2013

Cougar genetic variation and gene flow in a heterogeneous landscape

Matthew J. (Matthew James) Warren
Western Washington University

Follow this and additional works at: https://cedar.wwu.edu/wwuet

Part of the Environmental Sciences Commons

COUGAR GENETIC VARIATION AND GENE FLOW IN A HETEROGENEOUS LANDSCAPE

By

Matthew Warren

Accepted in Partial Completion
Of the Requirements for the Degree
Master of Science

Kathleen L. Kitto, Dean of the Graduate School

ADVISORY COMMITTEE

Chair, Dr. David Wallin

Dr. Andrew Bunn

Dr. Kenneth Warheit

MASTER'S THESIS

In presenting this thesis in partial fulfillment of the requirements for a master's degree at Western Washington University, I grant to Western Washington University the nonexclusive royalty-free right to archive, reproduce, distribute, and display the thesis in any and all forms, including electronic format, via any digital library mechanisms maintained by WWU.

I represent and warrant this is my original work, and does not infringe or violate any rights of others. I warrant that I have obtained written permissions from the owner of any third party copyrighted material included in these files.

I acknowledge that I retain ownership rights to the copyright of this work, including but not limited to the right to use all or part of this work in future works, such as articles or books.

Library users are granted permission for individual, research and non-commercial reproduction of this work for educational purposes only. Any further digital posting of this document requires specific permission from the author.

Any copying or publication of this thesis for commercial purposes, or for financial gain, is not allowed without my written permission.

<div align="right">

Matthew J. Warren

July 12, 2013

</div>

COUGAR GENETIC VARIATION AND GENE FLOW IN A HETEROGENEOUS LANDSCAPE

A Thesis
Presented to
The Faculty of
Western Washington University

In Partial Fulfillment
Of the Requirements for the Degree
Master of Science

by
Matthew J. Warren
July 2013

ABSTRACT

Management of game species requires an understanding not just of population abundance, but also the structure of and connections between populations. Like other large-bodied carnivores, the cougar (*Puma concolor*) exhibits density –dependent dispersal and is capable of long-distance movement; in the absence of barriers to movement, these traits should lead to high connectivity between individuals and a lack of genetic differentiation across areas of continuous habitat. Previous research has suggested that cougar movement may be influenced by landscape variables such as forest cover, elevation, human population density, and highways. I assessed the population structure of cougars (*Puma concolor*) in Washington and southern British Columbia by examining patterns of genetic variation in 17 microsatellite loci, and the contribution of landscape variables to this genetic variation.

I evaluated population structure using genetic clustering algorithms and spatial principal components analysis. I quantified the effect of distance on genetic variation by calculating the correlation between the genetic distance and geographic distance between every pair of individuals, as well as the spatial autocorrelation of genetic distances. To compare the observed pattern of genetic differentiation with that which would arise solely from isolation by distance, I simulated allele frequencies across the study area where the cost to movement between individuals was proportional to the distance between them. I also evaluated the support for evidence of male-biased dispersal in allele frequencies. Bayesian clustering analyses identified four populations in the study area, corresponding to the Olympic Peninsula, Cascade Mountains, northeastern Washington and Blue Mountains; these clusters were supported by patterns of genetic differentiation revealed with spatial PCA.

Although I found a significant relationship between the geographic and genetic distance between individuals, simulated allele frequencies displayed no meaningful spatial pattern of differentiation, suggesting that male dispersal would be adequate within the scale of the study area to prevent genetic isolation from occurring if the only factor to affect dispersal was geographic distance.

While cougars are capable of long-distance dispersal movements, dispersal in heterogeneous landscapes may be mediated by the resistance of the landscape to movement. I derived resistance surfaces for forest canopy cover, elevation, human population density and highways based on GIS data and estimated the landscape resistance between pairs of individuals using circuit theory. I quantified the effect of the resistance to movement due to each landscape factor on genetic distance using multiple regression on distance matrices and boosted regression tree analysis. Both models indicated that only forest canopy cover and the geographic distance between individuals had an effect on genetic distance, with forest cover exhibiting the greatest relative influence.

The boundaries between the genetic clusters I found largely corresponded with breaks in forest cover, showing agreement between population structure and landscape variable selection. The greater relative influence of forest cover may also explain why a significant relationship was found between geographic and genetic distance, yet geographic distance alone could not explain the observed pattern of allele frequencies. While cougars inhabit unforested areas in other parts of their range, forested corridors appear to be important for maintaining population connectivity in the northwest.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

Chapter 1

Chapter 2

LIST OF FIGURES

Chapter 1

Chapter 2

Cougar spatial genetic variation and population structure in Washington and southern British Columbia

INTRODUCTION

The cougar (*Puma concolor*) is an apex predator in Washington, where it is managed as a game species, and sport hunting is currently used as the primary mechanism of population control (WDFW 2011). Depending on the scale of hunting pressure, however, cougar population density may not be affected, as high mortality in one area may be offset by increased immigration of subadult males from surrounding areas (Robinson et al. 2008). Effective management of cougars therefore requires a better understanding of population structure across the state, specifically whether the state's cougars comprise a single, panmictic population, or function as a metapopulation.

Variation in selectively-neutral regions of the genome, such as microsatellites, can be used to infer population structure. Like other large-bodies carnivores, cougars are capable of dispersing across long distances, generally resulting in continuous populations across areas of suitable habitat (Anderson et al. 2004). Genetic differentiation may then result as a function of distance, where individuals separated by greater distances are less closely related than individuals in close proximity to each other; this phenomenon is referred to as isolation by distance (Wright 1943). Cougars typically disperse upon becoming independent from their mother, between 1 and 2 years of age (Logan and Sweanor 2010). Subadult males are more likely than subadult females to disperse, and generally travel farther than females (Logan and Sweanor 2000; Sweanor et al. 2000); in Washington, the average male dispersal distance is

190 – 250 km (R. Beausoleil, personal communication). This propensity for long-distance dispersal may be enough to prevent a pattern of isolation by distance from being observed in Washington, however isolation by distance has been reported in cougars from California and Wyoming (Ernest et al. 2003; Anderson et al. 2004).

Inbreeding in small, isolated populations can result in a loss of heterozygosity, and over time deleterious alleles may accumulate in the population, resulting in losses of fitness. Evidence of inbreeding depression was clearly seen in Florida panthers (*Puma concolor coryi*), an isolated subspecies of the cougar, including sperm defects, undescended testicles, and heart defects (Roelke et al. 1993). The shrinking population and low reproduction success ultimately led U.S. Fish and Wildlife biologists to translocate eight female cougars from Texas to South Florida, to reverse decades of inbreeding (Pimm et al. 2006). One of only three previously identified genetic bottlenecks in North American cougars comes from the Olympic Peninsula (Culver et al. 2000), and the low level of heterozygosity observed in Olympic cougars led Beier (2010) to suggest that reintroductions may be needed to ward off inbreeding depression. Connectivity between individuals on the Olympic Peninsula and the nearby Cascade Mountains remains unresolved.

Discrete populations can be identified and delimited using genetic clustering algorithms, which assign individuals to clusters by minimizing Hardy-Weinberg and linkage disequilibria within groups (Guillot et al. 2005). Allele frequencies in a population at Hardy-Weinberg equilibrium do not change from one generation to the next, allowing genotype frequencies to be estimated based on the squared sum of the frequency of alleles. The population's heterozygosity, which refers to having two different alleles at a particular locus,

can then be calculated, and provides an indication of the amount of genetic variability in the population. Hardy-Weinberg equilibrium is an idealized state where natural selection exerts no influence on any locus being considered, there is no mutation or migration of individuals into the population, population size is infinite and mating is random. Linkage disequilibrium is the non-random association between alleles at two or more different loci, and can be an indication of non-random mating or population structure within an area (Freeland 2006).

The primary aim of this study was to describe patterns of genetic variation across the state, in order to reveal the underlying population structure in the region. This was accomplished using cluster analysis and spatial principal components analysis (PCA). Based on these analyses, I sought to clarify the status of Olympic Peninsula cougars, and reevaluate their isolation from the rest of Washington and British Columbia. I also looked for genetic evidence of male-biased dispersal. Finally, to compare the spatial genetic variation which would arise solely from isolation by distance with the observed variation, I simulated allele frequencies across the study area where distance between individuals was the only cost to movement.

METHODS

Sample collection and genotyping

The Washington Department of Fish and Wildlife (WDFW) collected 612 muscle or tissue samples from cougars across the state of Washington between 2003 and 2010. Additionally, 55 samples were obtained from the British Columbia Ministry of Forests, Lands and Natural Resource Operations. Samples were taken from animals that were killed by hunters, removed

in response to public safety concerns or livestock depredation, or from live research subjects. All genotyping was performed by the WDFW molecular genetics laboratory in Olympia, Washington. DNA was extracted from blood and tissue samples using DNeasy blood and tissue isolation kits. Polymerase chain reaction was used to amplify 18 previously characterized microsatellite markers (Menotti-Raymond and O'Brien 1995; Culver 1999; Menotti-Raymond et al. 1999). PCR products were visualized with an ABI3730 capillary sequencer (Applied Biosystems) and sized using the Gene-Scan 500-Liz standard (Applied Biosystems). Locations of each animal were plotted based on hunter descriptions and added to a point shapefile; locations are considered accurate within 10 km (Figure 1).

I checked for amplification and allele scoring errors using Microchecker version 2.2.3 (van Oosterhout et al. 2004). I used Genepop version 4.1 (Rousset 2008) to test for deviations from Hardy-Weinberg and linkage equilibria; alpha was adjusted using a simple Bonferroni correction to accommodate multiple tests (Rice 1989).

Study area

The study area included all of Washington state and a portion of southern British Columbia (Figure 1). It was comprised of ten ecoregions: Columbia Mountains/Northern Rockies, North Cascades, Eastern Cascades slopes and foothills, Blue Mountains, Pacific and Nass Ranges, Strait of Georgia/Puget Lowland, Coast Range, Willamette Valley, Thompson-Okanogan Plateau, and Columbia Plateau (Wiken et al. 2011). Elevation ranged from 0 to 4,392 m above sea level. Human population density varied considerably across the study

area, ranging from roadless wilderness to the metropolitan areas of Seattle, Tacoma, and

Spokane, WA.

Cluster analysis

I used two Bayesian clustering programs to explore patterns of population structure within

the study area. Geneland version 3.3.0 (Guillot et al. 2005) estimates the number of clusters

within the global population and assigns individuals to clusters by minimizing Hardy-

Weinberg and linkage disequilibria within groups. The geographic coordinates of each

individual are included in their prior distributions (Guillot et al. 2005). I used the spatial

model with null alleles and uncorrelated allele frequencies. The uncertainty attached to the

coordinates for each individual was 10 km, a maximum of 10 populations was assumed, and

10^6 iterations were performed, of which every 100th observation was retained.

I used Structure version 2.3.4 (Falush et al. 2003) without prior location information

to see if patterns of cluster assignment changed when based solely on allele frequencies. I

used the admixture model with correlated allele frequencies, a burn-in period of 10^5

repetitions followed by 10^6 Markov Chain Monte Carlo repetitions. Both Structure and

Geneland assume that discrete subpopulations exist in the study area, and that allele

frequencies in these subpopulations are in Hardy Weinberg and linkage equilibria.

Spatial PCA

Clustering algorithms are designed primarily to identify discrete groups of individuals,

therefore I also used spatial PCA to detect clinal population structure. Spatial PCA is a

modified version of PCA where synthetic variables maximize the product of an individual's principal component score, based on allele frequencies, and Moran's *I*, a measure of spatial autocorrelation (Jombart et al. 2008). Spatial autocorrelation is calculated between neighboring points as defined by a connection network. I generated a Gabriel graph to define this connection network, which connects two sample points only if a circle drawn between those points does not include any others. A network based on a Gabriel graph has fewer connections than one based on a Delaunay triangulation, which may connect distant points on the edge of the network, and more connections than one based on a minimum spanning tree (Legendre and Legendre 2012). Unlike Geneland and Structure, spatial PCA does not make assumptions regarding Hardy Weinberg and linkage equilibria, so the results of this analysis are not subject to the same issues of interpretability as those of the above clustering programs. Spatial PCA breaks spatial autocorrelation into global structure, where neighbors are positively autocorrelated, and local structure, where neighbors are negatively autocorrelated (Jombart et al. 2008).

Descriptive statistics

I calculated total number of alleles, mean number of alleles per locus, number of private alleles, expected heterozygosity (Nei 1987), and observed heterozygosity for each population cluster identified by the Geneland analysis using Microsatellite Toolkit version 3.1.1 (Park 2001). To compare genetic differentiation between clusters I calculated pairwise estimates of F_{ST} (Weir and Cockerham 1984) using Genepop version 4.1. I also used GENALEX v. 6.4 to estimate inbreeding coefficients (F_{IS}) for each cluster (Peakall and Smouse 2006).

Sex-biased dispersal

I looked for evidence of male-biased dispersal in genotype frequencies using Monte-Carlo permutation tests of three common population genetic statistics: assignment index (AI_C), fixation index (F_{ST}), and inbreeding coefficient (F_{IS}) in Fstat version 2.9.3.2 (Goudet 2001; Goudet et al. 2002). I calculated the mean of each test statistic separately for each sex, and then the difference between the means for each sex; genotypes were then randomly assigned a sex and the difference between the means was recalculated. All tests were conducted with 1,000 permutations, and significance was assumed at $P \leq 0.05$. The AIc calculates how likely a given genotype is to occur within that subpopulation, based on subpopulation-specific allele frequencies (Goudet et al. 2002); I used the results of Geneland clustering to define subpopulations. Since male cougars are more likely to disperse than females, males were expected to have a lower mean AIc value, because the genotypes of immigrants have a lower probability of occurring in a subpopulation than those of residents. Variance of AIc values should be highest for males, due to higher immigration. I also tested for differences in genetic differentiation between the sexes, in terms of F_{ST} and F_{IS}. F_{ST} is a pairwise measure of the genetic differentiation between subpopulations, relative to total genetic variance. Immigration results in greater mixing of alleles, therefore F_{ST} was expected to be lower in males. F_{IS}, on the other hand, describes the fit of genotype frequencies to Hardy-Weinberg conditions; higher immigration rates for male cougars imply that within a single subpopulation, males will be from multiple subpopulations. This subpopulation structure, referred to as the Wahlund effect, manifests itself as a reduction in observed heterozygosity, and should result in higher F_{IS} values for males (Goudet et al. 2002).

