

Isolation by distance, resistance and/or clusters? Lessons learned from a forest-dwelling carnivore inhabiting a heterogeneous landscape

ARITZ RUIZ-GONZALEZ,*†‡ SAMUEL A. CUSHMAN,§ MARÍA JOSÉ MADEIRA,*†
ETTORE RANDI‡¶ and BENJAMÍN J. GÓMEZ-MOLINER*†

*Department of Zoology and Animal Cell Biology, University of the Basque Country, UPV/EHU, C/Paseo de la Universidad 7, 01006 Vitoria-Gasteiz, Spain, †Systematics, Biogeography and Population Dynamics Research Group, Lascaray Research Center, University of the Basque Country, UPV/EHU, Avda. Miguel de Unamuno, 3, 01006 Vitoria-Gasteiz, Spain, ‡Conservation Genetics Laboratory, National Institute for Environmental Protection and Research, ISPRA, Via Cà Fornacetta 9, 40064 Ozzano dell'Emilia, Bologna, Italy, §U.S. Forest Service, Rocky Mountain Research Station, 2500 S. Pine Knoll Dr., 86001 Flagstaff, AZ, USA, ¶Department 18/Section of Environmental Engineering, Aalborg University, Sohngårdsholmsvej 57, 9000 Aalborg, Denmark

Abstract

Landscape genetics provides a valuable framework to understand how landscape features influence gene flow and to disentangle the factors that lead to discrete and/or clinal population structure. Here, we attempt to differentiate between these processes in a forest-dwelling small carnivore [European pine marten (*Martes martes*)]. Specifically, we used complementary analytical approaches to quantify the spatially explicit genetic structure and diversity and analyse patterns of gene flow for 140 individuals genotyped at 15 microsatellite loci. We first used spatially explicit and nonspatial Bayesian clustering algorithms to partition the sample into discrete clusters and evaluate hypotheses of 'isolation by barriers' (IBB). We further characterized the relationships between genetic distance and geographical ('isolation by distance', IBD) and ecological distances ('isolation by resistance', IBR) obtained from optimized landscape models. Using a reciprocal causal modelling approach, we compared the IBD, IBR and IBB hypotheses with each other to unravel factors driving population genetic structure. Additionally, we further assessed spatially explicit indices of genetic diversity using sGD across potentially overlapping genetic neighbourhoods that matched the inferred population structure. Our results revealed a complex spatial genetic cline that appears to be driven jointly by IBD and partial barriers to gene flow (IBB) associated with poor habitat and interspecific competition. Habitat loss and fragmentation, in synergy with past overharvesting and possible interspecific competition with sympatric stone marten (*Martes foina*), are likely the main factors responsible for the spatial genetic structure we observed. These results emphasize the need for a more thorough evaluation of discrete and clinal hypotheses governing gene flow in landscape genetic studies, and the potential influence of different limiting factors affecting genetic structure at different spatial scales.

Keywords: Bayesian analysis, isolation by cluster, isolation by distance, isolation by landscape resistance, landscape genetics, *Martes martes*, reciprocal causal modelling

Received 18 March 2015; revision received 15 September 2015; accepted 16 September 2015

Introduction

The genetic structure of natural populations is influenced by the spatial heterogeneity of the landscapes in which they exist, which affects dispersal and, consequently, gene flow (Luque *et al.* 2012; Wagner & Fortin 2013). An important issue in landscape genetics is whether changes in landscape features create barriers to gene flow and lead to discrete genetic structure, or whether clinal population structure dominates as a result of isolation by distance and isolation by resistance processes (Meirmans 2012; Landguth & Schwartz 2014). In natural populations, substructuring of individuals can be caused by isolation by distance (IBD), landscape resistance (IBR) and/or barriers to dispersal (IBB) (Cushman *et al.* 2006; Cushman & Landguth 2010; van Strien *et al.* 2015). Multiple factors acting at various spatial scales affect the spatial distribution of genetic variation and rates of gene flow across a landscape (Anderson *et al.* 2010). These factors can include biotic (e.g. forest structure, presence of prey, predators or sympatric competitors), abiotic (e.g. rivers, topography, climatic conditions) and/or anthropogenic features (e.g. roads, urban areas, agriculture). Thus, emergent patterns of gene flow are the result of the interactions between structural landscape connectivity and how organisms respond to landscape structure (Luque *et al.* 2012). In this context, assessing spatial genetic structure is crucial for defining appropriate conservation and management units, and for maintaining gene flow and genetic diversity (Segelbacher *et al.* 2010; Luque *et al.* 2012).

The multidisciplinary nature and rapid growth of landscape genetics (Manel & Holderegger 2013) have led to multiple approaches to examine the relationships between individual-based genetic measures and spatially explicit landscape variables (Sork & Waits 2010; Wagner & Fortin 2013; Hall & Beissinger 2014). Most landscape genetic approaches largely focus on describing and mapping populations using Bayesian clustering algorithms (Guillot *et al.* 2009) and/or on identifying factors that influence rates and patterns of gene flow testing isolation by landscape resistance (IBR; McRae 2006; Cushman *et al.* 2006), and deducing from these which features restrict or promote movements of individuals.

A discrete population structure can easily be mistaken for a pattern of IBD (Guillot *et al.* 2009; Meirmans 2012; Landguth & Schwartz 2014) or IBR. However, the reverse is also true, as IBD-IBR processes create clines in allele frequencies that can lead to spurious identification of discrete genetic clustering (Blair *et al.* 2012; Meirmans 2012; Cushman *et al.* 2014; Landguth & Schwartz 2014). Indeed, recent studies have shown that the

presence of IBD can confound the identification of barriers using clustering or IBB methods, and may result in overestimation of population subdivision (Frantz *et al.* 2009; Tucker *et al.* 2014) and/or identification of artificial clusters (Blair *et al.* 2012; Cushman *et al.* 2014). In this context, Guillot *et al.* (2009) suggested comparing the outputs of spatially explicit and nonspatial clustering algorithms in addition to testing for the presence of IBD to infer genetic structuring that could be empirically explained. Furthermore, Cushman & Landguth (2010) suggested using analytical frameworks that allow simultaneous competition and separation of IBB, IBD and IBR hypotheses. However, very few landscape genetics studies have rigorously evaluated and jointly tested the effect of clinal (IBD and IBR) and discrete structuring (but see Cushman & Landguth 2010; Cushman *et al.* 2014; Yang *et al.* 2015), to avoid misinterpretation of spatial genetic diversity and differentiation and properly assess the validity of spurious clusters emerging as artefacts of clinal spatial genetic processes (e.g. Frantz *et al.* 2009; Cushman *et al.* 2014).

Another important issue in landscape genetics is delimiting 'populations' to calculate genetic diversity indices (Shirk & Cushman 2011, 2014). Although many populations are structured by IBD or IBR, the most commonly used approach to calculate genetic diversity indices is based on 'populations' derived from clustering approaches and/or arbitrary definitions, assuming an island model of population structure (Wright 1943). This clearly presents a mismatch between clinal population structure and discrete measures of genetic diversity, and spatially explicit tools are now available to overcome this limitation by estimating genetic diversity continuously across potentially overlapping genetic neighbourhoods that match the population structure, whether discrete or clinal (Shirk & Cushman 2011, 2014).

Forest-dependant species are useful for evaluating the genetic consequences of habitat fragmentation on spatial genetic structure due to their sensitivity to human-induced landscape transformation (e.g. Broquet *et al.* 2006; Segelbacher *et al.* 2008). For all *Martes* species, forested habitats are key features (Proulx *et al.* 2004; Schwartz *et al.* 2012). In Europe, the genus *Martes* is represented by the European pine marten (*M. martes*) and stone marten (*M. foina*) (Proulx *et al.* 2004). The European pine marten is a woodland-dwelling small carnivore that occurs from northern Iberian Peninsula to western Siberia (Proulx *et al.* 2004). Their occurrence patterns are affected by the percentage of woodland cover, forest patch size, food resources, sex, age class and habitat fragmentation levels and are generally considered a closed canopy forest-associated species (Zalewski & Jędrzejewski 2006; Mergey *et al.* 2011).

Given their strong associations with high forest structural complexity, the species is particularly sensitive to human influences on their habitats, including habitat loss and landscape-scale effects of habitat fragmentation (Pereboom *et al.* 2008; Mergey *et al.* 2011; Ruiz-González *et al.* 2014). Nonetheless, they have also been recently reported in fragmented landscapes characterized by small forest fragments within agricultural landscape matrices (Pereboom *et al.* 2008; Balestrieri *et al.* 2010, 2014; Mergey *et al.* 2011), suggesting they may not be as forest dependent as previously believed. The species has experienced a large range contraction in the last century due to overharvesting (Helldin 2000; Proulx *et al.* 2004), deforestation and forest fragmentation (Kurki *et al.* 1998; Brainerd & Rolstad 2002; Pereboom *et al.* 2008) and is either threatened or rare in many countries (Proulx *et al.* 2004). This population decline is likely to have affected population structure at different spatial scales (Schwartz *et al.* 2012), but also may have amplified competitive interactions with the widespread and sympatric stone marten, which expresses more behavioural plasticity, a broader habitat niche and less vulnerability to anthropogenic changes and human presence (Virgós *et al.* 2012). Competitive interactions between these sympatric carnivores, which have contrasting ecological preferences (stenotopic vs. synantropic) and inverse demographic trends (Proulx *et al.* 2004; Virgós *et al.* 2012), may affect population genetic structure through competition for the same resources and synergistically the amplification of the effects of other biotic or abiotic factors.

Previous studies of the pine marten have provided information about the genetic variability and structure of several European populations (reviewed in Schwartz *et al.* 2012). Nevertheless, information about the genetic variability and population structure of the Iberian populations at the southern edge of the species distribution is negligible, with previous findings suggesting that gene flow is facilitated by natural vegetation, and is resisted by anthropogenic land cover types and roads (Ruiz-González *et al.* 2014). Thus, further studies are needed to investigate the potential interactions of habitat fragmentation, interspecific competition and past population bottlenecks in affecting the genetic diversity and differentiation of the pine marten in the Iberian Peninsula but also a reliable detection of clinal vs. discrete genetic structure.

We had two main objectives in our study. Our first objective was to characterize the spatial genetic structure and diversity of the forest-dependant European pine marten and identify potential barriers and/or moderators of gene flow. Our second objective is to disentangle discrete [isolation by barriers, (IBB)] and/or clinal population structure [isolation by distance (IBD);

gradients of landscape resistance (IBR)] hypotheses by competing different models within a reciprocal causal modelling framework (Cushman *et al.* 2013c), using the pine marten as a case study.

As habitat specialist species are usually considered highly vulnerable to habitat fragmentation (Devictor *et al.* 2008), we predicted that, as a forest-dwelling species inhabiting a highly fragmented and heterogeneous landscape, pine martens would present a strong spatial genetic structure pattern, which could be influenced by different limiting factors at different spatial scales (Cushman *et al.* 2013b). Moreover, given that past overharvesting and potential interspecific competition with the sympatric stone marten have likely reduced species range and connectivity between populations in the northern Iberian peninsula, we expected that both factors could be also synergistically affecting pine marten gene flow.

Consequently, we hypothesise that pine marten genetic diversity in a heterogeneous landscape will follow a complex spatial genetic cline that could be related to a combined effect of isolation by distance (IBD) and landscape resistance (IBR) but also due to the presence of putative barriers (IBB) to gene flow, related to past demographic events and current biotic and abiotic interactions.

Materials and methods

Study area

The Basque Country and Navarre are located in the northern Iberian Peninsula, bordering the Cantabrian Sea to the north and situated between the Pyrenees and the Cantabrian mountains (Fig. 1). The study area comprises parts of the Atlantic and Mediterranean biogeographical regions, with the exception of the northeastern area of Navarre, which belongs to the Alpine region. The Basque Country comprises an area of 7235 km² and has a high population density of 298 inhabitants per square kilometre (IGN 2008) and a dense road network. The Basque Country is characterized by a mosaic of remnant patches of natural and semi-natural vegetation composed of deciduous oak and beech forests in a mosaic of agricultural and urban land uses. Natural forest covers approximately 28% of its area, forestry plantations 29%, nonwooded mountains 24%, cultivated land 14%, and urban land and infrastructures 5.7%. In contrast, Navarre comprises an area of 10 390 km² and has a low population density of 60 inhabitants per square kilometre (IGN 2008). The area is dominated by an unfragmented natural forest system, concentrated in the north, while cultivated land and urban areas are located primarily in the south.

