

Spatial scaling and multi-model inference in landscape genetics: *Martes americana* in northern Idaho

Tzeidle N. Wasserman · Samuel A. Cushman ·
Michael K. Schwartz · David O. Wallin

Received: 2 March 2010 / Accepted: 16 August 2010 / Published online: 29 August 2010
© Springer Science+Business Media B.V. 2010

Abstract Individual-based analyses relating landscape structure to genetic distances across complex landscapes enable rigorous evaluation of multiple alternative hypotheses linking landscape structure to gene flow. We utilize two extensions to increase the rigor of the individual-based causal modeling approach to inferring relationships between landscape patterns and gene flow processes. First, we add a univariate scaling analysis to ensure that each landscape variable is represented in the functional form that represents the optimal scale of its association with gene flow. Second, we use a two-step form of the causal modeling approach to integrate model

selection with null hypothesis testing in individual-based landscape genetic analysis. This series of causal modeling indicated that gene flow in American marten in northern Idaho was primarily related to elevation, and that alternative hypotheses involving isolation by distance, geographical barriers, effects of canopy closure, roads, tree size class and an empirical habitat model were not supported. Gene flow in the Northern Idaho American marten population is therefore driven by a gradient of landscape resistance that is a function of elevation, with minimum resistance to gene flow at 1500 m.

Keywords Landscape genetics · Scale dependency · Causal modeling · American marten · Population connectivity · Gene flow

Electronic supplementary material The online version of this article (doi:[10.1007/s10980-010-9525-7](https://doi.org/10.1007/s10980-010-9525-7)) contains supplementary material, which is available to authorized users.

T. N. Wasserman (✉)
School of Forestry, Northern Arizona University,
P.O. Box 1508, Flagstaff, AZ 86011, USA
e-mail: tnw23@nau.edu

S. A. Cushman · M. K. Schwartz
USDA Forest Service, Rocky Mountain Research Station,
800 E Beckwith, Missoula, MT 59801, USA
e-mail: scushman@fs.fed.us

M. K. Schwartz
e-mail: mkschwartz@fs.fed.us

D. O. Wallin
Huxley College of the Environment, Western Washington
University, Bellingham, WA 98225, USA
e-mail: wallin@cc.wvu.edu

Introduction

Landscape genetics is a synthetic field combining spatial population genetics with landscape ecology. Landscape genetics explicitly quantifies the effects of landscape composition, configuration, and matrix quality on spatial patterns in neutral and adaptive genetic variation and underlying microevolutionary processes (Manel et al. 2003; Storfer et al. 2007; Holderegger and Wagner 2008; Balkenhol et al. 2009). Recent landscape genetic approaches largely focus on describing and mapping populations (e.g. Pritchard et al. 2000; Dupanloup et al. 2001; Francois

et al. 2006) and on identifying factors that influence rates and patterns of gene flow within and between populations (e.g. Coulon et al. 2004, 2006; Cushman et al. 2006; McRae and Beier 2007; Schwartz et al. 2009).

Many past landscape genetic studies have used relatively simple null-hypothesis testing, such as testing for the presence of a barrier, rather than comparing competing hypotheses involving more complex landscape effects (Holderegger and Wagner 2008; Balkenhol et al. 2009). This may lead to misinterpretations of the pattern-process relationship governing gene flow (Cushman and Landguth 2010a). Commonly, population and landscape genetic studies have used *F*-statistics (Wright 1943), assignment tests or Bayesian clustering (Pritchard et al. 2000; Corander et al. 2003; Francois et al. 2006) to relate genetic differences among predefined subpopulations, propose interpopulation distance relationships (Michels et al. 2001), identify putative movement barriers (Manni et al. 2004; Funk et al. 2005), or correlations with landscape features (Vitalis and Couvet 2001; Spear et al. 2005). Once discrete subpopulations have been identified, post hoc analyses are performed, correlating observed genetic patterns with interpopulation distance or putative movement barriers (e.g. Proctor et al. 2005).

Assuming the existence of discrete subpopulation structure, followed by application of methods designed to detect such structure, and by post-hoc analysis to identify potential causes for the inferred population structure is a perilous path of inference (Cushman and Landguth 2010a). Methods to delineate discrete populations are known to identify boundaries even in continuously structured populations (Schwartz and McKelvey 2009); observing a coincidence between a landscape feature and a putative subpopulation boundary does not confirm the role of that feature in creating spatial genetic structure.

Individual-based analyses relating landscape structure to genetic distance across complex landscapes enables rigorous evaluation of multiple alternative hypotheses relating landscape structure to gene flow. The predominant analytical approach to directly associate landscape patterns with gene flow processes is based on pair-wise calculation of cost distances, using least cost paths (e.g. Krist and Brown 1994; Walker and Craighead 1997) or multi-path circuit

approaches (McRae 2006). These pair-wise cost distances among individuals across a landscape resistance model are then correlated with pair-wise genetic distances among the same individuals with methods such as Mantel and partial Mantel tests (Mantel 1967; Smouse et al. 1986).

There has been discussion in the literature about the appropriateness of Mantel testing in landscape genetics. Raufaste and Rousset (2001) questioned the use of partial Mantel tests in micro-evolutionary studies. Subsequently, Castellano and Balletto (2002) attempted to rehabilitate the use of the partial Mantel test in genetic analysis. Recently, Legendre and Fortin (2010) have clarified the case. They showed that Raufaste and Rousset (2001) raised a valid point about a situation requiring a particular permutation procedure, but made unwarranted claims that partial Mantel tests are a biased testing procedure, while Castellano and Balletto (2002) attempted to refute this, but advocated an inappropriate testing procedure. Legendre and Fortin (2010) noted that distance-based regression approaches, such as the Mantel test, have lower power than traditional linear models and tend to underestimate the true magnitude of a relationship. They concluded that Mantel and partial Mantel testing is an appropriate framework when hypotheses are explicitly defined in terms of distances. Recently, Cushman and Landguth (2010a) evaluated the power of partial Mantel testing in a causal modeling framework (Legendre and Troussellier 1988; Legendre 1993) and found that the method performs well in identifying the drivers of genetic distances in complex landscapes and rejecting incorrect and/or correlated alternatives.

