MOLECULAR ECOLOGY

Molecular Ecology (2011) **20**, 1092–1107 doi: 10.1111/j.1365-294X.2010.04944.x

Why replication is important in landscape genetics: American black bear in the Rocky Mountains

R. A. SHORT BULL,* S. A. CUSHMAN,† R. MACE,‡ T. CHILTON,‡ K. C. KENDALL,§ E. L. LANDGUTH,*† M.K. SCHWARTZ,† K. MCKELVEY,† FRED W. ALLENDORF* and G. LUIKART*¶**

**Division of Biological Sciences, University of Montana, Missoula, MT 59812, USA; †USDA Forest Service Rocky Mountain Research Station, 800 East Beckwith, Missoula, MT 59801, USA; ‡Montana Department of Fish, Wildlife and Parks, 490 North Meridian Road, Kalispell, MT 59901, USA, §U.S. Geological Survey-Northern Rocky Mountain Science Center, Glacier Field Station, Glacier National Park, West Glacier, MT 59936, USA, ¶CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, Universidade do Porto, Campus Agrário de Vairão, 4485-661 Vairão, Portugal, **Flathead Lake Biological Station, University of Montana, Polson, MT, 59860, USA*

Abstract

We investigated how landscape features influence gene flow of black bears by testing the relative support for 36 alternative landscape resistance hypotheses, including isolation by distance (IBD) in each of 12 study areas in the north central U.S. Rocky Mountains. The study areas all contained the same basic elements, but differed in extent of forest fragmentation, altitude, variation in elevation and road coverage. In all but one of the study areas, isolation by landscape resistance was more supported than IBD suggesting gene flow is likely influenced by elevation, forest cover, and roads. However, the landscape features influencing gene flow varied among study areas. Using subsets of loci usually gave models with the very similar landscape features influencing gene flow as with all loci, suggesting the landscape features influencing gene flow were correctly identified. To test if the cause of the variability of supported landscape features in study areas resulted from landscape differences among study areas, we conducted a limiting factor analysis. We found that features were supported in landscape models only when the features were highly variable. This is perhaps not surprising but suggests an important cautionary note – that if landscape features are not found to influence gene flow, researchers should not automatically conclude that the features are unimportant to the species' movement and gene flow. Failure to investigate multiple study areas that have a range of variability in landscape features could cause misleading inferences about which landscape features generally limit gene flow. This could lead to potentially erroneous identification of corridors and barriers if models are transferred between areas with different landscape characteristics.

Keywords: connectivity, gene flow, habitat fragmentation, landscape genetics, landscape resistance modelling, noninvasive sampling, *Ursus americanus*

Received 12 January 2010; revision received 3 September 2010; accepted 2 November 2010

Introduction

Habitat connectivity and fragmentation are landscape-level processes that affect population structure, dynamics, and evolution (Fischer & Lindenmayer 2007). Investigations of landscape-scale processes are often carried out on relatively large spatial scales, and require large amounts of time and money. As a result, research rarely assesses fragmentation patterns and processes at the landscape-level in replicated studies (Stutchbury 2007). For example, in a review of several hundred studies of habitat fragmentation, McGarigal & Cushman (2002) found that less than 5% of published papers

Correspondence: Ruth A. Short Bull,
E-mail: ruth.shortbull@umontana.edu

© 2011 Blackwell Publishing Ltd

reported any sort of replication of representative landscapes and very few had sufficient sample size for statistical inference at the landscape-level. Replication in landscape ecology can be defined as spatial replication or multiple spatial comparisons (Segelbacher *et al.* 2010).

Landscape genetics is a recently developed research approach that combines landscape ecology and population genetics for testing the relative influence of different landscape features on genetic population structure and gene flow (Manel *et al.* 2003; Storfer *et al.* 2007). The genetic characteristics of individuals sampled across landscapes allows identification of populations, localization of genetic discontinuities (barriers or contact zones), and quantification of the relative influence of different landscape features on gene flow. Landscape genetics has also been used in identifying and evaluating connectivity and corridors (Dixon *et al.* 2006; Epps *et al.* 2007; Cushman *et al.* 2008; Schwartz *et al.* 2009; Li *et al.* 2010). Holderegger & Wagner (2008) concluded landscape genetics can potentially infer functional connectivity at spatial scales and for species that other ecological techniques such as radio tracking, global positioning system technology, and mark–recapture in animals cannot.

Landscape genetics studies of habitat fragmentation and connectivity need spatial replication to test the generality of inferences about how gene flow is influenced by certain landscape features (Holderegger & Wagner 2008; Segelbacher *et al.* 2010). Replication in landscape genetics refers to replication of the study unit (i.e. the landscape itself) (Holderegger & Wagner 2008). Replication in ecology and landscape genetics is not as highly controlled as in laboratory studies; however, comparison of a fragmented landscape and highly connected landscape can be considered as one treatment. Although some examples of landscape level replication exist (Orrock *et al.* 2006; Peakall & Lindenmayer 2006; Born *et al.* 2008), to our knowledge no landscape genetic studies have included replication of multiple sampled landscapes.

Individual-based, landscape genetic analysis of population connectivity is particularly powerful as a means to quantify habitat fragmentation effects on population structure because it directly associates patterns of genetic relatedness between individuals with cost distances (i.e. cost or resistance on movement) between these individuals on a number of alternative explanatory models (e.g. Cushman *et al.* 2006). Importantly, individual-based landscape genetic analyses using causal modelling (i.e. modelling using simple and partial Mantel correlation coefficients to evaluate the degree of support for alternative hypotheses of causality; Cushman *et al.* 2006) appear to have high power to correctly identify driving processes and reject incorrect alternative models (Cushman & Landguth 2010). This approach facilitates comparison of a range of alternative hypotheses, such as isolation by Euclidean distance, isolation by barriers, and isolation by landscape resistance in a single formal multiple-hypothesis testing framework (Balkenhol *et al.* 2009).

Cushman *et al.* (2006) was one of the first studies to use this multiple-hypothesis testing framework. They evaluated 110 alternative hypotheses related to the effects of landscape structure on gene flow in a black bear population in northern Idaho. Their analysis compared support for 108 landscape resistance models, isolation by Euclidean distance, and isolation by a landscape barrier. They identified forest cover and elevation as strong predictors of gene flow with roads as a potential, but equivocally supported, feature influencing gene flow. They concluded that gene flow in the north Idaho black bear population was most highly correlated with continuous forest cover at middle elevations, and found no independent support for IBD or landscape barriers (i.e. partial Mantel tests for IBD that remove effects of landscape were not significant). The resistance map they developed from the one Idaho site (Cushman *et al.* 2006) was used to map potentially important movement routes across a very large area of western Montana (Cushman *et al.* 2008). The validity of extrapolation of landscape genetic results to broader regions requires demonstration of the generality of inferences obtained from a particular study landscape, for example, by conducting landscape genetic analysis across multiple study areas.

Many factors, such as the number of loci and individuals sampled, need to be carefully considered when designing landscape genetic studies. Through simulations, Murphy *et al.* (2008) observed a greater increase in power from increasing sample size of individuals than increasing the number of loci used in landscape genetics analysis. The effect of the number of loci on the landscape genetics results has not been evaluated with empirical data.

Our general goal was to improve understandings of how landscape features influence population structure and gene flow in black bears in a range of study areas with different landscapes in the Rocky Mountains of northern Idaho and western Montana. For this, we used an identical landscape genetic modelling approach and a similar suite of alternative models (i.e. models of landscape resistance from combinations of different landscape features) as Cushman *et al.* (2006) to black bear populations in 12 different study areas of varying landscape composition, variability, and complexity. The second goal was to evaluate the usefulness of the reduction in number of loci to assess confidence in

© 2011 Blackwell Publishing Ltd

conclusions about which landscape features influence gene flow. Finally, we assessed the effect of the variability of landscape features on their identification as important factors on gene flow and began to identify thresholds of variation in landscape features necessary to influence gene flow of black bears.

Materials and methods

Study areas

Both the Montana Fish, Wildlife and Parks (MFWP) and the U.S. Geological Survey (USGS) provided genotypes derived from samples collected from traps and rub trees from 11 different study areas during 2001–2008. Collectively, these 11 study areas consist of *c.* 32 124 km^2 in western Montana and include the Swan Valley, Glacier National Park (GNP), the Rocky Mountain East Front, the Yaak, Cabinet, Garnet, Gallatin, Big and Little Snowys, Whitefish, Pioneer, and Salish Mountains (Fig. 1). The size of the study areas ranged from 842 to 6574 km^2, with elevation ranging from 554 to 3231 meters (see Table 1 for mean elevations of study areas). We also included the results from the north Idaho study area from Cushman *et al.* (2006) as our 12th replicate landscape.