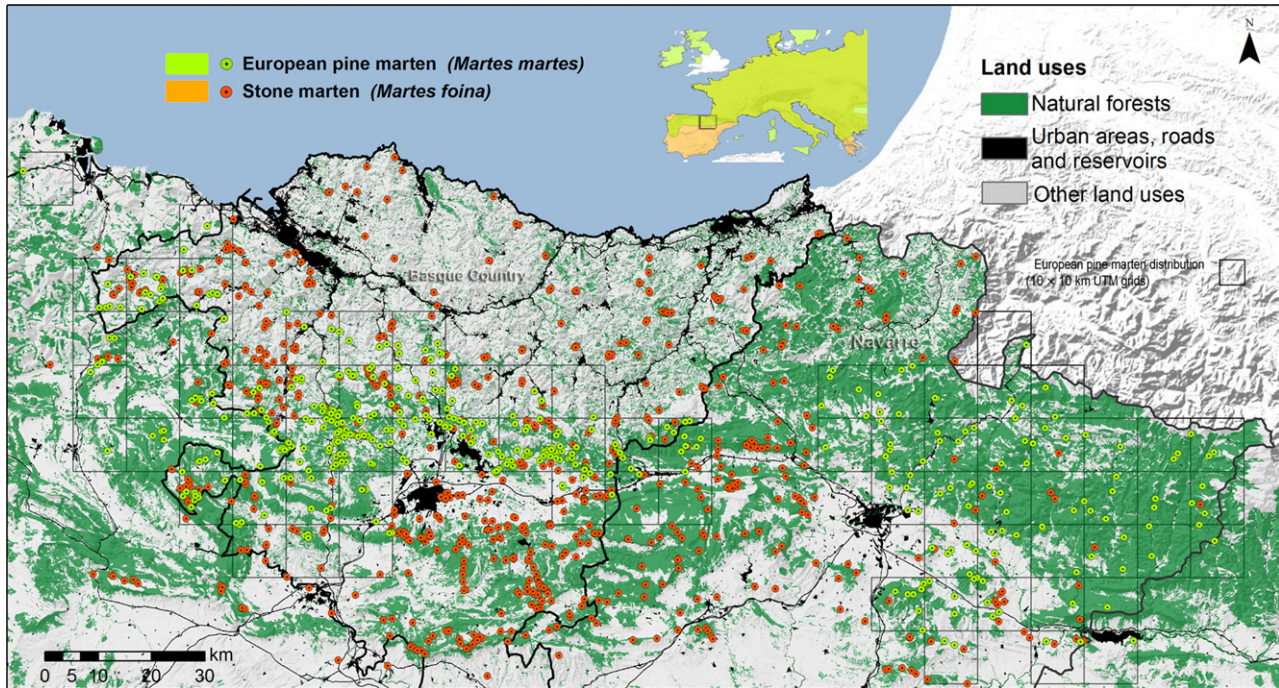


Fig. 1 Study area, located in the regions of the Basque Country and Navarre, northern Spain. Green and orange circles represent the geographical locations of samples genetically identified as *Martes martes* and *Martes foina*, respectively (faecal samples: $n = 670$; tissue samples: $n = 183$). Natural forest appears shaded in green, urban areas, roads and man-made reservoirs in black and other land uses in grey. The European distributions of *M. martes* and *M. foina* are shown in green and orange, respectively.

Natural forest covers nearly 36% of its area, forestry plantations 6%, nonwooded mountains 10%, cultivated land 46%, and urban land and infrastructures 1.4%. The study area plays an important natural link between the Cantabrian Mountains and the Pyrenees and is considered of strategic importance for the conservation of ecological connectivity in southwestern Europe (Mallarach *et al.* 2010). Moreover, the southern edge limit of a number of temperate forest species (e.g. *Glis glis*, *Myodes glareolus*), including the focal pine marten, is located in the study area (Mitchell-Jones *et al.* 1999). Thus, this region has high importance for biodiversity conservation, as distributional limits are thought to be particularly important as long-term stores of genetic diversity and hot spots for speciation (Hampe & Petit 2005).

Noninvasive genetic sampling: mtDNA species identification and individual identification from microsatellite genotyping

We used noninvasive scat sampling to collect genetic samples from the sympatric *Martes* sp. (*Martes martes* and *Martes foina*) in the study area between 2004 and 2011. Additionally, fresh tissue specimens from road-killed pine martens were also included. We first used a mtDNA molecular method for an effective genetic

identification of sympatric marten species following the method described in the study by Ruiz-González *et al.* (2008), to assess the spatial distribution of both species in the study area. Second, all the faecal samples genetically identified as pine marten were genotyped using a multiplex panel of 15 variable microsatellite loci following methods in Ruiz-González *et al.* (2013). Full methodological details about noninvasive genetic sampling are described in Appendix S1 (Supporting information).

Bayesian clustering analysis

We used spatial and nonspatial Bayesian clustering models to estimate the number of populations (K) and individual population memberships using STRUCTURE v.2.2 (Pritchard *et al.* 2000; Falush *et al.* 2003, 2007) and GENELAND v.3.3.2 (Guillot *et al.* 2005a,b; Guillot & Santos 2009). Blair *et al.* (2012), in an assessment of multiple methods to detect linear barriers to gene flow, found that these two clustering methods performed very well, with GENELAND the most powerful method tested.

GENELAND v 3.3.2 (Guillot *et al.* 2005a,b; Guillot & Santos 2009) was run through an extension of program R v.2.11.1 (R Development Core Team 2011) under the correlated allele frequency model with spatial uncertainty in spatial locations fixed at 1 km and using the filtering for the presence of null alleles. We allowed K

to vary from 1 to 10 for 50 independent runs each consisting of 500 000 MCMC iterations, a thinning of 100, maximum rate of the Poisson process fixed to 150, maximum number of nuclei in the Poisson–Voronoi tessellation fixed to 450, a maximum rate of the Poisson process fixed to 150 and burn-in of 100 000 in the post-processing. MCMC convergence was assessed by comparing the number of populations across replicate runs, with the mean posterior density used as a criterion to choose the best run (Guillot *et al.* 2005a,b; Guillot & Santos 2009).

The population structure was also estimated using the program *STRUCTURE* v. 2.2 (Pritchard *et al.* 2000; Falush *et al.* 2003, 2007) assuming population admixture and correlated allele frequencies within populations. Simulations were run using a burn-in period of 10^5 sweeps followed by 10^6 MCMC interactions. K was allowed to vary from 1 to 10 and 20 independent simulations were run for each K to check for consistency in the results. We determined the most likely number of clusters by examining both the log probability of the data [$\ln \Pr(X|K)$] and the ΔK statistic following Evanno *et al.* (2005). For the identified K value, we used *CLUMPP* v. 1.1.2 (Jakobsson & Rosenberg 2007) to determine the population assignment probability of each individual across all simulations. Once the true K was selected, the fractional membership of each individual in each cluster (q), averaging q over the 20 runs, was plotted on a map of the study region to assess geographical congruence of the clusters and contrast the results obtained from the spatially explicit *GENELAND* results.

Population genetic analysis of the inferred clusters: genetic diversity, migration and bottleneck detection

We summarized genetic variation through the number of alleles per locus (A) and expected (H_e) and observed (H_o) heterozygosities using *GENETIX* v 4.05.2 (Belkhir *et al.* 2004) for each of the inferred clusters (based on spatially explicit *GENELAND* results) and for the full data set. Accounting for differences in sample size, the program *HP-RARE* (Kalinowski 2005) was used to calculate allelic richness, with a minimum sample size of $n = 20$ gene copies. Estimates of pairwise linkage disequilibrium for each pair of loci in each population and deviation from Hardy–Weinberg equilibrium (HWE) were tested using the exact test implemented in *GENEPOP* version 4.0 (Rousset 2008). Statistical significance was evaluated by running a Markov chain Monte Carlo (MCMC) consisting of 10 000 batches of 10 000 iterations each, with the first 10 000 iterations discarded before sampling (Guo & Thompson 1992).

We calculated pairwise F_{ST} values (Weir & Cockerham 1984) and tested for pairwise genetic differences

among clusters identified by Bayesian models using a randomization procedure implemented in *FSTAT* v 2.9.3.2 (Goudet 1995; Goudet 2001). Moreover, analysis of molecular variance (*AMOVA*) was also conducted to test the significance of the inferred population structure with the software *ARLEQUIN* v 3.1 (Excoffier *et al.* 2005) at three levels: among populations (considering Navarre and the Basque Country), among individuals within populations and within individuals. *MICRO-CHECKER* software (Van Oosterhout *et al.* 2004) was used to check for potential scoring errors and the presence of null alleles.

In addition, past bottleneck events and estimates of effective migrants (N_m) among clusters within the last few generations were assessed using *BOTTLENECK* version 1.2.02 (Cornuet & Luikart 1996; Piry *et al.* 1999) and a Bayesian model implemented in *MIGRATE* software version 3.0.3 (Beerli & Felsenstein 2001; Beerli 2006), respectively. *BOTTLENECK* analyses were performed under the three microsatellite mutational models available: infinite allele model (IAM), stepwise (SMM) and two-phase model of mutation (TPM) (with 95% stepwise mutation). Recent migration analyses were conducted following the recommendations of Beerli (2004). We first did an initial run on our data set using *FST* to find the start parameters, and we used the output of the initial run as the start parameters of our second run. Because there were only minor differences between the outputs from the first and the second runs, we presented only the output from the second run.

Spatial autocorrelation analysis

We assessed fine-scale spatial genetic structure (SGS) using spatial autocorrelation analyses. Spatial autocorrelation analyses (i.e. the analyses of genetic relatedness between pairs of individuals as a function of geographical distance) were conducted using *SPAGED1* 1.2 (Hardy & Vekemans 2002). These were performed on the 15 polymorphic loci with kinship coefficients (F_{ij}) (Loiselle *et al.* 1995). As suggested by Vekemans & Hardy (2004), F_{ij} was chosen as a pairwise estimator of genetic relatedness, as it is a relatively unbiased estimator with low sampling variance. The slope (b) of this linear regression does not depend on an arbitrary choice of distance classes and, therefore, provides a good estimator of the degree of SGS at this scale (Hardy & Vekemans 2002). Because there is no consensus regarding how to generate distance classes, we used the recommendations of Hardy & Vekemans (2002); i.e. in all cases, more than 50% of all individuals were represented at least once in each distance interval, and the coefficient of variation of the number of times each individual was represented was less than one). Thus, to illustrate the pattern of spatial autocorrelation, the number of spatial distance

classes was set to 22, leading to a minimum of 148 pairwise comparisons per distance class. A jackknife procedure over loci was used to estimate standard errors for each distance class, and 10 000 randomizations of individual spatial locations were performed to test for the overall spatial structure (Hardy & Vekemans 2002).

IBD, IBR and IBB: competing models within a reciprocal causal modelling framework

Construction of IBD, IBR and IBB models. We proposed different alternative models governing gene flow: (i) a simple model of isolation by distance (IBD), (ii) isolation by barriers (IBB) obtained from both clustering approaches (i.e. GENELAND and STRUCTURE) and (iii) three different landscape resistance models previously optimized for the study species at the Basque Country level (Ruiz-González *et al.* 2014). Therefore, we proposed six alternative models [i.e. IBD; IBB (GENELAND and STRUCTURE) and IBR (LandFb_100, LandFb_50 and EN; see Table 1 for further details) that were later competed within a reciprocal causal modelling framework (Cushman *et al.* 2013c). Full methodological details for model construction and evaluation are reported in Appendix S2 (Supporting information).

Reciprocal causal modelling: (partial) Mantel correlations between IBD, IBR and clustering (IBB) hypotheses. We used reciprocal causal modelling (Cushman *et al.* 2013c) to compete the alternative hypotheses governing gene flow (i.e. IBD, IBB and IBR) and identify whether any hypothesis is supported independently of all alternative hypotheses. Cushman & Landguth (2010) found that the inherent high correlation among alternative models results in a high risk of spurious correlations using simple Mantel tests. Several refinements, including causal modelling (Cushman *et al.* 2006), have been developed to reduce the risk of affirming spurious correlations and to assist model selection. However, Cushman *et al.* (2013c) showed these still suffer from elevated type I error rates. Consequently, Cushman *et al.* (2013c) proposed 'reciprocal causal modelling', which they showed greatly lessens type I error rates observed by Meirmans (2012) and Amos *et al.* (2012) in landscape genetic analysis, allowing rigorously testing for the joint and independent effects of alternative resistance models (e.g. Yang *et al.* 2013; Castillo *et al.* 2014; Ruiz-González *et al.* 2014).