Spatial autocorrelation

To determine the scale at which spatial patterns are detectable in allele frequencies, I

constructed a Mantel correlogram using GENALEX v. 6.4 (Peakall and Smouse 2006). The

correlogram was based on a geographic distance matrix derived from the Euclidean distance

between the coordinates of each individual, and the genetic distance between individuals,

calculated as Peakall and Smouse's r (Peakall and Smouse 2006). I used Sturges' Rule to

determine the number of distance classes (*D*) to use in the correlogram based on the range of

pairwise distances (*R*) and the sample size (*n*):

$$D = \frac{R}{1 + \log_2 n}$$

(Sturges 1926). This resulted in 10 distance classes which were each 60 km wide. This

analysis was then repeated separately for each sex.

Isolation by distance

I tested for an effect of geographic distance on genetic distance using a simple Mantel test in

the ecodist package for R with 1,000 permutations (Goslee and Urban 2007). Euclidean

distances between the coordinates for every pair of individuals were calculated in R. I used

principal components analysis (PCA) of allele frequencies to calculate genetic distances

between individuals; I created a distance matrix in R derived from the first principal

component scores for each individual (Patterson et al. 2006; Shirk et al. 2010).

Simulation of allele frequencies

To compare the observed pattern of spatial genetic variation with that which would arise

under a scenario of isolation by distance, I used CDPOP version 1.2.08 (Landguth and

Cushman 2010) to simulate allele frequencies at 17 microsatellite loci across the study area.

Initial allele frequencies were randomized and the simulation was run for 6,000 years to

approximate genetic exchange during the late Holocene. The population began with 3,000

individuals; this number was based on the most recent estimate of juveniles and adults in

Washington (WDFW 2011), and extrapolated to include kittens and individuals in south-

central British Columbia. In addition to the locations of the 667 samples described above, I

generated 2,333 random locations in forested regions of the study area to serve as starting

locations for potential mates. I used a Euclidean distance matrix as the cost matrix for mating

movement and dispersal, where mating was random for pairs within 55 km of each other,

based on the average female home range size in southeastern British Columbia (Spreadbury

et al. 1996). Dispersal for both sexes was defined as the inverse square of the cost distance,

with thresholds based on average dispersal distances: 32 km for females (Anderson et al.

1992; Sweanor et al. 2000; Maehr et al. 2002), and 220 km for males (R. Beausoleil, personal

communication). Females began reproducing at 2 years of age and the number of offspring

was drawn from a Poisson distribution with a mean of 3. Age-specific mortality rates were

taken from published values in a lightly-hunted population in central Washington (Cooley et

al. 2009). I repeated the above clustering and sPCA analyses on the ending allele frequencies

for comparison with the observed data. The relationships between simulated genetic distance

and geographic distance, as well as observed genetic distance, were assessed using simple Mantel tests as described above.

RESULTS

Genotyping

I detected significant homozygote excess at 16 loci when all individuals were pooled into a single population, which could have resulted from the presence of null alleles or genetic structure in the study area, due to the Wahlund effect. Estimated frequencies of null alleles were ≤ 5.1% for all but one locus (FCA293, 13.5%). Geneland clustering revealed multiple populations in the study area (described below). After separating individuals into the clusters indicated by Geneland, the estimated frequency of null alleles at locus FCA293 was still greater than 10% in two of four clusters, therefore this locus was dropped and all subsequent analyses were based on the remaining 17 loci (Table 1).

Eight loci were out of HWE after Bonferroni correction for multiple tests, suggesting that there is not a single, panmictic population in the study area. Concurrent with HWE testing, I detected significant departures from linkage equilibrium in 55 of 152 pairwise comparisons between loci after Bonferroni correction. Seven of 17 loci occur on separate chromosomes or linkage groups and should be considered independent (Menotti-Raymond et al. 1999), while one locus, FCA166, has yet to be mapped. After separating individuals into clusters identified by Geneland, no consistent patterns of linkage or Hardy-Weiberg disequilibria between clusters remained. All 17 retained loci were polymorphic, with between 2 and 9 alleles per locus and 91 total alleles globally.

Cluster analysis

Support was highest for four populations in the study area in the Geneland simulations (Figure 2). The clusters corresponded roughly with the Blue Mountains in southeastern Washington, northeastern Washington, western Washington following the Cascade Mountains, and the Olympic Peninsula (Figure 3). Cluster 1 was geographically isolated and lay across the Columbia River Basin from clusters 2 and 3 (Figure 4a). The boundary between clusters 2 and 3 corresponded with the Okanogan Valley (Figure 4b and 4c), and cluster 3 was separated from cluster 4 by Puget Sound and the I-5 corridor (Figure 4d).

I also clustered the samples using the STRUCTURE program without location information and the number of clusters set to 4. Greater spatial overlap between clusters could be seen in the Structure assignment results compared with those from Geneland (Figure 5). The barplot of probability of population membership shows a sharply defined Olympic Peninsula cluster (Figure 5, in yellow), while the other three clusters transition gradually from one to the next, with several individuals in each cluster having mixed membership in multiple clusters.

sPCA

The first two global sPCA axes explained most of the spatial genetic variation (Figure 7a), and were well differentiated from all other axes (Figure 7b); therefore only these two axes were retained. Additionally, the test for global structure, a Monte Carlo randomization test using 999 permutations, was highly significant ($max(t) = 0.016$, $P = 0.001$; Fig. 8). Local sPCA axes explained little spatial genetic variation (Figure 7a) and were poorly differentiated

from each other (Figure 7b); no evidence of local structure was found (max(t) = 0.0028, P =

0.74; Figure 8b).

The first global sPCA axis displayed strong east-west genetic differentiation across

the study area; the strongest separation between neighboring samples was found along the

Okanogan Valley and edge of the Columbia River Basin (Figure 9). The second global sPCA

axis clearly separated out individuals on the Olympic Peninsula and in the Blue Mountains

from the rest of the state, as well as showing a weak east-west gradient in genetic similarity

in northeastern Washington, coinciding approximately with the Columbia River (Figure 10).

This does not imply that cougars on the Olympic Peninsula and in the Blue Mountains are

closely related to each other, rather that they are strongly differentiated from their nearest

neighbors.

Descriptive statistics

The total and mean number of alleles was highest in the northeast and Cascades clusters,

which also had the highest sample sizes (Table 2). Both expected and observed

heterozygosity were far lower in the Olympic cluster than in all other clusters, indicating

lower genetic diversity in Olympic cougars, and possibly greater isolation of this cluster

(Table 2). Population differentiation (F_{ST}) increased with distance between clusters;

differentiation was lowest between the northeast and Cascades clusters, and highest between

the Olympic and Blue Mountain clusters (Table 3). The geographically adjacent Olympic and

Cascades clusters showed a surprising degree of differentiation ($F_{ST} = 0.145$), in accord with

greater isolation of the Olympic cluster.

Spatial autocorrelation

I detected significant positive spatial autocorrelation in allele frequencies in the first three distance classes, and significant negative spatial autocorrelation in distance classes 4 through 8; this indicates that spatial autocorrelation is positive up to 180 km, and negative beyond 180 km (Figure 6a). The results beyond 480 km should not be interpreted due to high variances resulting from low sample sizes. When samples were separated by sex, the results for males did not differ from those of all samples combined (Figure 6b). For female samples, the 95% confidence interval for the 180 km distance class included 0, indicating that positive spatial autocorrelation was detected only up to 120 km (Figure 6c). This suggests that spatial autocorrelation of allele frequencies occurs over a smaller spatial scale for female cougars.

Sex-biased dispersal

The mean assignment index (AIc) value was significantly lower for males than for females, as would be expected with male-biased dispersal ($P = 0.029$; Table 4). The variance in AIc values was higher for male cougars, however this difference was not significant (Table 3). Also in keeping with male-biased dispersal, male cougars exhibited significantly higher F_{IS} values than female cougars ($P = 0.003$; Table 4). F_{ST} values were lower, though not significantly, for male cougars (Table 4).

Isolation by distance

Genetic distance was positively correlated with geographic distance, however this relationship was fairly weak ($r = 0.33$, $P = 0.001$). Log-transforming geographic distance did

not strengthen the correlation. Low genetic distances were seen at a wide range of geographic distances, indicating that closely related individuals can be found hundreds of kilometers apart from each other (Figure 11).

Simulation of allele frequencies

Simulated genetic distance and geographic distance were not significantly correlated for randomized initial allele frequencies (r = -0.002, P = 0.82). There was a significant correlation between genetic and geographic distance for years 10 -2000, however this relationship never explained more than 2.5% of variation, and after year 2000 this relationship became nonsignificant (Figure 12). Geneland cluster analysis revealed a single population in the study area based on ending allele frequencies (Figure 13). Structure clustering split the proportion of samples equally between clusters for each value of K tested, which coupled with high variances indicated a single population as well. I retained only the first sPCA axis based on the screeplot of the eigenvalues (Figure 14), however there was no clear pattern in genetic differentiation observed in this axis (Figure 15). Simulated genetic distance and observed genetic distance were not significantly correlated (r = 0.023, P = 0.054).

DISCUSSION

My results strongly suggest that cougar populations in Washington and southern British Columbia are structured as a metapopulation, not a single, panmictic population. The results of Geneland clustering largely agreed with those of spatial PCA, showing four

clusters in the study area. The boundaries between these clusters are not sharply defined, as evinced by differences and overlap between clusters identified by Geneland and Structure. Mixed membership in multiple clusters, observed in both Geneland and Structure clustering results, as well as geographic separation between individuals belonging to the same cluster observed in the Structure results, suggest that limited gene flow has been maintained between clusters. Overall, the Structure results imply greater migration in the study area than those of Geneland; this is to be expected, as the algorithm underlying the Structure program is better suited to identifying migrants because it is based wholly on allele frequencies. Geneland clusters samples by breaking the study area into polygons consisting of individuals with similar allele frequencies, as such, a single migrant is more likely to be mixed in with individuals from that particular subpopulation (Guillot et al. 2005). Spatial PCA is also a powerful method for detecting migrants, which would be negatively spatially autocorrelated to neighboring samples, resulting in local, as opposed to global, structure (Jombart et al. 2008). Given that the permutation test for local structure was not significant, and the spatial pattern in allele frequencies detected by the Geneland analysis closely matched that revealed by sPCA, these two methods seem to have produced the most realistic representation of population structure in the study area.

State-wide analyses in Nevada (Musial 2009), Oregon (Andreasen et al. 2012) and California (Ernest et al. 2003) revealed spatially-structured cougar populations, however similar analyses in Wyoming (Anderson et al. 2004) and Utah (Sinclair et al. 2001) did not. Anderson et al. (2004) found less genetic differentiation between cougars in Wyoming than was observed in Washington, yet found a stronger relationship between genetic and

geographic distance (r=0.61, P=0.011). This suggests that although there was an isolation by distance effect, the sparsely-developed Wyoming landscape may be more permeable to movement than that of Washington. Given the lack of differentiation seen in Wyoming, dispersing subadult males may encounter more resistance due to territoriality of resident males in forested habitat than in less suitable, yet less densely-populated shrub-steppe areas.

In Utah, Sinclair et al. (2001) found little evidence of population structure, however this may have been due to sampling design and low sample size. Genetic structure was evaluated using F-statistics where populations were a priori defined by management units, which may not have held any biological relevance, and each unit consisted of only 5 individual samples.

Musial (2009) detected a genetic cline in Oregon cougars where the eastern foothills of the Cascades meet the high desert, separating the state into eastern and western clusters. This closely resembles the pattern of differentiation I observed in the first sPCA axis, and between the Cascades and northeastern clusters in Geneland and Structure clustering, aligning approximately with the Okanogan Valley. Musial (2009) attributed this isolation to unsuitable habitat, characterized by low slope and the lack of vegetative cover, between the eastern and western clusters. Habitat in the Okanogan Valley is similar to that of the clinal region in Oregon, however the width of this unforested corridor in Washington is far narrower, ranging from 17 – 36 km. There is no doubt that cougars are physically capable of crossing this valley, however the frequency which with they do so, the resistance they meet from territorial resident males on the other side, their susceptibility to hunting mortality while

crossing and attempting to establish a new home range, and their probability of successfully mating once across are all unknown.

The population clusters identified here correspond closely to existing Cougar Management Units (CMU) in Washington, with the exception of the Cascades cluster, which is currently divided into 5 CMUs (WDFW 2011). Each CMU has its own population objective and hunting regulations, however the lack of genetic differentiation observed between these 5 CMUs suggests that gene flow between them is high; achieving different population goals within these CMUs may be impractical, as mortality in one CMU may be offset by immigration from nearby units.

Cougars are typically managed at the state level, however this may not be an apposite scale for analysis, as political boundaries often have no ecological relevance. Management agencies could make the most of limited resources for genetic analysis through collaboration with agencies in adjacent states or provinces to establish a consistent sampling procedure and series of genetic markers, so that analyses do not have to stop at the state line. This study shows that cougar populations overlap the international border between northern Washington and southern British Columbia, and likely extend into Idaho and Oregon as well.