Identifying the effects of landscape structure on gene flow using neutral markers is explicitly a distance based problem. Because these markers are neutral, we would expect no association between the presence of a given allele and the environment at which the individual is located. Rather, differences among individuals in neutral markers reflect the degree of genetic isolation. Gene flow is mainly influenced by breeding and dispersal. For a given individual at a given location in a landscape, the probability of dispersing to and selecting a particular mate can be seen as a function of the distance to that mate in Euclidean or cost units (e.g. Landguth and Cushman 2010). Thus, analysis of gene flow in complex landscapes using neutral markers is

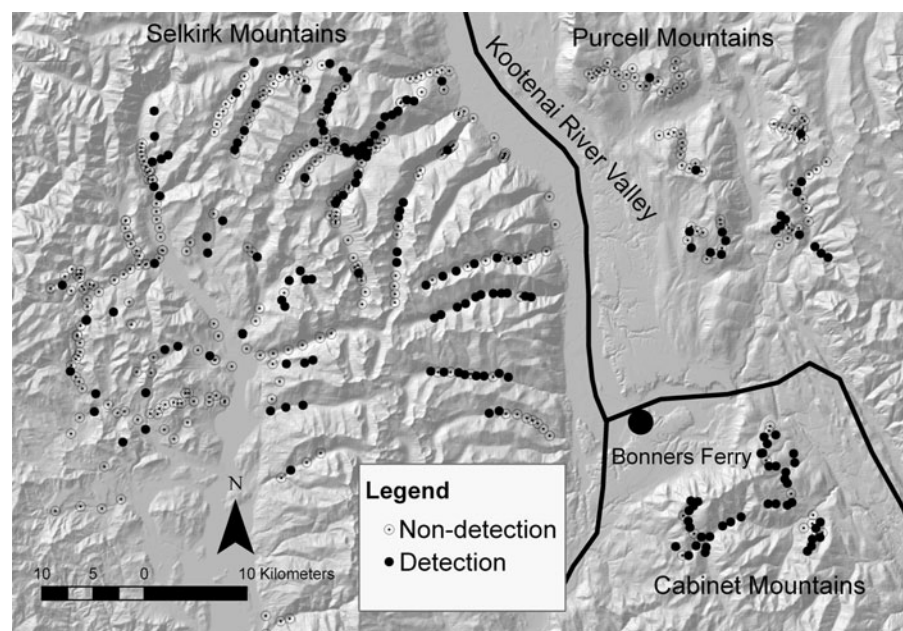
explicitly a distance based approach and the Mantel test is an appropriate method for analyzing such distance based problems.

The goal of this paper is to identify the processes governing gene flow in the American marten (*Martes americana*) in northern Idaho, USA (Fig. 1). The American marten is a habitat specialist that prefers mature and old growth forest types in the western United States (Buskirk and Powell 1994; Ruggiero et al. 1994). Marten populations are sensitive to forest fragmentation (Hargis 1996; Bissonette et al. 1997; Hargis et al. 1999) and depend on low-density winter snowpacks that provide protection from predators and access to subnivean hunting (Buskirk and Powell 1994). We utilize two extensions of the causal modeling approach in landscape genetics (Cushman et al. 2006). First, we add univariate scaling to ensure that each landscape variable is represented in the functional form that represents the optimal scale of its association with gene flow (Pérez-Espona et al. 2008; Wasserman 2008). This is important given the sensitivity of landscape genetic relationships to misspecification of the scale of landscape resistance models (Cushman and Landguth 2010b). Second, we use a two-step form of the causal modeling approach to integrate model selection with null hypothesis testing in individual-based landscape genetic analysis.

Methods

We used non-invasive hair snaring to collect genetic samples from *M. americana* in the Idaho Panhandle National Forest between 2003 and 2008 (Wasserman 2008). Identification of individual martens used nuclear DNA following methods in Riddle et al. (2003). DNA from one hair sample per hair-snare station was extracted using the Qiagen DNeasy Tissue kit (Qiagen Inc., Hilden, Germany) with modifications as outlined in Mills et al. (2001). Samples were genotyped at nine variable microsatellite loci according to Riddle et al. (2003) (Table S1). Deviations from Hardy–Weinberg proportions, heterozygote excess and deficiency were analyzed with GENEPOP 3.1d (Raymond and Rousset 1995). Genetic variability for each locus was estimated by calculating the mean number of alleles (A), observed heterozygosity (H_o), expected heterozygosity (H_e), and allelic richness with GENALEX 6 (Peakall and Smouse 2005). Probability of identity was calculated according to Evett and Weir (1998). In order to estimate gene flow between three mountain ranges in the study landscape of northern Idaho, USA (Selkirk, Purcell, and Cabinet mountain ranges); (Fig. 1), global F_{st} and F_{st} between pairs of mountain ranges were calculated with FSTAT 2.9.3 (Goudet 1995).

Fig. 1 Study area, with hair snare locations where marten were detected (*black solid circles*) and snares where marten were not detected (*open circles*). Putative barriers (valleys) are shown as *black lines*



An assignment test was conducted to evaluate population substructure related to mountain range (Selkirk, Purcell, or Cabinet) in GENALEX 6 (Peakall and Smouse 2005). Finally, a genetic distance matrix was calculated on all pairs of sampled martens using the Bray-Curtis percentage dissimilarity measure among individuals (Legendre and Legendre 1998).

Landscape resistance hypotheses

We developed landscape resistance surfaces based on four factors that have previously been reported to influence habitat selection by American martens (Taylor et al. 1993; Buskirk and Powell 1994; Phillips 1994; Bissonette et al. 1997; Hargis et al. 1999). These included elevation, roads, seral stage based on tree diameter at breast height (DBH; R1-VMAP; http://www.fs.fed.us/r1/gis/vmap_v06.html; Table S2), and percent canopy cover (1–100%; Fry et al. 2008). All of these base maps were re-sampled to 30 m pixel size.