The method of reciprocal causal modelling directly competes all alternative hypotheses based on relative support. The approach uses a pair of 'reciprocal' Mantel tests for each pair of alternative resistance hypotheses (Fig. S1, Supporting information). The first is a partial Mantel test which calculates the partial Mantel correlation between the first hypothesis (A) and genetic

Table 1 Resistance parameters for the reciprocal causal modelling of different hypothesis governing gene flow (i.e. IBD, IBR and IBB)

Land uses	IBR			
	IBD	Land_Fb ₅₀	Land_Fb ₁₀₀	EN
Forests	1	1	1	1
Forestry plantations	1	1	1	10
Scrubland	1	1	1	5
Agroforestry mosaics	1	1	1	20
Pastures and meadows	1	1	1	30
Rocks	1	50	100	40
Crops	1	50	100	60
Wetlands	1	100	100	100
National roads	1	200	200	200
Highways, urban areas, reservoirs and quarries	1	1000	1000	1000

We tested (i) IBD-isolation by distance; (ii) IBR-isolation by landscape resistance models optimized by Ruiz-González *et al.* (2014); Land_Fb100 and Land_Fb50: binary landscape resistance maps indicating that pine marten gene flow in is facilitated by forests, forestry plantations, scrubland, agroforestry mosaics and pastures and meadows and that crops have approximately 100 (LandFb_100) or 50 (LandFb_50) times higher resistance than optimal habitat and the existence of a barrier effect of national roads, highways, urban areas, reservoirs and quarries. Ecological network resistance map, EN: a resistance map analogous to that used in the design of the ecological network of the Basque Country (Gurrutxaga *et al.* 2010a); and 3) IBB-isolation by barriers or clusters (IBB_Geneland and IBB_Structure).

The IBB models are not shown, as that hypothesis does not assign resistance based on landscape conditions, but based on STRUCTURE and on GENELAND group membership.

distance (G), partialling out the second (B) ($G \sim A|B$). The second is a partial Mantel test which calculates the partial Mantel correlation between the second hypothesis and genetic distance, partialling out the first ($G \sim B|A$) (Fig. S1, Supporting information). The difference between the partial correlations of these two tests ($A|B-B|A$) is a measure of the relative support for hypothesis A relative to hypothesis B (e.g. Cushman *et al.* 2013b,c; Yang *et al.* 2013; Castillo *et al.* 2014) (Fig. S1, Supporting information). Specifically, if hypothesis A is correct, then the difference $A|B-B|A$ should be positive. Conversely, if hypothesis B is correct, then ($A|B-B|A$) should be zero or negative. The reciprocal causal modelling approach works by calculating a full matrix of all the $A|B-B|A$ differences between each pair of alternative hypothesis. For a model that is fully supported, all values in the column associated with that hypothesis would be positive, while all values in the row associated with that hypothesis would be negative. This would indicate that that hypothesis is supported independently of all alternative

hypotheses (positive values down the column) and that no alternative hypotheses are supported independently of it (negative values across the row).

In addition to the reciprocal causal modelling matrix (Cushman *et al.* 2013c), we calculated the raw form of the partial Mantel matrix that provide further insights when competing hypothesis could be jointly true (Castillo *et al.* 2014) [see Appendix S3 (Supporting information), with full methodological details of (partial) Mantel correlations and competing hypothesis within a reciprocal causal modelling framework (Cushman *et al.* 2013c)].

Spatially explicit indices of genetic diversity based on IBD

Software package sGD (Shirk & Cushman 2011) was used to estimate genetic diversity indices based on microsatellite genotypes and the pairwise distance matrix generated from IBD, as this model was identified as the dominant driver of gene flow of the pine marten at this geographical scale (see Results). This approach allows for the grouping of individuals based on spatially explicit genetic neighbourhoods that take into account the effects of a heterogeneous landscape on the genetic diversity of clinal populations. We applied a genetic neighbourhood diameter defined by the largest geographical distance that is positively and significantly correlated with genetic distance ($\alpha = 0.05$; Shirk & Cushman 2011). We identified this point using a Mantel correlogram characterizing the correlation of pairwise genetic distances among individuals (based on Bray–Curtis dissimilarity) across multiple ranges of pairwise Euclidean distances. The largest distance class retaining a positive Mantel correlation was chosen as the genetic neighbourhood diameter (D). The autocorrelation of pairwise genetic distances in the IBD model reached zero at a distance of 69 km. Thus, we set D for all the analysis as 69 km (Shirk & Cushman 2011) and the minimum population size to 20 individuals. We calculated five genetic diversity indices based on the codominant marker genotypes of all individuals in the neighbourhood using sGD: observed heterozygosity (H_o), Nei's gene diversity (H_s), Wright's inbreeding coefficient (F_{is}), allelic richness (A_r) and mean number of alleles.

Results

Noninvasive sample collection and species identification

Of 977 faecal samples collected from the entire study area, 179 were discarded because they were not fresh or

because they presumably belong to the same individual (samples separated by <1 km). About 670 of 798 analysed samples were classified as *Martes* sp. (*M. martes* and *M. foina*) based on genetic species identification results. Thus, unequivocal species identification was possible in 84% of the samples. We effectively identified 323 faecal samples as stone marten and 347 as pine marten.

Additionally, we obtained 74 tissue samples of pine marten and 109 of stone marten from road-kill animals. Thus, we obtained a total number of 421 samples of pine marten and 432 of stone marten. The geographical locations for the 670 correctly identified faecal samples and 183 tissue samples of both species in the study area are shown in Fig. 1. In this heterogeneous landscape, the stone marten is widely distributed across the whole study area, while the pine marten is restricted to the main forested areas.

Individual identification and genotype checking

Of 347 faecal samples identified as pine marten, 120 were not included in the microsatellite genotyping procedure. These samples correspond to the sampling period from 2004 to 2005, which were used for a first distribution assessment of sympatric martens in the study area and were not potentially fresh enough for microsatellite analysis (Ruiz-Gonzalez *et al.* 2013). Thus, 301 pine marten samples (227 faecal samples and 74 tissue samples) were used for microsatellite genotyping. The first quality-screening test, based on 4 replicates of four loci, was not passed by 103 noninvasive samples (45.4%), which were then discarded. The remaining 124 samples (54.6%) were amplified at the other 11 loci. After multiple-tube genotyping, 41 samples from this subset (33.06%) were also discarded because they showed <50% PCR success, or because of high failure rates. Full multilocus microsatellite genotypes were obtained for the remaining 83 samples (66.93% from the samples that passed the screening and 36.56% of the total samples analysed), all showing reliability score >0.95 (Miller *et al.* 2002).

The observed average error rates across loci were as follows: ADO = 0.178 and FA = 0.029. PID analysis showed that the set of 15 loci would produce an identical genotype with a probability of 9.83×10^{-12} , and with a probability of 2.09×10^{-5} for a full-sib, suggesting no 'shadow effect' (i.e. all the genotypes identify distinct individuals, and that matching genotypes were recaptures of the same individual) (Mills *et al.* 2000).

After a regrouping procedure, we identified 66 individual genotypes from faecal samples. The 74 tissue samples were correctly genotyped at 15 loci and provided new individuals. In total, we identified 157

genotypes that correspond with 140 different individuals. The number of times each individual was detected in the survey varied from 1 to 3 with a total number of 17 resamplings.

Bayesian clustering analyses

The GENELAND runs giving the highest average posterior probability suggested the presence of four geographically coherent genetic clusters in the study area. The modal population for each pixel of the study area inferred by GENELAND indicated that there is a clear spatial pattern with a west-to-east subdivision between the four inferred pine marten populations (Fig. 2). Individuals from Navarre (a highly forested area) formed a unique genetic cluster (NA) located in the eastern part of the study area. Individuals from the Basque Country (an area with low and high fragmentation of forest) formed three different clusters: eastern Basque (EB), central Basque (CB) and western Basque (WB) clusters (Fig. 2).

STRUCTURE provided consistent results over 20 replicated runs tested for each K . The log-likelihood values for each K [$L(K)$] increased quickly up to $K = 3$ and then reached an asymptote (Fig. 3a). Calculation of ΔK value (Evanno *et al.* 2005) produced a distinct apex

value (300.75) when $K = 3$ (Fig. 3a), implying the likely presence of three genetically distinct groups. The assignment of individuals to populations for $K = 3$ is presented in Fig. 3(b), and their spatial location is provided in Fig. 2.

The modal assignments by STRUCTURE approximately corresponded to those identified with spatial information in the GENELAND analyses (Fig. 3b and Fig. 2), with the exception of the WB cluster, which was not identified in STRUCTURE (Fig. 3b and Fig. 2). About 85.71% of the individuals (120 of 140) were assigned probabilistically to the same genetic cluster, that is N, EB and CB by both methods (taking into account the assignment of the individuals by STRUCTURE to the population for which the estimated membership was the highest). In the case of the 12 individuals (8.57%) identified as a different cluster by GENELAND (WB), these corresponded to individuals assigned by STRUCTURE to the EB or CB cluster and/or clearly admixed individuals ($q < 0.7$) and all of them were located in the western part of the study area (Fig. 3b and Fig. 2). The 8 individuals (5.71%) for which each method provided different population assignments were located between contact zones of different clusters (Fig. 3b and Fig. 2).

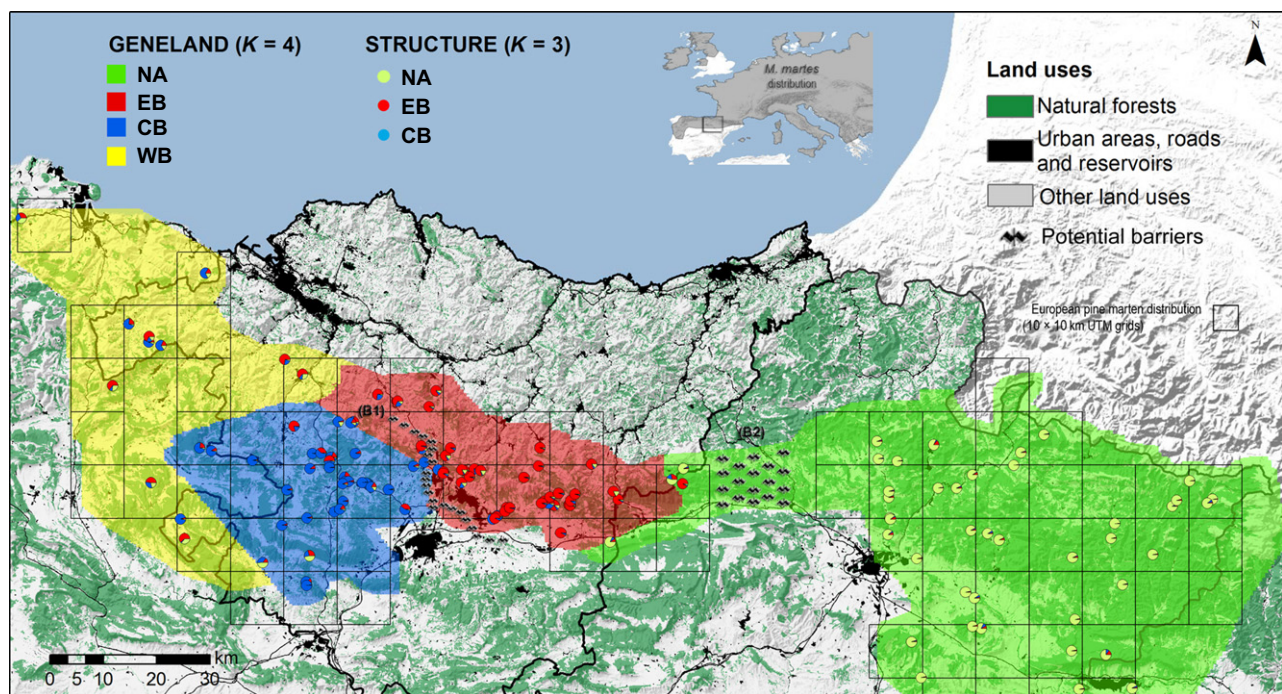


Fig. 2 Modal population for each pixel of the study area according to GENELAND ($K = 4$) and spatial distribution of the cluster membership coefficients according to STRUCTURE ($K = 3$) for each individual. The four genetic clusters are represented by different colours: Navarre (NA) cluster – green, eastern Basque (EB) cluster – red, central Basque (CB) cluster – blue, western Basque (WB) cluster – yellow. Structure results: each individual is represented by pie charts at the location where they were sampled ($n = 140$). The colours indicate the average membership coefficients for each individual to each of the three clusters uncovered by STRUCTURE (green, red and blue for cluster NA, EB and CB, respectively).

