MOLECULAR ECOLOGY

Molecular Ecology (2013) **22**, 5516–5530 doi: 10.1111/mec.12499

Optimizing the trade-off between spatial and genetic sampling efforts in patchy populations: towards a better assessment of functional connectivity using an individual-based sampling scheme

J. G. PRUNIER,*† B. KAUFMANN,† S. FENET,‡ D. PICARD,§ F. POMPANON,¶ P. JOLY† and J. P. LENA†

*Ecosphère, 3bis rue des Remises, 94100 Saint Maur des Fossés, France, †Université de Lyon, UMR5023 Ecologie des Hydrosystèmes Naturels et Anthropisés, Université Lyon 1, ENTPE, CNRS, Villeurbanne F-69622, France, ‡Université de Lyon, LIRIS, UMR5205, Université Lyon 1, F-69622, France, §GECCO Groupe Ecologie et Conservation des Vertébrés, Université d'Angers, 2 Boulevard Lavoisier, 49045 Angers Cedex 01, France, ¶Laboratoire d'Ecologie Alpine, Univ. Grenoble Alpes, CNRS, UMR 5553, 2233 Rue de la Piscine, 38041 Grenoble Cedex 9, France

Abstract

Genetic data are increasingly used in landscape ecology for the indirect assessment of functional connectivity, that is, the permeability of landscape to movements of organisms. Among available tools, matrix correlation analyses (e.g. Mantel tests or mixed models) are commonly used to test for the relationship between pairwise genetic distances and movement costs incurred by dispersing individuals. When organisms are spatially clustered, a population-based sampling scheme (PSS) is usually performed, so that a large number of genotypes can be used to compute pairwise genetic distances on the basis of allelic frequencies. Because of financial constraints, this kind of sampling scheme implies a drastic reduction in the number of sampled aggregates, thereby reducing sampling coverage at the landscape level. We used matrix correlation analyses on simulated and empirical genetic data sets to investigate the efficiency of an individual-based sampling scheme (ISS) in detecting isolation-by-distance and isolation-by-barrier patterns. Provided that pseudo-replication issues are taken into account (e.g. through restricted permutations in Mantel tests), we showed that the use of inter-individual measures of genotypic dissimilarity may efficiently replace interpopulation measures of genetic differentiation: the sampling of only three or four individuals per aggregate may be sufficient to efficiently detect specific genetic patterns in most situations. The ISS proved to be a promising methodological alternative to the more conventional PSS, offering much flexibility in the spatial design of sampling schemes and ensuring an optimal representativeness of landscape heterogeneity in data, with few aggregates left unsampled. Each strategy offering specific advantages, a combined use of both sampling schemes is discussed.

Keywords: CDPOP, *Ichthyosaura alpestris*, landscape genetics, Mantel tests, pairwise genetic distances, restricted permutations

Received 26 July 2012; revision received 8 August 2013; accepted 14 August 2013

Introduction

In the context of accelerating landscape fragmentation worldwide, the conservation of wildlife populations implies a better understanding of the movements of

Correspondence: Jérôme G. Prunier, Fax: (+33) 4 72 43 11 41; E-mail: jerome.prunier@gmail.com

individuals and genes across landscapes (Cushman 2006; Safner *et al.* 2011a). Landscape genetics has now emerged as an efficient approach to assessing the influence of landscape on gene exchanges among organisms: genetic data are collected across landscapes and then analysed to infer spatial genetic patterns considering various landscape features (Manel *et al.* 2003; Segelbacher *et al.* 2010).

Several authors recently highlighted the importance of sampling design in landscape genetic studies (Schwartz & McKelvey 2009; Anderson *et al.* 2010). The spatial distribution of collected genotypes, particularly the extent of the study area and the interval (or lag distance) between sampling sites, should coincide with the scale of spatial processes known to drive genetic patterns (Anderson *et al.* 2010; Cushman & Landguth 2010a), namely isolation-by-distance (IBD), isolation-by-barriers (IBB) and isolation-by-landscape resistance (IBR), to ensure an optimal coverage of landscape heterogeneity. This prerequisite may be easily met when studying continuously distributed organisms: with only $n = 1$ collected genotype per spot but with a large number P of sampling points, the individual-based sampling scheme ISS allows an optimal representativeness of landscape heterogeneity in data. On the other hand, when individuals are spatially clumped (e.g. pond-breeding amphibians), each aggregate is most often sampled and treated as a discrete population, following a conventional population-based sampling scheme PSS (Funk *et al.* 2005; Goldberg & Waits 2010): the number n of collected genotypes per aggregate is generally >20 (ideally 50; Kalinowski 2005; Broquet & Petit 2009), enabling the estimation of unbiased allelic frequencies.

This widespread approach stems from the predominance of the metapopulation paradigm (Olivieri *et al.* 1995; Hanski 1999) and theoretical models such as Wright's island model (1931), assuming restricted gene flow among local populations and extinction–recolonization events (Harrison 1991): in such models, each aggregate is presumed to constitute a distinct population. However, except for a few species, the relevancy of this paradigm to terrestrial landscapes is quite questionable (Baguette 2004). Indeed, many species tend to follow a patchy population model, with individuals distributed in spatially scattered aggregates and exhibiting higher dispersal rates than expected from the metapopulation paradigm (Harrison 1991; Smith & Green 2005; Fedy *et al.* 2008; Mayer *et al.* 2009). In such cases, the PSS may lead to potential bias due to the challenging delineation of putative population boundaries (Manel *et al.* 2003). Furthermore, the decrease in P in favour of n due to financial constraints has a direct impact on the sampling coverage of landscape heterogeneity: the PSS involves either a drastic reduction in the extent of the study area or an increase in the distance between sampled points, leaving many aggregates unsampled ('ghosts'; Beerli 2004; Broquet & Petit 2009; Lowe & Allendorf 2010).

Ideally, clumped organisms should be sampled following an ISS, the number n of collected genotypes per aggregate being chosen so as to optimally represent the local genetic characteristics of each aggregate ('genetic sampling effort') while allowing a sufficient number P of aggregates to be sampled ('spatial sampling effort'). Though suggested for years (Manel *et al.* 2003), this kind of sampling scheme is still particularly uncommon in empirical studies (but see Austin *et al.* 2011). Even though overlay methods (such as clustering algorithms; Pritchard *et al.* 2000; Chen *et al.* 2007; Jombart *et al.* 2008) usually rest upon individual-based multilocus genotypic data and are not contingent on the use of a PSS, they are most often envisioned alongside correlative analyses (Mantel 1967; Cushman *et al.* 2006) that are traditionally applied to interpopulation genetic distances based on allelic frequencies (Wright 1951; Cavalli-Sforza & Edwards 1967; Nei *et al.* 1983; Weir & Cockerham 1984; Ruzzante 1998).

Investigating the efficiency of an ISS compared with a PSS to detect specific genetic patterns using correlative analyses, we first asked whether interindividual measures of genotypic dissimilarity (a proxy assessment of genetic relatedness) may efficiently replace interpopulation genetic distances based on allelic frequencies. Second, we investigated the trade-off between 'genetic sampling effort' (increasing n) and 'spatial sampling effort' (increasing P) to characterize an individual-based optimal sampling scheme, optimizing the representativeness of both local genetic characteristics and landscape heterogeneity. In most landscape genetic studies, IBD is regarded as the null hypothesis, that is, as the standard process driving genetic differentiation among individuals or populations (e.g. Broquet *et al.* 2006; Emaresi *et al.* 2011). Competing hypotheses, such as IBB or IBR models, are then proposed to determine whether adding landscape variables may improve the predictive power of the standard IBD model (e.g. Cushman *et al.* 2006; Goldberg & Waits 2010). Using simulated multilocus genotypic data, we thus focused on the efficiency of the ISS in detecting IBD as well as IBB, a simple but common competing hypothesis (Cushman *et al.* 2006; Landguth *et al.* 2010; Safner *et al.* 2011b). In the case of IBB, we also examined how the efficiency of an ISS could be enhanced by concentrating the spatial sampling effort on the direct vicinity of the hypothetical barrier and leaving no unsampled aggregates, that is, using a targeted rather than a random sampling scheme (Anderson *et al.* 2010). IBD and IBB detection analyses were performed both in an 'ideal' homogeneous

landscape context and in a more realistic heterogeneous landscape context (IBR), as the spatially heterogeneous viscosity resulting from IBR might alter the detection power of both IBD and IBB. To this end, we used a real topographic map to simulate spatial heterogeneity in dispersal costs, as gradient surfaces are known to adequately represent continuous spatial processes such as dispersal (Cushman & Landguth 2010a). To illustrate the relevance of our simulation results, we also investigated the relative performance of both sampling schemes at detecting IBD patterns using an empirical data set on the alpine newt *Ichthyosaura alpestris*, a European pond-breeding amphibian.