Scaling landscape factors

Correctly representing the thematic resolution and form of the functional relationship between landscape features and genetic distance is essential to identifying pattern process relationships (e.g. Pérez-España et al. 2008; Cushman and Landguth 2010b). We produced a range of functionally scaled resistance surfaces for each of DBH, canopy cover, and elevation. DBH and percent canopy cover were evaluated over eight different levels that included linear and power functions of 0.2, 0.4, 0.6, 0.8, 2nd, 3rd, and 4th power. Landscape resistance predicted by these power functions ranged from strongly convex to strongly concave (Figures S1, S2). The power transformed percent canopy cover and DBH resistance maps were rescaled to 1–10 for analysis.

Elevation acted as a proxy for climate (e.g. snowpack) and vegetation composition. Landscape resistance due to elevation was modeled as a Gaussian function, on the expectation that martens have a unimodal optimum in movement ability with respect to elevation (Cushman et al. 2006). The form of the Gaussian function was defined on the basis of the optimum elevation and the standard deviation. The optimum elevation was assigned a minimum resistance of 1 and a maximum resistance of 10. We

evaluated marten response to a range of possible optimum elevations and a range of potential standard deviations of the Gaussian resistance function in order to find the combination of elevation and standard deviation to which marten genetic patterns are most strongly related. We tested nine levels of elevational optimum (1200–2000 m) and eight levels of standard deviation (300–1000 m) in 100 m increments, giving a factorial combination of 72 elevation/standard deviation models.

Two levels of roads were represented as categorical resistance functions. The first form of the roads layer included all maintained roads. The second included maintained roads plus abandoned, decommissioned or closed roads. Roads were classified as either major highways or other roads, including county and U.S. Forest Service roads that were both gated and un-gated. All non-road pixels were given a resistance value of 1, major highways a value of 10, and other roads a value of 5.

Univariate landscape-resistance modeling

For each scaled version of elevation, canopy closure, and DBH, we derived a matrix of least movement costs between all pairs of individuals using COST-DISTANCE in ARCGIS (ESRI 2003). We then compared the genetic distance among individuals with the cost of movement paths between them and identified the functions at which each factor had the strongest relationship with marten genetic structure. We ranked the scaled functions for each input landscape variable based on the partial Mantel test removing the effect of Euclidean distance. The function with the largest partial Mantel r value for each factor was chosen to be included in the first step causal modeling.

Two stage causal modeling

Once we identified which scale transformation of each variable was most supported, we extended the causal modeling framework (Legendre and Troussellier 1988; Legendre 1993; Cushman et al. 2006) to identify which of a number of alternative models was the best explanation of observed patterns of genetic distances across our study area. We proposed 16 alternative landscape hypotheses (Table 1). These included all combinations of the best scaled

Table 1 Definition of the 16 candidate models of landscape resistance in American marten evaluated in this study

Model acronym	Model definition
E	Gaussian function of elevation, mean 1500 m, 300 m standard deviation
D	Tree diameter at breast height to the 2nd power
R	Maintained roads
C	Canopy cover
CD	C + D
DR	D + R
CR	C + R
ED	E + D
ER	E + R
CD	C + D
ECR	E + C + R
ECD	E + C + D
ERD	E + R + D
CDR	C + D + R
ECRD	E + C + R + D
M	Habitat suitability model from Wasserman (2008)

Least cost distances were obtained from the Wasserman (2008) habitat suitability model by treating (1-probability of occurrence) as a resistance map and calculating cost distances as described in the text

univariate forms of elevation, DBH, canopy closure, and roads, as well as a habitat suitability model developed by the authors for this species in this study area using multi-scale logistic regression (Wasserman 2008). Habitat suitability is a frequently used surrogate for landscape resistance to movement (e.g. Coulon et al. 2004; Schwartz et al. 2009). Therefore, it is informative to compare its performance with a range of alternative landscape resistance models to test the hypothesis that landscape resistance is equivalent to habitat suitability.

The first step of the two-step causal modeling framework was to test each of these 16 alternative landscape resistance models against isolation by distance and isolation by barrier as described in Cushman et al. (2006). We identified three potential barriers to the movement of martens that potentially divided the regional population into three discrete, non-overlapping subpopulations. These barriers were (1) the Kootenai River trench between Bonners Ferry and the Canadian Border, (2) the Kootenai River valley upstream of Bonners Ferry to the Montana

Border, and (3) the Naples Valley south of Bonners Ferry to Sandpoint (Fig. 1). These three potential barriers separate the Selkirk, Purcell and Cabinet Mountains. We constructed a categorical model matrix predicting panmixia within each mountain range (Selkirk, Purcell, and Cabinet), but complete isolation between them (Legendre and Legendre 1998).

There were five sets of diagnostic Mantel and partial Mantel tests to complete this first step of the two-step causal modeling. These included: (1) simple Mantel tests between genetic distance and landscape resistances; (2) partial Mantel tests between genetic distance and landscape cost distances, partialling out the effects of Euclidean distance; (3) partial Mantel tests between genetic distance and landscape cost distance partialling out the effects of the barrier model; (4) partial Mantel tests between genetic distances and Euclidean distance, partialling out the effects of landscape resistance; (5) partial Mantel tests between genetic distance and the barrier model, partialling out the effects of landscape resistances. For a landscape resistance model to be supported in the first step of the two step causal modeling, tests (1)–(3) above had to be significant, and tests (4) and (5) had to be non-significant. All (partial) Mantel tests were conducted in ECODIST v1.1.3 (Goslee and Urban 2007) in R with 10,000 permutations (R Development Core Team 2007).