The results of the mean AIc and F_{ST} tests provided genetic evidence of male-biased dispersal, however the variance of AIc and F_{ST} tests were non-significant. Goudet et al. (2002) found that the variance of AIc test performs best when dispersal rates are <10%; the propensity for male cougars to disperse from natal areas may have resulted in low power for this particular test. Unlike AIc-based tests, spatial patterns in dispersal and genetic differentiation can diminish the power of the F_{ST} test, particularly when populations are

geographically distant (Goudet et al. 2002). The isolation by distance pattern observed in allele frequencies may have weakened my power to detect differences in F_{ST} between clusters. Furthermore, the extent of positive spatial autocorrelation was less for females than for males, consistent with shorter average dispersal distances for female cougars.

Cougars on the Olympic Peninsula do not appear to be as isolated as previously thought. The Olympic cluster had the lowest mean observed heterozygosity, 0.33, of the four clusters (Table 2); this value was similar to that found by Culver et al. (2000), 0.31, for Olympic cougars. The percentage of polymorphic loci for this cluster, however, was much higher in the present study, 94%, than was previously found (50%; Culver et al. 2000). This difference may be attributable to a disparity in sample size; Culver et al.'s (2000) analysis was based on only four samples, while the Olympic cluster in the present study was comprised of 26 individuals. The Olympic cluster also had the highest inbreeding coefficient (F_{IS}) of any cluster, at 0.078 (Table 2), yet this value was relatively low compared with those reported for small or isolated populations in California (0.03 - 0.20; Ernest et al. 2003) and the Intermountain West (0.036 – 0.227; Loxterman 2010). This evidence suggests that although the Olympic cougar population is small and relatively isolated from the rest of the state, genetic diversity is not as low as originally feared, and translocations do not appear to be necessary at this time.

The dispersal of young male cougars is likely responsible for maintaining population connectivity at the scale of the study area. Accordingly, the population structure observed would be expected to be the result of landscape features impeding dispersal. Boundaries between clusters corresponded with the unforested Columbia River Basin, the Okanogan

Valley, and the I-5 corridor and Puget Sound. None of these barriers appeared to completely preclude dispersal, however, given the overlap between clusters and occurrence of probable migrants.