Material and methods

Simulated data sets

Simulations were performed to test the ability of various sampling procedures to detect significant genetic patterns in patchy populations under different geneflow regimes. For this purpose, we used CDPOP (Landguth & Cushman 2010) to simulate over 100 nonoverlapping generations the genetic differentiation among 160 populations (referred to as 'aggregates') randomly placed in a 19 × 22 km area (Fig. 1). The maximal Euclidean distance between pairwise aggregates was 22 946 m; the mean distance between neighbouring aggregates was 1685 ± 587 m according to a Delaunay triangulation. Each aggregate was initiated with 30 individuals and kept at a constant size over generations. Genetic polymorphism was set to 10 microsatellite loci and 10 alleles per locus, genotypes being randomly assigned at the beginning of simulations (mean H_O: 0.90). The mutation rate was set to zero, as in Cushman & Landguth (2010b). Multiple mating was only allowed for males, and the litter size of paired animals was drawn according to a Poisson distribution with the mean set to three. Offspring sex was randomly assigned following a binomial distribution and an unbiased sex ratio. Dispersal was allowed only during the juvenile stage, and the dispersal distance of juveniles was drawn from a probability distribution inversely proportional to a linear function, with D_{max} the maximal dispersal cost distance that may be travelled (associated with a null probability; Landguth & Cushman 2010). D_{max} was fixed at 10%, 20% and 30% of the maximal Euclidean distance between aggregates (22 946 m) to simulate three levels of dispersal range (Low, Medium and High).

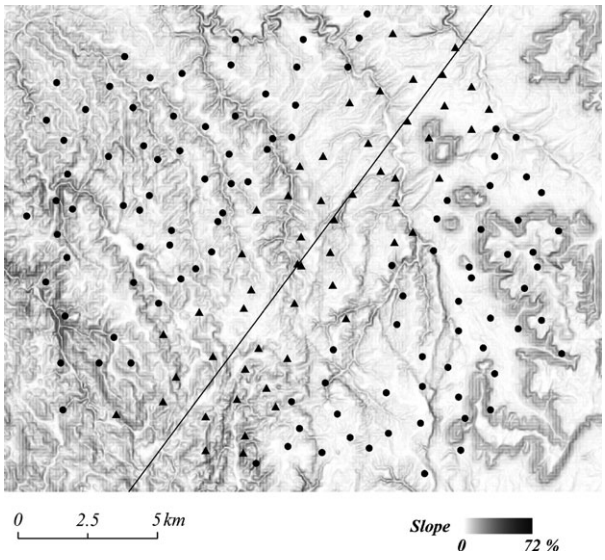

Fig. 1 Random localization of the 160 aggregates of 30 individuals, used to simulate genetic exchanges with CDPOP (Landguth & Cushman 2010). The barrier (black line) segregates individuals in two sets of 80 aggregates. Black triangles stand for aggregates located at less than 3000 m from the barrier.

Cost distances were first based on Euclidean distances between pairwise aggregates, resulting in three data sets with low interaggregate dispersal (D_{Low}), medium interaggregate dispersal (D_{Medium}) and high interaggregate dispersal (D_{High}; see Table 1). Preliminary analyses showed that no genetic pattern could ever be detected for $D_{max} \leq 5\%$, due to an excessive genetic drift, and for $D_{max} \geq 35\%$, due to an excessive homogenization of genotypes among aggregates (data not shown). Our simulated data sets thus covered a relevant range of geneflow regimes in patchy populations. All the tests for IBD detection were performed at the 20th generation, spatial genetic structure being at quasi-equilibrium in all data sets (see Appendix S1, Supporting information).

To investigate the efficiency of the various sampling procedures to detect a recent barrier to dispersal, we simulated three new data sets (D_{Low} + barrier, D_{Medium} + barrier and D_{High} + barrier) by placing an impermeable barrier to dispersal at the 20th generation, which segregated individuals into two sets of 80 aggregates. All the tests for IBB detection were realized at the 30th generation.

To test for the efficiency of an ISS at detecting IBD and IBB patterns in an IBR context, cost distances were also defined according to least-cost-paths computed between pairwise aggregates, resulting in six additional data sets (D_{Low} + IBR, D_{Medium} + IBR, D_{High} + IBR and D_{Low} + barrier + IBR, D_{Medium} + barrier + IBR and D_{High} + barrier + IBR). We used the MATLAB software-coding environment (Mathworks, Inc.) to compute least-cost-paths over a real gradient surface depicting slope (percent slope ranging from 0% to 72%; Fig. 1). Grid cell values, representing the cost of movement through each pixel, were parameterized according to

Table 1 Characteristics of simulated data sets. The maximal dispersal distance D_{max} was based on the homogeneous landscape and was defined as a proportion of the maximal Euclidean distance between pairwise aggregates (22 946 m). The resulting mean dispersal distances are to be compared with the mean Euclidean distance between neighbouring aggregates (1685 ± 587 m according to a Delaunay triangulation): the mean dispersal distances were lower in Low data sets, higher in High data sets and in a similar range of values in Medium data sets. Due to topography, the mean dispersal distances were slightly lower in IBR data sets than in data sets based on the homogeneous landscape

Landscape	Data set	D_{max} (%)	Mean dispersal distance (m)
Homogeneous	D_{Low}	2294.6 (10)	848
	D_{Low} + barrier		
	D_{Medium}	4589.2 (20)	2071
	D_{Medium} + barrier		
	D_{High}	6883.8 (30)	3018
	D_{High} + barrier		
Heterogeneous (IBR)	D_{Low} + IBR	2294.6 (10)	671
	D_{Low} + barrier + IBR		
	D_{Medium} + IBR	4589.2 (20)	1757
	D_{Medium} + barrier + IBR		
	D_{High} + IBR	6883.8 (30)	2707
	D_{High} + barrier + IBR		

the following linear function: $y = (2/72)x + 1$. Resulting grid cell values ranged from 1 to 3. Preliminary analyses showed that no genetic pattern could ever be detected for harsher viscosity (e.g. grid cell values ranging from 1 to 5), because of excessive genetic drift. The addition of spatially heterogeneous viscosity may lead to a slight decrease in dispersal range along with uneven aggregate accessibility ('spatial noise') due to specific landscape configuration.

Sampling procedures

For each simulated data set, we designed 110 random sampling procedures depending on both spatial and genetic sampling efforts. The spatial sampling effort was defined as the number P of sampled aggregates, with P ranging from 10 to 100: in IBD data sets, P aggregates were randomly selected across the landscape, while in IBB data sets, $P/2$ aggregates were randomly selected on either side of the barrier. The genetic sampling effort was defined as the number n of randomly sampled individuals per aggregate, with n ranging from 1 to 10 (ISS) or $n = 20$ (conventional PSS). For each IBB data set, we also designed a targeted sampling scheme, considering only the $P = 52$ aggregates located at less than 3000 m from the barrier (Fig. 1).

Genetic distances

For each sampling procedure, we computed a genetic distance matrix, a Euclidean distance matrix and, in IBB data sets, an effective distance matrix. To compute genetic distance matrices, we used either interindividual genetic distances (for $n \leq 10$ and $n = 20$) or interpopulation genetic distances (only for $n = 20$). Interindividual genetic distances were based on the Bray–Curtis percentage dissimilarity measure (B_c; Legendre & Legendre 1998). B_c is logically related to the shared allele distance (D_{SA}; Bowcock *et al.* 1994), but is directly calculated from pairwise allele frequency differences, rather than from the proportion of alleles not shared between individuals. As in Cushman *et al.* (2006), this metric was also highly correlated with the a_r metric (Rousset 2000; see Appendix S2, Supporting information) and was thus preferred for programming convenience. Interpopulation genetic distances were based on two common measures of genetic differentiation using allelic frequencies: Rousset's linearized F_{ST} ($F_{ST}/(1-F_{ST})$ (hereafter, denoted simply as F_{ST}; Rousset 1997), relying on the balance of gene flow and drift, and Nei's version of Cavalli-Sforza's chord distance D_a (Nei *et al.* 1983), not contingent on any theoretical assumption. To compute effective distance matrices, distances were set to 0 when two sampled aggregates were located on the same side of the barrier and to 1 when they were separated by the barrier (Epps *et al.* 2005).

Matrix correlation analyses

We based our statistical approach on matrix correlation analyses using simple and partial Mantel tests (Mantel 1967; Smouse *et al.* 1986; Legendre & Fortin 2010). Although the effectiveness of this statistical tool is questionable when used for model selection (Cushman *et al.* 2013; Graves *et al.* 2013), Mantel tests are suitable for IBD detection in the case of genetic mutation–migration–drift equilibrium (Guillot & Rousset 2013) and for IBB

detection in the case of a total barrier to gene flow (Jaquiéry *et al.* 2011; Graves *et al.* 2013; see Appendix S3, Supporting information). Usually, the Mantel statistic is tested by standard matrix permutations, that is, by randomly permuting the objects of one of the matrices (Legendre 2000). However, as soon as pairwise aggregates are associated with more than one distance measure, that is, when computing interindividual genetic distances with more than one sampled genotype per aggregate ($n > 1$), Mantel tests are to be performed with restricted permutations, as pseudo-replication issues may arise when clumped genotypes are considered independent. Indeed, restricted permutations will preserve the correlation among individuals belonging to the same aggregate (H_0: independence of aggregates) while standard permutations will dismantle the internal structure of aggregates by considering each individual as an independent object (H_0: independence of individuals). In restricted permutations, permuted objects are thus the aggregates (i.e. the P blocks of n individuals), rather than the individuals (Efron & Tibshirani 1993; Manly 2007). Ideally, the number n of sampled individuals per aggregate should be kept balanced from one aggregate to another to avoid systematic sampling bias; however, when randomly distributed through space, a slightly unbalanced sampling of genotypes is unlikely to influence the relevancy of matrix correlation analyses (see Appendix S4, Supporting information).