Field sampling

We sampled all study areas with hair traps constructed following the protocol of Woods *et al.* (1999). Hair traps consisted of double-strand, four-prong barbed wire encircling three to six trees or steel posts at a height of 50 cm. We poured scent lure, a mix of aged cattle blood and liquid from decomposed fish, on forest debris piled in the centre of the wire corral. We hung a canister with a small hole filled with cotton fibre saturated with lure or a cloth saturated with lure in a tree 4–5 m above the centre of the trap. Sampling sessions were 12–14 days. We collected hair from barbs, the ground near the wire, and the lure pile. All hairs from one set of barbs constituted a sample. We placed each hair sample in a paper envelope labelled with a unique number and stored hair samples on silica desiccant at room temperature.

Site selection for study areas 100, 102, 104, 130, 290, 301, 319 and 411 (Fig. 1) was coordinated by Montana Fish, Wildlife and Parks. In these study areas, hair traps were distributed across a 5 × 5-km grid. Site selection for study areas 103, 450, and Glacier National Park were coordinated by the U.S. Geological Survey based on systematically distributing hair traps using a grid of 7 × 7-km cells (Kendall *et al.* 2009). In GNP, we also

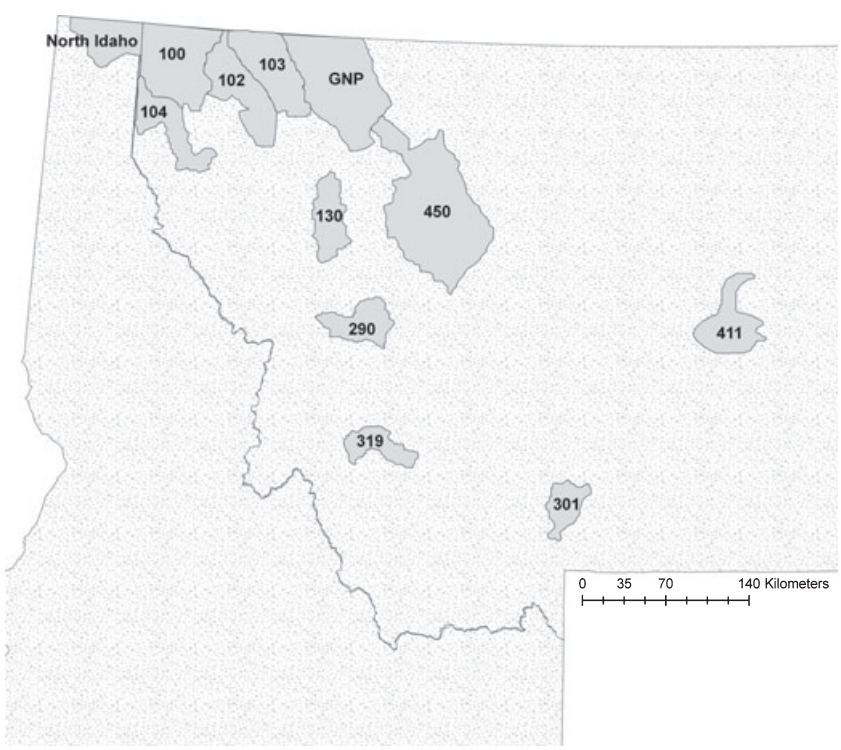

Fig. 1 The location of the study areas in Idaho and Montana: north Idaho (Cushman et al. 2006), 100 (the Yaak), 102 (Salish Mountains), 103 (Whitefish Mountains), 104 (Cabinet Mountains), 130 (Swan Valley), 290 (Garnet Mountains), 301 (Gallatin Mountains), 319 (Pioneer Mountains), 411 (Big and Little Snowys), 450 (Rocky Mountain East Front), and Glacier National Park (GNP).

© 2011 Blackwell Publishing Ltd

Table 1 Summary of landscape features for each of 11 study areas from Western Montana and North Idaho: area (km^2), mean elevation, standard deviation (SD) of elevation, correlation length of forest, correlation length of roads

Study area	Area (km^2)	Mean elevation	SD elevation	Correlation length forest	Correlation length road
100	3662	1333	294	20 797	20 303
102	4696	1388	223	21 163	19 602
103	2761	1577	327	21 465	14 377
104	3616	1294	380	25 672	6831
130	1605	1589	388	22 887	6343
290	1287	1571	234	13 227	10 340
301	842	2218	262	10 582	2712
319	1049	2236	286	12 783	3849
411	2168	1582	286	10 031	6553
450	3864	1786	312	23 462	15 643
GNP	6574	1646	380	31 165	14 058
NI	3000	1225	461	12 460	10 228

Correlation length is a landscape metric that quantifies the connectivity of a habitat class across the landscape and is equal to the distance (in meters) an organism placed into a random forest patch can move before encountering an edge.

collected hair during repeated visits to bear rubs using the protocols of Kendall *et al.* (2009).

Extraction and genotyping

Samples were analysed at the Wildlife Genetics International Laboratory, in Nelson, British Columbia, that specializes in low DNA quantity and quality samples, following standard protocols for noninvasive sampling (Woods *et al.* 1999; Paetkau 2003; Roon *et al.* 2005; Beja-Pereira *et al.* 2009). We analysed all samples with ≥1 guard hair follicle or five underfur hairs, and we used up to ten guard hairs plus under-fur when available.

DNA was extracted from hair samples using standard protocols established by the laboratory using QIAGEN DNeasy extraction kits. We genotyped the bears for each study area at a minimum of five microsatellite loci from a suite of twelve microsatellite loci: *G10B*, *G10H*, *G10IV*, *G10J*, *G10L*, *G10M*, *G10P*, *G10U*, *G10X*, *G1D*, *MU23* and *MU59*. Different loci in different study areas were used to maximize heterozygosity to increase the power for other independent studies of population estimation and relatedness that were conducted prior to this study. We have combined those datasets in this analysis. PCRs cycles were the following: for 1 min 20-s denaturation at 94 °C followed by 40 cycles of 94 °C for 20 s, 54 °C for 25 s, 72 °C for 10 s, with a final extension of 1 min 5 s at 72 °C and then cooling down to 4 °C.

The identification of unique individual multilocus genotypes followed a standard three-phase approach. Phase I involved an initial screening of all samples with the selected microsatellite markers. Species identification (to separate black bears from grizzly bears, *Ursus arctos*) was confirmed with an assignment test (Paetkau *et al.* 1995). The reference samples for calculating log genotype likelihoods were grizzly bear genotypes from either the southern Purcells or the Northern Divide Grizzly Bear Project (Kendall *et al.* 2009).

Phase II of the genotyping involved an attempt to fill in missing or weak data for samples that failed to produce reliable genotypes at three, four, or five markers during the initial screening. After a second pass at genotyping, samples with inadequate genotypes (<4 loci) were removed and not included in any further analysis. All pairs of remaining unique genotypes were subjected to exhaustive computerized comparison to check for similar genotypes that might be indicative of genotyping error. All pairs of genotypes that differed at fewer than three markers were scrutinized for possible error. All genotypes that differed from another genotype at just one marker were re-run (PCR, electrophoresis, and scoring).

Phase III involved re-analysing any pair of genotypes that differed at just one or two loci, following the published error-checking protocol established by Paetkau (2003). Once the genotyping was completed and checked for errors, individuals were defined by each unique genotype.

Data analysis

We estimated the number of alleles per locus (A), expected heterozygosity (H_E), heterozygote deficit (F_{IS}), and tested for Hardy–Weinberg proportions using the program GENALEX 6.1 (Peakall & Smouse 2006). We also tested for gametic (linkage) disequilibrium using the program GENEPOP 3.4 (Raymond & Rousset 1995).

Landscape genetic analysis

Genetic distance. For each individual we created five- to eight-locus genotypes with the following allelic encodings: a 0 for an absent allele, a 1 for a heterozygote or an allele found at one of two alleles, and a 2 for a homozygous individual with two copies of the same allele. This resulted in a matrix with one column for each allele in the pool of sampled bears and one row representing each individual bear genotype. Then inter-individual genetic distance was calculated following Bray–Curtis percentage dissimilarity measure (Legendre & Legendre 1998), to produce a matrix containing the genetic distances among all pairs of sampled bears. This calculation assumes the loci are independent and consistent with linkage disequilibrium results.

© 2011 Blackwell Publishing Ltd

Organizational models. A priori, we considered two potential drivers of genetic structure in these black bear populations, isolation by geographic distance (IBD) and isolation by landscape-resistance. Patterns of genetic structure among individuals can be correlated with landscape features by building resistance surfaces that assign different resistance-to-movement values to different landscape features (e.g. a high resistance-to-movement might be assigned to a known road or a body of water). Cells are given weights or 'resistance values' reflecting the presumed influence of each variable on movement of the species.