While I found a significant correlation between genetic distance and geographic distance, distance alone cannot explain the genetic structure observed. Allele frequencies simulated under a scenario of isolation by distance did not result in multiple genetic clusters or a clear spatial pattern of differentiation. Furthermore, if distance were the only factor influencing allele frequencies, then both north-south and east-west genetic clines should be apparent. North-south clines were notably absent, however, even in the Cascades cluster which covers over 480 km from the northern to southern tip. This distance is well over the 180 km threshold at which positive spatial autocorrelation was detectable, and nearly double the average male dispersal distance in Washington. The results of sPCA and Geneland and Structure clustering all indicate population structure within the study area, so some factor(s) other than or in addition to geographic distance must be driving this differentiation. Further analysis is needed to identify the landscape features which impede dispersal and isolate clusters from one another.

TABLES

Table 1. Number of alleles, expected heterozygosity (H_E) and observed heterozygosity (H_O) for 17 cougar microsatellite loci.

Locus	No. of alleles	He	Ho
FCA008	2	0.403	0.357
FCA026	5	0.483	0.417
FCA035	3	0.505	0.446
FCA043	3	0.658	0.582
FCA057	8	0.713	0.673
FCA082	7	0.717	0.671
FCA090	6	0.702	0.615
FCA091	7	0.691	0.649
FCA096	4	0.638	0.610
FCA126	4	0.354	0.348
FCA132	9	0.462	0.413
FCA166	5	0.558	0.485
FCA176	7	0.482	0.438
FCA205	7	0.709	0.647
FCA254	6	0.623	0.560
FCA262	3	0.262	0.239
FCA275	5	0.691	0.648

Table 2. Sample size, number of alleles, expected (H_E) and observed heterozygosity (H_O), and inbreeding coefficient (F_{IS}) for cougar population clusters.

Population	n	Total alleles	Mean alleles/locus	Private alleles	Mean H_E (SD)	Mean H_O (SD)	F_{IS}
Blue Mtns	32	126	3.71	0	0.568 (0.04)	0.534 (0.02)	0.033
Northeast	321	170	5.00	4	0.565 (0.03)	0.549 (0.01)	0.027
Cascades	288	172	5.06	5	0.535 (0.04)	0.498 (0.01)	0.066
Olympic	26	114	3.35	0	0.354 (0.06)	0.325 (0.02)	0.078

Table 3. Estimates of genetic differentiation (F_{ST}) between cougar population clusters.

	Blue Mtns	Northeast	Cascades
Northeast	0.094	--	--
Cascades	0.151	0.036	--
Olympic	0.310	0.205	0.145

Table 4. Permutation test results for sex-biased dispersal in cougars.

	n	Mean AIc	Variance AIc	F_{IS}	F_{ST}
Female	301	0.388	25.02	0.0297	0.0629
Male	308	-0.379	30.04	0.0761	0.0831
P-value		0.029	0.121	0.003	0.188

FIGURES
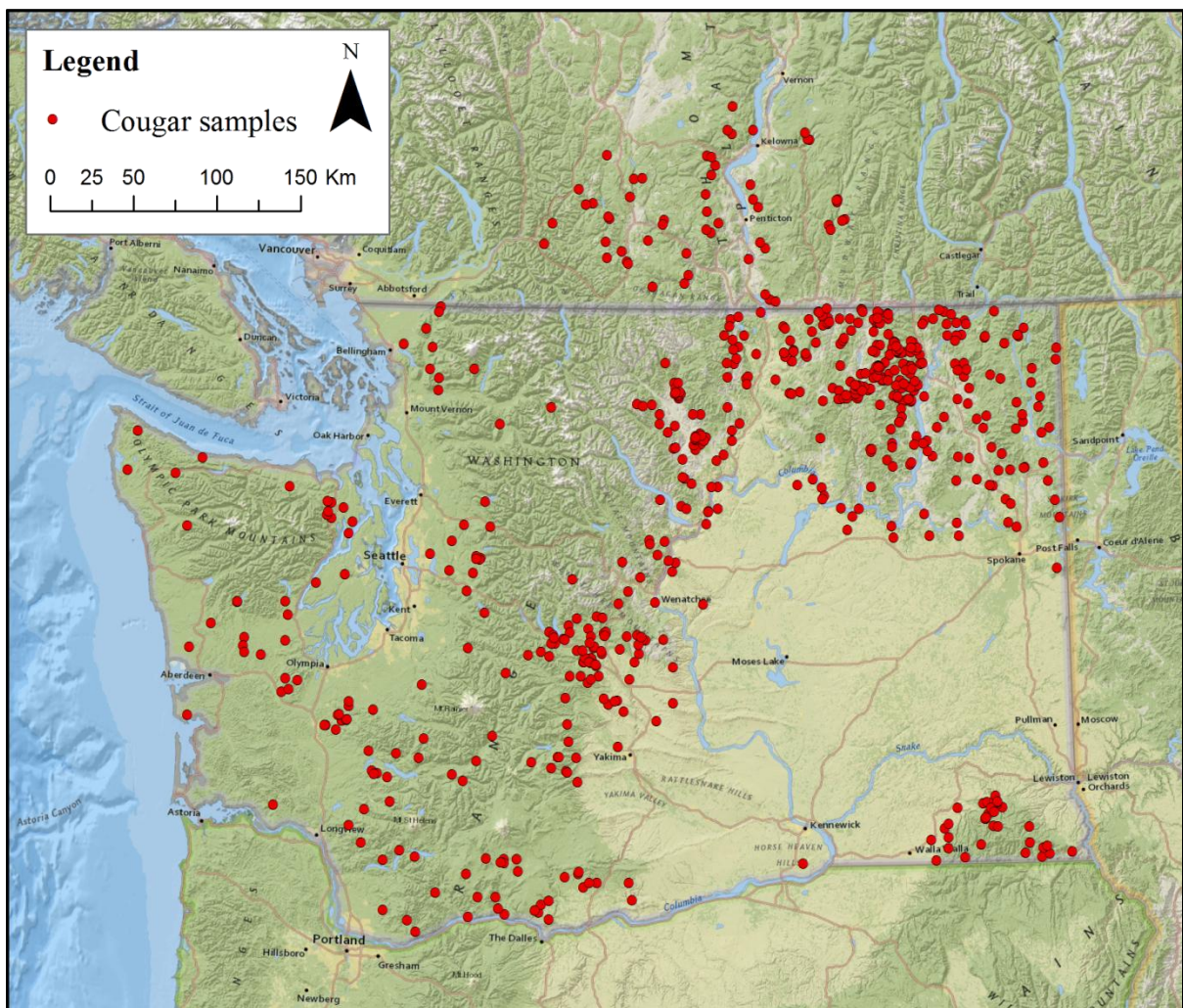
Figure 1. Locations of cougar genetic samples.
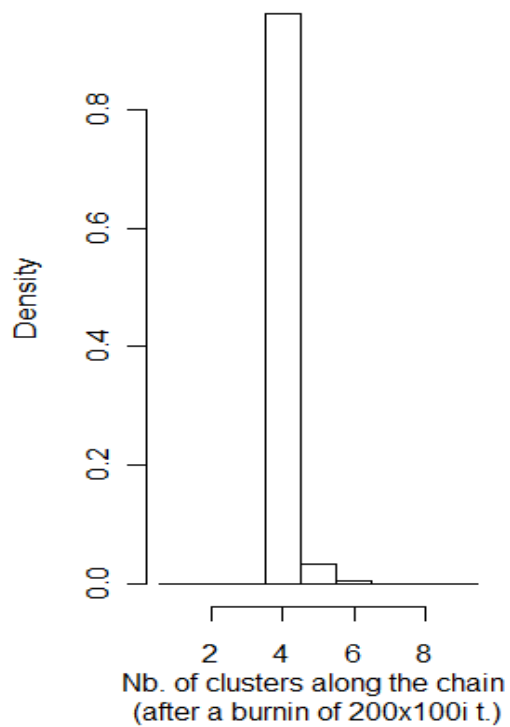
Figure 2. Number of population clusters simulated from the Geneland posterior distribution, after a burn-in of 200 iterations and a thinning interval of 100 iterations. The maximum a posteriori estimate is shown by the clear mode at 4 clusters.
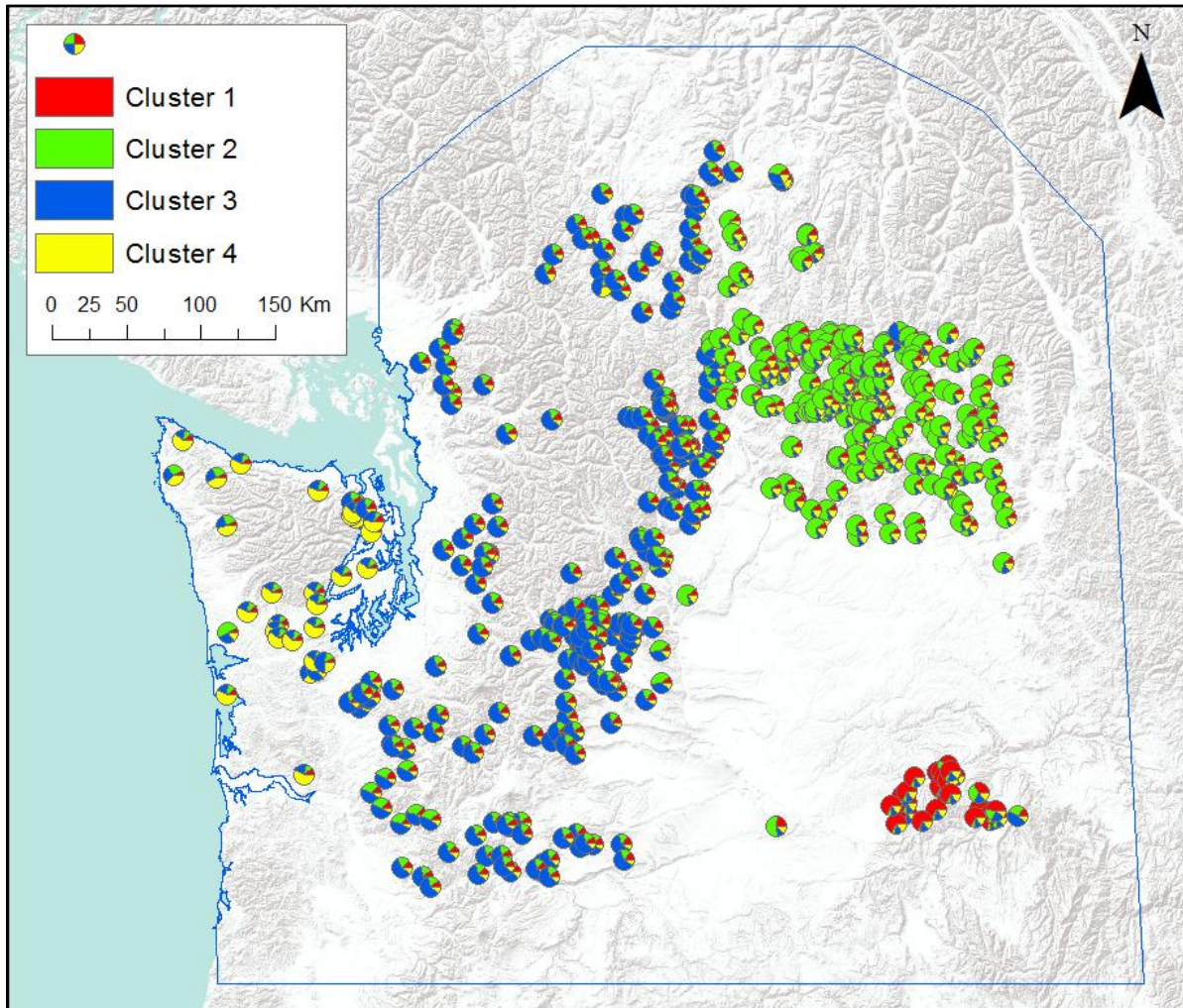
Figure 3. Posterior probability of membership in Geneland clusters.
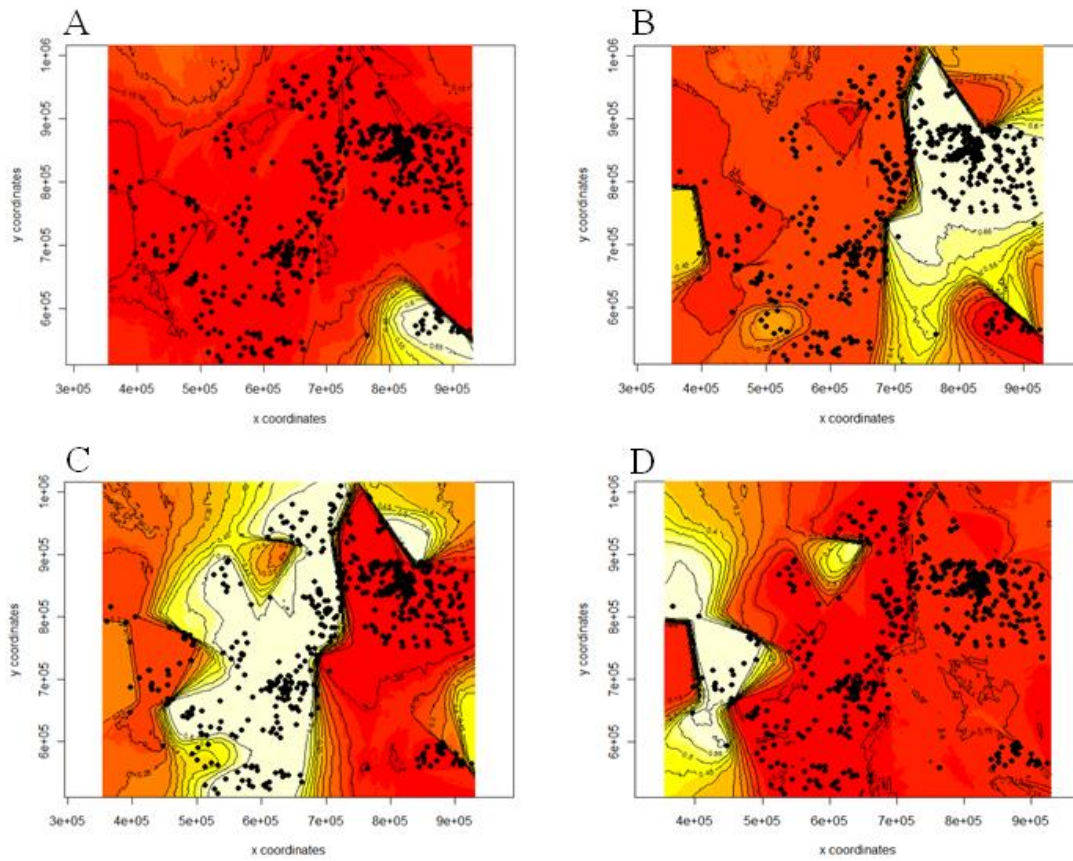
Figure 4. Boundaries between Geneland clusters. The posterior probability of Geneland cluster membership is shown in panels A-D, representing clusters 1-4, respectively, and lighter colors indicate a higher probability of membership to that cluster.
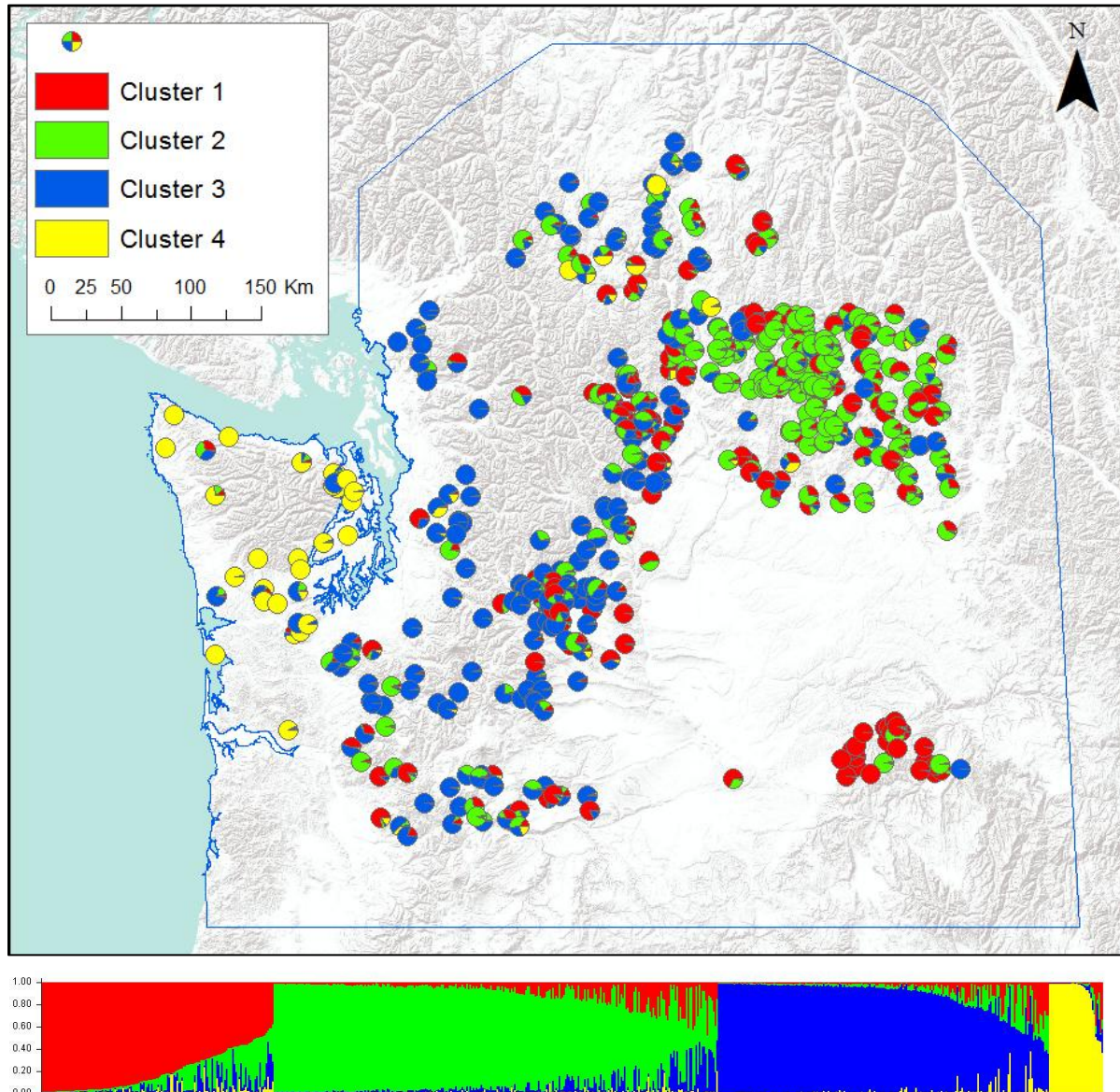
Figure 5. Prior probability of Structure cluster membership for all samples. The bar plot for K=4 is shown below.