In IBD detection, each genetic distance matrix was compared to the corresponding pairwise Euclidean distance matrix using a simple Mantel test with 5000 standard or restricted permutations. In IBB detection, each genetic distance matrix was compared to the corresponding effective distance matrix, after controlling for the effect of the Euclidian distance matrix, using partial Mantel tests with 5000 standard or restricted permutations. Partial Mantel tests were conducted by permuting residuals of genetic distances over Euclidean distances (null model IBD), as advised in Legendre (2000). When using restricted permutations, all within-population distances (that is genetic, Euclidean and effective distances between individuals from a same aggregate) were systematically removed from linearized semi-matrices, to fit with the calculation of Mantel statistics in conventional population-based approaches. All continuous variables (genetic and Euclidean distances) were log-transformed following the $D = \ln(d + 1)$ formula and standardized to meet linearity assumptions.

Each process (sampling genotypes, computing distance matrices and running Mantel tests) was repeated 100 times: a sampling procedure was considered efficient to detect a genetic pattern when at least 95% of repetitions led to a significant Mantel test (P-value ≤ 0.05). This approach allowed the visualization of the most parsimonious sampling procedures along an optimal detection threshold, defined as the minimum number nP of genotypes required to detect a significant genetic structure. Optimal detection thresholds were used to compare interindividual and interpopulation measures of genetic distances in PSSs ($n = 20$) and to investigate the trade-off between genetic and spatial sampling efforts in ISSs ($n \leq 10$).

All operations (sampling aggregates and individuals, computing genetic, Euclidian and effective distances and running Mantel tests with standard or restricted permutations) were automated in the MATLAB software-coding environment (Mathworks, Inc.). F_{ST} measures were calculated as in Fstat (Goudet 2001), following Weir & Cockerham (1984), and then linearized following Rousset (1997). D_a measures were calculated as in Genetix 4.03 (Belkhir *et al.* 2004). We checked the validity of MATLAB scripts by comparing F_{ST} and D_a measures in a few situations with F_{STAT} and Genetix 4.03 outputs, respectively, and Mantel correlation coefficients (r) and P-values obtained from standard permutations with ZT (Bonnet & Van de Peer 2002) outputs.

The Matlab script for simple and partial Mantel tests with restricted permutations is provided in Appendix S5 (Supporting information). Any further updates may be available at the following address: [http://www.jeromeprunier.eg2.fr/5.html]. A free software-efficient C implementation is also available at the following address: [http://liris.cnrs.fr/serge.fenet/Recherche/blockBasedMantelTest].

Empirical data set with both an ISS and a PSS in the alpine newt

The alpine newt (*I. alpestris*) is a widespread pond-breeding amphibian in northeastern France. This species benefits from extensive livestock farming, which allows the preservation of many artificial or natural ponds in pastures. It may be highly nomadic (Perret *et al.* 2003), with high gene flow and low genetic differentiation among subpopulations (Emaresi *et al.* 2011), suggesting a patchy population model. To compare the efficiency of various sampling procedures in detecting a significant genetic structure in this species, genetic data were collected during the 2010 breeding season according to both an ISS and a PSS in France (46.51°N, 5.17°E), in an area of approximately 20 × 25 km (Fig. 2). We prospected a total of 223 aquatic sites (e.g. artificial or natural ponds, flooded ruts, swamps, hereafter denoted as 'aggregates'). Each aggregate was prospected with a dip net for 30–60 mins (depending on site size and configuration) or until at least two alpine newts (a male and a female) were captured. Alpine newts were detected in 78 aggregates. Only six aggregates allowed the capture

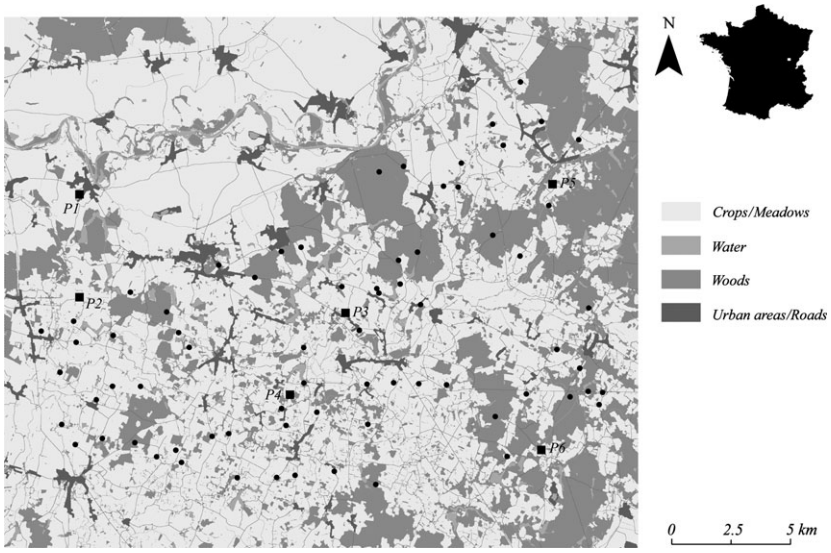

Fig. 2 Localization of alpine newts' genotypes in Burgundy, France (46.51°N, 5.17°E). Black squares stand for aggregates sampled following a PSS (n ≥ 23).

of at least 20 individuals: they were thus sampled following a conventional PSS with at least $n = 23$ sampled individuals ($P = 6$, $N = 159$ genotypes, Euclidean distance between pairwise aggregates ranging from 4168 m to 22 244 m; P_1–P_6 in Fig. 2). Only one individual per sex was sampled in the 72 other aggregates (ISS, $n \leq 2$, $N = 125$ genotypes, Euclidean distance between pairwise aggregates ranging from 209 m to 24 878 m). Nondestructive genetic samples were taken from each captured individual, using nonsterile buccal swabs (Broquet *et al.* 2007). DNA extraction, PCR amplifications and genotyping were performed using 12 microsatellite loci without null alleles (Table 2), following Prunier *et al.* (2012).

In the case of the PSS ($P = 6$), we computed two interpopulation genetic distance matrices using F_{ST} and D_a and an interindividual genetic distance matrix using B_c; in the case of the ISS ($P = 72$), we only computed an interindividual genetic distance matrix using B_c. These matrices were used in Matlab to perform spatial autocorrelation analyses through nondirectional Mantel correlograms (Smouse & Peakall 1999; Borcard & Legendre 2012), to determine the scale at which IBD occurred in this data set (Epperson 2003). For this purpose, Euclidean distance classes were defined every 5000 m, so that each distance class included at least one pair of aggregates, resulting in five binary matrices representing the membership of individuals or populations to the distance class being tested (with 0 for pairs belonging to the same distance class and 1 otherwise). With 72 sampled aggregates ensuring an optimal spatial coverage of the study area, the ISS allowed spatial autocorrelation analysis to be performed at a finer resolution: distance classes were thus also defined every 2500 m, resulting in 10 binary matrices. Each binary matrix was compared with the related genetic distance matrix using a simple Mantel test with 1000 standard or restricted permutations. We then plotted the Mantel correlation values over distance classes, with a 95% confidence interval determined by bootstrap resampling (1000 iterations). To test for the significance of each autocorrelogram, genetic and related Euclidean distance matrices were also used to perform simple Mantel tests with 9999 standard or restricted permutations over whole data.

Results

Relative performance of interindividual and interpopulation genetic distances

Whatever the spatial genetic pattern (IBD or IBB) and the level of spatial heterogeneity (with or without IBR), the use of metrics based on allelic frequencies in a PSS (D_a and F_{ST}, $n = 20$) led to very similar optimal detection thresholds (Figs 3 and 4). The minimum number nP of genotypes required to detect significant IBD patterns increased with the increase in interpatch movements (from D_{Low} to D_{High}) and was always slightly lower in IBR data sets. On the contrary, the minimum number nP of genotypes required to detect significant IBB patterns decreased with the increase in interpatch movements (from D_{Low} + barrier to D_{High} + barrier) and was always slightly higher in IBR data sets. No IBB pattern was ever detected in data sets with low interaggregate dispersal (D_{Low} + barrier and D_{Low} + barrier + IBR).

When based on exactly the same genetic data ($n = 20$), measures of relatedness between individuals (B_c) performed roughly as well as metrics based on allelic frequencies in all IBD and IBB situations. Interindividual genetic distances, though nonsignificant, led to better IBB detection than metrics based on allelic frequencies in D_{Low} + barrier data set.