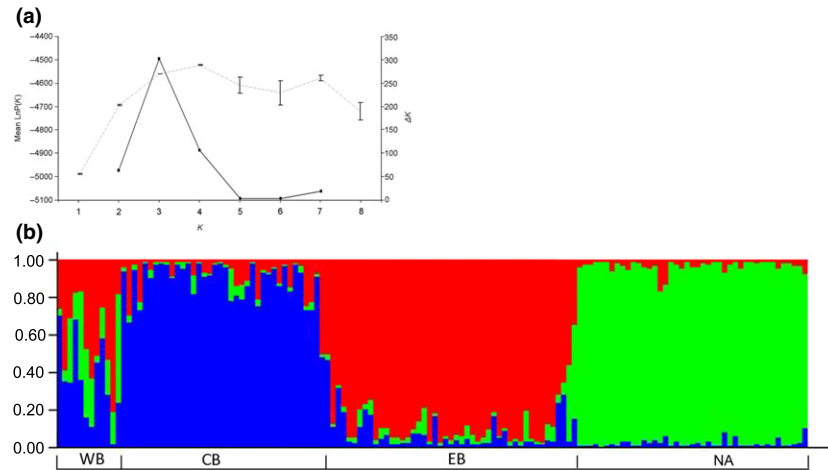


Fig. 3 STRUCTURE analysis results. (a) Estimation of the number of pine marten clusters (K) from 20 independent runs for $K = 1$ – 8 . Dotted line is the mean $\ln P(K)$ (Pritchard *et al.* 2000), and the solid line is the second-order rate of change (ΔK) (Evanno *et al.* 2005), inferring that $K = 3$. (b) Distribution of the three genetic clusters generated by STRUCTURE. The vertical lines are broken into coloured segments showing the proportion of each individual assigned to each of the inferred K . The three genetic clusters are presented by different coloured columns and are displayed according to east–west spatial location of individuals: Navarre (NA) cluster – green, eastern Basque (EB) cluster – red and central Basque (CB) cluster – blue. Letters at the bottom of the figure correspond to GENELAND inferred cluster composition for comparison. The western Basque (WB) GENELAND inferred cluster corresponds to 12 admixed individuals.

Overall, both Bayesian methods indicate that the entire *M. martes* population is separated into at least three subpopulations with a strong spatial pattern (Fig. 3b and Fig. 2) and with sharp borders between populations. Comparing landscape data with population boundaries (Fig. 2), it appears that gene flow barriers revealed by both Bayesian methods coincided with spatial distribution of anthropogenic habitat gaps (boundary between CB and EB cluster) and the presence of the sympatric stone marten (boundary between NA and EB cluster; Fig. 1 and Fig. 2).

Genetic diversity and HWE

The overall pine marten data set ($n = 140$) is not in HWE ($P < 0.001$) (Table 2). These results were consistent with the existence of population structure, which is to be expected at this broad geographical level (see below). Four tests (of 60) deviated from HWE proportions at $P = 0.05$ for the inferred populations. Loci Lut-435 diverged from HWE proportions in two populations, while markers MP-0188 and Mel-1 deviated in one of 4 populations. However, we observed no deviation from HWE in each of the inferred populations (Table 2). The clusters range in sample size from 12 to 47, with an average size of 35 individuals.

Thirty-one of 420 tests for LD (105 tests in each of four populations) were significant at $P = 0.05$. However, the significant LD tests involved different locus pairs

across the four populations, suggesting sampling effects rather than physical genetic linkage. The results of MICROCHECKER indicated that null alleles were apparently present at two loci: MP0188 and Ma19. However, the estimated frequency of the null allele at these loci occurred at a relatively low frequency (0.054–0.060). For all 140 pine martens, the average observed (H_o) and expected (H_e) heterozygosity values were 0.542 and 0.6191, respectively (Table 2). All 15 loci were variable with the total numbers of alleles ranging between 3 and 12 per locus. The mean number of alleles per locus in the 4 subpopulations ranged from 3.73 (WB) to 5.13 (NA), but allelic richness using rarefaction to standardize groups to the smallest sample size provided more similar results among populations with values ranging from 3.17 to 3.96 (Table 2). Similar levels of genetic diversity were observed within each of the 4 subpopulations (Table 2). Observed heterozygosities ranged from 0.511 in the CB subpopulation to 0.588 in the WB subpopulation. Fis values were positive but not significantly different from zero in each of the inferred clusters (Table 2).

Genetic differentiation, migration rates and population bottleneck

Genetic differentiation among clusters was also revealed by F_{ST} values and migration rates. All intercluster pairwise comparisons summarized by mean F_{ST} were sig-

Table 2 Genetic diversity indices of pine marten samples genotyped at 15 microsatellite loci

Cluster	<i>n</i>	<i>A</i>	<i>A_r</i>	<i>H_o</i>	<i>H_e</i>	HWE <i>P</i> value (SE)	Overall <i>F_{is}</i>
EB	47	4.200	3.44	0.5292	0.5621	0.0891 (0.0056)	0.0692
WB	12	3.8677	3.71	0.5878	0.5699	0.3991 (0.0192)	0.01298
CB	38	3.7333	3.17	0.5191	0.5338	0.3976 (0.0098)	0.04092
NA	43	5.1333	3.96	0.5608	0.5845	0.0559 (0.0047)	0.05245
Total	140	5.7333	5.66	0.5412	0.6191	0.0003 *(0.0002)	0.05130

Genetic clusters identified by Geneland: EB, eastern Basque; WB, western Basque; CB, central Basque; N, Navarre; *n*, number of individuals; *A*, mean number of alleles per locus, *A_r* allelic richness, *H_o*, the mean observed heterozygosity, *H_e*, expected heterozygosity, HWE, Hardy–Weinberg equilibrium *P* value and standard error and overall *F_{is}*. Significant values are marked with asterisk ($P < 0.001$).

nificant ($P < 0.05$, Table 3). *F_{ST}* values ranged from 0.056 to 0.170 (Table 2) and suggested low migration rates. The NA cluster was the most differentiated subpopulation with the highest *F_{ST}* value.

This result is in agreement with AMOVA analysis showing a high percentage of variation (11.07%) when considering Navarre and the Basque Country as the main groups in comparison with the 4.27% of variation among individuals within inferred clusters. Considering all the individuals, the percentage of variation was 84.66%, supporting the evidence of population substructure. MIGRATE analysis also indicated that the migration rates between each genetic cluster were low, with bidirectional migration rates ranging between 2.5 and 3.6 individuals per generation (Table 4). However, our results indicated higher emigration from the EB cluster to adjacent areas (i.e. CB and NA clusters) than from adjacent areas to the EB cluster. BOTTLENECK analysis found significant support ($P < 0.05$) for historical reductions in effective population size in each of the inferred clusters under all three microsatellite mutational models (i.e. IAM, SMM, TPM).

Spatial autocorrelation

Spatial autocorrelation analysis also suggested local genetic structure within the study area. The negative regression slope ($b = -0.0437$ SE = 0.00955) between kinship coefficient and logarithmic distance between

individuals was significant ($P < 0.001$). There was significant deviation from the population mean kinship estimate in the closest and most distant distance classes (Fig. 4). Positive values of kinship coefficient were found at short distances, meaning that neighbouring individuals had a higher genetic relatedness than random pairs of individuals, whereas negative values of kinship occurred at larger distances, indicating isolation by distance within the whole study area. The intercept of the correlogram with the X-axis was approximately 70 km (Fig. 4), suggesting that within this distance, individual pine martens are more related than on average in the population.

Spatially explicit indices of genetic diversity

Using the Mantel correlogram, autocorrelation of pairwise genetic distances reached zero at a distance of 69 km and is thus similar to that obtained by spatial autocorrelation analysis using F_{ij} (Fig. 4). sGD mapping of spatial patterns of genetic diversity revealed a trend towards generally increasing observed heterozygosity, gene diversity and number of alleles an allelic richness from west to east across the study area (Fig. 5). All the genetic diversity statistics from sGD analysis revealed a sharp break in spatial genetic diversity coinciding with the distribution gap between the Basque Country and

Table 3 *F*-statistic (*F_{ST}*) tests for pairwise population differentiation (GENELAND groups) based on microsatellite loci frequencies

Cluster	EB	WB	CB	NA
EB	—	0.05684*	0.07981*	0.11379*
WB	—	—	0.07480*	0.07753*
CB	—	—	—	0.17262*

* $P < 0.05$.

Table 4 Bidirectional estimates of gene flow (Nm) between pine marten clusters as calculated with MIGRATE

	Migration from			
	EB	WB	CB	NA
Migration to				
EB	—	2.51806	2.5973	2.86636
WB	2.84265	—	3.62889	2.85025
CB	3.09339	2.52047	—	2.65878
NA	3.17622	2.51639	2.6798	—

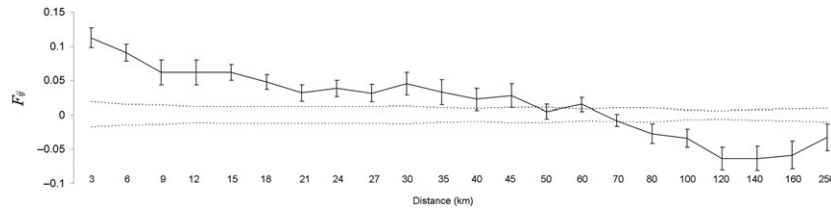


Fig. 4 Average kinship coefficients, F_{ij} , between pairs of pine martens individuals plotted against the Ln geographical distance. Dashed lines represent 95% confidence intervals for F_{ij} under the null hypothesis that genotypes are randomly distributed. Significant deviation of all distance classes from the population mean ($P < 0.001$) was observed with the exception of distance classes of 50 and 70 km.

Navarre pine marten populations (Fig. 2 and Fig. 5). Overall, the lowest genetic diversity indices (H_o , H_e , A_r and A) and the highest inbreeding coefficient (F_{is}) values are located in the Basque Country area (i.e. EB and CB clusters) showing a spatially complex pattern and contrasting with the more homogeneous pattern obtained across the more continuous Navarre population (i.e. NA cluster). Interestingly, sGD also showed an apparent difference in diversity indices values between CB and EB clusters. In addition, summary statistics from sGD analysis showed dramatic differences in genetic diversity values from CB to WB (Fig. 5). Thus, sGD analysis reinforces the boundaries identified by spatially explicit and nonexplicit Bayesian clustering approaches.

Competing IBD, IBR and IBR

The relative support of the reciprocal causal modelling matrix showed that IBD is the strongest overall single hypothesis (Fig. 6a) and is supported independently of IBB and IBR alternative hypotheses. However, the raw form of the partial Mantel matrix (Fig. 6b), more appropriate when competing hypothesis that could be jointly true, showed that IBD and IBB (i.e. isolation by STRUCTURE and GENELAND groupings) are supported independent of all other hypotheses and are supported independently of each other. (partial) Mantel correlation results obtained for each model for both the log-transformed and the untransformed distances provided similar results suggesting a consistent pattern (Fig. 6). All the (partial) Mantel tests were statistically significant ($P < 0.001$).

Discussion

An important issue in population and evolutionary genetics is whether changes in landscape features create barriers to gene flow and lead to discrete population structure or whether clinal population processes create isolation by distance or isolation by resistance patterns of genetic differentiation (Meirmans 2012; Landguth &

Schwartz 2014). In this study, we found that the genetic diversity and differentiation of a forest-dependant small carnivore species inhabiting a heterogeneous landscape follow a complex spatial genetic cline that is related both to isolation by distance and putative habitat and interspecific barriers. As a result, there is both clinal variation in genetic diversity (decreasing from east to west) and differentiation (increasing with distance and more rapidly across 'cluster' boundaries), which produces strong genetic differentiation over short geographical distances.