(a)

(b)

(c)
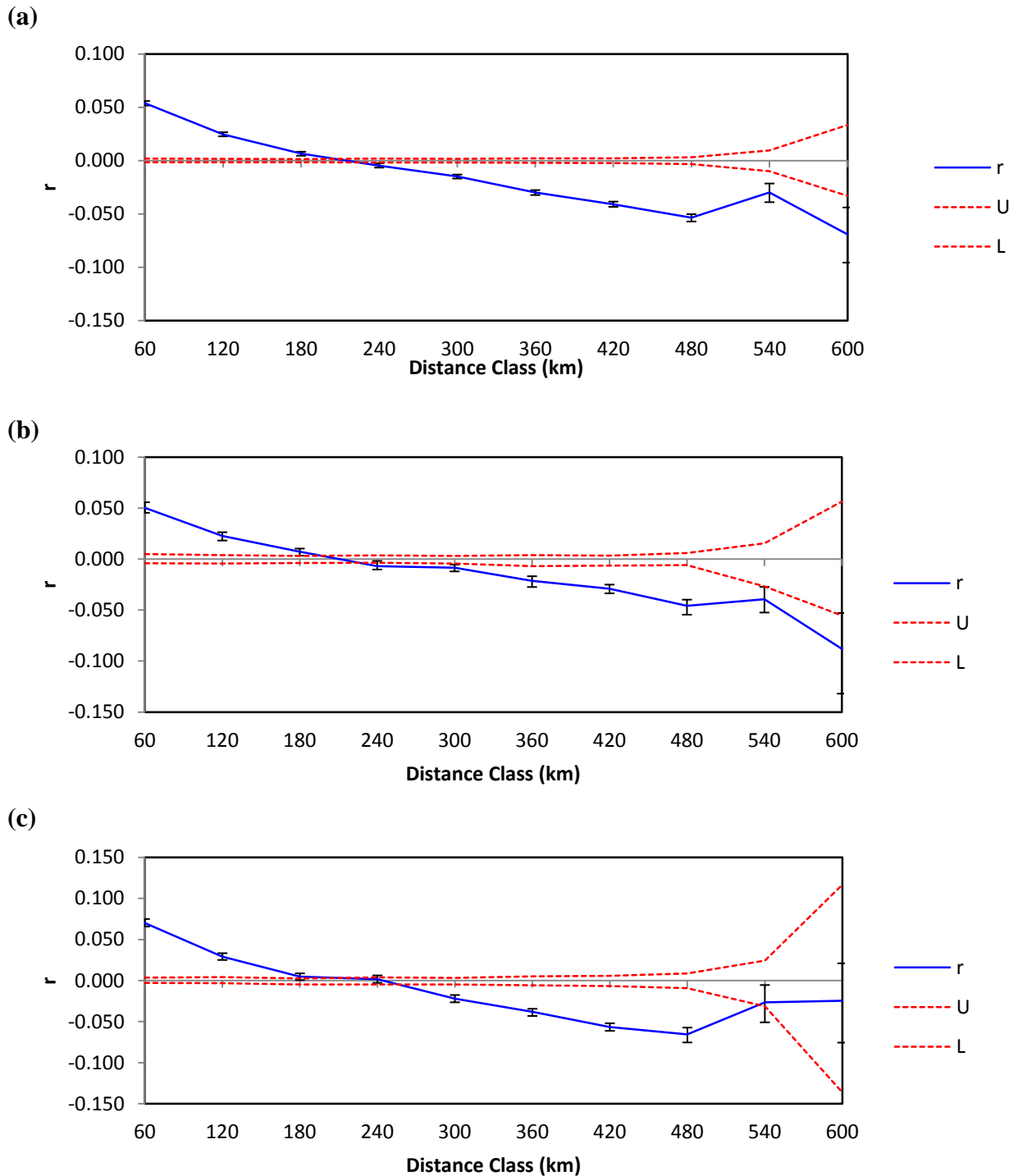
Figure 6. Mantel correlogram showing spatial autocorrelation of allele frequencies for **(a)** all samples, **(b)** male samples and **(c)** female samples. The dashed red lines represent the upper (U) and lower (L) 95% confidence limits of the null hypothesis that there is no spatial structure present in the dataset. The value of Mantel's correlation (r) is shown on the Y axis.
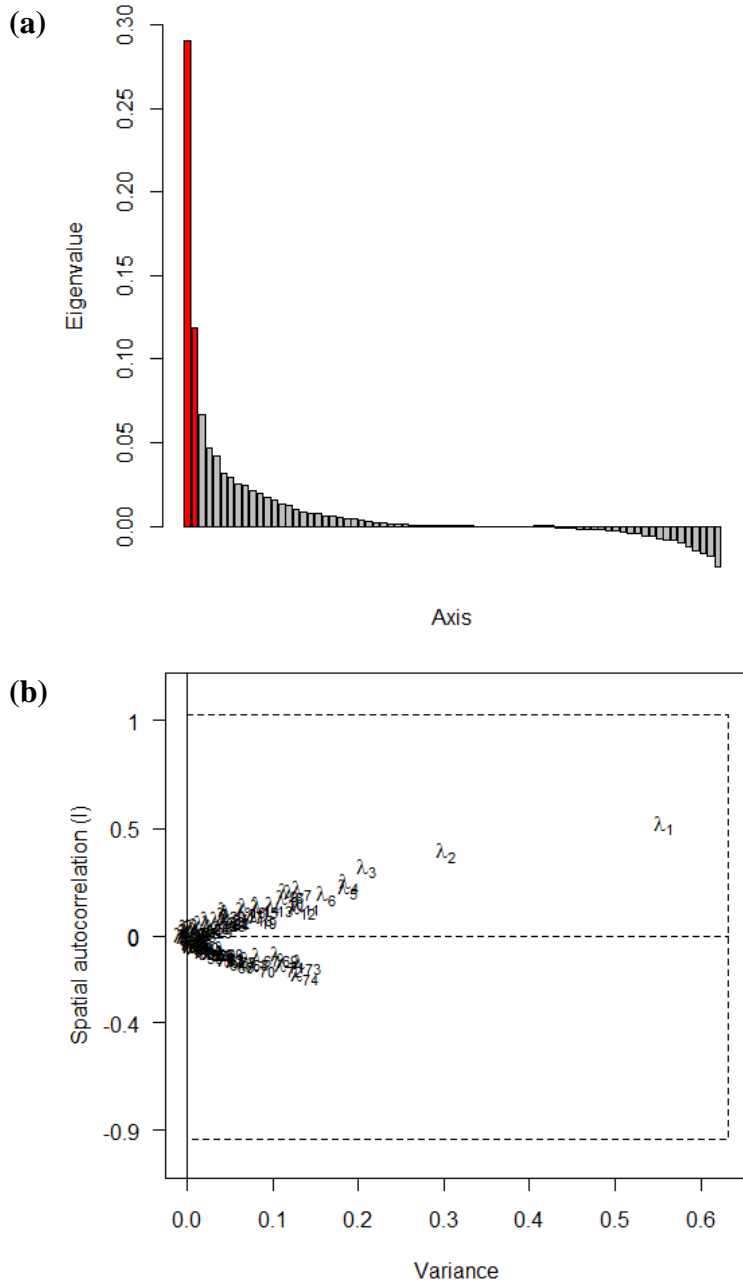
Figure 7. **(a)** sPCA eigenvalues; the first two global axes (on left, in red) were retained while no local axes (on right) were retained. **(b)** Scree plot of the spatial and variance components of the sPCA eigenvalues. Axes 1 and 2 (denoted by λ_1 and λ_2) were well differentiated from all others, therefore only these two were retained.
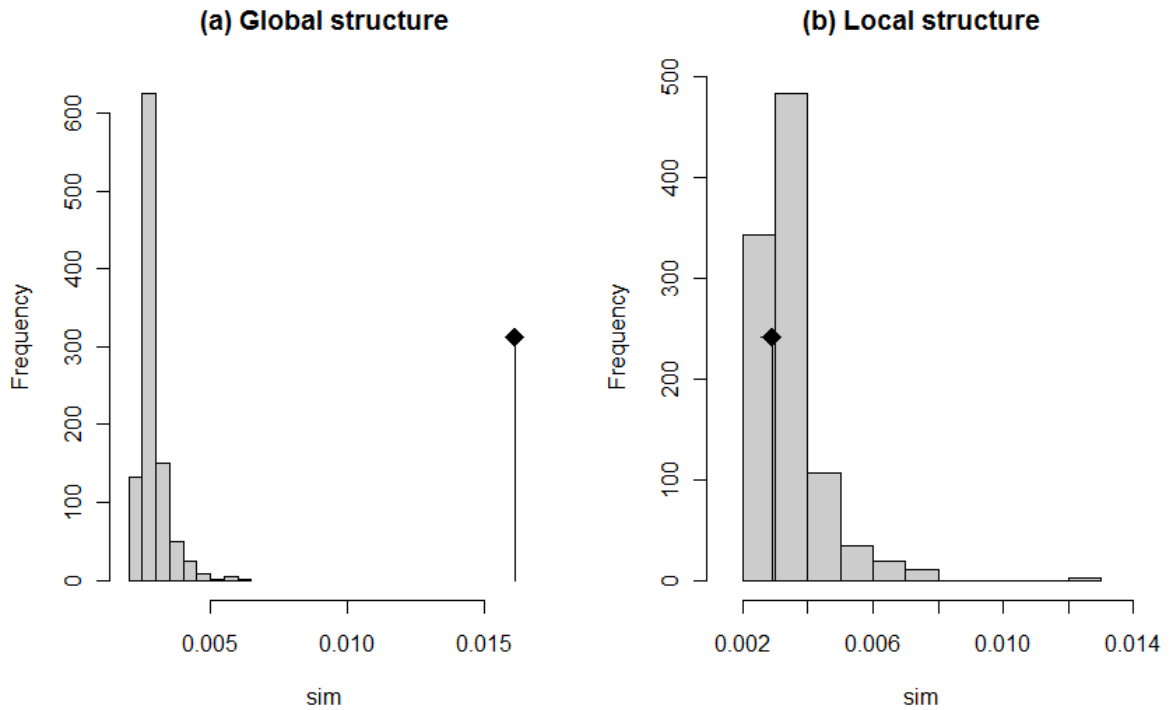
Figure 8. Spatial PCA Monte-Carlo permutation test results for global structure **(a)** and local structure **(b)**, using 999 permutations. The location of the test statistic, max(t), is represented by a black diamond. Significant global structure, or positive spatial autocorrelation, was detected (max(t) = 0.016, P = 0.001). Local structure, or negative spatial autocorrelation, was not detected (max(t) = 0.0028, P = 0.74).
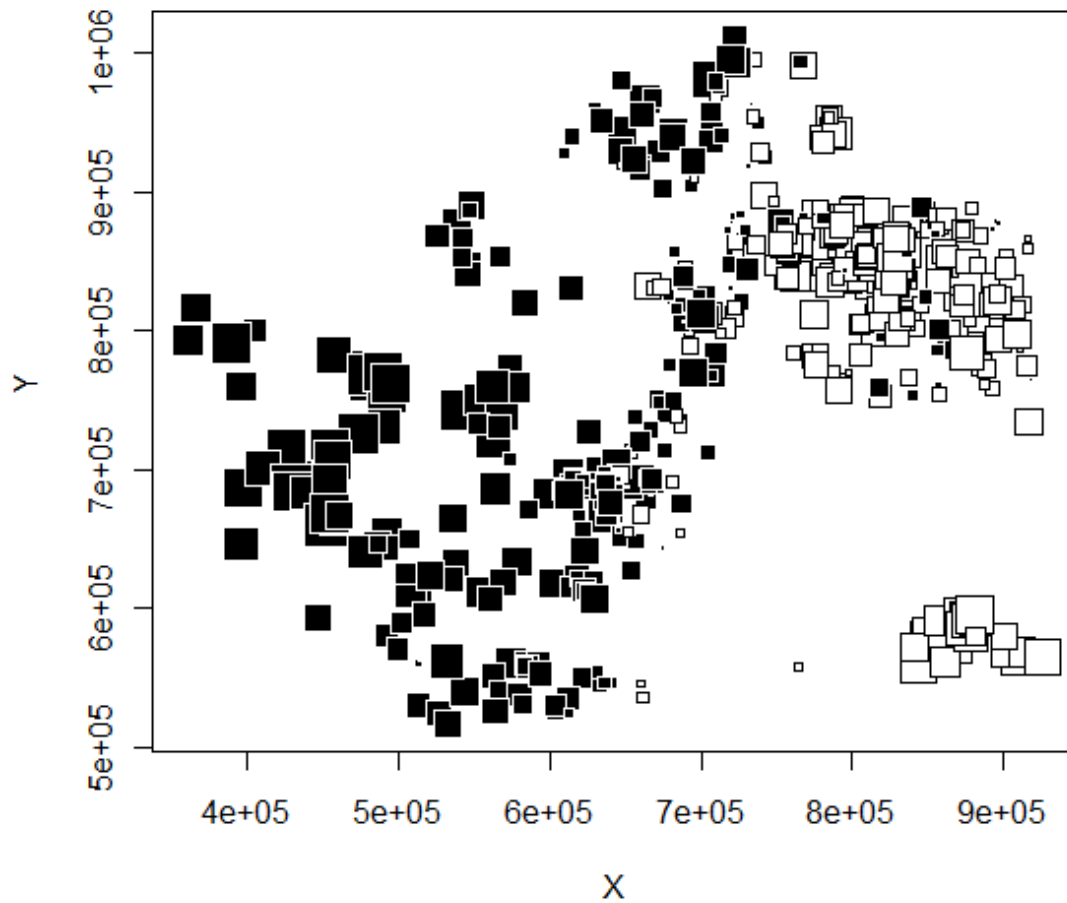
Figure 9. First global sPCA axis. Genetic similarity is represented by color and size of squares, where squares of different color are strongly differentiated from each other, while squares of similar color but different size are weakly differentiated. Geographic coordinates in UTM's are shown on the X and Y axes.

