Alapetite, A., Cornuet, J.M., Paetkau, D., Baudouin, L., Estoup, A., 2004. GENECLASS2: a software for genetic assignment and first-generation migrant detection

Available at 〈http://www.montpellier.inra.fr/CBGP/softwares/〉. J. Hered. 95, 536–539.

Pontier, D., Natoli, E., 1996. Male reproductive success in domestic cat (*Felis catus* L.): a case history. Behav. Process. 37, 85–88.

Pritchard, J.K., Wen, W., 2003. Documentation for Structure software: Version 2. Available from 〈http://pritch.bsd.uchicago.edu〉.

Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. Genetics 155, 945–959.

Randi, E., Pierpoli, M., Beaumont, M., Ragni, B., Sforzi, A., 2001. Genetic identification of wild and domestic cats (*Felis silvestris*) and their hybrids using Bayesian clustering methods. Mol. Biol. Evol. 18, 1679–1693.

Rannala, B., Mountain, J.L., 1997. Detecting immigration by using multilocus genotypes. Proc. Natl. Acad. Sci. USA 94, 9197–9201.

Riley, S.P.D., Pollinger, J.P., Sauvageot, R.M., York, E.C., Bromley, C., Fuller, T.K., Wayne, R.K., 2006. A southern California freeway is a physical and social barrier to gene flow in carnivores. Mol. Ecol. 15, 1733–1741.

Rosenberg, N.A., Burke, T., Elo, K., Feldman, M.W., Freidlin, P.J., Groenen, M.A.M., Hillel, J., Maki-Tanila, A., Tixier-Boichard, M., Vignal, A., Wimmers, K., Weigend, S., 2001. Empirical evaluation of genetic clustering methods using multilocus genotype from 20 chicken breeds. Genetics 159, 699–713.

Rosenberg, N.A., Pritchard, J.K., Weber, J.L., Cann, H.M., Kidd, K.K., Zhivotovsky, L.A., Feldman, M.W., 2002. Genetic structure of human populations. Science 298, 2381–2385.

Rosenberg, N.A., Mahajan, S., Ramachandran, S., Zhao, C., Pritchard, J.K., Feldman, M.W., 2005. Clines, clusters, and the effect of study design on the inference of human population structure. Plos Genet. 1, 0660–0671.

Rousset, F., 1997. Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. Genetics 145, 1219–1228.

Say, L., 2000. Système d'appariement et succès de reproduction chez le chat domestique (*Felis catus*). consequences sur la distribution de la variabilité génétique. Ph.D. Thesis, University of Lyon, 160pp.

Say, L., Pontier, D., Natoli, E., 1999. High variation in multiple paternity of domestic cat (*Felis catus* L.) in relation to environmental conditions. Proc. R. Soc. Biol. Ser. B 266, 2071–2074.

Say, L., Bonhomme, F., Desmarais, E., Pontier, D., 2003. Microspatial genetic heterogeneity and gene flow in stray cats (*Felis catus* L.): a comparison of coat colour and microsatellite loci. Mol. Ecol. 12, 1669–1674.

Slatkin, M., 1985. Gene flow in natural populations. Ann. Rev. Ecol. Syst. 16, 393–430.

Smouse, P.E., Peakall, R., 1999. Spatial autocorrelation analysis of individual multiallele and multilocus genetic structure. Heredity 82, 561–573.

Tabor, R., 1983. The Wild Life of the Domestic Cat. Arrow Books, London.

Waples, R.S., 1998. Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. J. Hered. 89, 438–450.

Weir, B.S., Cockerham, C.C., 1984. Estimation of *F*-statistics for the analysis of population structure. Evolution 38, 1358–1370.

Wilson, G.A., Rannala, B., 2003. Bayesian inference of recent migration rates using multilocus genotypes. Genetics 163, 1177–1191.

Wright, S., 1943. Isolation by distance. Genetics 28, 139–156.

Wright, S., 1951. The genetical structure of populations. Ann. Eugenet. 15, 323–354.

Wright, S., 1978. Evolution and the Genetics of Population, Variability within and among Natural Populations. University of Chicago Press, Chicago, IL.

Xémar, V., 1997. Le chat errant urbain. Contrôle des populations et état sanitaire. Thesis, Veterinary Thesis, University of Lyon, 136pp.

Yamane, A., 1998. Male reproductive tactics and reproductive success of the group living (*Felis catus*). Behav. Process. 43, 239–249.