We selected 35 landscape-resistance models, representing combinations of three landscape features: elevation, roads, and forest cover (Cushman *et al.* 2006). These landscape features are known to be important to black bear movement and denning (Brody and Pelton 1898; Lyons *et al.* 2003; Mitchell & Powell 2003; Gaines *et al.* 2005; Mitchell *et al.* 2005); and influence gene flow (Cushman *et al.* 2006). Resistance of these features was modelled across four levels for elevation and three levels for roads and forest. The four levels for the feature elevation (E), consisted of a null model (EN), in which there was no penalty for elevation in the resistance surface, and three inverse-Gaussian resistance models, with minimum resistance of 1 at 500 (EL), 1000 (EM), and 1500 m (EH) elevation above sea level, respectively, 500-m standard deviation, and maximum resistance of 10. These three levels reflect a range of potential relationships between resistance to movement and elevation, with increasing resistance to gene flow at elevations higher and lower than the minima, with a maximum resistance of 10 times that of the minima achieved at asymptote (Cushman *et al.* 2006). Similarly, three levels of the forest cover feature were modelled. The first level was the null model (FN) in which forest cover had no effect in the resistance surface. The remaining two levels were models in which we posited that landscape resistance is minimum in closed canopy forest and linearly increases in nonforest cover types. In the forest high (FH) level we stipulated high relative resistance for crossing nonforest cover types, representing a condition where an individual bear strongly favours movement through forest, whereas in the forest low (FL) level nonforest classes have lower landscape resistance. Finally, three levels for the roads (R) were used, consisting of a null model (RN) where there was no effect for resistance of roads, a model with relatively strong effect of roads on resistance (RH), and a model with relatively lower effect of roads on resistance (RL). Isolation by Euclidean distance was included as a 36th model.

The landscape resistance models corresponding to each feature and level were combined into the 35 landscape-resistance models by addition. These hypotheses were represented by maps with cell values equal to the hypothetical resistance of each cell to gene flow. Forest cover data layers were derived using the GAP analysis program. Roads were mapped as a raster, including the two classes: major highways and other roads using TIGER 1997 (http://www.census.give/geo/www/tiger/). Elevation was mapped in meters and the layers were derived from 30-m digital elevation model (DEM). Before analysis, the base maps were re-sampled to a 90-m pixel size and rectified to a Universal Transverse Mercator projection.

Cost models. A matrix of movement costs among all pairs of individual bears in each study area was then computed based on least-cost distance. When an individual was sampled at more than one location we used the first location recorded in the dataset. We used ArcGIS COSTDISTANCE (ESRI Corp., Redlands, CA, USA) to calculate the least-cost distance from the location the individual bear sampled to every other bear's location across each of the 36 resistance surfaces. The cost matrix for the IBD hypothesis was created from the Euclidean distances based on UTM coordinates between all pairs of bears.

Mantel tests. The most widely used method to associate genetic structure with landscape features involves the use of Mantel tests (Mantel 1967) to correlate genetic distances with geographic distance or with alternative ecological distances that test hypotheses of the effect of landscape structure on gene flow (e.g. Broquet *et al.* 2006; Cushman *et al.* 2006; Schwartz *et al.* 2009a,b; Cushman & Landguth 2010; Storfer *et al.* 2010). We used partial Mantel tests (Smouse & Chakraborty 1986) within a causal modelling framework (Legendre 1993; Cushman *et al.* 2006; Cushman & Landguth 2010) to test the 36 resistance hypotheses for the influences of landscape features on gene flow. This framework has been shown to have high power to identify the drivers of gene flow and reject incorrect, correlated alternative hypotheses (Cushman & Landguth 2010).

The partial Mantel test measures the residual association between two dissimilarity matrices, after removing the association with a third dissimilarity matrix. In this study, we report partial Mantel test results, after factoring out the influence of Euclidean distance. This tests for a significant relationship between genetic distances and landscape resistance after accounting for (removing) the effects of the IBD null hypothesis. For each study area, we also partialled out effects of landscape from the Euclidean distance model to test for any independent support of isolation by Euclidean distance. All Mantel tests were conducted using the library ECODIST

© 2011 Blackwell Publishing Ltd

version 1.1.3 (Goslee & Urban 2007) in the statistical software package R (R Development Core Team 2007).

Effect of number of loci

The number of loci varied among study areas from five to nine. Since the majority of our study areas used six loci, we began by testing the effect of the reduction of loci down to six loci on the consistency of our results for study areas with more than six loci (north Idaho, 319, and 411). We conducted Mantel and partial Mantel tests using the genetic distance matrices for all subsets of loci down to six loci for our study areas. Then we further tested the effect of the reduction of loci on the consistency of results for our study areas with six loci. We conducted Mantel and partial Mantel tests using genetic distance matrices for the study areas with six loci (102, 103, 104, 290, 301, 450, and GNP) using all subsets of five loci from the six total loci. We compared the landscape features identified as influencing gene flow of each subset with the landscape features identified for the original data containing all loci.

Variation of landscape features

A priori we developed three hypotheses related to when the effects of particular landscape features (i.e. elevation, forest, and roads) on gene flow will be detected in a given landscape. Our first hypothesis was that elevation will be a landscape feature identified as influencing gene flow in study areas that have a relatively high variance of elevation. In study areas where there is little variation in elevation we posited that elevation would not be related to genetic differentiation among individual bears, as elevation would not limit gene flow where topography is relatively flat. We further posited that the optimal elevation at which resistance to gene flow was minimized would vary across western Montana in relation to regional climate patterns and mean elevation of the study area. Specifically, we hypothesized that gene flow would be maximal at middle elevation (Cushman *et al.* 2006, mean 1000 m, SD 500 m) in study areas located in the north, whereas gene flow would be maximal at high elevation (mean 1500 m, SD 500 m) in the southern part of the greater study area. This is because of regional climate patterns, in which precipitation and snow pack are highest in the northern portion of the study area and lowest in the southern part, resulting in similar biophysical conditions occurring at higher elevations in study areas in the south. We tested the first part of this hypothesis by conducting *t*-tests of the difference in mean standard deviation of elevation between study areas which elevation was in a landscape feature identified as influencing gene flow and those where elevation was not identified as influencing gene flow. We tested the second part of our hypothesis by conducting *t*-tests using the latitudes of the study areas with mid-elevation or high elevation identified as landscape features influencing gene flow.

Our second hypothesis was forest cover will be a landscape feature identified as influencing gene flow in study areas where forest is fragmented or has limited continuity. In study areas where forest is extensive and unfragmented we posited that there will be no relationship between gene flow and forest cover. Forest cover will not limit gene flow of a forest dependent species in landscapes that are continuously forested. In contrast, in landscapes where forests are fragmented we would expect gene flow of a forest dependent organism to be highly related to forest cover as a limiting factor. We tested this hypothesis by conducting *t*-tests evaluating the significance of the differences in the correlation length (McGarigal *et al.* 2002; Cushman *et al.* 2010) of forest cover between study areas in which forest was a landscape feature identified as influencing gene flow and study areas where it was not identified. Correlation length is a landscape metric that quantifies the connectivity of a habitat class across the landscape and is equal to the distance an organism placed into a random forest patch can move before encountering an edge. Correlation length is calculated using FRAGSTATS (McGarigal *et al.* 2002) on the reclassified forest cover map used to derive the forest cover resistance layers described above.

Our last hypothesis was that roads will be in a landscape feature identified as influencing gene flow in study areas that are highly dissected by extensive road networks. Where roads are extensive and highly fragment the landscape, we would expect them to limit gene flow. In contrast, where roads are few and do not dissect the landscape we posited that there should be no relationship between roads and gene flow, even if the species strongly avoids crossing roads. We tested this hypothesis by conducting *t*-tests evaluating the significance of differences in the correlation length (McGarigal and Marks 2002; Cushman *et al.* 2010a,b) of roads between study areas with roads identified as a landscape feature influencing gene flow and study areas with roads not identified as influencing gene flow.