Figure 10. Second global sPCA axis. Genetic similarity is represented by color and size of squares. Geographic coordinates in UTM's are shown on the X and Y axes.
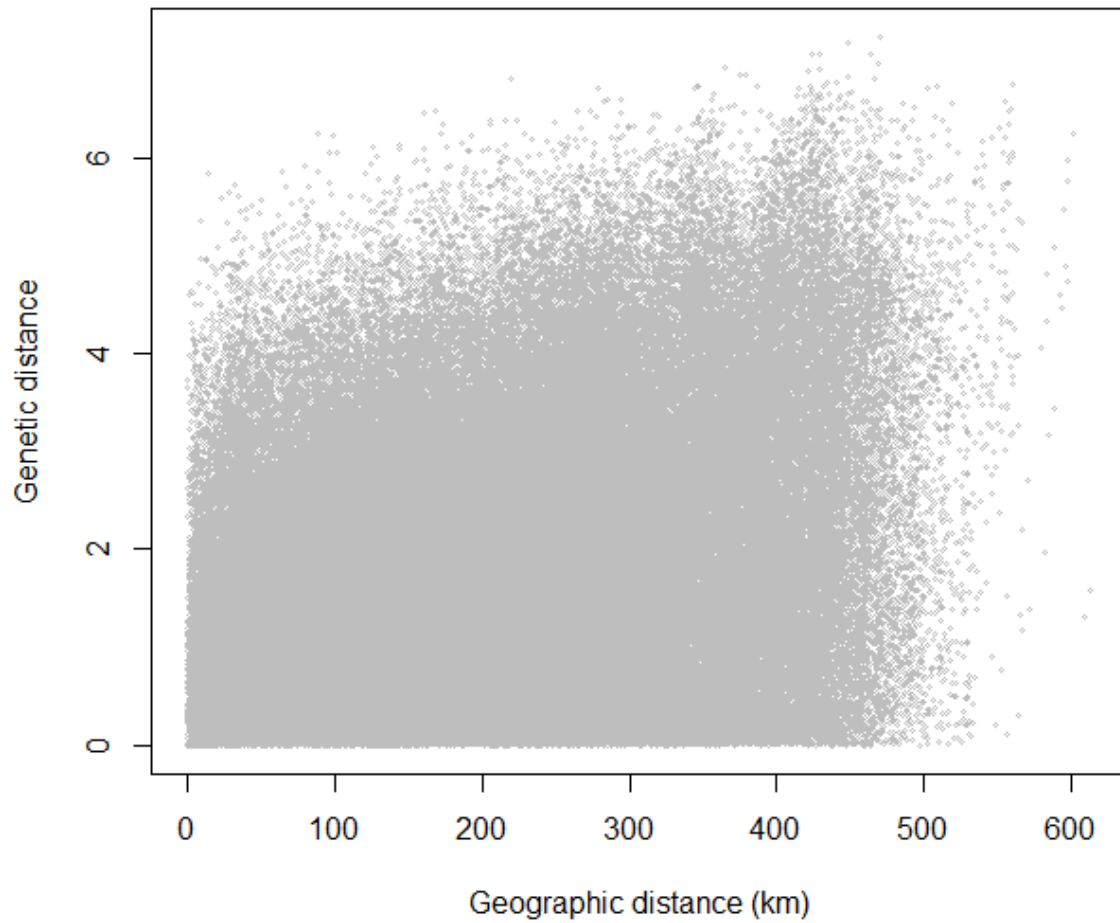
Figure 11. Relationship between geographic and genetic distance. PCA-based genetic distance was derived from the first principal component scores of allele frequencies, and geographic distance was measured as the Euclidean distance between pairs of coordinates for each individual.
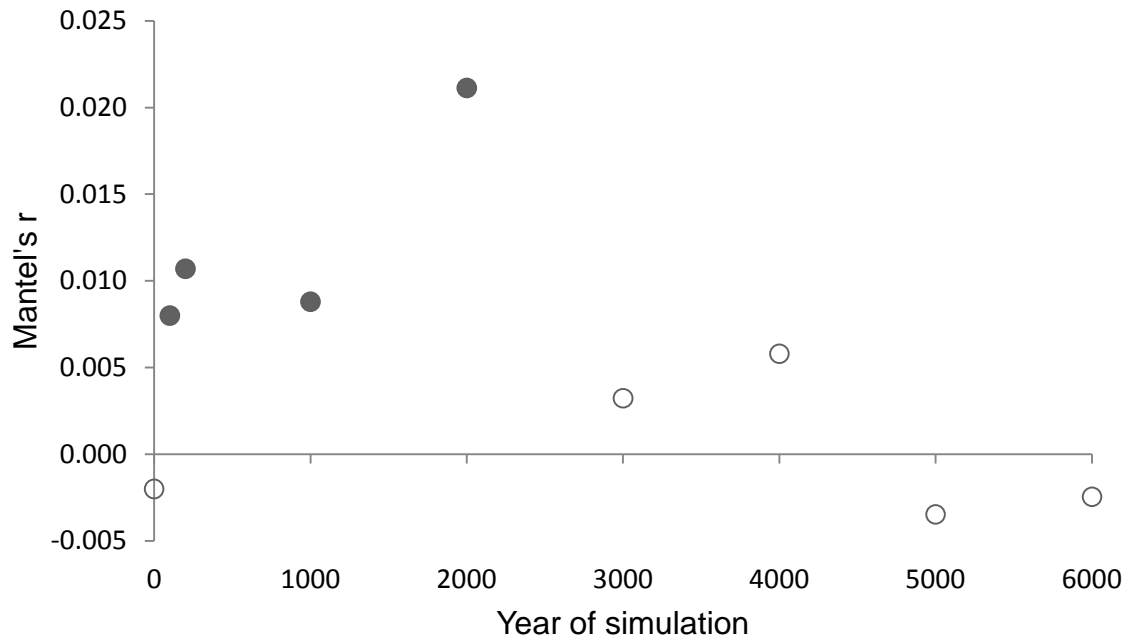
Figure 12. Correlation between genetic and geographic distance for simulated allele frequencies. Closed circles show correlations which were significant at α = 0.05; open circles were not significant.
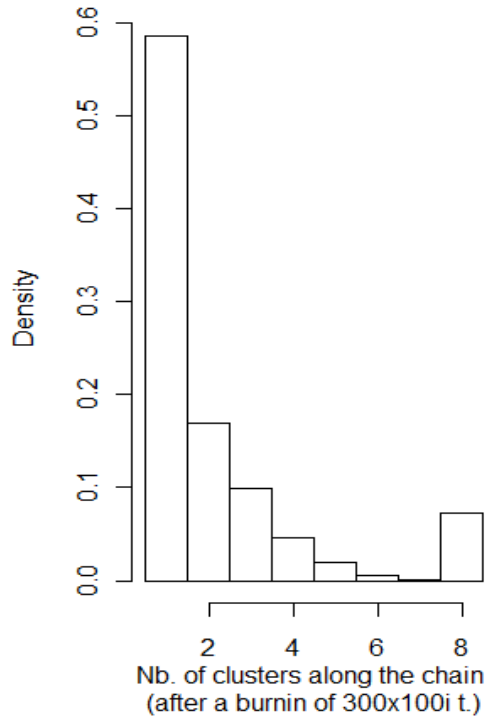
Figure 13. Number of population clusters simulated from the Geneland posterior distribution for simulated allele frequencies, after a burn-in of 200 iterations and a thinning interval of 100 iterations. The maximum a posteriori estimate is shown by the clear mode at 1 cluster.
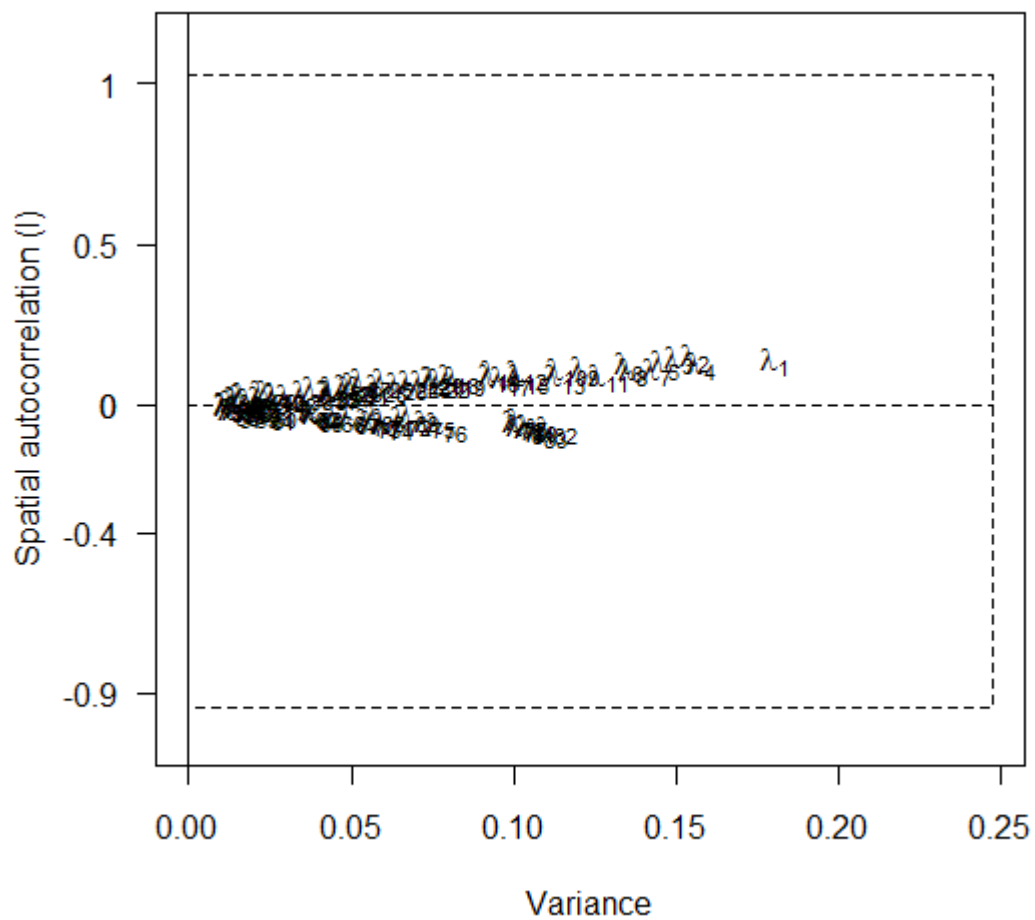
Figure 14. Scree plot of the spatial and variance components of the sPCA for simulated allele frequencies. Axis 1 was well differentiated from all others, therefore only this axis was retained.
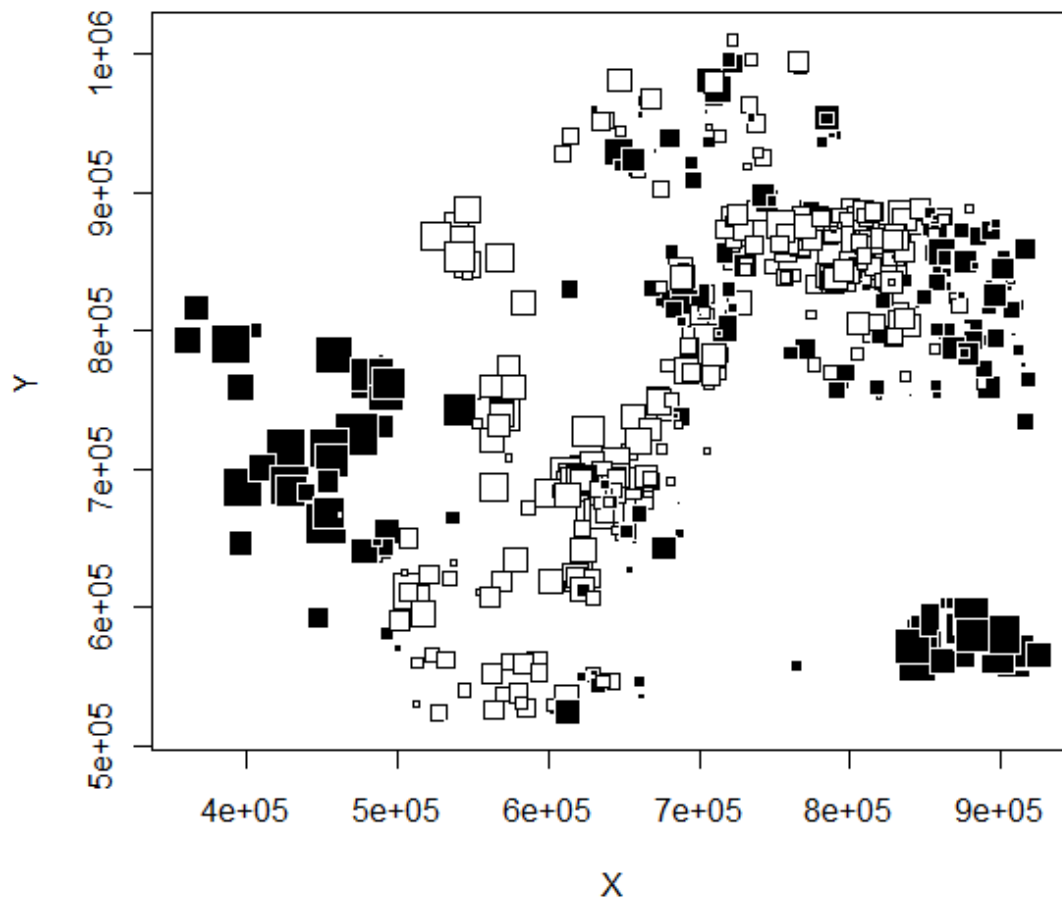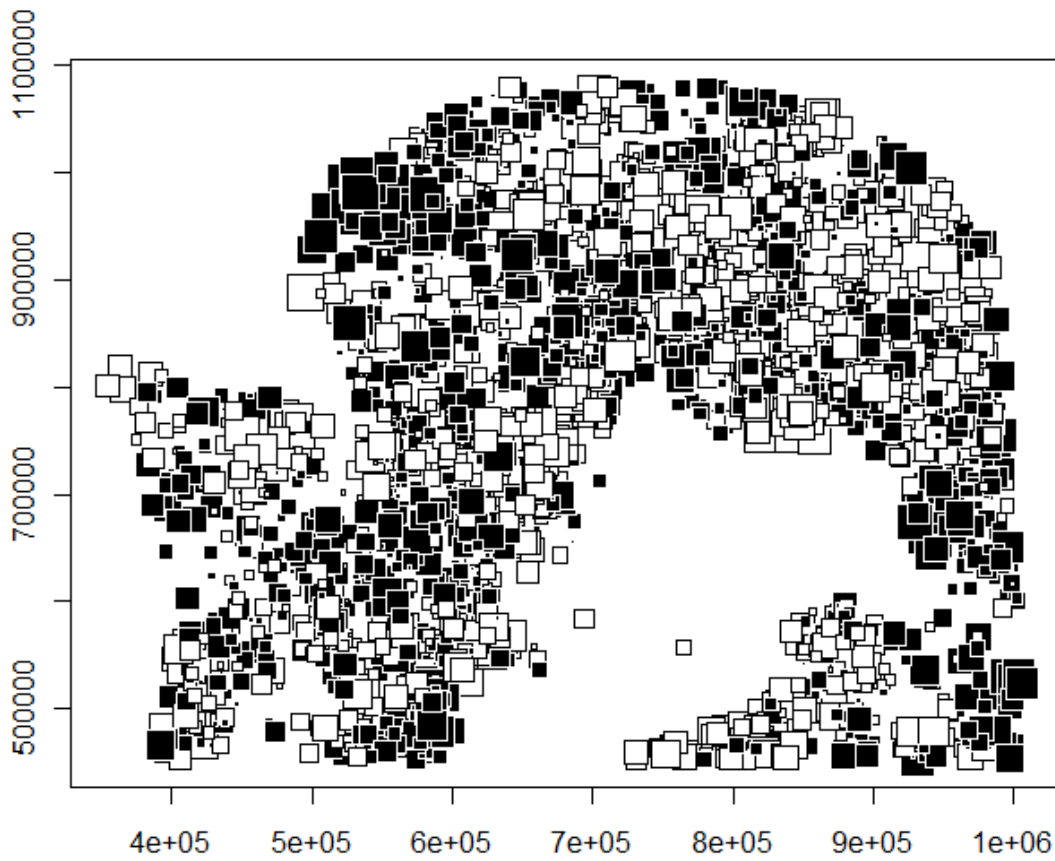
Figure 15. First global sPCA axis for simulated allele frequencies. Genetic similarity is represented by color and size of squares. Geographic coordinates in UTM's are shown on the X and Y axes.

CHAPTER 2

**Cougar gene flow in a heterogeneous landscape**

INTRODUCTION

In heterogeneous landscapes, dispersal may be facilitated or impeded by the resistance to movement inherent to the landscape matrix (Verbeylen et al. 2003). The cougar's reclusive nature and sparse distribution across the landscape present challenges to studying dispersal in this species; much of what is known about habitat use during dispersal comes from small samples of radio-tagged individuals (Beier 1995; Sweanor et al. 2000). However, radio telemetry studies are of limited use because many dispersal events are unsuccessful in the sense that animals die before they reproduce. Unlike radio-telemetry studies, the use of genetic data can provide more information about dispersal in that the genetic structure of a population is a reflection only of successful dispersal events – those that have resulted in successful reproduction – that have occurred over the past few generations (Cushman et al. 2006). Genetic distance, or relatedness, can therefore serve as a proxy for dispersal and can be used to gauge the degree of connectivity between populations.

The primary driver of gene flow in cougars is the dispersal of subadults away from natal areas following independence from their mother between one and two years of age (Logan and Sweanor 2010). Male cougars are more likely to leave their natal area than female cougars, and males generally disperse greater distances than females (Logan and Sweanor 2010). However, in areas of reduced or no hunting pressure gender differences may become trivial (Newby 2011). Factors shown to influence cougar movement include

elevation, slope, terrain ruggedness, landcover, forest cover, high-speed paved roads, human development, and proximity to water (Beier 1995; Dickson and Beier 2002; Dickson et al. 2005; Kertson et al. 2011; Newby 2011). Dispersing subadults in the Rocky Mountains used habitat types similar to those used by resident adults (Newby 2011).

In western Washington, radio-collared cougars selected low elevation areas (Kertson et al. 2011). While there may be some differences between daily movements and dispersal movements, Newby (2011) also found selection for low elevations in dispersing subadults in the Rocky Mountains. Similarly, cougar space use in southern California was highest in canyon bottoms (Dickson and Beier 2007).

Although cougars may cross open areas, they spend the majority of their time in forests with a developed understory, which provides stalking cover and concealment of food caches (Logan and Irwin 1985; Beier 1995). Cougar space use in the Rocky Mountains and western Washington has been positively correlated with forest cover (Kertson et al. 2011; Newby 2011).

Cougars may make use of dirt roads while traveling, however high-speed paved roads pose a serious mortality risk (Taylor et al. 2002; Dickson et al. 2005). Previous studies have provided evidence that highways can reduce gene flow in cougars and other large mammals (McRae et al. 2005; Riley et al. 2006; Balkenhol and Waits 2009; Shirk et al. 2010).

While they have been documented crossing urban areas, most cougars avoid areas of human habitation (Stoner and Wolfe 2012). Additionally, cougars were less likely to use areas lit by artificial street lighting than those that were not (Beier 1995). Cougar space use

has been negatively correlated with residential density in western Washington (Kertson et al. 2011).

The emerging field of landscape genetics focuses on the use of genetic distance between individuals, based on allele frequencies, to evaluate alternative hypotheses regarding landscape features that may influence gene flow (Manel et al. 2003; McRae 2006). For each hypothesis, a landscape resistance surface is derived from GIS data layers. A matrix of "resistance distances" between every pair of individuals is then generated (Spear et al. 2010) using either least cost paths (Cushman et al. 2006) or, more recently, Circuitscape, which uses circuit theory to model all possible dispersal pathways across the landscape (McRae 2006). Circuitscape has the advantage, over least cost path analysis, that it more realistically accounts for the presence of multiple dispersal pathways and the effect of the width of dispersal pathways. The relationship between genetic distance and resistance distance for a given landscape variable can then be tested. Permutation tests are required to determine statistical significance because of the interdependence of elements of a distance matrix (Legendre and Legendre 2012).