Spatial genetic structure: clinal and/or hierarchical clusters?

Our data did not support the hypothesis that pine martens in the study area exist as a single, panmictic population. When we considered all sampled individual pine martens, we detected a significant deviation from random mating, indicating that *M. martes* were not distributed in a random fashion and were instead associated with genetically defined clusters. Subpopulation structuring of pine martens at the regional level is confirmed when looking for genetic partitions using Bayesian methods, spatial autocorrelation and sGD analysis and further confirmed by the significant F_{ST} values and low migration rates. Levels of genetic differentiation among the clusters represented by pairwise F_{ST} confirmed moderate to high levels of population differentiation (0.056–0.113) and were significantly higher than similar measures estimated among other pine marten populations sampled across much wider geographical areas in Europe (Kyle *et al.* 2003; Pertoldi *et al.* 2008; Mergey *et al.* 2012). Kyle *et al.* (2003) found an overall F_{ST} value of 0.18 (range: 0.016–0.330), comprising comparisons between geographically distant European populations (e.g. F_{ST} value between Italy and Germany 0.044), while Mergey *et al.* (2012) reported F_{ST} values ranging between 0.047 and 0.052 for populations separated by more than 130 km. Thus, genetic differentiation reported in this study suggested relatively strong population structuring, taking into account the

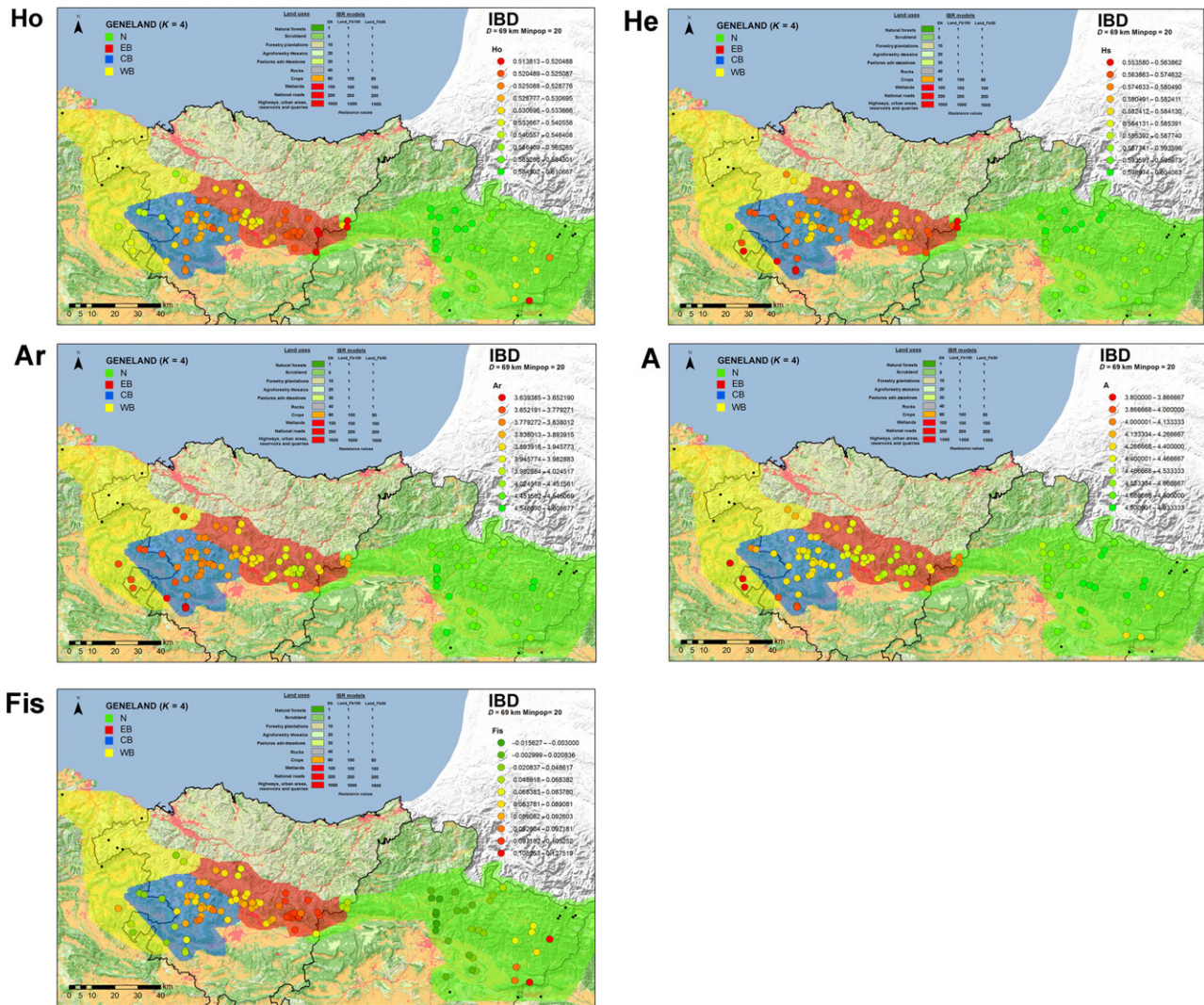


Fig. 5 Observed heterozygosity (H_o), Nei's gene diversity (H_s), allelic richness (A_r), number of Alleles (A), Nei's gene diversity (H_s) and the inbreeding coefficient (F_{is}) calculated using sGD for IBD model. Modal population for each pixel of the study area according to GENELAND ($K = 4$) is shown. Land-use data and related resistance values are indicated according to the different top IBR models identified by Ruiz-González *et al.* (2014) (i.e. EN, Land_FB100 and Land_Fb50). Black dots represent individuals without data as they do not fit the minimum population size of $n = 20$ within the 69-km neighbourhood.

regional scale of the study area and the contiguous nature of the inferred clusters.

The results generated independently by spatially explicit and nonspatial Bayesian approaches suggested the presence of at least three genetic units, with a coherent spatial pattern. Even though the number of clusters varied between approaches (i.e. GENELAND, $K = 4$ and STRCTURE $K = 3$), the location of, and membership within, each of the inferred clusters was largely concordant between methods. Both spatial and aspatial Bayesian clustering methods inferred mainly the same three genetic clusters (i.e. NA, EB and CB) and have a clear convergence between the assignments of the individuals to each cluster (120 of 140 individuals analysed

were assigned probabilistically to the same genetic cluster), giving a high degree of confidence to the proposed clustering solution (Guillot *et al.* 2009).

A certain amount of nonconvergence between different Bayesian clustering methods thus appears to be relatively frequent (Frantz *et al.* 2009; Francois & Durand 2010; Blair *et al.* 2012). In recent applications of spatially explicit and implicit models together to the same data set, when more than one model is used, the studies often report consensual results (Fontaine *et al.* 2007; Latch *et al.* 2008; Liu *et al.* 2009), but there are interesting exceptions (Rowe & Beebe 2007; Ball *et al.* 2010; Francois & Durand 2010; Frantz *et al.* 2012). The additional cluster exclusively identified by GENELAND (i.e.

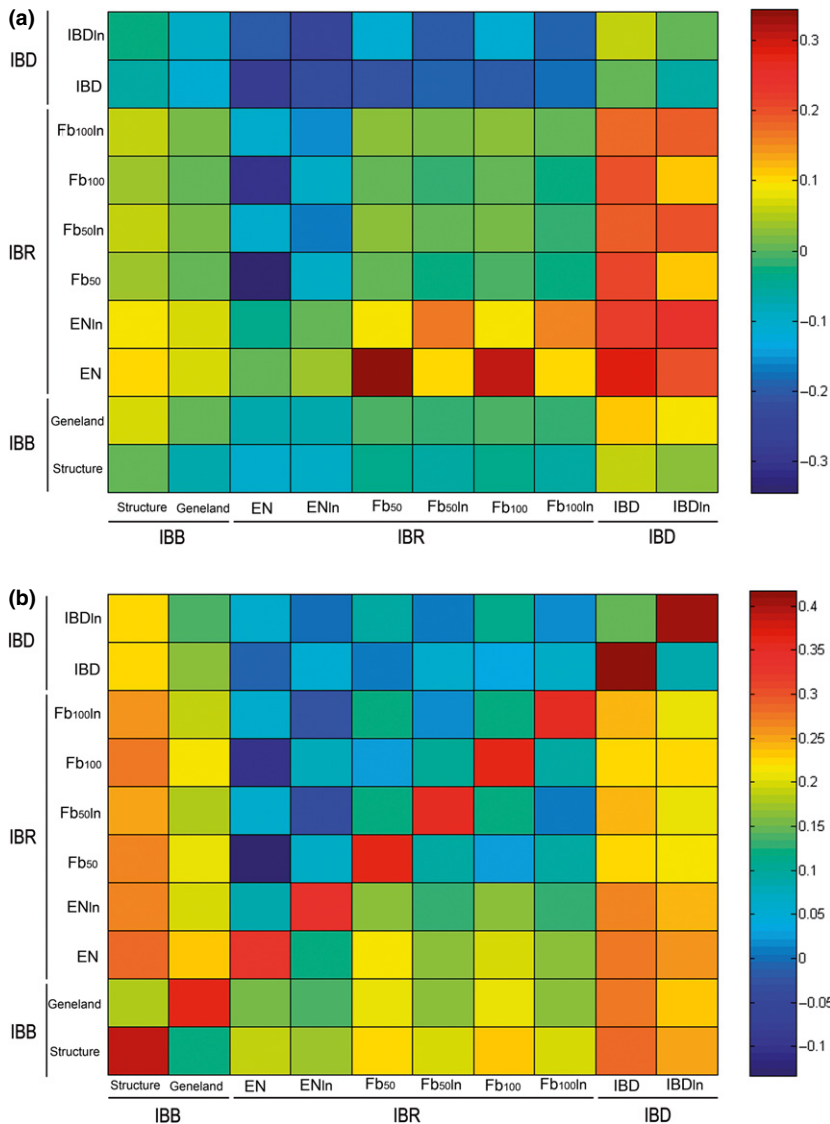


Fig. 6 Reciprocal causal modelling results. Results of (reciprocal) causal modelling on the untransformed and log-transformed (cost) distances for IBD, IBB and IBR models. Columns indicate focal models, and rows indicate alternative models. (a) The colour gradient from blue to red indicates relative support for the focal model independent of the alternative model (e.g. focal model | alternative model – alternative model | focal model is positive). A fully supported model would have all positive values in the vertical dimension (e.g. that model is supported independently of all other models), and all negative values in the horizontal dimension (no other model is supported independently of the focal model) *only* IBD. (b) Partial correlations (Mantel r) of the column model partialling out the row model. The diagonal shows the simple Mantel r value of the focal model. All (partial) Mantel test are significant ($P < 0.001$).

WB) can be interpreted as an example of increased power in spatially structured populations as has been previously documented (Blair *et al.* 2012). However, Francois & Durand (2010) showed that nonadmixture models are not sufficiently robust to the inclusion of admixed individuals in the sample, thus leading to an incorrect assessment of population genetic structure in many generic cases. Consequently, taking into account the high proportion of admixed individuals encompassed within the WB cluster and the low number of individuals ($n = 12$), the validity of the fourth inferred GENELAND cluster should be interpreted with caution and needs to be further evaluated with an increased sample size.

As noted by the authors of the Bayesian algorithms used here (Pritchard *et al.* 2000; Guillot *et al.* 2005b) and recently shown by empirical and simulation studies (Frantz *et al.* 2009; Schwartz & McKelvey 2009; Hobbs

et al. 2011; Blair *et al.* 2012; Aurelle & Ledoux 2013), deviations from random mating not caused by barriers to gene flow (i.e. spatial autocorrelation and IBD) may have impacts on the accurate detection and interpretation of genetic structure. In this study, to test for the presence of spurious genetic clusters and evaluate the relative support for discrete vs. clinal structure, we jointly tested the effect of clinal (i.e. IBD and IBR) and/or discrete clustering (i.e. IBB) using a reciprocal causal modelling approach (Cushman *et al.* 2013c, 2014). Our results suggest a complex clinal pattern of increasing isolation by distance, with steep gradients coinciding with the Bayesian cluster boundaries. The independent support of IBD and both IBB models (i.e. GENELAND and STRUCTURE) within the reciprocal causal modelling framework indicates nonisotropic clinal structure with steep gradients of increasing differentiation in areas corresponding to the cluster boundaries.