Results

Genetic diversity and disequilibrium

Mean expected heterozygosity ranged from 0.67 (study area 411) to 0.84 (study area 104) with a grand mean = 0.803. The overall mean number of alleles per locus was 11 (range = 5–14; Table 2). The results for F_{IS}

© 2011 Blackwell Publishing Ltd

Table 2 Genetic summary statistics for eleven study areas in Western Montana, plus North Idaho (NI): (*N*) Number of individuals sampled, Number (#) of loci genotyped, mean observed heterozygosity (H_O), mean expected heterozygosity (H_E), and mean number of alleles per locus

Study area	*N*	# of loci	Mean F_{IS}	H_O	H_E	Mean alleles
100	160	5	0.019	0.811	0.824	11.6
102	131	6	0.013	0.813	0.824	11.5
103	196	6	-0.019	0.844	0.828	11.8
104	148	6	-0.009	0.852	0.841	13.5
130	132	5	-0.040	0.859	0.827	12.2
290	62	6	-0.063	0.839	0.789	9.8
301	60	6	-0.029	0.819	0.797	8.3
319	43	8	-0.038	0.843	0.812	10.1
411	72	7	-0.023	0.685	0.669	5.3
450	166	6	0.017	0.796	0.810	11.5
GNP	508	6	-0.006	0.828	0.823	13.0
NI	146	9	0.020	0.779	0.795	10.0
Grand mean	152	6.3	-0.013	0.814	0.803	10.7

were near 0 suggesting no cryptic substructure (Wahlund effect) or excessive genotyping error (e.g., allelic dropout). Significant departures from H–W proportions ($P < 0.01$) were found at four loci (*MU59*, *G10J*, *G1A* and *G10X*), one in each of four populations. Gametic (linkage) disequilibrium was significant ($P < 0.01$) at five pairs of loci. Only one pair (*G10J* and *G10L*) were in disequilibrium in more than one population (411 and 450).

Landscape genetic analysis

Five of the 11 Montana study areas had statistically significant landscape resistance models ($P < 0.05$; partial Mantel removing IBD) (102, 103, 104, 301, 319). None of these areas had the same most-supported landscape resistance model as the Idaho study area from Cushman *et al.* (2006), and all five areas had a different most-supported landscape resistance model.

The landscape feature of forest (high forest or low forest cover) was a landscape feature identified as influencing gene flow in three study areas (301, 319, Idaho). Elevation (elevation high, or elevation middle) was a landscape feature identified as influencing gene flow in three study areas (103, 104, Idaho). Roads (roads high or roads low) were a landscape feature identified as influencing gene flow in four study areas (102, 103, 301, Idaho) as a high resistance path.

When a partial Mantel test was conducted to test for IBD after removing landscape effects, 10 of the 11 study areas were nonsignificant for IBD when the most-supported landscape resistance model's landscape distance was partialled out (Table 3). IBD was statistically significant in seven of the 11 study areas using a simple Mantel test for correlation of genetic distance to Euclidean distance.

Effect of the number of loci

When we conducted Mantel and partial Mantel tests using the genetic distance matrices created using all subsets of seven of the eight loci in study area 319, we usually obtained the same most-supported landscape resistance model; Seven of the eight subsets produced the same significant most-supported landscape resistance model (FH: forest high) (Table 4). These subsets consistently (100%) produced forest at the high level as a feature within the most-supported landscape resistance model. In the subsets ($n = 28$) of six loci, our original most-supported landscape resistance model occurred 71% of the time and was still significant. FH occurred within the most-supported landscape resistance models of these subsets 86% of the time.

When we conducted Mantel and partial Mantel tests from seven loci down to six loci for study area 411, the original most-supported landscape resistance model from seven loci (RL: roads low) was the most common (57%) most-supported landscape resistance model in the subsets of six loci.

For the Idaho study area from Cushman *et al.* (2006), we found less consistency in the occurrence of the most-supported landscape resistance model (FHEMRH: forest high, mid-elevation, roads high). In the subsets ($n = 9$) of eight loci, the original most-supported landscape resistance model was not produced. However, both EM and FH were factors in 89% of the most-supported landscape resistance models from this group of subsets. In the subsets of seven loci ($n = 36$), the original most-supported landscape resistance model occurred 3% of the time with the FH resistance model having the most support occurring at a frequency of 42% of the time. FH occurred within the most-supported landscape resistance models from the subsets 67% of the time. EM occurred within the most-supported landscape resistance models 33% of the time from these subsets. The last factor from the original most-supported landscape resistance model, RH, occurred within the subsets only 9% of the time. In the subsets of six loci ($n = 84$), the original most-supported landscape resistance model occurred only 8% of the time with FH occurring most frequently with an occurrence of 30% of the time. Similar to the previous subsets, this subset had FH occurring most frequently (62%), EM occurring second most frequently (26%), and RH occurring the least frequently (8%) within the most-supported landscape resistance models.

Subsets of five loci (from six total loci) for our last group of seven study areas produced similar results to

© 2011 Blackwell Publishing Ltd

Table 3 Results of landscape resistance modelling and isolation by distance

Study area	Model	Mantel *r*	*P*-value	Study area	Model	Mantel *r*	*P*-value	Study area	Model	Mantel *r*	*P*-value	Study area	Model	Mantel *r*	*P*-value
100	EMFH	0.033	0.088	104	EH	0.06	0.023	301	FLRH	0.117	0.018	450	EHRL	0.056	0.061
	EMFL	0.033	0.099		FHEHRL	0.056	0.025		FHRH	0.115	0.023		EHRH	0.055	0.066
	FLRH	0.023	0.115		EHRL	0.057	0.030		FLRL	0.094	0.032		EHRL	0.054	0.069
	FLEMRH	0.030	0.127		FLEHRL	0.056	0.030		FHRL	0.091	0.044		EH	0.052	0.076
	RH	0.031	0.131		EHRH	0.056	0.031		FHEHRH	0.086	0.061		RL	0.054	0.088
	IBD_{PM}	−0.030	*0.885*		IBD_{PM}	−0.054	*0.963*		IBD_{PM}	−0.112	*0.977*		IBD_{PM}	−0.029	*0.810*
	IBD_{SM}	0.025	0.101		IBD_{SM}	0.005	0.422		IBD_{SM}	0.062	0.058		IBD_{SM}	0.066	0.007
102	**RL**	0.059	0.028	130	FHEMRH	0.024	0.135	319	FH	0.132	0.005	GNP	**RL**	0.028	0.094
	RH	0.059	0.036		FHEMRL	0.023	0.141		FHRL	0.124	0.010		RH	0.027	0.096
	ELRH	0.035	0.115		FLEMRL	0.022	0.147		FHRH	0.113	0.014		EHRL	0.016	0.181
	FLELRH	0.033	0.135		FLEMRH	0.023	0.149		FL	0.109	0.018		EHRH	0.016	0.183
	EL	0.032	0.141		EMFH	0.025	0.150		FLRL	0.105	0.024		EH	0.016	0.189
	IBD_{PM}	−0.016	*0.702*		IBD_{PM}	−0.007	*0.585*		IBD_{PM}	−0.118	*0.979*		IBD_{PM}	*0.035*	*0.001*
	IBD_{SM}	0.046	0.015		IBD_{SM}	0.058	0.016		IBD_{SM}	0.079	0.044		IBD_{SM}	0.035	0.001
103	EMRL	0.038	0.041	290	RH	0.032	0.263	411	**RL**	0.055	0.168				
	EMRH	0.040	0.047		RL	0.032	0.272		RH	0.055	0.183				
	FHEMRH	0.035	0.059		ELFH	0.029	0.284		ELFL	0.001	0.488				
	FLEMRH	0.035	0.059		FHELRL	0.029	0.289		ELFH	−0.009	0.565				
	EM	0.037	0.061		FHELRH	0.027	0.292		EL	−0.010	0.576				
	IBD_{PM}	−0.029	*0.896*		IBD_{PM}	*0.014*	*0.392*		IBD_{PM}	*0.006*	*0.455*				
	IBD_{SM}	0.035	0.031		IBD_{SM}	0.097	0.001		IBD_{SM}	0.078	0.076				

Top five most supported landscape models for each study area with Mantel *r* statistic and *P*-value for the partial Mantel comparing landscape resistance models partialling out the effect of Euclidean distance. The most frequent top model (RL: roads low) among study areas is in bold. Isolation by distance (IBD_{SM}) is from a simple Mantel test. Isolation by distance (IBD_{PM}, values in italics) is a partial Mantel test comparing genetics to Euclidean distance after partialling out the effect of the top landscape resistance model for each study area. The two levels of the 'forest' landscape factor (F) are represented as follows: FH, high-resistance to nonforest; FL, moderate-resistance to nonforest. The two levels of the 'roads' landscape factor (R) are represented as follows: RH, high resistance due to roads; RL, low resistance due to roads. The three levels of the landscape factor of elevation are represented as follows: EL, minimum resistance at low elevation; EM, minimum resistance at mid-elevation; EH, minimum resistance at high elevation.

our study areas with more than six loci. For the four study areas with a significant landscape resistance model ($P < 0.05$) using all six loci, most subsets (three of four study areas) of five loci had a significant ($P < 0.05$) most-supported landscape resistance model that produced the same most-supported landscape resistance model as with all six loci (Table 5). For the three study areas with a less significant landscape resistance model ($P > 0.05$), only one study area had the most-supported landscape resistance model identical to that with all six loci.