The most common approach to relating landscape resistance to genetic distance has been to use partial Mantel tests, however this method has been criticized for having inflated Type I error rates (Raufaste and Rousset 2001; Guillot and Rousset 2013) and performing poorly in distinguishing between multiple correlated distance measures (Balkenhol et al. 2009). Multiple regression on distance matrices has proven more accurate than Mantel tests in simulation studies (Balkenhol et al. 2009), and, unlike the Mantel test, the scale of resistance between the response and explanatory variables does not need to be known

beforehand. A linear relationship between variables is still assumed under multiple regression on distance matrices, and multicollinearity must be checked for. An alternative to linear regression, boosted regression tree analysis is a recently developed machine learning technique that can explain the relative influence of independent variables on a response variable, and is appropriate for nonlinear data (Elith et al. 2008).

My primary objective was to identify the landscape variables which influence gene flow in cougars. I generated resistance surfaces across the study area for four candidate variables based on previous studies of cougar movement and dispersal: elevation (Dickson and Beier 2007; Kertson et al. 2011; Newby 2011), forest canopy cover (Logan and Irwin 1985; Beier 1995; Kertson et al. 2011; Newby 2011), human population density (Beier 1995; Kertson et al. 2011), and highways (Taylor et al. 2002; Dickson et al. 2005; McRae et al. 2005). I estimated resistance between all pairs of cougar sample locations on each of these resistance surfaces using Circuitscape (McRae 2006). Finally, I evaluated the relative influence of these factors on gene flow by examining the relationship between genetic distance and resistance using two different statistical approaches: multiple regression on distance matrices and boosted regression trees.

METHODS

Study area

The study area included all of Washington state, a portion of southern British Columbia, the western edge of Idaho and the northern edge of Oregon (Figure 1). It was comprised of ten ecoregions: Columbia Mountains/Northern Rockies, North Cascades, Eastern Cascades slopes and foothills, Blue Mountains, Pacific and Nass Ranges, Strait of Georgia/Puget

Lowland, Coast Range, Willamette Valley, Thompson-Okanogan Plateau, and Columbia

Plateau (Wiken et al. 2011). Elevation ranged from 0 to 4,392 m above sea level. Human

population density varied considerably across the study area, ranging from roadless

wilderness to the metropolitan areas of Seattle and Tacoma, WA, Vancouver, BC , Spokane,

WA, and the northern edge of Portland, OR.

Sample collection and genotyping

See chapter 1 methods for sample collection and genotyping. Samples from southeastern

Washington, referred to as the Blue Mountain cluster in chapter 1, were excluded from

landscape resistance analysis due to their geographic isolation and the artificial barriers

imposed by the boundaries of the study area, i.e. when calculating landscape conductance

due to forest canopy cover with Circuitscape, current would be forced to travel across

unforested areas of the study area, when in reality dispersing cougars could follow forested

corridors outside of the study area to reach the Blue Mountains. After this cluster was

removed a total of 633 individual samples remained (Figure 1).

Landscape resistance surfaces

I generated landscape resistance surfaces using data layers for elevation, forest canopy cover,

human population density and highways. All GIS layers were projected in a modified Albers

projection (see WHCWG 2010). The untransformed raw values of each layer were rescaled

from 0 to 1 by dividing each cell by the maximum value for that layer; this was done to

standardize resistance estimates and allow for evaluation of the relative importance of each

factor. Circuitscape treats 0 values as no data, therefore I added 1 to each cell, resulting in all layers being scaled from 1 to 2. The resolution of each layer was reduced to 300 m^2 by aggregating cells based on the average cell value to maintain practical Circuitscape computation times. All sample points were at least 70 km from the map boundary, except where boundaries coincided with actual barriers to dispersal, such as Puget Sound; this buffer was used to minimize the risk of overestimating resistance near map edges (Koen et al. 2010).

Elevation

U.S. elevation data was taken from the National Elevation Dataset (USGS 2012). Rasters were downloaded as tiles and mosaicked together. Canadian elevation data came from Terrain Resource Information Management Digital Elevation Model (Crown Registry and Geographic Base 2012). U.S. and Canadian elevation layers were mosaicked together, however some gaps were left along the international border. Gaps were filled in by creating a mask over the problem area and calculating the focal mean for a 5 by 5 rectangle around each cell within the mask (Figure 2).

Forest canopy cover

Forest canopy cover data was downloaded from the Washington Wildlife Habitat Connectivity Working Group (WHCWG 2010). U.S. forest canopy cover was based on Landsat imagery from 1999-2003. Canadian forest canopy cover was based on Landsat imagery from 2000. Forest canopy cover in this dataset was classified into four broad ranges

(nonforest, 0-40%, 40-70%, and 70-100% canopy cover). Each category was reclassified as the median of its range (Figure 3).

Human population density

A residential density layer was downloaded from WHCWG (2010). U.S. residential density was based on census data from 2000. Although more recent census data was available, the 2000 census data may more realistically represent human impacts on cougar populations during the 2001-10 timeframe during which the genetic samples were collected since the genetic structure of the population reflects dispersal and mating events over the past several generations. Canadian residential density was based on census data from 2001. Residential density was classified into ranges based on acres per housing unit; I reclassified each category as the median of its range (Figure 4).

Highways

Rasters for freeways, major highways and secondary highways were downloaded from WHCWG (2010). U.S. roads were based on 2000 U.S. census TIGER roads and the Washington state DNR GIS transportation data layer. Canadian roads were based on Digital Road Atlas data for British Columbia (Figure 5). The final raster was reclassified according to annual average daily traffic volumes for each category of highway (WSDOT 2012).

Resistance to gene flow

I calculated pairwise resistance estimates for each landscape variable between every pair of individuals using Circuitscape version 3.5.8 (McRae et al. 2008). Circuitscape uses circuit theory algorithms to calculate the resistance cost for an individual moving between two points, in this case the coordinates of each genetic sample, based on a user-supplied resistance surface. Landscape resistance is likened to electrical current, allowing for multiple pathways of dispersal, with narrow dispersal corridors presenting higher resistance than wide corridors (McRae 2006). Elevation, human population density and highway traffic volume were run as resistance surfaces, while forest canopy cover was run as a conductance surface, where conductance is simply the reciprocal of resistance (McRae and Shah 2011). Regardless of whether the input is a resistance or conductance surface, the output is always a resistance estimate. I used an eight neighbor, average resistance/conductance cell connection scheme for each grid.

Multiple regression on distance matrices

I used multiple regression on distance matrices (Legendre et al. 1994) to evaluate the relationships between PCA-based genetic distance (see chapter 1) and resistance estimates for each landscape variable. While multiple regression on distance matrices produces coefficients and R^2 values identical to those produced with ordinary multiple regression, significance must be determined using permutation tests because the individual elements of a distance matrix are not independent from one another (Legendre et al. 1994). In order to evaluate the contribution of geographic distance alone, I also included a pairwise distance

matrix based on the Euclidean distance between the coordinates for each genotyped individual, generated using the Ecodist package in R (Goslee and Urban 2007). Each resistance distance matrix was included as a term in a linear model, where genetic distance was the response variable:

$$G \sim R_E + R_F + R_P + R_H + R_G$$

where G = Genetic distance, R_E = Resistance due to elevation, R_F = Resistance due to the reciprocal of forest canopy cover, R_P = Resistance due to human population density, R_H = Resistance due to highways, and R_G = Resistance due to geographic distance

Resistance estimates were z-transformed to standardize partial regression coefficients. P-values were derived from 1,000 random permutations of the response (genetic distance) matrix. All regression modeling was performed using the Ecodist package in R (Goslee and Urban 2007). To remove variables which did not contribute significantly to model fit, I used forward selection with a *P*-to-enter value of 0.05 (Balkenhol et al. 2009).

Geographic distance is a component of all resistance estimates, therefore some correlation was expected between resistance estimates for each landscape variable. Like other forms of linear regression, uncorrelated independent variables are an assumption of multiple regression on distance matrices. I calculated pairwise correlations between all resistance distance matrices using Mantel tests with the Ecodist package in R (Goslee and Urban 2007); I used the Pearson correlation method and significance was based on 1,000 permutations. As a complement to correlation analysis, I calculated the variance inflation factor for each resistance estimate using the Companion to Applied Regression (car) package in R (Fox and

Weisberg 2011); a variance inflation factor greater than 10 generally indicates that terms in a model are too highly correlated (Marquardt 1970).

Boosted regression trees

Regression trees are a nonparametric alternative to linear regression analysis; regression trees repeatedly split the response data into two groups based on a single variable, while trying to keep the groups as homogeneous as possible. The number of splits in the response, referred to as the size of the tree, can be determined by cross-validation, where a sequence of regression trees built on a random subset of the data is used to predict the response of the remaining data. The optimal tree size has the smallest error between the observed and predicted values. When numeric data is split, all values above the split value are placed in one group, while all values below the split value are placed in the other group, making only the rank order of the data important. Monotonic transformations of the explanatory variables, therefore, have no effect on the results of a regression tree; this is particularly advantageous in a landscape genetic framework, where the functional relationships between the response and explanatory variables are rarely known (De' Ath and Fabricius 2000). Simple regression trees are generally used as an exploratory tool to detect patterns in the data, and are well-suited to the noisy nature of landscape genetic relationships (Storfer et al. 2007).

Boosting algorithms aim to reduce the loss in predictive performance of a final model, referred to as deviance, by averaging or reweighting many models. Boosted regression trees minimize deviance by adding, at each step, a new tree that best reduces prediction error. The first regression tree reduces deviance by the greatest amount. The

second regression tree is then fitted to the residuals of the first tree and could contain different variables and split points than the first, and so on. Each tree becomes a term in the model, and the model is updated after the addition of each successive tree and the residuals recalculated (Elith et al. 2008). To avoid overfitting the model, the learning process is usually slowed down by shrinking the contribution of each tree; this learning rate, generally ≤ 0.1, is multiplied by the sum of all trees to yield the final fitted values (De'ath 2007). Stochasticity is introduced to the process by randomly selecting a fraction of the training set, called the bag fraction, to build each successive tree. The relative influence of each predictor variable is measured by the number of splits it accounts for weighted by the squared improvement to the model, averaged over all trees (Elith et al. 2008).

The model with the lowest deviance based on cross-validation consisted of 1,100 regression trees and a learning rate of 0.05. Given that geographic distance is a component of every resistance estimate I chose not to model interactions between predictor variables. In order to accurately model resistance, I constrained conductance/resistance due to forest canopy cover, human population density and highways to increase monotonically with genetic distance. I used the R package tree (Ripley 2013) for simple regression tree analysis and the packages gbm (Ridgeway 2013) and gbm.step (Elith et al. 2008) for boosted regression tree analysis.

RESULTS

Data distribution and trends

Visual inspection of the data indicated that although there was a fair bit of noise, forest canopy cover and geographic distance had a roughly linear relationship with genetic distance, while human population density was potentially logarithmically related to genetic distance (Figure 6). There did not appear to be a strong relationship between either elevation or highways and genetic distance (Figure 6).

Multiple regression on distance matrices

In the global model, resistance due to the reciprocal of forest canopy cover and geographic distance were the only two significant variables, and both variables had a positive relationship with genetic distance (Table 1). Following forward selection, again only these two variables were found to be significant, and the final model explained 14.8% of the variation in PCA-based genetic distance (Table 2). The null hypothesis that there was no relationship between any explanatory variable and PCA-based genetic distance was rejected ($F = 17{,}388.4$, $P = 0.001$; Table 2).

As expected, nearly all resistance estimates were significantly correlated with each other, except for elevation, which was only correlated with geographic distance (Table 3). All Mantel r values were < 0.75 (Table 3). The most highly correlated resistance surfaces were forest canopy cover and human population density (Mantel $r = 0.74$, $P = 0.001$). In contrast to the Mantel test results, all variance inflation factor coefficients were < 4, suggesting that multicollinearity was not a problem (Table 4).

Boosted regression trees

The single regression tree model had two splits, the first on forest canopy cover and the second on geographic distance, forming a total of three groups (Figure 7). This means that when resistance due to the reciprocal of forest canopy cover is greater than 1.2 (the unitless measure of resistance generated by Circuitscape), which was slightly higher than the mean resistance for this variable, data was placed into the first group, which had a mean PCA-based genetic distance of 3.2. To put these numbers into perspective, genetic distance ranged from 0 to 7.5, with a mean of 1.7. When resistance due to the reciprocal of forest canopy cover was less than 1.2, geographic distance became important (Figure 7). When geographic distance was greater than 159.7 km, data was placed into the second group, which had a mean genetic distance of 1.8. Individuals separated by less than 159.7 km fell into the third group, which had the lowest mean genetic distance at 1.3. In other words, when forest canopy cover resistance was high, suggesting that individuals were separated by unforested areas, individuals shared few alleles. When individuals were within forested areas, genetic distance was a function of geographic distance.

The boosted regression tree model explained 19.2% of the deviance in PCA-based genetic distance. Of the explained deviance, resistance due to the reciprocal of forest canopy cover had the highest relative influence on the model (53.6%), followed by geographic distance (31.8%), human population density (8.9%), elevation (3.0%), and highways (2.8%; Figure 8). There was a negative marginal effect of resistance due to the reciprocal of forest cover at low resistances, and a positive marginal effect at high resistances (Figure 9). In other words, high percent canopy cover led to a decrease in genetic distance, and low percent

canopy cover led to an increase in genetic distance. Geographic distance appeared to have an approximately linear relationship with genetic distance up to 400 km; results beyond 400 km are highly variable and questionable, due to the small number of comparisons at this distance range (Figure 9). Resistance due to human population density and highways were flat across much of their range, which is in keeping with their low estimates of relative influence. Resistance due to elevation appeared to have an inverse Gaussian relationship with genetic distance, however the relative influence of this variable was very low, which cautions against drawing inferences based on this weak relationship (Figure 9).