The second step of the two-step causal modeling approach tested whether the model that ranked highest in support across the five diagnostic tests in the first step was significantly more supported than the other models also supported in the first step and whether any of these other models had any significant residual explanatory ability after accounting for the effects of the most supported, top model. To accomplish this we conducted two sets of additional diagnostic tests: (1) partial Mantel tests between genetic distance and the top model, partialling out the effect of each alternative model in turn; (2) partial Mantel tests between genetic distance and each alternative model in turn, partialling out the effect of the top model. For the top model to be affirmed as the only hypothesis supported, all of the tests in (1) had to be significant and none of the tests in (2) had to be significant. This would show that this top model has statistically significant ability to predict genetic distances after removing the effects of all competing

alternative models, and that none of the competing alternative models have significant ability to predict genetic distances after removing the effects of the top model.

Results

Genetic results

We detected marten at 322 of 622 stations surveyed. Genotypes for 118 stations were obtained, and 117 unique marten genotypes were identified. Within nine variable microsatellite loci, the number of alleles ranged from four to ten alleles per locus, with a total of 57 alleles. Observed heterozygosity was less than expected heterozygosity in six out of nine loci (Table S1). However, tests of departure from Hardy–Weinberg equilibrium by mountain range failed to identify any locus that consistently departed from expectation (GENALEX Peakall and Smouse 2005). Pairwise F_{st} was 0.004 between Purcell and Cabinet mountain ranges, 0.036 between Purcell and Selkirk mountain ranges, and 0.037 between Cabinet and Selkirk mountain ranges. This indicated little differentiation between the three putative populations separated by valleys. Assignment tests showed poor assignments to origin mountain ranges, with a global misassignment rate of 21%. Selkirk martens were correctly assigned to their mountain range 89% of the time, while 55% of Purcell martens and 48% of Cabinet martens were misassigned. Genetic distance was significantly correlated with Euclidean distance ($r = 0.1681$, $P = 0.0001$) and the barrier model ($r = 0.1814$, $P = 0.0001$).

Scaled univariate landscape resistance models

Elevation 1500 m with a standard deviation of 300 m was the best supported elevation model ($r = 0.2058$, $P = 0.0001$; Figure S3). The highest ranked of eight scale functions canopy closure was landscape resistance as the canopy closure to the second power (Mantel $r = 0.1839$, $P = 0.0001$; Figure S4). A linear function of DBH was identified as the most supported function (Mantel $r = 0.2034$, $P = 0.0001$; Figure S5). The highest ranked road resistance hypothesis represented was landscape resistance as a function of

crossing maintained roads (Mantel $r = 0.1699$, $P = 0.0001$; Figure S6).

First step causal modeling

Eight of the 16 alternative scaled models were supported and eight rejected in the first step of causal modeling (Table 2). These eight supported models had statistically significant ability to explain genetic distance among individual martens after partialling out the effects of Euclidean distance and the barrier model. Conversely, Euclidean distance and the barrier model had no independent explanatory ability after partialling out the effects of each of the eight supported models. Elevation was the best supported of the eight supported models based on Mantel r and P -value. Elevation was therefore chosen as the top model for the second step of causal modeling.

Second step causal modeling

The second step causal modeling analysis tested whether the top model, (Elevation 1500_300) had significant ability to explain genetic distances among individual martens after partialling out the effects of the other seven models supported in the first causal modeling step. The results indicated that elevation was significantly related to genetic distance after removing the effects of all seven of the alternative models, except for the model Elevation + DBH (Table 3). It is not surprising that this latter test was not significant ($P = 0.102$) as the Elevation + DBH model was highly correlated with the top model. None of the seven alternative models were significant after partialling out the elevation effect, and all of their mantel r values were negative. This indicated that after removing the effect of elevation, none of these alternative models provided significant explanation of genetic distances. This series of causal modeling thus confirmed that Elevation1500_300 was the model that best predicts gene flow in the American marten in northern Idaho (Fig. 2), and that alternative hypotheses involving isolation by distance, geographical barriers, canopy closure, roads, DBH and an empirical habitat model are not supported.

Table 2 Results of first step of causal modeling of landscape resistance on genetic distance in American marten

Model	$G \sim L$, R	$G \sim L$, p	$G \sim LiD$, R	$G \sim LiD$, p	$G \sim LiB$, R	$G \sim LiB$, R	$G \sim DiL$, R	$G \sim DiL$, p	$G \sim BiL$, R	$G \sim BiL$, p	Causal model supported
E	0.207	<i>0.0001</i>	0.131	<i>0.0006</i>	0.101	<i>0.0180</i>	−0.048	<i>0.9110</i>	−0.010	<i>0.6083</i>	Y
D	0.190	<i>0.0001</i>	0.102	<i>0.0067</i>	0.093	<i>0.0240</i>	−0.049	<i>0.8990</i>	0.039	<i>0.2031</i>	Y
CD	0.188	<i>0.0001</i>	0.102	<i>0.0080</i>	0.067	<i>0.0460</i>	−0.056	<i>0.9227</i>	0.045	<i>0.1451</i>	Y
C	0.184	<i>0.0001</i>	0.089	<i>0.0198</i>	0.065	<i>0.0405</i>	−0.047	<i>0.8733</i>	0.057	<i>0.0915</i>	Y
CDR	0.181	<i>0.0001</i>	0.076	0.0656	0.070	<i>0.0469</i>	−0.041	<i>0.8429</i>	0.059	<i>0.1079</i>	N
DR	0.180	<i>0.0001</i>	0.069	0.0886	0.072	<i>0.0500</i>	−0.038	<i>0.8276</i>	0.061	<i>0.1161</i>	N
ECD	0.180	<i>0.0001</i>	0.132	<i>0.0009</i>	0.086	<i>0.0220</i>	−0.076	<i>0.9780</i>	0.064	<i>0.0685</i>	Y
H	0.179	<i>0.0001</i>	0.076	<i>0.0294</i>	0.054	0.0595	−0.042	<i>0.8527</i>	0.060	<i>0.0950</i>	N
ECRD	0.178	<i>0.0001</i>	0.125	<i>0.0012</i>	0.079	<i>0.0330</i>	−0.075	<i>0.9764</i>	0.066	<i>0.0725</i>	Y
CR	0.178	<i>0.0001</i>	0.074	<i>0.0342</i>	0.069	0.0580	−0.045	<i>0.8768</i>	0.066	<i>0.0720</i>	N
ED	0.177	<i>0.0001</i>	0.133	<i>0.0009</i>	0.092	<i>0.0215</i>	−0.065	<i>0.8784</i>	0.069	<i>0.0650</i>	Y
ERD	0.176	<i>0.0001</i>	0.127	<i>0.0015</i>	0.082	<i>0.0265</i>	−0.065	<i>0.8869</i>	0.070	<i>0.0675</i>	Y
CE	0.173	<i>0.0001</i>	0.136	<i>0.0007</i>	0.056	0.0740	−0.074	<i>0.9767</i>	0.076	<i>0.0505</i>	N
ECR	0.173	<i>0.0001</i>	0.130	<i>0.0008</i>	0.083	<i>0.0260</i>	−0.077	<i>0.9801</i>	0.076	0.0350	N
R	0.170	<i>0.0001</i>	0.036	0.1704	0.051	0.0760	−0.027	<i>0.7563</i>	0.082	0.0290	N
ER	0.170	<i>0.0001</i>	0.128	<i>0.0006</i>	0.089	<i>0.0290</i>	−0.065	<i>0.8869</i>	0.084	0.0310	N