Given the small tag size associated with most of these studies, we were not able to fully explore the ramifications associated with choosing models based on short tags. However, we believe that re-sampling larger tags can provide significant understandings concerning the relative importance of various factors and the stability of the most highly supported model. For example, forest is the only factor to dominate in all genetic subsets for the Idaho study area. Given the modelling design, each factor occurs in one-thirds of the applied models. Although both elevation and roads are in the best supported model given nine loci, in subsets, roads is present less than expectation in the best supported models for all subsets (Table 5). Thus, the subsampling suggests that for this study area, although RH is in the best nine-locus model, that roads, unlike forest, is not generally supported by the genetic data.

Variation of landscape features

We tested whether the features included in the most-supported landscape resistance models for different study areas were related to variation of a landscape feature in the given landscape. Study areas containing high variation (SD) in elevation were significantly more likely to have elevation as a landscape feature identified as influencing gene flow ($P = 0.019$, Table 6). In addition, they had on average 52% greater SD of elevation than study areas where elevation was not included in the most-supported landscape resistance model.

We also hypothesized that the level of elevation in the most-supported landscape resistance model would be

© 2011 Blackwell Publishing Ltd

Table 4 Summary of the effect of the number of loci on the top landscape resistance models identified for study areas with more than six loci

(a)

	7 loci		6 loci	
411	Percent	*P*-value	Percent	Mean *P*-value
RL	100	0.168	57	0.068
RH	0	N/A	29	0.280
ELFL	0	N/A	14	0.248

(b)

	8 loci		7 loci		6 loci	
319	Percent	*P*-value	Percent	Mean *P*-value	Percent	Mean *P*-value
FH	100	0.005	88	0.011	71	0.018
FHRH	0	N/A	12	0.029	4	0.069
FHRL	0	N/A	0	N/A	7	0.004
FLRH	0	N/A	0	N/A	7	0.008
RH	0	N/A	0	N/A	7	0.008
EMFH	0	N/A	0	N/A	4	0.030

(c)

	9 loci		8 loci		7 loci		6 loci	
NI	Percent	*P*-value	Percent	Mean *P*-value	Percent	Mean *P*-value	Percent	Mean *P*-value
EMFH	0	N/A	67	0.032	14	0.020	8	0.042
EMFL	0	N/A	11	0.001	8	0.030	2	0.010
FH	0	N/A	11	0.028	42	0.027	30	0.018
FHEMRL	0	N/A	11	0.014	0	N/A	1	0.169
FL	0	N/A	0	N/A	8	0.042	15	0.056
FHELRL	0	N/A	0	N/A	5	0.095	12	0.046
FLEMRL	0	N/A	0	N/A	5	0.001	1	0.116
EHFH	0	N/A	0	N/A	3	0.001	0	N/A
ELFL	0	N/A	0	N/A	3	0.017	1	0.083
EMRH	0	N/A	0	N/A	3	0.028	2	0.001
FHEMRH	100	0.011	0	N/A	3	0.014	2	0.009
FLELRH	0	N/A	0	N/A	3	0.045	1	0.026
FLELRL	0	N/A	0	N/A	3	0.035	0	N/A

Results from the partial Mantel tests using the genetic distance matrices created using all subsets of loci down to six loci for study areas 411, 319, and North Idaho (NI). Percent is the percent of occurrence (%) of the top landscape resistance model for all subsets of loci down to six loci. The most supported landscape resistance model with the full set of loci for each study area is in bold. (a) Results from study area 411. (b) Results from study area 319. (c) Results from North Idaho.

EM (mid-elevation), as in Cushman *et al.* (2006) when study landscapes were in the north, and would be EH when study landscapes were in the south. Landscapes in which EM was present in the most-supported landscape resistance model were on average 135 km farther north than landscapes where EH was in the most-supported landscape resistance model ($P = 0.055$; Table 5).

Study areas in which forest cover was a factor identified as influencing gene flow had on average 48% lower correlation length of forest than study areas in which forest was not in the most-supported landscape resistance model. The *t*-test of the differences in correlation length of forest between landscapes in which forest was in the most-supported landscape resistance model and those where it was not was highly significant ($P = 0.001$, Table 6).

There was a very large difference in the correlation length of roads between study areas in which roads

© 2011 Blackwell Publishing Ltd

Table 5 Summary of the effect of the number of loci on the most supported landscape resistance models from study areas with six loci

Study area	6 loci		5 loci		Study area	6 loci		5 loci	
102	Percent	*P*-value	Percent	Mean *P*-value	301	Percent	*P*-value	Percent	Mean *P*-value
RL	100	0.028	50	0.021	**FLRH**	100	0.018	67	0.021
RH	0	N/A	33	0.027	FHRH	0	N/A	33	0.041
FLEHRH	0	N/A	17	0.088					
Study area	6 loci		5 loci		Study area	6 loci		5 loci	
103	Percent	*P*-value	Percent	Mean *P*-value	450	Percent	*P*-value	Percent	Mean *P*-value
EMRL	100	0.041	0	N/A	**RH**	100	0.064	0	N/A
EMFL	0	N/A	100	0.261	RL	0	N/A	50	0.148
					EHRH	0	N/A	33	0.047
					EHRL	0	N/A	17	0.022
Study area	6 loci		5 loci		Study area	6 loci		5 loci	
104	Percent	*P*-value	Percent	Mean *P*-value	GNP	Percent	*P*-value	Percent	Mean *P*-value
EH	100	0.023	50	0.047	**RL**	100	0.094	33	0.073
EHFL	0	N/A	16	0.108	RH	0	N/A	33	0.111
EHRL	0	N/A	16	0.008	EHRL	0	N/A	17	0.022
FHEHRL	0	N/A	16	0.027	EH	0	N/A	17	0.122
Study area	6 loci		5 loci						
290	Percent	*P*-value	Percent	Mean *P*-value					
RH	100	0.263	0	N/A					
ELFH	0	N/A	33	0.330					
EMFL	0	N/A	17	0.118					
RL	0	N/A	17	0.188					
FHELRH	0	N/A	17	0.013					
FL	0	N/A	17	0.241					

Results from the partial Mantel tests using the genetic distance matrices created using all subsets of five loci (from the six) for study areas 102, 103, 104, 290, 301, 450, and GNP. Percent is the percent of occurrence (%) of the most supported landscape resistance model. The most supported model from the original six loci for each study area is in bold.

were present in the most-supported landscape resistance model and study areas where they were not. The correlation length of roads was on average 120% greater in study areas in which roads appear in the most-supported landscape resistance model than in study areas where they do not ($P = 0.089$, Table 6).

Discussion

The degree to which different landscape features vary in a given landscape may lead to different statistical inferences about which landscape features influence gene flow and movement, even if the species has a globally consistent response to landscape structure. Therefore, landscape-level 'replication' of landscape-genetic research is essential to assess if we can generalize species' habitat requirements for gene flow. Replication provides a means to evaluate whether there is consistency in the landscape–genetic relationship across multiple landscapes, and to evaluate different alternative explanations of observed differences in landscape–genetic relationships among the different landscapes. Replication could also prevent misleading interpretations that a landscape feature (e.g. forest) is not important for a species, for example when the feature is minimally variable (e.g. continuous forest) across a single study area. Such a misleading interpretation is possible for any statistical inference: where if a factor is not substantially variable, there is no effect of the factor.

© 2011 Blackwell Publishing Ltd

Table 6 Results of *t*-tests and power for testing hypotheses regarding the landscape features of elevation, forest, and roads

	Mean 1	Mean 2	SD 1	SD 2	*P*-value	Effect size	Power
Elevation	257	389	32	68	0.019	52%	0.442
W to E	177038	164211	62395	100283	0.449	93 km	0.067
N to S	344544	542344	221637	334500	0.055	135 km	0.09
Forest	22767	11942	5706	2108	0.001	48%	0.354
Roads	11729	5339	6421	4051	0.089	120%	0.125

Power is the likelihood of obtaining a significant statistical test if the true difference between means is as large as that observed, given the observed standard deviations (SD). Mean 1 of elevation is the average of the elevation SD (m) of all study areas that do not contain elevation within their top model. Mean 2 of Elevation is the average of the elevation SD (m) of all study areas that contain elevation within their top model. Mean 1 of W to E is the average longitude of study areas with mid-elevation as a component of their top model. Mean 2 of W to E is the average longitude of the study areas with high elevation as a component of their top model. Mean 1 of N to S is the average latitude of study areas with mid-elevation as a component of their top model. Mean 2 of N to S is the average latitude of study areas with high elevation as a component of their top model. Mean 1 of Forest is the mean correlation length of study areas that do not contain forest as a component of their top model. Mean 2 of Forest is the mean correlation length of study areas with forest as a component of their top model. Mean 1 of Roads is the mean correlation length of roads in study areas that contain roads as a component of their top model. Mean 2 is the mean correlation length of roads in study areas that do not contain roads as a component of their top model. All *P*-values are from one-tailed tests.