Table 2 Main characteristics of empirical genetic data in the alpine newt *Ichthyosaura alpestris*. For each locus (Garner *et al.* 2003; Prunier *et al.* 2012), number of alleles (*A*; in brackets, effective number of alleles), expected and observed heterozygosity (H_E and H_O) are given for the six aggregates that were sampled following a conventional PSS (P_1–P_6, *N* ranging from 23 to 30) and for the 72 aggregates that were sampled following an ISS. In the latter case, the $N = 125$ genotypes were pooled in a unique cluster (Ind)

Clusters	*N*	*A*	H_E	H_O	*A*	H_O	H_E
		CopTa1 (Prunier *et al.* 2012)			CopTa11 (Prunier *et al.* 2012)		
P_1	29	3	0.572	0.552	6	0.691	0.759
P_2	25	3	0.541	0.640	7	0.673	0.760
P_3	28	3	0.614	0.714	5	0.745	0.821
P_4	23	3	0.577	0.522	6	0.736	0.739
P_5	24	3	0.430	0.417	4	0.741	0.708
P_6	30	3	0.581	0.467	7	0.695	0.600
Ind	125	3	0.573	0.616	9	0.751	0.736
Total	284	3 (3.00)			9 (7.25)		
		CopTa2 (Prunier *et al.* 2012)			CopTa12 (Prunier *et al.* 2012)		
P_1	29	3	0.673	0.586	3	0.346	0.345
P_2	25	3	0.615	0.800	2	0.150	0.160
P_3	28	3	0.662	0.786	2	0.195	0.214
P_4	23	3	0.598	0.478	2	0.496	0.478
P_5	24	4	0.691	0.542	2	0.467	0.458
P_6	30	3	0.608	0.733	2	0.440	0.500
Ind	125	3	0.627	0.576	3	53.803	53.000
Total	284	4 (3.08)			3 (2.54)		
		CopTa3 (Prunier *et al.* 2012)			CopTa13 (Prunier *et al.* 2012)		
P_1	29	7	0.755	0.759	3	0.522	0.552
P_2	25	7	0.681	0.640	4	0.543	0.480
P_3	28	6	0.659	0.607	4	0.622	0.679
P_4	23	5	0.592	0.696	4	0.590	0.652
P_5	24	8	0.769	0.792	4	0.495	0.625
P_6	30	5	0.695	0.600	4	0.597	0.500
Ind	125	10	0.676	0.688	4	0.575	0.504
Total	284	11 (7.93)			5 (3.90)		
		CopTa4 (Prunier *et al.* 2012)			CopTa14 (Prunier *et al.* 2012)		
P_1	29	4	0.598	0.483	3	0.515	0.655
P_2	25	3	0.561	0.600	3	0.575	0.640
P_3	28	3	0.538	0.536	3	0.568	0.643
P_4	23	4	0.600	0.522	3	0.527	0.435
P_5	24	3	0.526	0.500	3	0.613	0.500
P_6	30	3	0.605	0.600	3	0.514	0.500
Ind	125	5	0.587	0.504	3	72.936	77.000
Total	284	5 (4.06)			3 (3.00)		
		CopTa8 (Prunier *et al.* 2012)			Ta1Ca1 (Garner *et al.* 2003)		
P_1	29	2	0.242	0.276	3	0.249	0.276
P_2	25	2	0.040	0.040	4	0.541	0.560
P_3	28	4	0.532	0.643	4	0.316	0.321
P_4	23	3	0.300	0.261	4	0.468	0.522
P_5	24	5	0.483	0.500	4	0.488	0.583
P_6	30	3	0.310	0.233	3	0.391	0.400
Ind	125	6	0.376	0.352	5	0.418	0.392
Total	284	6 (4.40)			7 (4.23)		
		CopTa10 (Prunier *et al.* 2012)			Ta1Caga4 (Garner *et al.* 2003)		
P_1	29	9	0.811	0.828	15	0.915	0.931
P_2	25	10	0.811	0.800	15	0.934	0.960
P_3	28	8	0.792	0.821	17	0.932	0.893
P_4	23	12	0.855	0.783	18	0.954	1.000
P_5	24	9	0.801	0.583	16	0.927	0.958
P_6	30	12	0.866	0.767	20	0.936	0.900
Ind	125	24	0.886	0.816	35	0.942	0.920
Total	284	25 (16.15)			37 (24.86)		

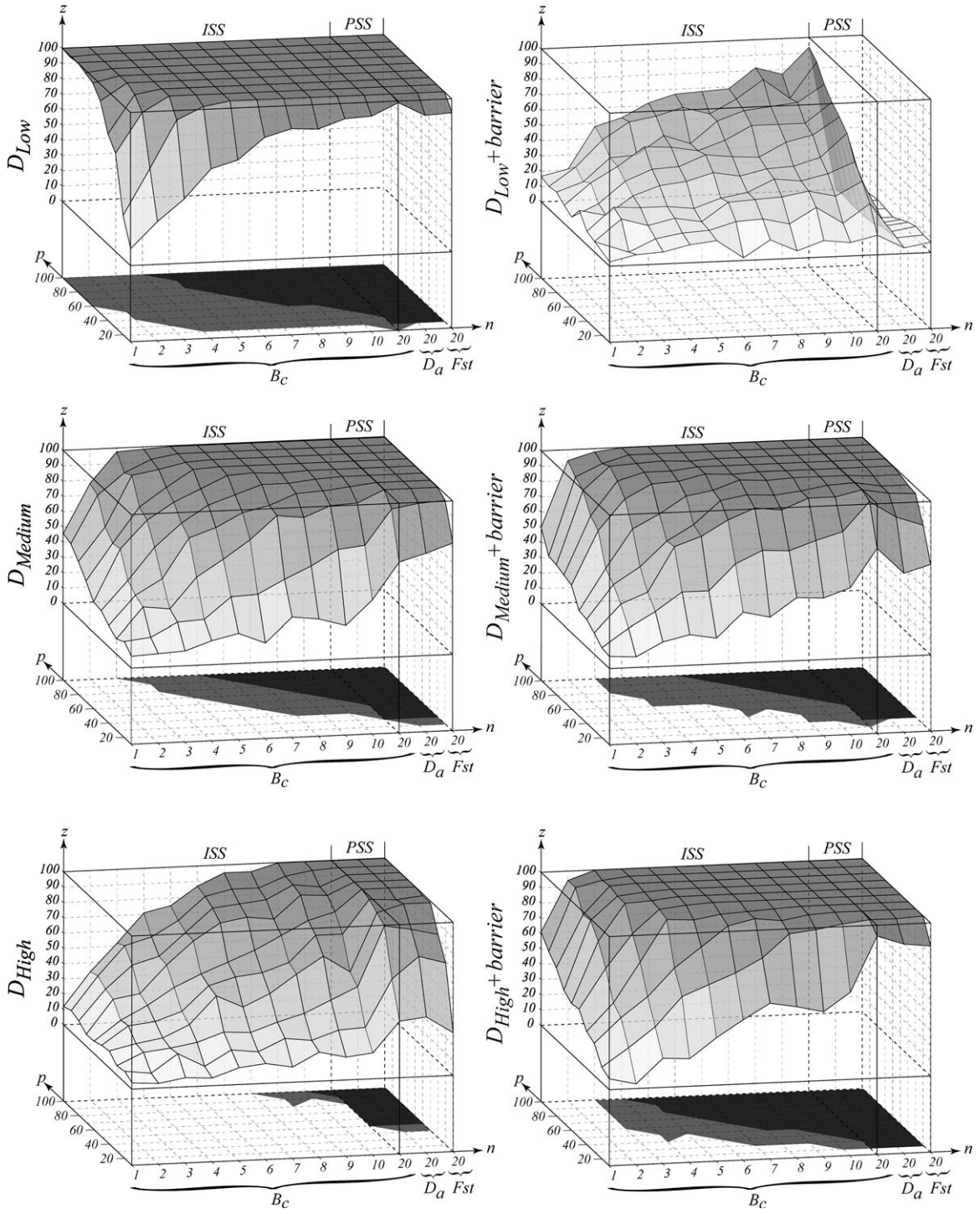

Fig. 3 Genetic structure detection in D_{Low}, D_{Medium} and D_{High} (IBD) and D_{Low} + barrier, D_{Medium} + barrier and D_{High} + barrier (IBB), expressed in percentage of the total number of significant replicates (*z*-axis), according to the spatial sampling effort (number *P* of selected populations; *y*-axis), the genetic sampling effort (number *n* of subsampled individuals per population; *x*-axis) and the measure used to calculate pairwise genetic distances (B_c: measures of relatedness between individuals using the Bray–Curtis coefficient; D_a and F_{ST}: measures of genetic differentiation between populations based on allelic frequencies using Nei's version of Cavalli-Sforza's chord distance and Rousset's linear F_{ST}; see text for details). Light grey surfaces at the bottom of graphs enable a better visualization of optimal detection thresholds, that is, most parsimonious sampling procedures leading to a significant detection at 95%; dark grey surfaces at the bottom of graphs enable a better visualization of sampling procedures requiring a number *nP* of sampled genotypes at least as high as the less parsimonious population-based sampling scheme (PSS).