Model definitions according to Table 1. There are five Mantel tests comprising the first-step causal modeling: (1) $G \sim L$ —simple Mantel test between the candidate model and genetic distance; (2) $G \sim LiD$ —partial Mantel test between the candidate model and genetic distance, partialling out Euclidean distance; (3) $G \sim LiB$ —partial Mantel test between the candidate model and genetic distance, partialling out the barrier hypothesis; (4) $G \sim DiL$ —partial Mantel test between the Euclidean and genetic distance, partialling out the candidate model; (5) $G \sim BiL$ —partial Mantel test between the barrier hypothesis and genetic distance, partialling out the candidate model. For a candidate model to be supported tests (1), (2) and (3) must be significant, while tests (4) and (5) must be non-significant. Mantel tests meeting each criterion are italicized

Table 3 Results of second step causal modeling of landscape resistance effects on genetic distance in American marten

Model	$G \sim El$, r	$G \sim El$, p	$G \sim lE$, r	$G \sim lE$, p
ECRD	0.0820	0.0182	−0.0431	0.8747
CD	0.0958	0.0142	−0.0387	0.8284
C	0.1015	0.0090	−0.0335	0.8107
ED	0.0488	<i>0.1020</i>	−0.0247	0.7491
ECD	0.0631	0.0451	−0.0287	0.7720
D	0.0931	0.0124	−0.0412	0.8450
ERD	0.0780	0.0216	−0.0465	0.9065

The first column indicates the candidate models (Table 1). (1) $G \sim El$ —partial Mantel test between genetic distance and the elevation model, partialling out the candidate model; (2) $G \sim lE$ —partial Mantel test between genetic distance and the candidate model, partialling out the elevation model. For a candidate model to pass the second step of causal modeling test (1) must be significant, and test (2) must be non-significant. Only one candidate model passes test (ED), and none pass test (2). This indicates that none of the seven candidate models is supported independently of the elevation model



Fig. 2 Landscape resistance to gene flow of American marten in northern Idaho is a Gaussian function of elevation (model Elevation 1500_300), i.e. resistance to gene flow for marten in this study area is minimum at 1500 m in elevation and increases as a Gaussian function with a 300 m standard deviation. Resistance ranges from 1 (black) at 1500 m, to 10 (white)

Discussion

Two-step causal modeling and inference in landscape genetics

The results of F_{st} and assignment tests suggested low to moderate differentiation in the studied marten population. However, these approaches have lower power to detect landscape effects than Mantel r (Landguth et al., in press) and assume division into internally panmictic subpopulations (Balkenhol et al. 2009). In contrast, individual-based landscape resistance approaches take no such assumption and provide flexibility to test a range of alternative hypotheses to identify the drivers of gene flow and reject incorrect alternatives (Cushman and Landguth 2010a).

We implemented two refinements of the causal modeling approach that improved our ability to infer the causes driving gene flow. First, a two-step causal modeling approach improved a power of the method. The first step identified a set of alternative models that explained gene flow better than either isolation by distance or isolation by barriers. Thus, there was no independent support for isolation by distance or isolation by large agricultural valleys in the studied marten population. The second step causal modeling allowed us to determine if one of the eight remaining models could be affirmed as the driving factor and if the other seven contending models could be rejected. This second step demonstrated that landscape resistance to gene flow in American marten was mainly affected by elevation, with optimum gene flow 1500 m.

Our causal modeling approach enabled us to definitively reject the isolation by distance and barrier hypotheses and confirmed that the elevation model was strongly and independently supported compared to alternative landscape resistance hypotheses. However, if we had just tested models of geographic distance or isolation by barrier, we would have erroneously concluded that a model of isolation by barriers was the primary driver of genetic structure in *M. americana* in northern Idaho. Many past studies have stopped when a model of isolation by distance or barrier was supported, and may therefore often have missed the true drivers of genetic structure (Holderegger and Wagner 2008; Balkenhol et al. 2009; Cushman and Landguth 2010a).

A surprising number of landscape genetic studies have evaluated a single landscape resistance

hypothesis relative to a null model of isolation by distance or isolation by barriers (e.g. Coulon et al. 2004; Broquet et al. 2006; Schwartz et al. 2009). This has been shown to produce high risk of inferential error due to spurious correlations with alternative but untested hypotheses (Cushman and Landguth 2010a). The results of this study reaffirm this risk. All of the alternative landscape resistance hypotheses were highly significantly supported relative to the null hypotheses of isolation by distance and isolation by barriers. A priori selection of a single landscape resistance model would have yielded highly significant support, but would have lead to inferential error, as only one of the 16 alternative models tested was independently supported. The observation of high correlation between multiple alternative resistance models in other species (e.g. Cushman et al. 2006; Shirk et al. 2010) suggests that reliable inference of the factors driving gene flow requires formal evaluation of support among a set of realistic candidate models, while evaluating support for a single landscape resistance model in isolation provides a weak basis for inference.