By testing IBB, IBD and IBR using reciprocal causal modelling (Cushman *et al.* 2013c), we were able to rigorously test for the joint and independent effects of alternative clinal and discrete hypotheses of genetic structure, which is an improvement over most previous studies of genetic differentiation which did not formally compete clinal and discrete alternative hypotheses (but see Cushman & Landguth 2010; Cushman *et al.* 2014), and typically struggle with high levels of type I error and issues of affirming the consequent (Cushman *et al.* 2013c).

Limiting factors acting at different spatial scales

The fact that IBD and IBB outperformed the IBR models optimized in the study by Ruiz-González *et al.* (2014) could reflect different limiting factors acting at different spatial scales (Cushman *et al.* 2013b). Specifically, the IBR model was optimized in the Basque Country, which is characterized by relatively low extent and high fragmentation of forest, while the larger extent of the current study area includes large areas of highly connected habitat in the Navarre region. Cushman *et al.* (2013b) showed that when the landscape is highly fragmented, isolation by resistance is detectable independently of isolation by distance, but as habitat extent increases and fragmentation declines, the ability to independently detect effects of IBR declines. The limiting factor issue in landscape genetics has been thoroughly described in Cushman *et al.* (2013b), and this appears to be what is happening in this study. Similarly, Wasserman *et al.* (2012) used simulation modelling to show that when suitable American pine marten (*Martes americana*) habitat was not highly fragmented, statistical models were not able to detect the effects of habitat fragmentation on gene flow because of the limiting factor issue described by Cushman *et al.* (2013b). Therefore, the fact that none of the IBR hypotheses were supported independently of IBD or IBB in this study should be interpreted within the context of previous work which showed that pine marten gene flow is indeed affected by habitat fragmentation in portions of the range that are highly fragmented (Ruiz-González *et al.* 2014). Indeed, this has been demonstrated by the presence of sharp genetic discontinuities largely mirroring habitat gaps between populations and previous landscape genetic findings, suggesting an important landscape effect of human-dominated landscapes in the Basque Country area (Ruiz-González *et al.* 2014). Additionally, the pine marten population is experiencing range expansion in some areas of the northern Iberian Peninsula, which may produce a time lag effect leading to weak or undetectable landscape genetic relationships during the range expansion event (Anderson *et al.* 2010; Reding *et al.*

2013). Based on simulations, Landguth *et al.* (2010) found that distance and landscape effects can be detected almost immediately in spatial genetic patterns (1–15 generations), but do not reach equilibrium for several hundred generations. Thus, a temporal lag in the build-up of a genetic response together with the effects of the limiting factor issue (Cushman *et al.* 2013b) may be contributing to the inability to disentangle a landscape effect from a pure distance and/or barrier effect. Similar results have been recently obtained by Reding *et al.* (2013), in an expanding recolonizing bobcat population in a landscape that might not be limiting gene flow. Recently, Cushman (2015) described a number of challenges in the landscape genetic analysis of expanding populations, including nonequilibrium between landscape and genetic structure, the limiting effects of congeneric competitors and possible differences in factors limiting gene flow across the expanding population. All three of these issues are identified as possible explanations of the results reported here. As suggested by Cushman (2015), future work should seek to confirm and separate these processes by employing simulation modelling (e.g. Landguth *et al.* 2010) and/or comparative mensurative replicated field experiments (i.e. sampling is conducted in a structured way and is blocked into units corresponding to variables of interest; e.g. McGarigal & Cushman 2002; Short Bull *et al.* 2011; Shirk *et al.* 2014).

Spatially explicit evaluation of genetic diversity

Diversity indices are typically calculated from the multilocus genotypes of all individuals sampled within discretely defined habitat patches, larger regional extents and/or based on clustering approaches (Shirk & Cushman 2011, 2014). In this study, we compared diversity indices calculated for putative clusters and across the population in a continuous moving-window approach using sGD (Shirk & Cushman 2011). We found that diversity indices based on clustering approach were intermediate between the highest and lowest values calculated with sGD (Table 2, Fig. 5). However, the discrete population approach did not capture the clinal attributes of population structure. Thus, the sGD analysis provided a more accurate measure of genetic diversity patterns that are hidden using standard population genetic approaches that assume discrete population structure.

Spatially explicit sGD analysis clearly showed a complex spatial pattern of genetic diversity that reinforces the existence of clines of varying steepness across the study area, with gradients in genetic diversity which steepen at the clustering boundaries. This reinforces the presence of dispersal barriers in these locations, but

reinterprets their meaning as areas of relatively more rapid change within an overall continuous and clinal pattern of genetic differentiation and diversity. This is a welcome finding, because while clustering suggests population subdivision in these areas, spatially explicit genetic diversity analysis showed that these subdivisions are characterized by reduced genetic diversity, reflecting low population density and gene flow. Interestingly, the lowest genetic diversity and highest FIS values were found in the fragmented landscape of the Basque Country area (i.e. CB, EB, WB clusters), with an increased genetic diversity eastwards, in the continuous forest system of the Pyrenees (i.e. NA cluster).

Putative barriers to gene flow: synergic effects of overharvesting, habitat fragmentation and landscape structure

Different and nonexclusive hypotheses can explain the significant genetic structure observed in pine marten populations at the regional scale. First, overharvesting driven by their valuable fur greatly depleted pine marten populations during 20th century across northern Spain (López-Martín 2007). The three primary genetic consequences of overharvesting are the alteration of population structure, loss of genetic variation and evolution resulting from selection (Harris *et al.* 2002; Allendorf *et al.* 2008). Second, habitat fragmentation and deforestation may have had important consequences for population genetic structure, given the species' high dependence on forest habitats and strong sensitivity to habitat loss (Schwartz *et al.* 2012). Therefore, our results suggested that overharvesting processes in combination with deforestation and forest fragmentation (López-Martín 2007) could be the main drivers of the population decline and range contraction of pine marten populations in the study area, thus leading to (i) recurrent population bottlenecks detected in all the inferred clusters, (ii) the strong spatial genetic structure observed between geographically proximal localities and (iii) very limited migration between adjacent subpopulations. Consistent with these findings, the genetic discontinuity found between CB and EB, evidenced by the Bayesian clustering and spatially explicit indices of genetic diversity, clearly overlapped with the geographical location of an area characterized by a high density of roads, the presence of large man-made reservoirs and the high presence of human urbanization (Fig. 2). Indeed, this area has been previously identified as one of the critical connectivity areas in the framework of the Basque connectivity network, specifically designed for forest species (Gurrutxaga *et al.* 2010a,b). Moreover, recent research in the Basque Country has shown that gene flow of pine marten is facilitated through natural vegetation and is

diminished in human-dominated land uses such as agricultural, urban areas and linear anthropogenic barriers such as highways (Ruiz-González *et al.* 2014). On the other hand, we found that the Navarre population, which inhabits a region characterized by extensive and well-connected forest, is represented by a single genetic cluster, suggesting that in this area, habitat fragmentation does not create deviance from isolation by distance. These results are in agreement with previous studies suggesting that forest fragmentation may often be one of the main factors structuring present-day genetic diversity in forest-dwelling species (Broquet *et al.* 2006; Segelbacher *et al.* 2008; Liu *et al.* 2009).

In addition to habitat fragmentation and overharvesting, our results suggest that interspecific competition between sympatric marten species (Goszczyński *et al.* 2007; Virgós *et al.* 2012) could also contribute to the genetic structuring of pine marten populations. In contrast to the pine marten, which have experienced widespread population declines and are either threatened or rare in many European countries (Proulx *et al.* 2004), the range and density of stone martens has increased in many countries (Proulx *et al.* 2004; Virgós *et al.* 2012). It has been hypothesized that these inverse demographic trends may have amplified interspecific competition, potentially leading to further reduction in the pine marten's range (i.e. stenotopic species) and the geographical expansion of the stone marten's (i.e. synanthropic species) (Goszczyński *et al.* 2007; Virgós *et al.* 2012), which shows more behavioural plasticity, a broader habitat and climate niche (Virgós *et al.* 2012), and being less vulnerable to anthropogenic changes and human presence (Herr *et al.* 2009; Virgós *et al.* 2012). Spatial distribution data obtained through a long-term noninvasive genetic monitoring provided support to this hypothesis, showing an overall spatial segregation pattern, with the pine marten confined to remnant forest habitats and the stone marten inhabiting a wide range of landscapes, from forested areas to human-dominated landscapes. Interestingly, we found that the genetic discontinuity located between the EB and NA clusters corresponded to an area where the pine marten distribution shows a clear gap in spite of the presence of suitable forest habitat (Fig. 1 and Fig. 2). Indeed, the distribution gap matched with a pattern of local abundance of the more synanthropic stone marten. Consequently, the colonization of new areas by the pine marten may be significantly reduced in areas where the stone marten is well established, thus acting as a potential biotic barrier to gene flow between pine marten populations which could synergistically amplify the effects of other abiotic factors. These results provide an interesting insight into the potential effects of interspecific competition on dispersal and genetic structure in sympatric carnivores

that should be further explored. An ongoing work on multiscale habitat suitability modelling (Vergara *et al.* in prep.) will allow more accurately assessing the interspecific relationships of martens in relation to environmental and landscaping variables.

Implications for conservation and connectivity restoration

Our results show that the demographic bottleneck in the 20th century, which resulted from the combined effects of habitat loss and persecution, has impacted genetic variation of the pine marten. Moreover, these processes gave rise to an increased range contraction and spatial differentiation between remnant populations inhabiting remnant forested mountainous habitats of the Basque Country that perfectly matched the core areas of the identified pine marten clusters. In addition, our study identified a clear distributional gap between the Basque Country populations and those occupying the Pyrenean range (i.e. Navarre) that corresponds to a major genetic disjunction. If the species continues to follow the recent pattern of range recovery in the study area (pers. obs.), which is also reported in other pine marten populations (e.g. Balestrieri *et al.* 2010; O'Mahony *et al.* 2012), this could re-establish genetic exchange between some isolated populations identified in our study. However, the identified barriers for dispersal could seriously hamper this recolonization process. The resistance surfaces and genetic diversity mapping we produced may be instrumental in identifying areas where management intervention, such as reintroductions or assisted migration, would be warranted. Additionally, the populations under study, inhabiting at the southwesternmost frontier of the species range, are of special conservation relevance as they are restricted to a very narrow ecological niche and associated small population sizes and high population fragmentation, which are further threatened by ongoing climatic warming.

For effective pine marten conservation on the Iberian Peninsula, securing dispersal between disjunct habitat patches appears to be the major challenge. Indeed, ensuring connectivity is of high priority in the study area (Gurrutxaga *et al.* 2010a,b) as this region is considered of strategic importance for the conservation of ecological connectivity in southwestern Europe (Mallarach *et al.* 2010), given its critical location as a linkage between the Cantabrian and the Pyrenees mountains. Thus, given severe limitations to connectivity among remnant populations (Ruiz-González *et al.* 2014) in interaction with an abundant and widespread potential competitor (i.e. *Martes foina*), additional conservation measures should explicitly consider population

processes operating at the landscape scale to ensure population connectivity (Cushman *et al.* 2013a). Improving habitat quality and connectivity to mitigate the identified barriers will not only improve the long-term persistence of the pine marten but also favour connectivity for other forest-associated species (Gurrutxaga *et al.* 2010a). The future coexistence and relative abundances of the two martens in forest habitats will thus depend on the mode of forest management as well as on the existence of effective corridors connecting neighbouring forest patches (Goszczyński *et al.* 2007; Gurrutxaga *et al.* 2010a). Additional fine-scale demographic and genetic surveys of *Martes* spp. will be necessary in the next years to effectively monitor the conservation status of sympatric marten populations.

Conclusions

Very few studies have rigorously evaluated and jointly tested the effects of clinal (IBD and IBR) and discrete structuring (IBB) using different comparable analytical approaches (Cushman & Landguth 2010; Cushman *et al.* 2014), and fewer still have employed spatially explicit indices of genetic diversity (Shirk & Cushman 2011) to describe and quantify complex patterns of genetic diversity. To our knowledge, this study is the first to disentangle discrete (IBB) and/or clinal population structure (IBD and/or IBR) in a forest-dwelling mesocarnivore. By directly competing clinal and discrete alternative models in an individual-based landscape genetic framework, we demonstrated that there is a complex clinal structure with areas of strong local differentiation associated with distribution gaps, areas of low population density, habitat fragmentation and abundance of a potential congeneric competitor. In composite, our results have shown that pine martens inhabiting a fragmented landscape at the southern edge of species distribution do not belong to a single panmictic population, but rather form a complex cline that is represented jointly by IBD and putative habitat and interspecific barriers producing strong genetic differentiation over short geographical distances.