Variability among study areas

There is notable variability in the influence of different landscape features among the 12 study areas, which taken at face value suggests that elevation, roads, and forest cover often influence gene flow in this species, but are inconsistent predictors for different landscapes. None of the five statistically significant study areas had the same most-supported landscape resistance model. Explaining this apparent variability across landscapes assists in obtaining a generalized understanding of the pattern–process relationships governing gene flow in the American black bear. Future analyses could use statistical models (e.g. multivariate regression models) to identify the landscape features with the strongest influence on gene flow among study areas.

Most of our study areas (11 of 12) yielded landscape resistance models (partialling out IBD) that explained genetic distance between individuals better than the IBD model (partialling out the effects of landscape), which suggests landscape resistance is a stronger predictor of genetic structure of black bears than the null hypothesis of IBD (as in Cushman *et al.* 2006). In the causal modelling framework (Legendre 1993; Cushman *et al.* 2006; Cushman & Landguth 2010), the only way we would have strong support for IBD is if it is significantly supported when partialling out the most-supported landscape resistance models.

Partialling out landscape (when testing for IBD) showed a lack of independent statistical support for IBD. Failure to compare the IBD hypothesis with the stronger landscape resistance hypotheses in these landscapes could lead to incorrect conclusion that isolation by Euclidean distance is the main process driving gene flow in these landscapes (Legendre *et al.* 2002; Murphy *et al.* 2008). This error would be an example of affirming the consequent in landscape genetics described by Cushman & Landguth (2010), in which multiple logically exclusive hypotheses might have strong spatial correlation with the true driving process and failure to compare multiple models could lead to erroneous conclusions. These findings support the importance of testing multiple alternative hypotheses and in particular of testing landscape resistance hypotheses against a biologically meaningful null model of IBD (Antolin *et al.* 2006; Neville *et al.* 2006; Holderegger & Wagner 2008; Balkenhol *et al.* 2009; Liu *et al.* 2009).

Number of loci

Relatively little is known about how variability in number of loci analysed affects reliability or power to detect correct underlying processes in landscape genetic analysis, although previous simulation studies suggest power increased more rapidly by adding loci than by adding spatial locations (Murphy *et al.* 2008). We conducted Mantel and partial Mantel tests using genetic distance matrices created from subsets of loci in three study areas with more than six loci down to six loci, which was the average number of loci used across study areas. The effect of the number of loci differed among these three study areas. In study area 319, the results suggested little effect on the consistency of model support. In study area 319, forest cover was predicted to be an important facilitator of gene flow in all of the most-supported landscape resistance models identified when we used genetic distance matrices for seven loci and 93% of the models identified at six loci. Our results revealed some apparent instability in model support in study area 411 and in the Idaho study area. Study area

© 2011 Blackwell Publishing Ltd

411 did not have a model with statistical support ($P > 0.10$) with seven loci so it is logical the reduction of loci would not result in a landscape resistance model with statistical support.

In the Idaho study area (Cushman *et al.* 2006) the reduction of the number of loci appeared to result in less stability in the pattern of support for the most-supported landscape resistance models. This inconsistency was primarily driven by models with the roads feature having nearly equivocal support (RH, RL, and RN had very similar support) and roads dropping out of the most-supported landscape resistance models in the subsets. Nonetheless, support for forest cover and elevation was consistent in the Idaho study area when the number of loci was reduced. The results from the Idaho study area seem to indicate that there is high consistency across the number of loci used in identifying the importance of forest as contributing to the genetic structure. Similarly, Cushman *et al.* (2006) identified middle elevation as an important predictor of gene flow. The subsets of eight loci reaffirmed the relationship with middle elevation 87% of the time. However, the identification of middle elevation as important to genetic connectivity dropped dramatically in the subsets of seven and six loci. This seems to suggest that the features that dominate the landscape-genetic pattern process relationship (e.g. forest in the Idaho study area) will usually be consistently identified in analyses of fewer loci. However, it also suggests loss of power to detect the effects of weaker predictors, such as middle elevation and roads in the Idaho study area. Some model instability in north Idaho may have resulted from some individuals having less than nine loci genotyped. In the eleven Montana study areas, nearly all individuals had complete genotypes with no missing data. Missing data might influence model stability and may be a possibility for future research to assess.

For the study areas with six loci, we used genetic distances matrices derived from the subsets of six to five loci and observed study areas with statistical support in the original dataset almost consistently (three of four study areas; 102, 104, 301) resulted in the same significant most-supported landscape features as influencing gene flow. This demonstrates remarkable stability, suggesting limited sensitivity to the number of loci. These results may also suggest a less arbitrary threshold for statistical support than the commonly accepted ($P < 0.05$) to ($P < 0.04$), which produced more stable support for models when the loci were reduced.

Variation of landscape features

There are several potential explanations of the variability among landscapes in terms of the most-supported landscape resistance models. One explanation could be it is possible that variability in which landscape resistance models were supported among study landscapes is related to whether or not each landscape feature is variable and limits gene flow in a given landscape. For example, consider a situation where one is correlating gene flow of a species that is completely dependent upon forest with landscape structure in a landscape that is completely covered in forest. In this situation forest is a necessary element of the species' habitat and its occurrence and movement are totally dependent upon it. However, forest would not appear in a model predicting movement because forest is not limiting in a landscape that is completely covered in forest. Thus, it is possible for a critical dependence upon certain landscape features to be invisible to analysis depending upon whether this landscape element limits movement.

A second explanation may be that while there is a general relationship between spatial genetic structure of this species and landscape features the relatively small sample sizes of individuals in some landscapes and few sampled loci (five to seven) result in imprecision and low power such that we fail to identify the correct underlying process in many landscapes. For example, the low support for landscape resistance models in GNP might result from the hair snares being farther apart (7 km) and the higher density of bears than in other study areas, could lead to less sampling of closely related bears on adjacent home ranges and thus lower power to detect correlations between genetic distance and landscape distance.

We will focus our consideration on the first of these possibilities. Differences in supported models may result when certain landscape features do no limit gene flow in certain landscapes due to their extent or pattern, but do limit gene flow in other landscapes. A priori, we formalized three hypotheses related to this expectation.

Hypothesis 1: Variability in elevation

Our first hypothesis was that elevation will not be a landscape feature identified as influencing gene flow in landscapes where topography is relatively flat and elevation is not highly variable. The reasoning was, even if elevation is highly related to gene flow in American black bear its effect will not be detectible in landscapes that have little variability in elevation because in such landscapes there will be very little difference in movement cost as a function of elevation among individuals. Our results were fully consistent with this hypothesis. When analysis was restricted only to landscapes containing resistance models supported at a Mantel P value of less than 0.05, the effects size was 52% and was statistically significant ($P = 0.019$).

© 2011 Blackwell Publishing Ltd

We began to identify thresholds of variation necessary to have an observable influence on gene flow. The three study areas that we identified elevation as a landscape feature influencing gene flow had SDs in elevation above 300 m. The remaining three study areas had SDs in elevation less than 300 m. These results are all consistent with our hypothesis that elevation will be a landscape feature identified as influencing gene flow only when it is limiting to gene flow, and that it will be limiting only when there is a relatively high variability in elevation across a study area.

1a: Mid-elevation vs. high elevation

The second part of our first hypothesis was that we would expect middle elevation (EM) to be in models including elevation in landscapes to the northern parts of the full study area, and high elevation (EH) to be in models including elevation in landscapes in the southern parts of the study area. The reasoning behind this was that as one moves south lower and upper tree lines move upward in elevation, such that similar ecological conditions occur at higher elevations in the south than the north. This higher snowpack in the northern part of the study area also is related to the lower location of the upper tree lines in the north than the south.

Our hypotheses of mid-elevation (EM) occurring as a landscape feature identified as influencing gene flow in the north and becoming (EH) as you move south was statistically significant and the effects size was 197 km. This is a large effect size given the scale of our entire study area which is *c* 250 km across. The optimal elevation for gene flow is lower in the northern portion of the study area and higher in the south. This shows that nonstationarity in relationships between landscape structure and gene flow across broad geographical extents (e.g. Cushman *et al.* 2010b), which has important implications for conducting broad-scale landscape genetic analyses.

Hypothesis 2: Fragmentation of forest

Our second hypothesis was that we expected that forest cover would be a landscape feature identified as influencing gene flow for landscapes in which forest was highly fragmented and would not be included in landscapes that had low forest fragmentation. The reasoning was that even if forest cover is essential and nonforest is highly impermeable to gene flow, this relationship would only be detectible in landscapes where limited forest extent or substantial forest fragmentation limits gene flow. In landscapes where forest cover does not limit gene flow, such as landscapes that are continuously forested, there would be no statistical relationship between forest connectivity and gene flow across the landscape. Our analysis provided strong statistical support for this hypothesis. We expected that the correlation length of forest would be significantly lower in landscapes in which forest cover was a landscape feature identified as influencing gene flow than in landscapes in which it was not. We observed a large difference between means in the direction we expected. Effects size was 48% which reflects large differences in the connectivity of forest (Neel *et al.* 2004). These differences were highly statistically significant, despite very low power resulting from a small sample size.