Importance of scaling landscape resistance

Scale-dependence has rarely been evaluated in landscape genetics (but see Pérez-Espona et al. 2008; Wasserman 2008; Cushman and Landguth 2010b; Shirk et al. 2010). It is important to correctly match the scale of each driving variable to the response process (Wiens 1989). Mismatches in scale may result in failure to observe a relationship between pattern and process when one exists, the observation of a spurious or distorted relationship, or error in evaluation of the significance or effects size of a pattern-process relationship (e.g. Thompson and McGarigal 2002). Recently, Cushman and Landguth (2010b) used an individual-based, spatially explicit simulation model to explore scale dependencies in landscape genetic inference. Their results indicate little sensitivity in individual-based landscape genetic inference to variation in the grain of input landscape variables. In contrast, they demonstrated extremely high sensitivity to misspecification of landscape resistances. The same pattern of sensitivity to specification of landscape resistance has also been observed in empirical studies (Pérez-Espona et al. 2008; Shirk et al. 2010). Therefore, correct representation of the scale and

functional form of each resistance hypothesis is critical for reliable inference in landscape genetics, which was the primary motivation behind the scaling approach used here.

The scaling conducted in this paper is focused on empirically optimizing the landscape resistance of input variables comprising alternative landscape resistance hypotheses (Pérez-Espona et al. 2008; Wasserman 2008). Our results provide an empirical verification of the sensitivity to landscape definition as shown in the simulation analysis of Cushman and Landguth (2010b). Specifically, our results demonstrate that incorrect specification of a resistance function may result in error in attributing the strength and nature of the relationships between landscape composition, resistance to gene flow and genetic distances among individuals across landscapes. For example, this study identified elevation as the dominant driver of gene flow. However, an incorrect choice of the functional form of the elevation model might have lead to the incorrect conclusion that elevation was not an important predictor of gene flow. Indeed, choice of the elevation model with mean 1200 m and standard deviation 400 m instead of the optimally scaled model (mean 1500, 300 m std) would result in rejection of the elevation model in the first step of causal modeling (Figure S3). This shows that optimizing the scaling of resistance models may often be essential for reliable inference in individual-based landscape genetics.

Insights into marten ecology and population structure

Some marten populations separated by distances of several hundred kilometers appear genetically undifferentiated (Kyle and Strobeck 2003). For example, Broquet et al. (2006) found that *M. americana* in northwestern Ontario had very weak correlation between genetic distance and geographic or ecological distance across large spatial extents, with Mantel *r*-values nearly an order of magnitude lower than those of this study. However, the above studies are from boreal landscapes where elevation, climate and major vegetation types are constant across large extents. Failure to detect correlations between landscape variables and genetic distances in landscapes in which the driving variables are uniform and unvarying does not demonstrate that the species does not

rely on suitable landscape conditions to maintain population connectivity. Relationships between landscape variables and gene flow will only be detected if the pattern of those variables on the landscape limits gene flow. In our study area the highly varying topography led to strong spatial variation in landscape resistance that was highly correlated with marten genetic distances.

In addition, our results suggest that habitat suitability may not be a reliable proxy for predicting landscape effects on gene flow. Wasserman (2008) found that marten occurrence in the study area is highly dependent on forest fragmentation and road density at broad spatial scales and presence of old-growth forest at fine spatial scales. This sensitivity to forest fragmentation and occurrence of late seral forest is consistent with other studies of American marten habitat associations (Buskirk and Powell 1994; Ruggiero et al. 1994; Hargis 1996; Bissonette et al. 1997; Chapin et al. 1998; Hargis et al. 1999). However, genetic distances were not independently related to any of these factors, indicating that habitat selection and gene flow of American marten may be driven by different factors at different scales. This may not be surprising, as habitat selection reflects the behavior of individual organisms to maximize fitness within home ranges, while gene flow is driven by mating and dispersal events. This highlights the importance of not assuming that a known habitat relationship optimally reflects the landscape features governing gene flow.

Finding that gene flow of American marten in northern Idaho is predominately driven by movement resistance along elevational gradients has several important implications for understanding the biology of this species and enabling prediction of population connectivity under current and potential future landscape change scenarios (e.g. Cushman et al. 2008; Schwartz et al. 2009). In our study area, elevation is closely associated with snowpack and gradients of forest vegetation composition and structure. Precipitation ranges from a mean of more than 1778 mm on the highest peaks to less than 762.0 mm within the rain shadow of the Selkirk Mountains, and average winter snow depth ranges from over 300 cm above 1500 m to less than 40 cm below 1200 m. Deep persistent snow pack is a critical habitat element for American marten. It excludes predators (e.g. *Canis latrans*), and provides high-quality hunting conditions via subnivean space (Buskirk and Powell 1994). In

our study area Subalpine fir (*Abies lasiocarpa*) and Engelmann spruce (*Picea engelmannii*) are codominant above 1500 m, and a diverse mixed forest of Western Larch (*Larix occidentalis*), Western Red cedar (*Thuja plicata*), and other conifer dominates between 1300 and 1500 m (Evans and Cushman 2009). These forest types have been found to be strong predictors of occurrence of marten in the study area (Wasserman 2008). Thus our observed relationship between elevation and gene flow probably reflects a combination of the effects of snowpack and vegetation gradients on dispersal.

The above results suggest that population connectivity of American marten in the northern Rocky Mountains may be vulnerable to climate change. The optimal resistance model identified in this study is closely related to areas of high winter snowpack and moist montane forest. Climate change is predicted to result in large increases in winter temperature in the northern Rocky Mountains (IPCC 2007). This is likely to result in substantial decrease in the depth of average winter snowpacks and migration of forest communities towards higher elevation. In consequence, this would greatly increase the average resistance of the landscape to martens and might result in increased isolation of remnant populations in high elevation mountains separated by the deep river valleys.