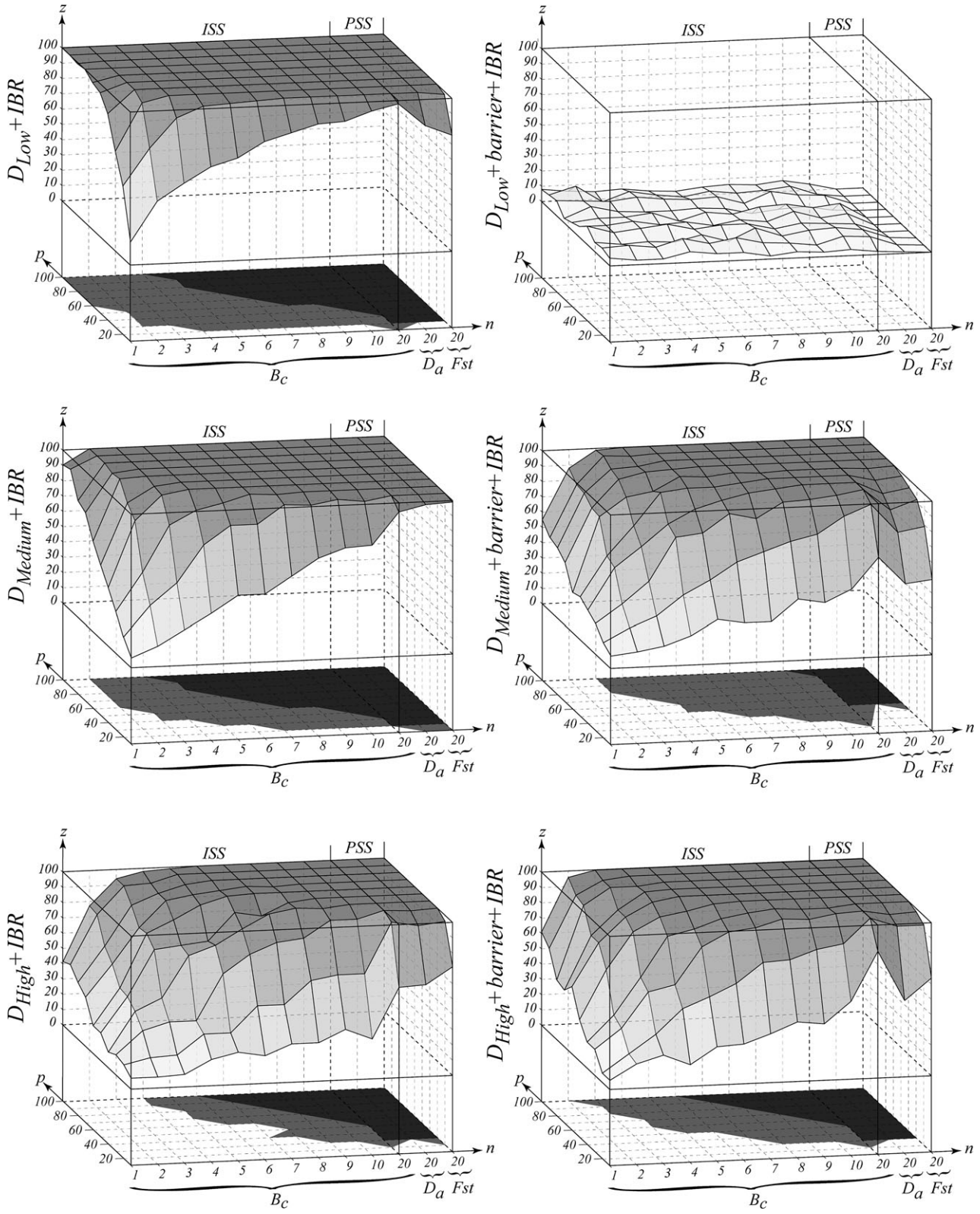

Fig. 4 Genetic structure detection in D_{Low} + IBR, D_{Medium} + IBR and D_{High} + IBR (IBD) and D_{Low} + barrier + IBR, D_{Medium} + barrier + IBR and D_{High} + barrier + IBR (IBB). See Fig. 3 for details.

Relative performance of random sampling schemes

When considering the minimum number nP of genotypes required to detect a significant genetic structure, the use of an ISS with interindividual genetic distances (B_c, $n \leq 10$) provided IBD and IBB optimal detection thresholds at least as low as the use of a PSS, whatever the data set and the level of spatial heterogeneity (Figs 3 and 4). ISSs were particularly parsimonious for IBD detection in D_{Low} and D_{Medium} data sets and for IBB detection in all situations except in D_{Low} + barrier + IBR data set. Though nonsignificant in D_{Low} + barrier data set, the use of measures of relatedness led to better IBB detection than any metric based on allelic frequencies.

Whatever the data set, there was a trade-off between the number n of sampled individuals per aggregate and the number P of sampled aggregates: detection thresholds increased either with the increase in n to the expense of P or conversely, with the increase in P at the expense of n. Except in D_{High}, D_{Low} + barrier and D_{Low} + barrier + IBR data sets, three to four genotypes per aggregate were sufficient to detect significant genetic structures.

Relative performance of targeted sampling schemes

When using a targeted sampling scheme, IBB detection was significant as soon as $n \geq 3$, in all data sets except D_{High} + barrier + IBR, the latter requiring the sampling of at least four genotypes per aggregate (Fig. 5). When considering the total number nP of sampled genotypes, the targeted ISS was far more parsimonious than the targeted PSS, although both designs outperformed the random sampling schemes (Figs 3 and 4).

Relative performance of ISS and PSS to detect an IBD pattern in the alpine newt

When using interpopulation genetic distances in a PSS ($P = 6$), neither Mantel tests nor spatial autocorrelation analyses allowed the detection of any IBD pattern (F_{ST}: $r = 0.284$, *P-value* = 0.174; D_a: $r = 0.301$, P-value = 0.159; Fig. 6a). When using interindividual measures of relatedness in a PSS ($P = 6$), significant genetic relatedness between pairwise individuals was only detected for the first 5 km (Fig. 6b), the global Mantel test being close to significance (B_c: $r = 0.07$, P-value = 0.054). On the contrary, the use of an ISS (72 sampled aggregates) led to the detection of a highly significant IBD pattern (B_c: $r = 0.073$, P-value $<10^{-4}$). Spatial autocorrelation analyses were significant at both spatial resolutions, with individuals less than 10 km apart showing significant genetic similarity (Fig. 6c–d). As highlighted by Legendre (2000), the numerical value of the Mantel statistic required to reach significance was low, particularly in individual-based approaches. The confidence intervals determined by bootstrap resampling were much smaller with interindividual genetic distances (Fig. 6b, c, d) than with interpopulation genetic distances.

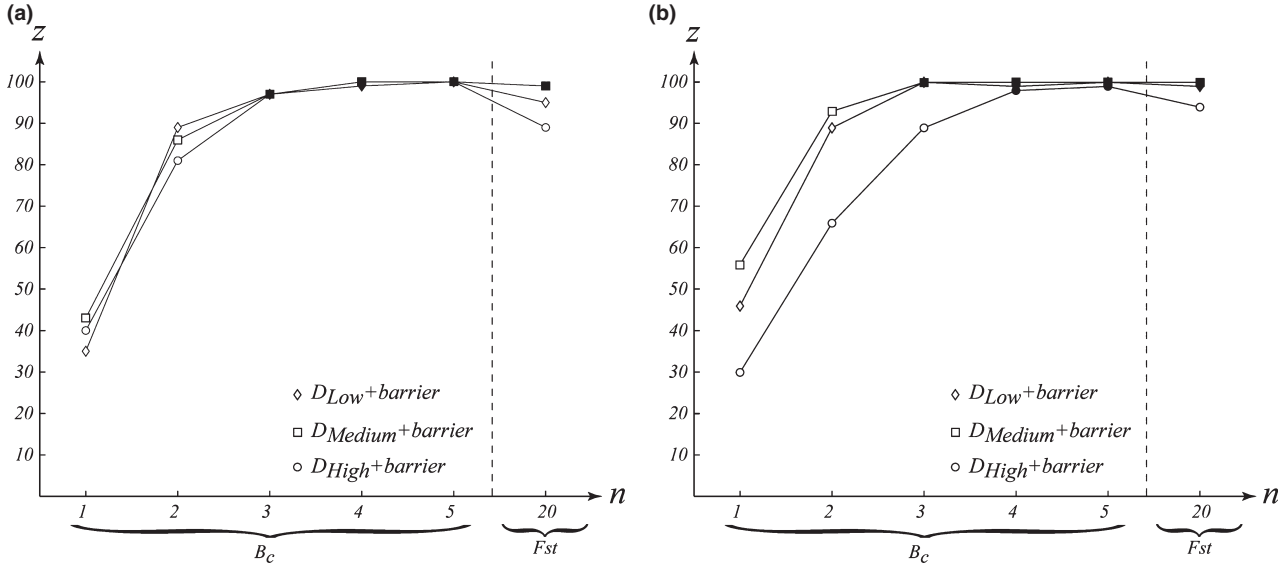

Fig. 5 IBB detection following a targeted sampling scheme (52 aggregates localized at less than 3000 m from the barrier), expressed in percentage of the total number of significant replicates (*z*-axis), according to the number n of subsampled individuals per population (*x*-axis), in the homogeneous landscape (a) or with IBR (b). Pairwise genetic distances were based on measures of relatedness between individuals for $n \leq 5$ (using B_c) or on allelic frequencies for $n = 20$ (using F_{ST}). A filled symbol indicates that the test was significant in more than 95% of the replicates.