The data reported in this study suggested that habitat loss and fragmentation, in synergy with past overharvesting and interspecific competition with the stone marten, are likely the main factors responsible for the reduced range, recurrent bottlenecks and low migration rates in pine marten populations, giving rise to the observed spatial genetic structure. Several landscape features (i.e. unforested areas, urbanized areas, roads and man-made reservoirs) act as moderators of gene flow because of their high resistance to pine marten migrations (Ruiz-González *et al.* 2014), and hence, their cumulative effect likely led to the differentiation of the

inferred genetic units. Overall, these findings are of interest to the wide audience of scientists interested in landscape and evolutionary genetics, emphasizing the need for a more thorough evaluation of different hypothesis governing gene flow (i.e. IBD, IBR and IBB) in landscape genetic studies, and the potential influence of different limiting factors affecting spatial genetic structure at different spatial scales. Future work exploring the interaction of range expansion and limiting factors should be undertaken to further explore and explain the joint clinal and clustered structure of pine marten populations in northern Spain (e.g. Cushman 2015).

Acknowledgements

This study has been funded by the Basque Government through the Research group on 'Systematics, Biogeography and Population Dynamics' (Ref. IT317-10; GIC10/76; IT575/13), the SAIOTEK research programme (Ref: S-PE11UN028) and by the University of the Basque Country (UPV-EHU) and the Department of Environment, Territorial Planning, Agriculture and Fisheries (Basque Government) through IKT S.A under the University-Enterprise Research programme (Ref. UE07/02). Ruiz-González holds a Post doc fellowship awarded by the Dept. of Education Universities and Research of the Basque Government (Ref. DKR-2012-64). Several samples analysed in this study have been obtained in the framework of different carnivore surveys funded by regional or national administrations (Spanish Ministry of Environment, Regional Governments of Navarre and Aragon, Alava Provincial Council). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript. The authors wish to thank the following persons and institutions for supplying some of the tissue and faecal samples used in this study: Patricia Lizarraga, Ricardo Gutierrez and Laura Elorza (CRF Martioda—Alava Provincial Council); the technical staff and rangers of the Natural Parks of Alava (Juan Carlos Ortíz, Ricardo Ortíz, María Elena García, Kepa García, Sonia Benítez, Elisabeth Cabanillas, Iker Ayala, Arantza Puente, Jesús Gómez, Iñaki Martínez, Lidia Lacha, Jokin Sáez de Camara); Rangers from Alava (Asier Martínez, Tomás Landa, Javier Pinedo, Mario Corral) and Bizkaia (Igor Aginako, Julio Ruiz Guijarro, Ignacio Martínez, Eneko Díaz, Juan Manuel Pérez de Ana, David Vado, Ander Eguia) Provincial Councils; Joseba Carreras (Alava Provincial Council); Jokin Larumbe (Navarre Government); Dr. Fermín Urrea (GAVRN); Asun Gómez (TRAGSATEC); Dr. Madis Podra (Asoc. Visión Europeo); Javier Sesma and Txema Fernandez (IKT); Nerea Ruiz de Azua (EKOS, S.L.);

Gorka Belamendia (CEA); Felipe Canales and Miguel Ángel Campos (Consultora de Recursos Naturales, S.L); Enrique Arberas; Oskar Bardiñ; Javier López de Luzuriaga; Dr. Iñigo Zuberogoitia (ICARUS); Dr. Marta Barral (NEIKER); Rebeca Pérez; Haizea Aguirre; Pablo Pérez; Amaia García de Albéniz; Alvaro Osés; and Gerardo Domínguez. Especial thanks to Maria Vergara, Luis Javier Chueca and Oihana Razkin (UPV/EHU) and to the staff of the Conservation Genetics Laboratory (ISPRA) for their valuable help at different stages of the study. We are also indebted to M. Gurrutxaga (Dept. of Geography, UPV-EHU) for his support with GIS data.

References

- Allendorf FW, Englund PR, Luikart G, Ritchie PA, Ryman N (2008) Genetic effects of harvest on wild animal populations. *Trends in Ecology & Evolution*, **23**, 327–337.
- Amos JN, Bennett AF, Mac Nally R *et al.* (2012) Predicting landscape-genetic consequences of habitat loss, fragmentation and mobility for multiple species of woodland birds. *PLoS One*, **7**, e30888.
- Anderson CD, Epperson BK, Fortin M-J *et al.* (2010) Considering spatial and temporal scale in landscape-genetic studies of gene flow. *Molecular Ecology*, **19**, 3565–3575.
- Aurelle D, Ledoux J-B (2013) Interplay between isolation by distance and genetic clusters in the red coral *Corallium rubrum*: insights from simulated and empirical data. *Conservation Genetics*, **14**, 705–716.
- Balestrieri A, Remonti L, Ruiz-Gonzalez A *et al.* (2010) Range expansion of the pine marten (*Martes martes*) in an agricultural landscape matrix (NW Italy). *Mammalian Biology*, **75**, 412–419.
- Balestrieri A, Remonti L, Ruiz-González A *et al.* (2014) Distribution and habitat use by pine marten *Martes martes* in a riparian corridor crossing intensively cultivated lowlands. *Ecological Research*, **30**, 153–162.
- Ball MC, Finnegan L, Manseau M, Wilson P (2010) Integrating multiple analytical approaches to spatially delineate and characterize genetic population structure: an application to boreal caribou (*Rangifer tarandus caribou*) in central Canada. *Conservation Genetics*, **11**, 2131–2143.
- Berli P (2004) Effect of unsampled populations on the estimation of population sizes and migration rates between sampled populations. *Molecular Ecology*, **13**, 827–836.
- Berli P (2006) Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. *Bioinformatics*, **22**, 341–345.
- Berli P, Felsenstein J (2001) Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. *Proceedings of the National Academy of Sciences of the United States of America*, **98**, 4563–4568.
- Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (2004) Genetix 4.02, Logiciel sous windows pour la génétique des populations.
- 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.

- Brainerd SM, Rolstad J (2002) Habitat selection by Eurasian pine martens *Martes martes* in managed forests of southern boreal Scandinavia. *Wildlife Biology*, **8**, 289–297.
- Broquet T, Johnson C, Petit E *et al.* (2006) Dispersal and genetic structure in the American marten, *Martes americana*. *Molecular Ecology*, **15**, 1689–1697.
- Castillo JA, Epps CW, Davis AR, Cushman SA (2014) Landscape effects on gene flow for a climate-sensitive montane species, the American pika. *Molecular Ecology*, **23**, 843–856.
- Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*, **144**, 2001–2014.
- Cushman SA (2015) Pushing the envelope in genetic analysis of species invasion. *Molecular Ecology*, **24**, 259–262.
- Cushman SA, Landguth EL (2010) 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. *The American Naturalist*, **168**, 486–499.
- Cushman SA, Lewis JS, Landguth EL (2013a) Evaluating the intersection of a regional wildlife connectivity network with highways. *Movement Ecology*, **1**, 1–12.
- Cushman SA, Shirk AJ, Landguth EL (2013b) Landscape genetics and limiting factors. *Conservation Genetics*, **14**, 263–274.
- Cushman SA, Wasserman TN, Landguth EL, Shirk AJ (2013c) Re-evaluating causal modeling with mantel tests in landscape genetics. *Diversity*, **5**, 51–72.
- Cushman SA, Max T, Meneses N *et al.* (2014) Landscape genetic connectivity in a riparian foundation tree is jointly driven by climatic gradients and river networks. *Ecological Applications*, **24**, 1000–1014.
- Devictor V, Julliard R, Jiguet F (2008) Distribution of specialist and generalist species along spatial gradients of habitat disturbance and fragmentation. *Oikos*, **117**, 507–514.
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611–2620.
- Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics*, **1**, 47–50.
- Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, **164**, 1567–1587.
- Falush D, Stephens M, Pritchard JK (2007) Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes*, **7**, 574–578.
- Fontaine MC, Baird SJE, Piry S *et al.* (2007) Rise of oceanographic barriers in continuous populations of a cetacean: the genetic structure of harbour porpoises in Old World waters. *BMC Biology*, **5**, 30.
- Francois O, Durand E (2010) Spatially explicit Bayesian clustering models in population genetics. *Molecular Ecology Resources*, **10**, 773–784.
- Frantz A, Cellina S, Krier A, Schley L, Burke T (2009) Using spatial Bayesian methods to determine the genetic structure of a continuously distributed population: clusters or isolation by distance? *Journal of Applied Ecology*, **46**, 493–505.
- Frantz AC, Bertouille S, Eloy MC *et al.* (2012) Comparative landscape genetic analyses show a Belgian motorway to be a gene flow barrier for red deer (*Cervus elaphus*), but not wild boars (*Sus scrofa*). *Molecular Ecology*, **21**, 3445–3457.
- Goszczyński J, Posłuszny M, Pilot M, Gralak B (2007) Patterns of winter locomotion and foraging in two sympatric marten species: *Martes martes* and *Martes foina*. *Canadian Journal of Zoology*, **85**, 239–249.
- Goudet J (1995) FSTAT (Version 1.2): a computer program to calculate F-statistics. *Journal of Heredity*, **86**, 485–486.
- Goudet J (2001) FSTAT: A program to estimate and test gene diversities and fixation indices. Version 2.9.3, updated from Goudet 1995. Available from <http://www.unil.ch/izea/software/fstat.html>.
- Guillot G, Santos F (2009) A computer program to simulate multilocus genotype data with spatially autocorrelated allele frequencies. *Molecular Ecology Resources*, **9**, 1112–1120.
- Guillot G, Estoup A, Mortier F, Cosson JF (2005a) A spatial statistical model for landscape genetics. *Genetics*, **170**, 1261–1280.
- Guillot G, Mortier F, Estoup A (2005b) GENELAND: a computer package for landscape genetics. *Molecular Ecology Notes*, **5**, 712–715.
- Guillot G, Leblois R, Coulon A, Frantz AC (2009) Statistical methods in spatial genetics. *Molecular Ecology*, **18**, 4734–4756.
- Guo SW, Thompson EA (1992) Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics*, **48**, 361–372.
- Gurrutxaga M, Lozano PJ, del Barrio G (2010a) GIS-based approach for incorporating the connectivity of ecological networks into regional planning. *Journal for Nature Conservation*, **18**, 318–326.
- Gurrutxaga M, Lozano PJ, Del Barrio G (2010b) Assessing highway permeability for the restoration of landscape connectivity between protected areas in the Basque Country, Northern Spain. *Landscape Research*, **35**, 529–550.
- Hall LA, Beissinger SR (2014) A practical toolbox for design and analysis of landscape genetics studies. *Landscape Ecology*, **29**, 1487–1504.
- Hampe A, Petit RJ (2005) Conserving biodiversity under climate change: the rear edge matters. *Ecology Letters*, **8**, 461–467.
- Hardy OJ, Vekemans X (2002) SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes*, **2**, 618–620.
- Harris RB, Wall WA, Allendorf FW (2002) Genetic consequences of hunting: what do we know and what should we do? *Wildlife Society Bulletin*, **30**, 634–643.
- Helldin JO (2000) Population trends and harvest management of pine marten *Martes martes* in Scandinavia. *Wildlife Biology*, **6**, 111–120.
- Herr J, Schley L, Roper TJ (2009) Socio-spatial organization of urban stone martens. *Journal of Zoology*, **277**, 54–62.
- Hobbs GI, Chadwick EA, Bruford MW, Slater FM (2011) Bayesian clustering techniques and progressive partitioning to identify population structuring within a recovering otter population in the UK. *Journal of Applied Ecology*, **48**, 1206–1217.
- IGN (2008) Base cartográfica numérica BCN200.
- Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, **23**, 1801–1806.
- Kalinowski ST (2005) HP-RARE 1.0: a computer program for performing rarefaction on measures of allelic richness. *Molecular Ecology Notes*, **5**, 187–189.