Acknowledgments This research was primarily supported by the U.S. Forest Service Rocky Mountain Research Station, the Idaho Department of Fish and Game, and Western Washington University, Huxley College of the Environment. We especially thank Jim Hayden of Idaho Fish and Game for his support and the RMRS Wildlife Genetics Lab in Missoula, MT. We also thank the two anonymous reviewers and Rolf Holderegger for their helpful insights and comments on earlier drafts of this manuscript.

References

- Balkenhol N, Gugerli F, Cushman SA, Waits LP, Coulon A, Arntzen JW, Holderegger R, Wagner HH, Arens P, Campagne P, Dale VH, Nicieza AG, Smulders MJM, Tedesco E, Wang H, Wasserman TN (2009) Identifying future research needs in landscape genetics: where to from here? *Landscape Ecol* 24:455–463
- Bissonette JA, Harrison DJ, Hargis CD, Chapin TG (1997) The influence of spatial scale and scale-sensitive properties on habitat selection by American marten. In: Bissonette JA (ed) *Wildlife and landscape ecology*. Springer, New York, pp 368–385
- Broquet T, Ray N, Petit E, Fryxell JM, Burel F (2006) Genetic isolation by distance and landscape connectivity in the American marten (*Martes americana*). *Landscape Ecol* 21:877–889
- Buskirk SW, Powell RA (1994) Habitat ecology of fishers and American martens. In: Buskirk SW, Harestad AS, Raphael MG, Powell RA (eds) *Martens, sables, and fishers*. Cornell University Press, Ithaca, pp 283–296
- Castellano S, Balletto E (2002) Is the partial Mantel test inadequate? *Evolution* 56:1871–1873
- Chapin TG, Harrison DJ, Katnik DD (1998) Influence of landscape pattern on habitat use by American marten in an industrial forest. *Conserv Biol* 12:96–227
- Corander J, Waldmann P, Sillanpää MJ (2003) Bayesian analysis of genetic differentiation between populations. *Genetics* 163:367–374
- Coulon A, Cosson JF, Angibault JM, Cargnelutti B, Galan M, Morellet N, Petit E, Aulagnier S, Hewson AJM (2004) Landscape connectivity influences gene flow in a roe deer population inhabiting a fragmented landscape: an individual-based approach. *Mol Ecol* 13:2841–2850
- Coulon A, Guillot G, Cosson GF, Angibault JMA, Aulagnier S, Cargnelutti B, Galan M, Hewison AJM (2006) Genetic structure is influenced by landscape features: empirical evidence from a roe deer population. *Mol Ecol* 15: 1669–1679
- Cushman SA, Landguth EL (2010a) Spurious correlations and inference in landscape genetics. *Mol Ecol* 19:3592–3602
- Cushman SA, Landguth EL (2010b) Scale dependent inference in landscape genetics. *Landscape Ecol* 25:967–979
- Cushman SA, McKelvey KS, Hayden J, Schwartz MK (2006) Gene-flow in complex landscapes: testing multiple models with causal modeling. *Am Nat* 168:486–499
- Cushman SA, McKelvey KS, Schwartz MK (2008) Using empirically derived source-destination models to map regional conservation corridors. *Conserv Biol* 23:368–376
- Dupanloup I, Schneider S, Excoffier L (2001) A simulated annealing approach to define genetic structure of populations. *Mol Ecol* 58:2021–2036
- ESRI (2003) *ARCGIS*. Environmental Systems Research Incorporated, Redlands
- Evans JS, Cushman SA (2009) Gradient modeling of conifer species using random forests. *Landscape Ecol* 24:673–683
- Eveitt IW, Weir BS (1998) *Interpreting DNA evidence*. Sinauer, Sunderland
- Francois O, Ancelet S, Guillot G (2006) Bayesian clustering using hidden Markov random fields in spatial population genetics. *Genetics* 174:805–816
- Fry JA, Coan MJ, Homer CG, Meyer DK, Wickham JD (2008) Completion of the national land cover database (NLCD) 1992–2001 land cover change retrofit product. *USGS OF* 2008–1379
- Funk CW, Blouin MS, Corn PS, Maxell BA, Pilliod DS, Amish S, Allendorf FW (2005) Population structure of Columbia spotted frogs (*Rana luteiventris*) is strongly affected by the landscape. *Mol Ecol* 14:483–496
- Goslee SC, Urban DL (2007) The ecodist package for dissimilarity-based analysis of ecological data. *J Stat Softw* 22(7):1–19
- Goudet J (1995) *FSTAT* (version 1.2): a computer program to calculate F-statistics. *J Hered* 86:485–486