Similar to elevation, we attempted to find thresholds of correlation length required to have an observable influence on black bear gene flow. The three study areas with forest identified as a landscape feature influencing gene flow were also the study areas with the lowest correlation lengths. The three study areas with correlation lengths of at least 20 000 m (less fragmentation) did not identify forest as a landscape feature influencing gene flow.

This has important implications for landscape genetic analyses. It is likely that forest cover is an essential component of habitat for American black bears and is likely essential to promote gene flow. However, our results indicate that landscape genetic analyses in many landscapes would fail to detect this relationship. In several of our study landscapes forest cover is high and forest fragmentation is low. It is likely that gene flow across these landscapes is not related to patterns in forest cover, as forest extent and fragmentation are not limiting to movement and dispersal. This does not mean that forest cover is not important, only that it is not limiting. This is an important case of where a relationship with a necessary resource is not detectible because it is not limiting and therefore does not structure the response variable. Landscape genetic analysis in the landscapes where forest is not limiting would not identify forest as an important driver of gene flow. From this it would be tempting to incorrectly conclude that forest cover is not important to black bear gene flow. This would be a logical error of denying the antecedent (Cushman & Huettmann 2010; Murphy *et al.* 2010) which commonly results from misinterpretation of statistical tests in which a model term with low variation it might have no statistical signal (Sokal & Rohlf 1995). This is one of the most important findings of this analysis, and highlights the importance of careful statistical interpretation and of landscape-level replication across a broad range of study landscapes to determine the features that limit gene flow and under what circumstances of landscape structure they become limiting.

© 2011 Blackwell Publishing Ltd

Hypothesis 3: Fragmentation by roads

The inclusion of roads in the most-supported landscape resistance model is predicted in this hypothesis to only occur when roads are at a sufficient density to limit gene flow. The correlation length of roads was 120% higher in study areas in which roads were included in the most-supported landscape resistance model than in study areas in which roads were not in the most-supported landscape resistance model. While this difference is only marginally significant due to high variability, the mean difference of 120% is a very large effects size, and is highly consistent with our expectation. The large number of study areas in which roads was included in a most-supported landscape resistance model suggests that roads are often an important limiting factor to gene flow in black bears. Our analysis suggests that the correlation length of roads in a landscape is related to whether or not roads limit gene flow. Another consideration is we modelled roads as having a resistance against gene flow; however, roads can also serve as movement and dispersal corridors (Balkenhol & Waits 2009).

Synthesis

Given all the above, this study has produced novel findings that have contributed to understandings of black bear ecology, population genetic structure, and gene flow. Using genetic data and individual-based modelling, our study has re-affirmed previous findings (e.g. radio collar data, Cushman & Lewis 2010) of the importance of landscape features such as middle elevation and forest cover for black bear movement. Our study has also evaluated the effect of the number of loci on landscape genetic study results, suggesting that six to eight loci ($H_E \approx 0.80$) might be sufficient if model support is strong ($P < 0.04$), confirming observations from simulated data (Murphy *et al.* 2008). Lastly, through examining the variation in landscape features within each of multiple study areas, we were able to begin to establish thresholds of variation in landscape features necessary to influence gene flow of black bears. For example, if SD in elevation is greater than *c.* 300 m, then elevation appears to influence gene flow.

Limitations and future research

A limitation of this research might be the relatively small number of loci used in the landscape analysis. However, the loci were highly polymorphic and thus have relatively high power to estimate important parameters such as interindividual genetic distance. Future research could include more loci and test for outlier loci, e.g. because of selection (Schwartz *et al.* 2009a,b). In addition, as in any landscape genetic study, there could be a lag time for a landscape signal to develop in the genetic data. Thus very recent landscape changes might not yet be detectable. Future research is needed to quantify the time lag until new barriers become detectable (e.g. Murphy *et al.* 2008; Landguth *et al.* 2010a,b), as well as to quantify the time until ancient historical barrier signals disappear (e.g. Landguth *et al.* 2010b). Future research should test a wider range of resistance models, conduct more extensive model optimization, assess the effect of scale or study area size on stability of support, and carefully quantify effects of noise (e.g. subsampling loci to assess most-supported landscape resistance model stability) vs. landscape signal (i.e. landscape variation).

Conclusion

Conducting studies in different landscapes can help achieve a general understanding of the relative influence of different landscape features on gene flow. This is crucial to understand how landscapes and landscape change can influence a species' ecology and evolution and thus influence management to maintain connectivity. Our results within 12 study areas generally support previous work which shows that gene flow in American black bear is facilitated by forest cover at optimal elevations, whereas nonforest cover and roads can impede gene flow. Our research suggests that using subsets of a full suite of loci can help assess support for landscape genetic models; we recommend future researchers use subsampling of loci to assess confidence in inferences about which features influence gene flow. Our study suggests that failure to study multiple landscape areas could lead to erroneous conclusions about which landscape features generally limit gene flow, and suggests ways to avoid erroneous conclusions. Failure to observe an effect of a given landscape feature in a landscape genetic analysis (e.g. Type I error) does not necessarily show that the feature is not critically related to gene flow. Further, we suggest that even critical landscape features will present strong relationships with genetic differentiation only when their pattern within a given landscape is substantially variable and thus limiting to gene flow. Conclusions that a certain landscape feature is (or is not) important for gene flow or substructure could be specific to a certain landscape or study area. Future research is needed to characterize the limiting factor relationships we describe and further quantify thresholds of variation in elevation, fragmentation of forest and extensiveness of road networks where these landscape conditions begin to influence gene flow in this species.

© 2011 Blackwell Publishing Ltd

Acknowledgements

We thank the Montana Fish, Wildlife and Parks, as well as the U.S. Geological Survey for coordinating and collecting hair samples, which also includes the hundreds of employees and volunteers. We also thank the following agencies that provided other support: Blackfeet Nation; Confederated Salish and Kootenai Tribes; Montana Department of Natural Resources and Conservation; National Park Service; Northwest Connections; United States Bureau of Land Management; United States Fish and Wildlife Service; and Wildlife Genetics International. Financial support was provided by the U.S. Geological Survey and United States Forest Service. We also thank the bears that allowed this research to be possible. F.W.A., G.L., and R.S. were partially supported by a grant from the U.S. National Science Foundation Grant DEB 074218. G.L. also was supported by an NSF grant (DEB-1067613) and the Portuguese-American Science Foundation, CIBIO-UP, and research grant PTDC/BIA-BDE/65625/2006 from the Portuguese Science Foundation (FCT).

References

Andren H (1994) Effects of habitat fragmentation on birds and mammals in landscapes with different proportions of suitable habitat – a review. *Oikos*, **71**, 355–366.

Antolin MF, Savage LT, Eisen RJ (2006) Landscape features influence genetic structure of black-tailed prairie dogs (*Cynomys ludovicianus*). *Landscape Ecology*, **21**, 867–875.

Balkenhol N, Waits LP (2009) Molecular road ecology: exploring the potential of genetics for investigating transportation impacts on wildlife. *Molecular Ecology*, **18**, 4151–4164.

Balkenhol N, Gugerli F, Cushman SA *et al.* (2009) Identifying future research needs in landscape genetics: where to from here? *Landscape Ecology*, **24**, 455–463.

Beja-Pereira A, Bricker BJ, Chen S *et al.* (2009) DNA genotyping suggests recent brucellosis outbreaks near Yellowstone National Park originate from elk. *Journal of Wildlife Diseases*, **45**, 1174–1177.

Born C, Hardy OJ, Chevallier MH *et al.* (2008) Small-scale genetic structure in Central African rainforest tree species *Aucoumea klaineana*: a stepwise approach to infer the impact of limited gene dispersal, population history and habitat fragmentation. *Molecular Ecology*, **17**, 2041–2050.

Brody AJ, Pelton MR (1989) Effects of roads on black bear movements in western North Carolina. *Wildlife Society Bulletin*, **17**, 5–10.

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

Coulon A, Guillot G, Cosson JF *et al.* (2006) Genetic structure is influenced by landscape features: empirical evidence from a roe deer population. *Molecular Ecology*, **15**, 1669–1679.

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

Cushman SA, Huettmann F (2009) Ecological knowledge, theory and information in space and time. In: *Chapter 1: Spatial Complexity, Informatics and Wildlife Conservatio* (eds Cushman SA, Huettman F). pp. 3–18, Springer, Tokyo.

Cushman SA, Huettmann F (2010) Ecological knowledge, theory and information in space and time. In: *Spatial Complexity, Informatics and Wildlife Conservatio* (eds Cushman SA, Huettman F). pp. 3–18, Springer, Tokyo.