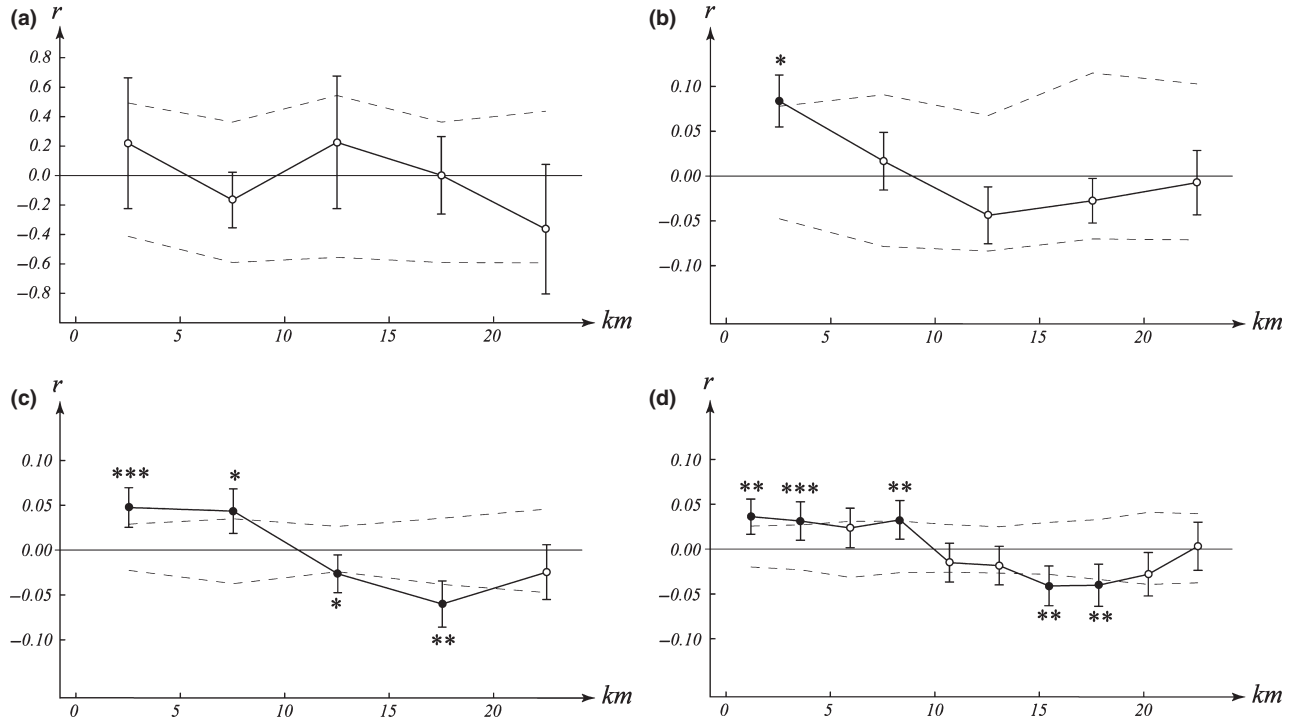

Fig. 6 Spatial patterns of genetic similarity among aggregates sampled following a PSS (a–b) and an ISS (c–d). For the PSS, we used (a) interpopulation genetic distances based of allelic frequencies (F_{ST}; the use of D_a led to a very similar pattern) and (b) interindividual measures of relatedness (B_c). For the ISS, we used only interindividual measures of relatedness (B_c) with distance classes defined every 5000 m (c) and every 2500 m (d). *r*: standard Mantel correlation with 1000 standard or restricted permutations. Error bars bound the 95% confidence interval about *r* as determined by bootstrap resampling. Upper and lower confidence limits (dotted line) bound the 95% confidence interval about *r* under the null hypothesis of no spatial structure. ∗: *P*-value <0.05; ∗∗: *P*-value <0.01; ∗∗∗: *P*-value <0.001.

Discussion

Overall, our study shows that the interindividual measure of genotypic dissimilarity, a proxy assessment of genetic relatedness, is a convenient surrogate for allelic frequencies when investigating landscape connectivity, whatever the spatial distribution and dispersal range of organisms and regardless of the presence of uncontrolled landscape heterogeneity. More importantly, it paves the way to adapting spatial sampling schemes to the objectives of landscape connectivity studies. In particular, our results indicate that the ISS enables an efficient trade-off between genetic sampling effort, allowing an accurate capture of genetic information at the spot level, and spatial sampling effort, allowing an optimal representativeness of landscape heterogeneity in data.

As expected by theory (Wright 1943), reducing the dispersal range of organisms increased the intensity of IBD patterns in the homogeneous landscape (Fig. 3): in this situation, distant populations rarely exchange migrants and drift apart over time, while neighbouring populations show higher genetic similarity. On the contrary, reducing the dispersal range of organisms reduced the detection power of IBB patterns, as less mobile individuals are less prone to encounter the barrier than individuals with high dispersal abilities (Safner *et al.* 2011a; Blair *et al.* 2012). Introducing topographic obstacles (i.e. IBR) slightly altered the detection power of both IBD and IBB patterns in all but the data set with low interaggregate dispersal, in which case the IBB pattern was no longer detected (D_{Low} + barrier + IBR; Fig. 4). This slight effect of IBR, observed in most situations, was probably due to the overall decrease in the dispersal range, rather than to the spatial heterogeneity in dispersal *per se* (i.e. spatial noise due to uneven aggregate accessibility), both phenomena being concomitant with the introduction of a topographic resistance to dispersal in addition to IBD (see Table 1). Simulations focusing on spatial noise only (i.e. with the maximal dispersal range D_{max} being based on the friction map rather than on the homogeneous map; data not shown) confirmed this point, as well as the strong negative effect of spatial heterogeneity on the detection power of IBB patterns in organisms with low dispersal abilities. This latter result may be due to the properties of the topographic map, with relief and hydrographic

network resulting in highly resistant lines segregating several clusters of aggregates (see Fig. 1), and affecting IBB signal-to-noise ratio when analyses are performed on the scale of the whole map. Nevertheless, even in such a situation, the signal-to-noise ratio may be enhanced by using a targeted sampling procedure. Indeed, IBB detection was clearly improved when focusing on the landscape feature of interest even in data sets with low interaggregate dispersal (Fig. 5), that is, when sampled aggregates were selected to avoid both ghost populations and large gaps relative to the putative barrier to gene flow (Anderson *et al.* 2010).

Both simulated and empirical data sets showed that interindividual measures of genotypic dissimilarity, such as the Bray–Curtis index B_c, could efficiently replace interpopulation genetic distances based on allelic frequencies. Interindividual genetic distances were roughly as efficient as interpopulation genetic distances when applied to the same sample (i.e. PSS) in all situations, and outperformed the use of allelic frequencies for IBB detection, especially in a data set with low interaggregate dispersal (see Fig. 2). Furthermore, our analyses indicate that using interindividual genetic distances substantially enhanced IBD detection in the alpine newt (see Fig. 6). While metrics based on allelic frequencies may suffer from a loss of resolution due to the averaging of genetic information over individuals (Kelly *et al.* 2010), measures of interindividual genetic relatedness are based on the averaging of the genetic information over alleles: the evolution of alleles co-occurrence through generations is thus likely to show less inertia than the evolution of allelic frequencies. One may therefore expect interindividual measures of genotypic dissimilarity to outperform measures based on allelic frequencies, especially in organisms with low dispersal abilities as suggested by our results regarding IBB detection (Fig. 3). Besides, provided that pseudo-replication issues are taken into account (e.g. using restricted permutations in Mantel tests), the use of interindividual genetic distances implies an increase in the size of distance matrices, potentially increasing the accuracy and the inferential power of correlative analyses (see Fig. 6a–b; Legendre & Fortin 2010). We thus argue that many unpublished empirical genetic studies based on a PSS may benefit from the use of interindividual measures of genetic dissimilarity instead of interpopulation measures of genetic differentiation, although further work will be required for a better understanding of differences and potential benefits of various interindividual metrics. In any case, the number n of genotypes should be kept balanced from one aggregate to one another, although a slightly and randomly unbalanced sampling is unlikely to affect matrix correlation analyses (see Appendix S4, Supporting information).

Not contingent on the use of allelic frequencies, the use of individuals as the operational unit also allows the number n of sampled genotypes per aggregate to be reduced in favour of a better coverage of landscape heterogeneity. For the same genotyping effort (i.e. nP), increasing P to the expense of n in an ISS, that is, decreasing the genetic sampling effort at the spot level in favour of a better sampling coverage of spots at the landscape level (Jaquiéry *et al.* 2011), was never less efficient than the conventional PSS. It was even more efficient in most situations: no more than ten individuals had to be sampled per aggregate to efficiently detect IBD and IBB genetic patterns, three or four individuals being sufficient in most situations (Figs 3 and 4). Correlative analyses may therefore gain from including sparsely inhabited spots that are usually discarded in conventional PSSs. Furthermore, by drastically increasing the number P of aggregates that may be sampled, the ISS offers a higher flexibility in spatial sampling designs. For instance, even though a targeted PSS was clearly more efficient than a random PSS, it was largely outperformed by the targeted ISS as the same conclusions were obtained with only a few sampled genotypes per aggregate (Fig. 5). In the same way, the ISS allowed a more exhaustive sampling and a decrease in the distance between sampled aggregates in the alpine newt, thereby providing a finer description of spatial pattern of genetic similarity among aggregates than when using a PSS (Fig. 6c–d). Individuals located less than 10 km apart showed significant genetic similarity. This pattern is highly relevant from a conservation perspective as it suggests that genes spread out according to a 'stepping-stone' model (Kimura & Weiss 1964; Fedy *et al.* 2008), that is over greater distances than the dispersal ability of the alpine newt might suggest (less than 1000 m a year; Smith & Green 2005; Kovar *et al.* 2009). More broadly speaking, the ISS may allow a better representativeness of landscape heterogeneity through an exhaustive sampling of aggregates over a larger study area, and/or over a number of independent replicates, allowing accurate inferences about genetic patterns (Anderson *et al.* 2010; Manel *et al.* 2010). In this respect, genetic data from an ISS may be particularly well fitted for use in mixed models (Clarke *et al.* 2002; Selkoe *et al.* 2010; Van Strien *et al.* 2012), providing that specific random effects are implemented to take into account pseudo-replication issues at various spatial scales. Indeed, preliminary analyses showed that mixed models and Mantel tests with restricted permutations yielded similar results in our study (see Appendix S3, Supporting information). As an example, collinearity between environmental variables, underlying the alternative resistance hypotheses that can be hypothesized from a same landscape, has recently been suggested as