- Hargis CD (1996) The influence of forest fragmentation and landscape pattern on American marten and their prey. PhD dissertation, Utah State University, Logan, Utah
- Hargis CD, Bissonette JA, Turner DL (1999) The influence of forest fragmentation and landscape pattern on American martens. *J Appl Ecol* 36:157–172
- Holderegger R, Wagner HH (2008) Landscape genetics. *Bio-science* 58:199–207
- IPCC (2007) Climate change 2007: the physical science basis. Contribution of working group I to the fourth assessment report of the intergovernmental panel on climate change. Cambridge University Press, Cambridge
- Krist FJ, Brown DG (1994) GIS modeling of paleo-indian period caribou migrations and viewsheds in northeastern lower Michigan. *Photogramm Eng Remote Sensing* 60: 1129–1137
- Kyle CJ, Strobeck C (2003) Genetic homogeneity of Canadian mainland marten populations underscores the distinctiveness of Newfoundland pine martens (*Martes americana atrata*). *Can J Zool* 81:57–66
- Landguth EL, Cushman SA (2010) CDPOP: an individual-based, cost-distance spatial population genetics model. *Mol Ecol Resour* 10:156–161
- Landguth EL, Cushman SA, Schwartz MK, McKelvey KS, Murphy M, Luikart G (in press) Quantifying the lag time to detect barriers in landscape genetics. *Mol Ecol*
- Legendre P (1993) Spatial autocorrelation: trouble or new paradigm? *Ecology* 74:1659–1673
- Legendre P, Fortin M-J (2010) Comparison of the Mantel test and alternative approaches for detecting complex multivariate relationships in the spatial analysis of genetic data. *Mol Ecol Resour* 10:831–844
- Legendre P, Legendre L (1998) Numerical ecology. Elsevier, Amsterdam
- Legendre P, Troussellier M (1988) Aquatic heterotrophic bacteria: modeling in the presence of spatial autocorrelation. *Limnol Oceanogr* 33:1055–1067
- Manel S, Schwartz MK, Luikart G, Taberlet P (2003) Landscape genetics: combining landscape ecology and population genetics. *Trends Ecol Evol* 18:189–197
- Manni F, Guerard E, Heyer E (2004) Geographic patterns of (genetic, morphologic, linguistic) variation: how barriers can be detected by using Monmonier's algorithm. *Hum Biol* 76:173–190
- Mantel N (1967) The detection of disease clustering and a generalized regression approach. *Cancer Res* 27: 209–220
- McRae BH (2006) Isolation by resistance. *Evolution* 60: 1551–1561
- McRae BH, Beier P (2007) Circuit theory predicts gene flow in plant and animal populations. *Proc Natl Acad Sci USA* 104:19885–19890
- Michels E, Cottenie K, Neys L, DeGalas K, Coppin P, DeMeester L (2001) Geographical and genetic distances among zooplankton populations in a set of interconnected ponds: a plea for using GIS modeling of the effective geographical distance. *Mol Ecol* 10:1929–1938
- Mills LS, Pilgrim K, Schwartz MK, McKelvey K (2001) Identifying lynx and other North American felids based on MtDNA analysis. *Conserv Genet* 1:285–289
- Peakall R, Smouse PE (2005) GENALEX 6: genetic analysis in EXCEL: population genetic software for teaching and research. *Mol Ecol Notes* 6:288–295
- Pérez-Espona S, Pérez-Barbería FJ, McLeod JE, Jiggins CD, Gordon IJ, Pemberton JM (2008) Landscape features affect gene flow of Scottish Highland red deer (*Cervus elaphus*). *Mol Ecol* 17:981–996
- Phillips DM (1994) Social and spatial characteristics, and dispersal of marten in a forest preserve and industrial forest. M.S. thesis, University of Maine, Orono, USA
- Pritchard JK, Stephens M, Peter D (2000) Inference of population structure using multilocus genotype data. *Genetics* 155:945–959
- Proctor MF, McLellan BN, Strobeck C, Barclay RMR (2005) Genetic analysis reveals demographic fragmentation of grizzly bears yielding vulnerability by small populations. *Proc R Soc B Biol Sci* 272:2409–2416
- R Development Core Team (2007) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, <http://www.R-project.org>
- Raufaste N, Rousset F (2001) Are partial Mantel tests adequate? *Evolution* 55:1703–1705
- Raymond M, Rousset F (1995) GENEPOP (version 1.2), population genetics software for exact tests and ecumenicism. *J Hered* 86:248–249
- Riddle A, Pilgrim KL, Mills LS, McKelvey KS, Ruggiero LF (2003) Identification of mustelids using mitochondrial DNA and non-invasive sampling. *Conserv Genet* 4:241–243
- Ruggiero LF, Aubrey KB, Buskirk J, Lyona ND, Zielinski WJ (1994) The scientific basis for conserving forest carnivores: American marten, fisher lynx, and wolverine in the western United States. U.S. Forest Service General Technical Report RM-254
- Schwartz MK, McKelvey KS (2009) Why sampling scheme matters: the effect of sampling scheme on landscape genetic results. *Conserv Genet* 10:441–452
- Schwartz MK, Copeland JP, Anderson NJ, Squires JR, Inman RM, McKelvey KS, Pilgrim KL, Waits LP, Cushman SA (2009) Wolverine gene flow across a narrow climatic niche. *Ecology* 90:3222–3232
- Shirk A, Wallin DO, Cushman SA, Rice C, Warheit K (2010) Inferring landscape effects on gene flow: a new multi-scale model selection framework. *Mol Ecol* 19:3489–3495
- Smouse PE, Long JC, Sokal RR (1986) Multiple regression and correlation extensions of the Mantel test of matrix correspondence. *Syst Zool* 35:627–632
- Spear SF, Peterson CR, Maticq M, Storfer A (2005) Landscape genetics of the blotched tiger salamander (*Ambystoma tigrinum melanostictum*). *Mol Ecol* 14:2553–2564
- Storfer A, Murphy MA, Evans JS, Goldberg CS, Robinson S, Spear SF, Dezzani R, Delmelle E, Vierling L, Waits LP (2007) Putting the 'landscape' in landscape genetics. *Heredity* 98:128–142
- Taylor PD, Fahrig L, Henein K, Merriam G (1993) Connectivity is a vital element of landscape structure. *Oikos* 68:571–573
- Thompson CM, McGarigal K (2002) The influence of research scale on bald eagle habitat selection along the lower Hudson River, New York (USA). *Landscape Ecol* 17: 569–586

- Vitalis R, Couvet D (2001) Estimation of effective population size and migration rate from one- and two-locus identity measures. *Genetics* 157:911–925
- Walker W, Craighead FL (1997) Analyzing wildlife movement corridors in Montana using GIS. In: Proceedings of the 1997 ESRI user conference
- Wasserman TN (2008) Habitat relationships and landscape genetics of *Martes americana* in northern Idaho. M.S. thesis, Western Washington University, Bellingham
- Wiens JA (1989) Spatial scaling in ecology. *Funct Ecol* 3: 385–397
- Wright S (1943) Isolation by distance. *Genetics* 28:114–138