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

Cushman SA, Lewis J (2010) Movement behavior explains genetic differentiation in American black bear. *Landscape Ecology*, **25**, 1613–1625.

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

Cushman SA, McKelvey KS, Schwartz MK (2008) Use of empirically derived source-destination models to map regional conservation corridors. *Conservation Biology*, **23**, 368–376.

Cushman SA, Compton BW, McGarigal K (2010a) Habitat fragmentation effects depend on complex interactions between population size and dispersal ability: modeling influences of roads, agriculture and residential development across a range of life-history characteristics. In: *Spatial Complexity, Informatics and Wildlife Conservatio* (eds Cushman SA, Huettman F). pp. 369–387, Springer, Tokyo.

Cushman SA, Littell J, McGarigal K (2010b) The problem of ecological scaling in spatially complex, nonequilibrium ecological systems. In: *Spatial Complexity, Informatics and Wildlife Conservatio* (eds Cushman SA, Huettman F). pp. 43–64, Springer, Tokyo.

Dixon JD, Oli MK, Wooten MC *et al.* (2006) Effectiveness of a regional corridor in connecting two Florida black bear populations. *Conservation Biology*, **20**, 155–162.

Dixon JD, Oli MK, Wooten MC *et al.* (2007) Genetic consequences of habitat fragmentation and loss: the case of the Florida black bear (*Ursus americanus floridanus*). *Conservation Genetics*, **8**, 455–464.

Epps CW, Wehausen JD, Bleich VC *et al.* (2007) Optimizing dispersal and corridor models using landscape genetics. *Journal of Applied Ecology*, **44**, 714–724.

Fischer J, Lindenmayer DB (2007) Landscape modification and habitat fragmentation: a synthesis. *Global Ecology and Biogeography*, **16**, 265–280.

Fontaine MC, Baird SJE, Piry S *et al.* (2007) Rise of oceanographic barriers in continuous populations of a cetacean: the genetic structure of harbor porpoises in Old World waters. *BMC Biology*, **5**, 30–46.

Gaines WL, Lyons AL, Lehmkuhl JF *et al.* (2005) Landscape evaluation of female black bear habitat effectiveness and capability in the North Cascades, Washington. *Biological Conservation*, **125**, 411–425.

Goslee SC, Urban DL (2007) The ecodist package for dissimilarity-based analysis of ecological data. *Journal of Statistical Software*, **22**, 1–19.

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

Kendall KC, Stetz JB, Boulanger J *et al.* (2009) Demography and genetic structure of a recovering grizzly bear population. *Journal of Wildlife Management*, **73**, 3–17.

Keyghobadi N (2007) The genetic implications of habitat fragmentation for animals. *Canadian Journal of Zoology*, **85**, 1049–1064.

© 2011 Blackwell Publishing Ltd

Kindall JL, Van Manen FT (2007) Identifying habitat linkages for American black bears in North Carolina, USA. *Journal of Wildlife Management*, **71**, 487–495.

Landguth EL, Cushman SA, Murphy MA *et al.* (2010a) Quantifying landscape connectivity: assessing lag time until barrier signals are detectable. *Molecular Ecology*, **19**, 4179–4191.

Landguth EL, Cushman SA, Murphy MA *et al.* (2010b) Relationships between migration rates and landscape resistance assessed using individual-based simulations. *Molecular Ecology Resources*, **10**, 854–862.

Legendre P (1993) Spatial autocorrelation – trouble or new paradigm? *Ecology*, **74**, 1659–1673.

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

Li H, Li D, Li T *et al.* (2010) Application of least-cost path model to identify a giant panda dispersal corridor network after the Wenchuan earthquake – case study of Wolong Nature Reserve in China. *Ecological Modelling*, **221**, 944–952.

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.

Lyons AL, Gaines WL, Servheen C (2003) Black bear resource selection in the northeast Cascades, Washington. *Biological Conservation*, **113**, 55–62.

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

Mantel N (1967) Ranking procedures for arbitrarily restricted observation. *Biometrics*, **23**, 65–78.

McGarigal K, Cushman SA (2002) Comparative evaluation of experimental approaches to the study of habitat fragmentation effects. *Ecological Applications*, **12**, 335–345.

McGarigal K, Cushman SA, Neel MC *et al.* (2002) *FRAGSTATS: Spatial Pattern Analysis Program for Categorical Maps*. Computer software program produced by the authors at the University of Massachusetts, Amherst. Available at the following web site: http://www.umass.edu/landeco/fragstats.html.

Mitchell MS, Powell RA (2003) Response of black bears to forest management in the southern Appalachian Mountains. *The Journal of Wildlife Management*, **67**, 692–705.

Mitchell FS, Onorato DP, Hellgren EC *et al.* (2005) Winter ecology of American black bears in a desert montane island. *Wildlife Society Bulletin*, **33**, 164–171.

Murphy MA, Evans JS, Cushman SA *et al.* (2008) Representing genetic variation as continuous surfaces: an approach for identifying spatial dependency in landscape genetic studies. *Ecography*, **31**, 685–697.

Murphy MA, Evans JS, Storfer A (2010) Quantifying *Bufo boreas* connectivity in Yellowstone National Park with landscape genetics. *Ecology*, **91**, 252–261.

Neel MC, McGarigal K, Cushman SA (2004) Behavior of class-level landscape metrics across gradients of class aggregation and area. *Landscape Ecology*, **19**, 435–455.

Neville HM, Dunham JB, Peacock MM (2006) Landscape attributes and life history variability shape genetic structure of trout populations in a stream network. *Landscape Ecology*, **21**, 901–916.

Orrock JL, Levey DJ, Danielson BJ *et al.* (2006) Seed predation, not seed dispersal, explains landscape-level abundance of an early-successional plant. *Journal of Ecology*, **94**, 838–845.

Paetkau D (2003) An empirical exploration of data quality in DNA-based population inventories. *Molecular Ecology*, **12**, 1375–1387.

Paetkau D, Calvert W, Stirling I, Strobeck C (1995) Microsatellite analysis of population-structure in Canadian polar bears. *Molecular Ecology*, **4**, 347–354.

Peakall R, Lindenmayer D (2006) Genetic insights into population recovery following experimental perturbation in a fragmented landscape. *Biological Conservation*, **132**, 520–532.

Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, **6**, 288–295.

Pease KM, Freedman AH, Pollinger JP *et al.* (2009) Landscape genetics of California mule deer (*Odocoileus hemionus*): the roles of ecological and historical factors in generating differentiation. *Molecular Ecology*, **18**, 1848–1862.

Raymond M, Rousset F (1995) GENEPOP (VERSION-1.2) Population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.

Roon DA, Waits LP, Kendall KC (2005) A simulation test of the effectiveness of several methods for error-checking non-invasive genetic data. *Animal Conservation*, **8**, 203–215.

Schwartz MK, McKelvey KS (2008) Why sampling scheme matters: the effect of sampling scheme on landscape genetic results. *Conservation Genetics*, **10**, 441–452.

Schwartz MK, Copeland JP, Anderson NJ *et al.* (2009a) Wolverine gene flow across a narrow climatic niche. *Ecology*, **90**, 3222–3232.

Schwartz MK, Luikart G, McKelvey KS *et al.* (2009b) Landscape genomics: A brief perspective. In:*Chapter 19: Spatial Complexity, Informatics and Animal Conservatio* (eds Cushman SA, Huettman F). pp. 19–42, Springer, Tokyo.

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

Smouse PE, Chakraborty R (1986) The use of restriction-fragment-length-polymorphisms in paternity analysis. *American Journal of Human Genetics*, **38**, 918–939.

Sokal RR, Rohlf FJ (1995) *Biometry: The Principles and Practice of Statistics in Biological Research*, Freeman Publishing, New York.

Storfer A, Murphy MA, Evans J *et al.* (2007) Putting the landscape in landscape genetics. *Heredity*, **98**, 128–142.

Storfer A, Murphy MA, Spear SF *et al.* (2010) Landscape genetics: where are we now? *Molecular Ecology*, **19**, 3496–3514.

Stutchbury BJM (2007) The effects of habitat fragmentation on animals: gaps in our knowledge and new approaches. *Canadian Journal of Zoology*, **85**, 1015–1016.

Telles MPC, Diniz-Filho JAF, Bastos RP *et al.* (2007) Landscape genetics of *Physalaemus cuvieri* in Brazilian Cerrado: correspondence between population structure and patterns of human occupation and habitat loss. *Biological Conservation*, **139**, 37–46.

Woods JG, Paetkau D, Lewis D *et al.* (1999) Genetic tagging of free-ranging black and brown bears. *Wildlife Society Bulletin*, **27**, 616–627.

© 2011 Blackwell Publishing Ltd