an important drawback when applying matrix correlation analyses to assess landscape connectivity from genetic data (Cushman *et al.* 2013; Guillot & Rousset 2013). Considering multiple independent replicates in a mixed modelling framework could help solve this difficulty as it may allow disentangling the potential impacts of collinearity between variables at the replicate level (Graham 2003; Cushman & Landguth 2010b). The advantages of an ISS may also benefit to alternative tools such as overlay methods based on genotypes rather than allelic frequencies (Manni *et al.* 2004; Crida & Manel 2007; Jombart *et al.* 2008). Although a higher sampling effort of genotypes per aggregate may provide more accurate genetic distances, one should note that increasing the number or the polymorphism of genetic markers may provide comparable benefits (Landguth *et al.* 2012). This alternative is now within our grasp, given the current advance of next-generation sequencing technologies (Segelbacher *et al.* 2010).

As already advocated (e.g. Cushman *et al.* 2006), ISS could be applied across broad landscapes in order to minimize a priori knowledge about the scale on which genetic pattern occurs as well as the relevant landscape features that impede gene flow. Both our simulations and our empirical test on a patchy distributed organism clearly support this point of view. However, our results also suggest that sampling more than one organism per spot should be of great benefit to the detection of spatial genetic patterns in most situations. The optimal sample size probably depends on both the size of aggregates and the actual dispersal rate between aggregates. According to our results, around ten per cent of the aggregate could be sufficient for this purpose. However, one should note that this sample size is far from the one required for the identification of critical demographic processes such as bottlenecks or inbreeding, or for the direct estimation of gene flow through the use of parentage analyses (Manel *et al.* 2005; Broquet & Petit 2009): PSSs are clearly still needed to precisely investigate the influence of connectivity on population functioning. Ideally, genetic field sampling schemes in patchy populations should be designed to benefit from the advantages of both strategies (Broquet & Petit 2009). One should first apply an ISS to extensive areas in order to provide a global preliminary picture of spatial processes likely to affect genetic structures and to investigate the relative influence of various landscape features using a multimodel inference approach. This first step could clear the way to implementing accurate population-based analyses using a PSS in the second phase. In this second step, sampling could be performed on targeted landscape features selected according to the results from previous individual-based analyses. Such an accurate PSS may then allow the use of assignment tests or coalescent methods, possibly providing insightful complementary information about landscape permeability to gene flow in clustered populations (Holderegger & Wagner 2008; Anderson *et al.* 2010).

Acknowledgements

This work was supported by the Agence Nationale de la Recherche et de la Technologie (ANRT, agreement no 2 / 2009), by two infrastructure managers (Réseau Ferré de France RFF and Autoroutes Paris-Rhin-Rhône APRR) and by a consulting company in ecological engineering (ECOSPHERE). The genetic sampling of alpine newts was conducted in accordance with French laws and with the approval of both the Préfecture de Saône-et-Loire and the Préfecture du Jura. We warmly thank E. Calonnier, A. Plaisance, G. Perez, J. Mangin and O. Grolet for their precious help and support in the field and laboratory. We gratefully acknowledge support from the CNRS/IN2P3 Computing Center (Lyon/Villeurbanne - France), for providing a significant amount of the computing resources needed for this work. We also thank E. Landguth for clues as to the use of CDPOP, K. Selkoe for her help with the use of mixed models, S. Manel and P. Giraudoux for insightful comments about results, and A. Corrigan and S. Béraud for proofreading. Finally, we thank Dr S. Spear and four anonymous reviewers for comments that improved this manuscript.

References

Anderson CD, Epperson BK, Fortin MJ *et al.* (2010) Considering spatial and temporal scale in landscape-genetic studies of gene flow. *Molecular Ecology*, **19**, 3565–3575.

Austin JD, Gorman TA, Bishop D (2011) Assessing fine-scale genetic structure and relatedness in the micro-endemic Florida bog frog. *Conservation Genetics*, **12**, 833–838.

Baguette M (2004) The classical metapopulation theory and the real, natural world: a critical appraisal. *Basic and Applied Ecology*, **5**, 213–224.

Beerli P (2004) Effect of unsampled populations on the estimation of population sizes and migration rates between sampled populations. *Molecular Ecology*, **13**, 827–836.

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (2004) *GENETIX 4.03, logiciel sous WindowsTM pour la génétique des populations*. Laboratoire Génome, Populations, Interactions CNRS UMR 5000, Université de Montpellier II, Montpellier, France.

Blair C, Weigel DE, Balazik M *et al.* (2012) A simulation-based evaluation of methods for inferring linear barriers to gene flow. *Molecular Ecology Resources*, **12**, 822–833.

Bonnet E, Van de Peer Y (2002) zt: a software tool for simple and partial Mantel tests. *Journal of Statistical software*, **7**, 1–12.

Borcard D, Legendre P (2012) Is the Mantel correlogram powerful enough to be useful in ecological analysis? A simulation study. *Ecology*, **93**, 1473–1481.

Bowcock AM, Ruizlinares A, Tomfohrde J *et al.* (1994) High-resolution of human evolutionary trees with polymorphic microsatellites. *Nature*, **368**, 455–457.

Broquet T, Petit EJ (2009) Molecular Estimation of Dispersal for Ecology and Population Genetics. *Annual Review of Ecology*

Evolution and Systematics, pp. 193–216. Annual Reviews, Palo Alto.

Broquet T, Ray N, Petit E, Fryxell JM, Burel F (2006) Genetic isolation by distance and landscape connectivity in the American marten (*Martes americana*). *Landscape Ecology*, **21**, 877–889.

Broquet T, Berset-Braendli L, Emaresi G, Fumagalli L (2007) Buccal swabs allow efficient and reliable microsatellite genotyping in amphibians. *Conservation Genetics*, **8**, 509–511.

Cavalli-Sforza LL, Edwards AWF (1967) Phylogenetic analysis: models and estimation procedures. *American Journal of Human Genetics*, **19**, 233–257.

Chen C, Durand E, Forbes F, Francois O (2007) Bayesian clustering algorithms ascertaining spatial population structure: a new computer program and a comparison study. *Molecular Ecology Notes*, **7**, 747–756.

Clarke RT, Rothery P, Raybould AF (2002) Confidence limits for regression relationships between distance matrices: estimating gene flow with distance. *Journal of Agricultural Biological and Environmental Statistics*, **7**, 361–372.

Crida A, Manel S (2007) WOMBSOFT: an R package that implements the Wombling method to identify genetic boundary. *Molecular Ecology Notes*, **7**, 588–591.

Cushman SA (2006) Effects of habitat loss and fragmentation on amphibians: a review and prospectus. *Biological Conservation*, **128**, 231–240.

Cushman SA, Landguth EL (2010a) Scale dependent inference in landscape genetics. *Landscape Ecology*, **25**, 967–979.

Cushman SA, Landguth EL (2010b) Spurious correlations and inference in landscape genetics. *Molecular Ecology*, **19**, 3592–3602.

Cushman SA, McKelvey KS, Hayden J, Schwartz MK (2006) Gene flow in complex landscapes: testing multiple hypotheses with causal modeling. *American Naturalist*, **168**, 486–499.

Cushman SA, Wasserman TN, Landguth EL, Shirk AJ (2013) Re-evaluating causal modeling with Mantel tests in landscape genetics. *Diversity*, **5**, 51–72.

Efron B, Tibshirani R (1993) *An Introduction to the Bootstrap*. Chapman & Hall, New York, London.

Emaresi G, Pellet J, Dubey S, Hirzel A, Fumagalli L (2011) Landscape genetics of the Alpine newt (*Mesotriton alpestris*) inferred from a strip-based approach. *Conservation Genetics*, **12**, 41–50.

Epperson BK (2003) *Geographical Genetics*. Princeton University Press, Princeton and Oxford.

Epps CW, Palsboll PJ, Wehausen JD *et al.* (2005) Highways block gene flow and cause a rapid decline in genetic diversity of desert bighorn sheep. *Ecology Letters*, **8**, 1029–1038.

Fedy BC, Martin K, Ritland C, Young J (2008) Genetic and ecological data provide incongruent interpretations of population structure and dispersal in naturally subdivided populations of white-tailed ptarmigan (*Lagopus leucura*). *Molecular Ecology*, **17**, 1905–1917.