- Kurki S, Nikula A, Helle P, Linden H (1998) Abundances of red fox and pine marten in relation to the composition of boreal forest landscapes. *Journal of Animal Ecology*, **67**, 874–886.
- Kyle CJ, Davison A, Strobeck C (2003) Genetic structure of European pine martens (*Martes martes*), and evidence for introgression with *M. americana* in England. *Conservation Genetics*, **4**, 179–188.
- Landguth EL, Schwartz MK (2014) Evaluating sample allocation and effort in detecting population differentiation for discrete and continuously distributed individuals. *Conservation Genetics*, **15**, 981–992.
- Landguth EL, Cushman SA, Schwartz MK *et al.* (2010) Quantifying the lag time to detect barriers in landscape genetics. *Molecular Ecology*, **19**, 4179–4191.
- Latch EK, Scognamiglio DG, Fike JA, Chamberlain MJ, Rhodes OE (2008) Deciphering ecological barriers to North American river otter (*Lontra canadensis*) gene flow in the Louisiana landscape. *Journal of Heredity*, **99**, 265–274.
- Liu ZJ, Ren BP, Wu RD *et al.* (2009) The effect of landscape features on population genetic structure in Yunnan snub-nosed monkeys (*Rhinopithecus bieti*) implies an anthropogenic genetic discontinuity. *Molecular Ecology*, **18**, 3831–3846.
- Loiselle BA, Sork VL, Nason J, Graham C (1995) Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). *American Journal of Botany*, **82**, 1420–1425.
- López-Martín J (2007) *Martes martes* (Linnaeus, 1758). In: *Atlas y Libro Rojo de los Mamíferos Terrestres de España* (eds Palomo L, Gisbert J, Blanco J), pp. 302–304. Dirección General de Biodiversidad-SECEM-SECEMU, Madrid.
- Luque S, Saura S, Fortin M-J (2012) Landscape connectivity analysis for conservation: insights from combining new methods with ecological and genetic data. *Landscape Ecology*, **27**, 153–157.
- Mallarach JM, Rafa M, Sargatal J (2010) Cantabrian Mountains-Pyrénées-Massif Central-Western Alps great mountain corridor. In: *Connectivity Conservation Management. A Global Guide* (eds Worboys GL, Francis WL, Lockwood M), pp. 269–279. Earthscan, London.
- Manel S, Holderegger R (2013) Ten years of landscape genetics. *Trends in Ecology & Evolution*, **28**, 614–621.
- McGarigal K, Cushman SA (2002) Comparative evaluation of experimental approaches to the study of habitat fragmentation effects. *Ecological Applications*, **12**, 335–345.
- McRae BH (2006) Isolation by resistance. *Evolution*, **60**, 1551–1561.
- Meirmans PG (2012) The trouble with isolation by distance. *Molecular Ecology*, **21**, 2839–2846.
- Mergey M, Helder R, Roeder J-J (2011) Effect of forest fragmentation on space-use patterns in the European pine marten (*Martes martes*). *Journal of Mammalogy*, **92**, 328–335.
- Mergey M, Larroque J, Ruet S *et al.* (2012) Linking habitat characteristics with genetic diversity of the European pine marten (*Martes martes*) in France. *European Journal of Wildlife Research*, **58**, 909–922.
- Miller CR, Joyce P, Waits LP (2002) Assessing allelic dropout and genotype reliability using maximum likelihood. *Genetics*, **160**, 357–366.
- Mills LS, Citta JJ, Lair KP, Schwartz MK, Tallmon DA (2000) Estimating animal abundance using noninvasive DNA sampling: promise and pitfalls. *Ecological Applications*, **10**, 283–294.
- Mitchell-Jones AJ, Amori G, Bogdanowicz W *et al.* (1999) *Atlas of European Mammals*. Academic Press, London.
- O'Mahony D, O'Reilly C, Turner P (2012) Pine marten (*Martes martes*) distribution and abundance in Ireland: a cross-jurisdictional analysis using non-invasive genetic survey techniques. *Mammalian Biology*, **77**, 351–357.
- Pereboom V, Mergey M, Villerette N *et al.* (2008) Movement patterns, habitat selection, and corridor use of a typical woodland-dweller species, the European pine marten (*Martes martes*), in fragmented landscape. *Canadian Journal of Zoology-Revue Canadienne De Zoologie*, **86**, 983–991.
- Pertoldi C, Barker SF, Madsen AB *et al.* (2008) Spatio-temporal population genetics of the Danish pine marten (*Martes martes*). *Biological Journal of the Linnean Society*, **93**, 457–464.
- Piry S, Luikart G, Cornuet JM (1999) BOTTLENECK: a computer program for detecting recent reductions in the effective population size using allele frequency data. *Journal of Heredity*, **90**, 502–503.
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.
- Proulx G, Aubry KB, Birks J *et al.* (2004) World distribution and status of the genus *Martes* in 2000. In: *Martens and Fishers (Martes) in Human-Altered Environments: An International Perspective* (eds Harrison DJ, Fuller AK, Proulx G), pp. 77–98. Springer-Verlag, New York City, New York.
- R Development Core Team (2011) *R: A Language and Environment for Statistical Computing*. R Development Core Team, Vienna, Austria. Available from <http://www.r-project.org/>.
- Reding DM, Cushman SA, Gosselink TE, Clark WR (2013) Linking movement behavior and fine-scale genetic structure to model landscape connectivity for bobcats (*Lynx rufus*). *Landscape Ecology*, **28**, 471–486.
- Rousset FDO (2008) GENEPOP '007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources*, **8**, 103–106.
- Rowe G, Beebe TJC (2007) Defining population boundaries: use of three Bayesian approaches with microsatellite data from British natterjack toads (*Bufo calamita*). *Molecular Ecology*, **16**, 785–796.
- Ruiz-Gonzalez A, Madeira MJ, Randi E, Urrea F, Gomez-Moliner BJ (2013) Non-invasive genetic sampling of sympatric marten species (*Martes martes* and *Martes foina*): assessing species and individual identification success rates on faecal DNA genotyping. *European Journal of Wildlife Research*, **59**, 371–386.
- Ruiz-González A, Rubines J, Berdiñ O, Gómez-Moliner BJ (2008) A non-invasive genetic method to identify the sympatric mustelids pine marten (*Martes martes*) and stone marten (*Martes foina*): preliminary distribution survey on the northern Iberian Peninsula. *European Journal of Wildlife Research*, **54**, 253–261.
- Ruiz-González A, Gurrutxaga M, Cushman SA *et al.* (2014) Landscape genetics for the empirical assessment of resistance surfaces: the European Pine Marten (*Martes martes*) as a target-species of a regional ecological network. *PLoS One*, **9**, 19.
- Schwartz MK, McKelvey KS (2009) Why sampling scheme matters: the effect of sampling scheme on landscape genetic results. *Conservation Genetics*, **10**, 441–452.
- Schwartz MK, Ruiz-González A, Pertoldi C, Masuda R (2012) *Martes* conservation genetics: assessing within species movements, units to conserve and connectivity cross ecological and evolutionary time. In: *Biology and Conservation of*

- Marten, Sables, and Fisher: A New Synthesis (eds Aubry KB, Zielinski WJ, Raphael MG, Proulx G, Buskirk SW), pp. 398–428. Cornell University Press, New York.
- Segelbacher G, Manel S, Tomiuk J (2008) Temporal and spatial analyses disclose consequences of habitat fragmentation on the genetic diversity in capercaillie (*Tetrao urogallus*). *Molecular ecology*, **17**, 2356–2367.
- Segelbacher G, Cushman S, Epperson B *et al.* (2010) Applications of landscape genetics in conservation biology: concepts and challenges. *Conservation Genetics*, **11**, 375–385.
- Shirk AJ, Cushman SA (2011) sGD: software for estimating spatially explicit indices of genetic diversity. *Molecular Ecology Resources*, **11**, 922–934.
- Shirk AJ, Cushman SA (2014) Spatially-explicit estimation of Wright's neighborhood size in continuous populations. *Frontiers in Ecology and Evolution*, **2**, 1–12.
- Shirk AJ, Raphael MG, Cushman SA (2014) Spatiotemporal variation in resource selection: insights from the American marten (*Martes americana*). *Ecological Applications*, **24**, 1434–1444.
- Short Bull RA, Cushman SA, Mace R *et al.* (2011) Why replication is important in landscape genetics: American black bear in the Rocky Mountains. *Molecular Ecology*, **20**, 1092–1107.
- Sork VL, Waits L (2010) Contributions of landscape genetics - approaches, insights, and future potential. *Molecular Ecology*, **19**, 3489–3495.
- van Strien MJ, Holderegger R, Van Heck HJ (2015) Isolation-by-distance in landscapes: considerations for landscape genetics. *Heredity*, **114**, 27–37.
- Tucker JM, Schwartz MK, Truex RL, Wisely SM, Allendorf FW (2014) Sampling affects the detection of genetic subdivision and conservation implications for fisher in the Sierra Nevada. *Conservation Genetics*, **15**, 123–136.
- Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, **4**, 535–538.
- Vekemans X, Hardy OJ (2004) New insights from fine-scale spatial genetic structure analyses in plant populations. *Molecular Ecology*, **13**, 921–935.
- Virgós E, Zalewski A, Rosalino L, Mergely M (2012) Habitat ecology of genus *Martes* in Europe: a review of the evidences. In: *Biology and Conservation of Martens, Sables, and Fishers. A New Synthesis* (eds Aubry KB, Zielinski WJ, Proulx G, Buskirk SW), pp. 255–266. Cornell University Press, New York.
- Wagner HH, Fortin M-J (2013) A conceptual framework for the spatial analysis of landscape genetic data. *Conservation Genetics*, **14**, 253–261.
- Wasserman TN, Cushman SA, Shirk AS, Landguth EL, Littell JS (2012) Simulating the effects of climate change on population connectivity of American marten (*Martes americana*) in the northern Rocky Mountains, USA. *Landscape Ecology*, **27**, 211–225.
- Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.
- Wright S (1943) Isolation by distance. *Genetics*, **28**, 114–138.
- Yang J, Cushman SA, Yang J, Yang M, Bao T (2013) Effects of climatic gradients on genetic differentiation of Caragana on the Ordos Plateau, China. *Landscape Ecology*, **28**, 1729–1741.
- Yang J, Cushman SA, Song X, Yang J, Zhang P (2015) Genetic diversity and drivers of genetic differentiation of Reaumuria soongorica of the Inner Mongolia plateau in China. *Plant Ecology*, **216**, 925–937.
- Zalewski A, Jędrzejewski W (2006) Spatial organisation and dynamics of the pine marten *Martes martes* population in Białowieża Forest (E Poland) compared with other European woodlands. *Ecography*, **29**, 31–43.

A.R.G. E.R. B.J.G.M. conceived and designed the experiments. A.R.G. performed the experiments. A.R.G. S.C. M.J.M. analysed the data. E.R. B.J.G.M. contributed reagents/materials/analysis tools. A.R.G. S.C. wrote the manuscript. A.R.G. S.C. M.J.M. E.R. B.J.G.M. revised the manuscript.

Data accessibility

Sample locations, microsatellite data and clustering assignment (complete genetic profiles at 15 microsatellite loci, geographical coordinates and Bayesian clustering assignment according to STRUCTURE and GENELAND for the 140 pine marten individuals), landscape resistance maps (raw ascii files of the optimal resistance models i.e. IBR_EN, IBR_LandFB_100 and IBR_LandFb_50), LCD matrices for IBD, IBR (IBR_EN, IBR_LandFB_100, IBR_LandFb_50) and IBB (IBB_STR and IBB_GEN) models and genetic distance (Rousset's a) matrix are accessible on Dryad Digital Repository, doi:10.5061/dryad.2k4b7.

Supporting information

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

Fig. S1 Support for hypothesis A relative to hypothesis B is determined using two partial Mantel tests. Test 1: $G \sim A|B$, produces correlation 1 in the diagram above. Test 2: $G \sim B|A$, produces correlation 2 in the diagram above. If hypothesis A is supported relative to hypothesis B, we would expect the value of 1–2 to be positive. In contrast, if hypothesis A is not supported relative to hypothesis B then we would expect the value of 1–2 to be negative.

Appendix S1 Full methodological details of noninvasive genetic sampling: mtDNA species identification and individual identification from microsatellite genotyping

Appendix S2 Methods for construction of IBD, IBR and IBB models

Appendix S3 Methods for reciprocal causal modelling including (partial) Mantel correlations between isolation by distance (IBD), landscape resistance (IBR) and clustering (IBB) hypotheses