Funk WC, Blouin MS, Corn PS *et al.* (2005) Population structure of Columbia spotted frogs (*Rana luteiventris*) is strongly affected by the landscape. *Molecular Ecology*, **14**, 483–496.

Garner TWJ, Schmidt BR, Hoeck P, Van Buskirk J (2003) Di- and tetranucleotide microsatellite markers for the Alpine newt (*Triturus alpestris*): characterization and cross-priming in five congeners. *Molecular Ecology Notes*, **3**, 186–188.

Goldberg CS, Waits LP (2010) Comparative landscape genetics of two pond-breeding amphibian species in a highly modified agricultural landscape. *Molecular Ecology*, **19**, 3650–3663.

Goudet J (2001) FSTAT, a Program to Estimate and Test Gene Diversities and Fixation Indices (Version 2.9.3). University of Lausanne, Dorigny.

Graham MH (2003) Confronting multicollinearity in ecological multiple regression. *Ecology*, **84**, 2809–2815.

Graves T, Beier P, Royle JA (2013) Current approaches using genetic distances produce poor estimates of landscape resistance to interindividual dispersal. *Molecular Ecology*, **22**, 3888–3903.

Guillot G, Rousset F (2013) Dismantling the Mantel tests. *Methods in Ecology and Evolution*, **4**, 336–344.

Hanski I (1999) *Metapopulation Ecology*. Oxford University Press, Oxford.

Harrison S (1991) Local extinction in a metapopulation context : an empirical evaluation. *Biological Journal of the Linnean Society*, **42**, 73–88.

Holderegger R, Wagner HH (2008) Landscape genetics. *BioScience*, **58**, 199–207.

Jaquiéry J, Broquet T, Hirzel AH, Yearsley J, Perrin N (2011) Inferring landscape effects on dispersal from genetic distances: how far can we go? *Molecular Ecology*, **20**, 692–705.

Jombart T, Devillard S, Dufour AB, Pontier D (2008) Revealing cryptic spatial patterns in genetic variability by a new multivariate method. *Heredity*, **101**, 92–103.

Kalinowski ST (2005) Do polymorphic loci require large sample sizes to estimate genetic distances? *Heredity*, **94**, 33–36.

Kelly RP, Oliver TA, Sivasundar A, Palumbi SR (2010) A method for detecting population genetic structure in diverse, high gene-flow species. *Journal of Heredity*, **101**, 423–436.

Kimura M, Weiss GH (1964) The stepping stone model of population structure and the decrease of genetic correlation with distance. *Genetics* **49**, 561–576.

Kovar R, Brabec M, Vita R, Bocek R (2009) Spring migration distances of some Central European amphibian species. *Amphibia-Reptilia*, **30**, 367–378.

Landguth EL, Cushman SA (2010) CDPOP: a spatially explicit cost distance population genetics program. *Molecular Ecology Resources*, **10**, 156–161.

Landguth EL, Cushman SA, Schwartz MK *et al.* (2010) Quantifying the lag time to detect barriers in landscape genetics. *Molecular Ecology*, **19**, 4179–4191.

Landguth EL, Fedy BC, Oyler-McCance SJ *et al.* (2012) Effects of sample size, number of markers, and allelic richness on the detection of spatial genetic pattern. *Molecular Ecology Resources*, **12**, 276–284.

Legendre P (2000) Comparison of permutation methods for the partial correlation and partial Mantel tests. *Journal of Statistical Computation and Simulation*, **67**, 37–73.

Legendre P, Fortin MJ (2010) Comparison of the Mantel test and alternative approaches for detecting complex multivariate relationships in the spatial analysis of genetic data. *Molecular Ecology Resources*, **10**, 831–844.

Legendre L, Legendre P (1998) *Numerical Ecology*, 2nd edn. Elsevier, Amsterdam.

Lowe WH, Allendorf FW (2010) What can genetics tell us about population connectivity? *Molecular Ecology*, **19**, 3038–3051.

Manel S, Schwartz MK, Luikart G, Taberlet P (2003) Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology & Evolution*, **18**, 189–197.

Manel S, Gaggiotti OE, Waples RS (2005) Assignment methods: matching biological questions techniques with appropriate techniques. *Trends in Ecology & Evolution*, **20**, 136–142.

Manel S, Joost S, Epperson BK *et al.* (2010) Perspectives on the use of landscape genetics to detect genetic adaptive variation in the field. *Molecular Ecology*, **19**, 3760–3772.

Manly BFJ (2007) *Randomization, Bootstrap and Monte Carlo Methods in Biology*, 3rd edn. Chapman & Hall/CRC, Boca Raton.

Manni F, Guerard E, Heyer E (2004) Geographic patterns of (genetic, morphologic, linguistic) variation: how barriers can be detected by using Monmonier's algorithm. *Human Biology*, **76**, 173–190.

Mantel N (1967) Detection of disease clustering and a generalized regression approach. *Cancer Research* **27**, 209–220.

Mayer C, Schiegg K, Pasinelli G (2009) Patchy population structure in a short-distance migrant: evidence from genetic and demographic data. *Molecular Ecology*, **18**, 2353–2364.

Nei M, Tajima F, Tateno Y (1983) Accuracy of estimated phylogenetic trees from molecular data. *Journal of Molecular Evolution*, **19**, 153–170.

Olivieri I, Michalakis Y, Gouyon PH (1995) Metapopulation genetics and the evolution of dispersal. *American Naturalist*, **146**, 202–228.

Perret N, Pradel R, Miaud C, Grolet O, Joly P (2003) Transience, dispersal and survival rates in newt patchy populations. *Journal of Animal Ecology*, **72**, 567–575.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

Prunier J, Kaufmann B, Grolet O *et al.* (2012) Skin swabbing as a new efficient DNA sampling technique in amphibians, and 14 new microsatellite markers in the alpine newt (*Ichthyosaura alpestris*). *Molecular Ecology Resources*, **12**, 524–531.

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

Rousset F (2000) Genetic differentiation between individuals. *Journal of Evolutionary Biology*, **13**, 58–62.

Ruzzante DE (1998) A comparison of several measures of genetic distance and population structure with microsatellite data: bias and sampling variance. *Canadian Journal of Fisheries and Aquatic Sciences*, **55**, 1–14.

Safner T, Miaud C, Gaggiotti O *et al.* (2011a) Combining demography and genetic analysis to assess the population structure of an amphibian in a human-dominated landscape. *Conservation Genetics*, **12**, 161–173.

Safner T, Miller MP, McRae BH, Fortin M-J, Manel S (2011b) Comparison of Bayesian clustering and edge detection methods for inferring boundaries in landscape genetics. *International Journal of Molecular Sciences*, **12**, 865–889.

Schwartz MK, McKelvey KS (2009) Why sampling scheme matters: the effect of sampling scheme on landscape genetic results. *Conservation Genetics*, **10**, 441–452.

Segelbacher G, Cushman SA, Epperson BK *et al.* (2010) Applications of landscape genetics in conservation biology: concepts and challenges. *Conservation Genetics*, **11**, 375–385.

Selkoe KA, Watson JR, White C *et al.* (2010) Taking the chaos out of genetic patchiness: seascape genetics reveals ecological and oceanographic drivers of genetic patterns in three temperate reef species. *Molecular Ecology*, **19**, 3708–3726.

Smith MA, Green DM (2005) Dispersal and the metapopulation paradigm in amphibian ecology and conservation: are all amphibian populations metapopulations? *Ecography*, **28**, 110–128.

Smouse PE, Peakall R (1999) Spatial autocorrelation analysis of individual multiallele and multilocus genetic structure. *Heredity*, **82**, 561–573.

Smouse PE, Long JC, Sokal RR (1986) Multiple-regression and correlation extensions of the mantel test of matrix correspondence. *Systematic Zoology*, **35**, 627–632.

Van Strien MJ, Keller D, Holderegger R (2012) A new analytical approach to landscape genetic modelling: least-cost transect analysis and linear mixed models. *Molecular Ecology*, **21**, 4010–4023.

Weir BS, Cockerham CC (1984) Estimating F-Statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.

Wright S (1931) Evolution in Mendelian populations. *Genetics*, **16**, 0097–0159.

Wright S (1943) Isolation by distance. *Genetics*, **28**, 114–138.

Wright S (1951) The genetical structure of populations. *Annals of Eugenics*, **15**, 323–354.

B.K., J.P. and J.P.L. designed the study; D.P., J.P. and J.P.L. collected empirical data; J.P. and S.F. performed modelling and simulating work; B.K., F.P., J.P., J.P.L. and P.J. analysed output data; J.P. and J.P.L. wrote the first draft of the manuscript and all authors contributed substantially to revisions; F.P., J.P.L. and P.J. ensured the work supervision.

Data accessibility

Simulated and empirical genetic data with spatial locations of individuals and Matlab script for Mantel tests with restricted permutations: Dryad Digital Repository. doi:10.5061/dryad.490p9

Supporting information

Additional supporting information may be found in the online version of this article.

Appendix S1 Evolution of genetic structure over generations in IBD simulated data sets.

Appendix S2 Choice of interindividual genetic distances.

Appendix S3 Comparison of mixed models and Mantel tests with restricted permutations.

Appendix S4 Influence of an unbalanced genetic sampling scheme on Mantel tests using restricted permutations.

Appendix S5 Matlab script for simple and partial Mantel tests with restricted permutations.