DISCUSSION

The results of multiple regression on distance matrices and boosted regression tree analysis both suggest that forest canopy cover has the strongest influence on gene flow, followed by geographic distance. These factors explain the genetic differentiation I observed in chapter 1 between the Cascades and northeastern clusters, which were separated by the unforested Okanogan Valley. Though the Blue Mountains cluster was not included in landscape resistance modeling, forest cover and geographic distance could both logically have contributed to the differentiation of this cluster from the others, as it is separated from them by the wide shrub-steppe expanse of the Columbia River Basin.

Geographic distance, or the combination of forest cover and geographic distance, could account for the differentiation observed in chapter 1 between the Olympic cluster and the northeastern and Blue Mountains clusters, but not between it and the adjacent Cascades cluster. The boundary between these two clusters is also mostly forested, so the lack of forest

cover does not explain the observed genetic differentiation either. Culver et al. (2000) found evidence for a genetic bottleneck in Olympic cougars, which Beier (2010) attributed to landcover change in the 19th and early 20th centuries, a severe reduction in population due to persecution, or to consistently low diversity due to the peninsular geography which limits dispersal. Other wildlife species on the Olympic peninsula have shown reduced genetic diversity compared with continental populations, including Cope's giant salamander (*Dicamptodon copeii*; Spear et al. 2011) and Roosevelt elk (*Cervus canadensis roosevelti*; Dratch 1983). It is likely that the three factors listed above worked in concert with each other, to further restrict gene flow to an area that had always exhibited reduced genetic diversity.

Human population density did not have a significant effect on genetic distance in the multiple regression on distance matrices model. It appeared to have a curvilinear relationship with genetic distance in both the plotted raw resistance estimates (Figure 6) and the partial dependence function computed during boosted regression tree analysis (Figure 9); this departure from linearity might have explained its lack of significance in linear regression, however this theory is not supported by the relative influence results from boosted regression tree analysis, where population density only accounted for 8.9% of the explained deviance. Furthermore, because resistance estimates are, in part, a function of the distance between individuals, the negative marginal effect observed for both human population density and highways at low resistances may be due to very low resistance between points which are geographically very close together. As the distance, and therefore resistance, between individuals increases, however, these variables can no longer explain the variation in genetic distance. While collared cougars have generally avoided densely-populated urban areas

(Kertson et al. 2011; Newby 2011), urban crossings have been reported, and may be facilitated by greenbelts or riparian corridors (Stoner and Wolfe 2012). This suggests that although cougars may avoid residential areas in their daily movements, they do not present enough of a barrier to dispersal to affect gene flow at large spatial scales. An additional consideration is that the configuration of urban areas within the study area may not have been adequate for testing the effects of population density, due to the majority of urban areas being located directly adjacent to the map boundary (Figure 4). Given that the metropolitan areas of Vancouver, Seattle, and Tacoma border the Strait of Georgia and Puget Sound, their potential to impede cougar dispersal can't be tested because no pairwise comparisons cross these areas (Figure 1).

Spear et al. (2010) stated that the greatest challenge in landscape genetics comes in parameterizing resistance surfaces; when using correlation analysis, such as Mantel tests, the scale of resistance and weighting relative to other factors must be specified a priori. In order to estimate these parameters, a range of values may be tested (Wasserman et al. 2010), or models may be optimized in an iterative process (Shirk et al. 2010). As the number of models being tested increases, however, so does the risk of Type I error, a risk that some consider too high under the Mantel test to begin with (Raufaste and Rousset 2001; Guillot and Rousset 2013). Additionally, as resistances are transformed during optimization, relationships that were not collinear in the raw data could become so, which is a problem when using linear correlation analysis, such as the commonly used Pearson product-moment correlation coefficient within the Mantel test. By using analysis methods that objectively calculate parameter estimates based on the raw raster values, such as multiple regression, these

quandaries can be circumvented, provided that multicollinearity is not present in the initial resistance estimates (Garroway et al. 2011).

I found congruence in variable selection among the multiple regression on distance matrices global model, reduced model after forward selection, and boosted regression trees model. While multiple regression on distance matrices assumes a linear relationship between variables, I reached the same conclusion using boosted regression trees, which makes no such assumptions. This suggests that any nonlinear relationships in my data did not have a strong effect on the results, however this may not be the case for all datasets, highlighting the need for exploratory analysis and inference based on multiple methods. Few landscape genetics studies have taken advantage of recent advances in machine learning techniques (but see Balkenhol 2009; Murphy et al. 2010), yet the flexibility in error distributions and quantification of relative influence inherent to methods such as boosted regression trees and random forests make them well-suited to exploring landscape genetic relationships. Further research should focus on significance testing or some other framework for variable selection using these methods, as well as the effects of spatial autocorrelation on explanatory power. Models can be validated by using multiple approaches and finding congruence in variable selection, as was done here, or by simulating allele frequencies based on resistance surfaces for selected variables (Landguth and Cushman 2010; Shirk et al. 2012).

Potential sources of error in this analysis included the imprecision associated with cougar sample coordinates, as well as nonuniform sample coverage across the study area. Coordinates for most genetic samples were based on hunter descriptions, and were accurate only to 10 km. Therefore, error could have been introduced into pairwise resistances at short

distances if capture sites were incorrectly placed. This was a random source of error, however, and should not have resulted in a systemic bias for any variable. Furthermore, Graves et al. (2012) found only a small reduction in the strength of landscape genetic relationships under a scenario of simulated spatial uncertainty. With regard to highways, all sample locations were on a known side of the highway, so there was no risk of placing a point on the wrong side of a linear feature. Cougar samples were obtained opportunistically, a necessity due to this species' reclusiveness and solitary life history. This irregular sampling design provides a wide range of distances for pairwise comparisons, but can undersample or oversample some areas (Storfer et al. 2007). Indeed, sample coverage is very poor in wilderness areas and national parks (Figure 1), due to lack of access or prohibitions against hunting. The one variable this could have affected meaningfully was highways, as most cougar samples were obtained in proximity to paved roads. Maletzke (2010) reported mean cougar home range sizes from 199 to 753 km^2 in Washington, depending on sex and hunting pressure, which suggests that the majority of cougars in the state have at least some exposure to highways in their daily movements. A bias toward proximity to highways in sample collection may not necessarily translate to a misrepresentative sample, then, if the home ranges of most cougars overlap one or more highways.

Multiple regression on distance matrices and boosted regression trees both highlighted the importance of forest canopy cover and geographic distance, however each model explained only 15% and 19% of the variation in genetic distance, respectively. Models based on pairwise dissimilarities between points, as in distance matrices, generally have less explanatory power than those based on variables measured at the points themselves

(Legendre and Fortin 2010). Model fit in this study was similar to that reported by other researchers working with vagile predators (Balkenhol 2009; Garroway et al. 2010), and was likely limited by the cougar's ecological niche as a habitat generalist. Clearly, however, other factors are influencing gene flow in northwestern cougars, factors that could include prey distribution and density, sport hunting, and intraspecies territoriality and social interactions. The influence of sport hunting on cougar gene flow is difficult to quantify, because it can both restrict dispersal, through direct mortality of immigrants, and encourage dispersal, when resident males are killed and dispersing subadults from other areas move in to take their place (Robinson et al. 2008).

Cooley et al. (2009) demonstrated that low hunting mortality in an area led to increased emigration of subadults, while Robinson et al. (2008) showed that heavy hunting can produce a population sink. Building on these conclusions, the results of this study suggest that hunting pressure, male territoriality and forest cover interact to shape gene flow across the landscape. This is further supported by the pattern of genetic differentiation observed in chapter 1 being largely explained by breaks in forest cover. The implication then for management of cougars at the state level is that forested corridors between source and sink populations are essential to maintaining landscape connectivity. Regional population stability depends on the ability of migrants to move from source to sink populations, and regions that lose this connectivity could see local population declines and genetic isolation. Indices of inbreeding described in chapter 1 do not warrant concern over inbreeding depression at this time; however, as fragmentation of forested lands continues, forested

corridors will become increasingly important in maintaining genetic connectivity and

population stability for cougars in the northwest.

TABLES

Table 1. Multiple regression on distance matrices results for the global model explaining landscape effects on PCA-based genetic distance. *P*-values are based on 1,000 random permutations of the genetic distance matrix.

	β	*P*
Intercept	1.696	0.001
Forest	0.346	0.001
Population	-0.014	0.760
Elevation	0.017	0.564
Highways	-0.029	0.380
Geographic distance	0.220	0.001

	Test statistic	*P*
R^2	0.149	0.001
F	7,003.1	0.001

Table 2. Multiple regression on distance matrices results for the final model explaining landscape effects on PCA-based genetic distance, following forward selection with a *P*-to-enter value of 0.05. *P*-values are based on 1,000 random permutations of the genetic distance matrix.

	β	*P*
Intercept	1.696	0.001
Forest	0.319	0.001
Geographic distance	0.229	0.001

	Test statistic	*P*
R^2	0.148	0.001
F	17,388.4	0.001

Table 3. Mantel *r* correlation values between Circuitscape resistance and geographic distance matrices. Significant results are shown in bold; significance was assessed at α = 0.05 and was based on 1,000 random permutations of one of the two matrices being compared.

	Population	Highways	Elevation	Geographic Distance
Forest	**0.740**	**0.310**	-0.089	**0.619**
Population		**0.387**	-0.240	**0.312**
Highways			-0.165	**0.150**
Elevation				**0.094**

Table 4. Variance inflation factor coefficients for landscape resistance and geographic distance matrices.

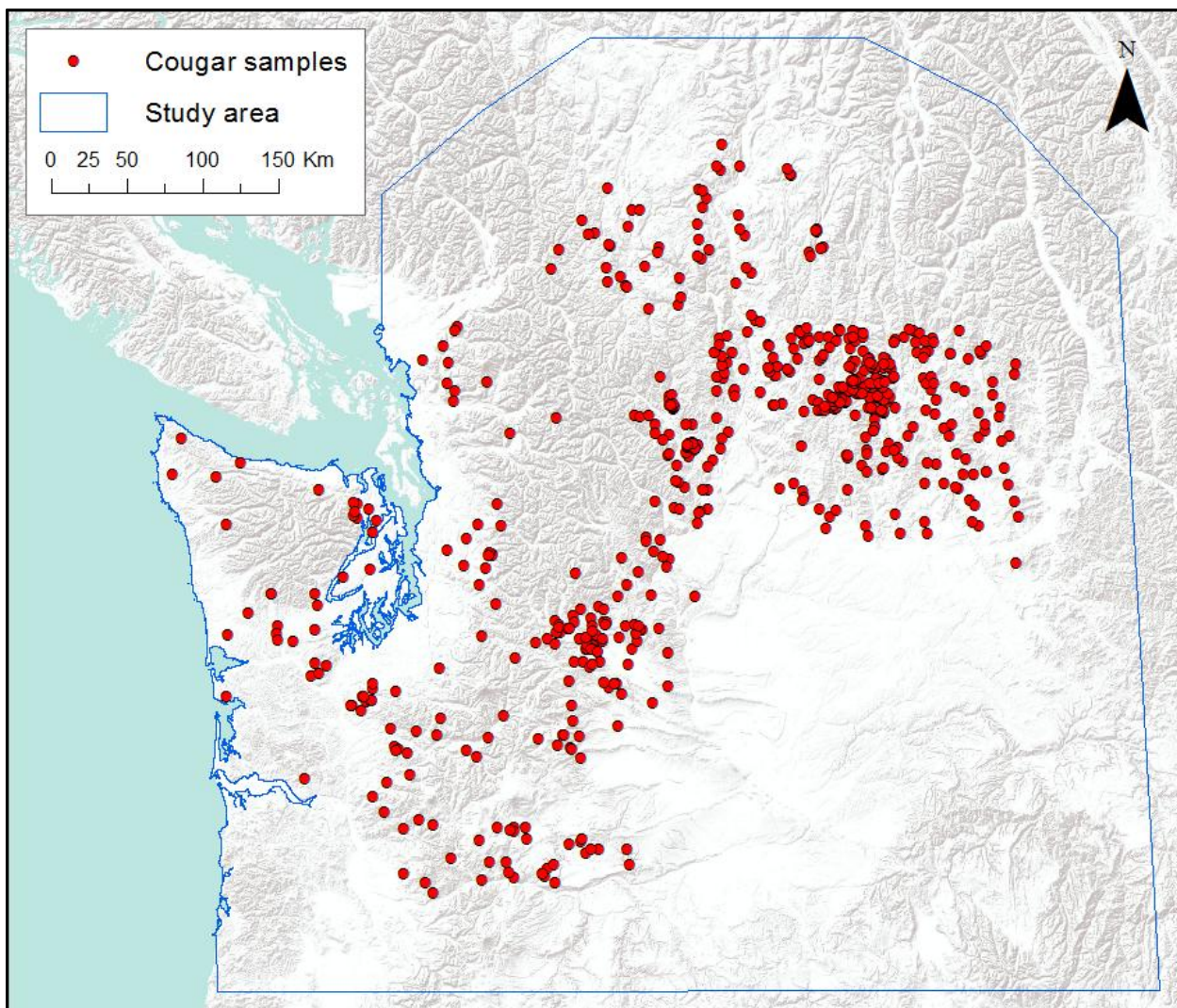|	*VIF*
Forest	3.50
Population	2.64
Highways	1.19
Elevation	1.11
Geographic Distance	1.79

FIGURES

Figure 1. Locations of cougar genetic samples, with samples from the Blue Mountains removed to avoid artificially inflating resistance estimates due to the configuration of forest cover and boundaries of the study area.
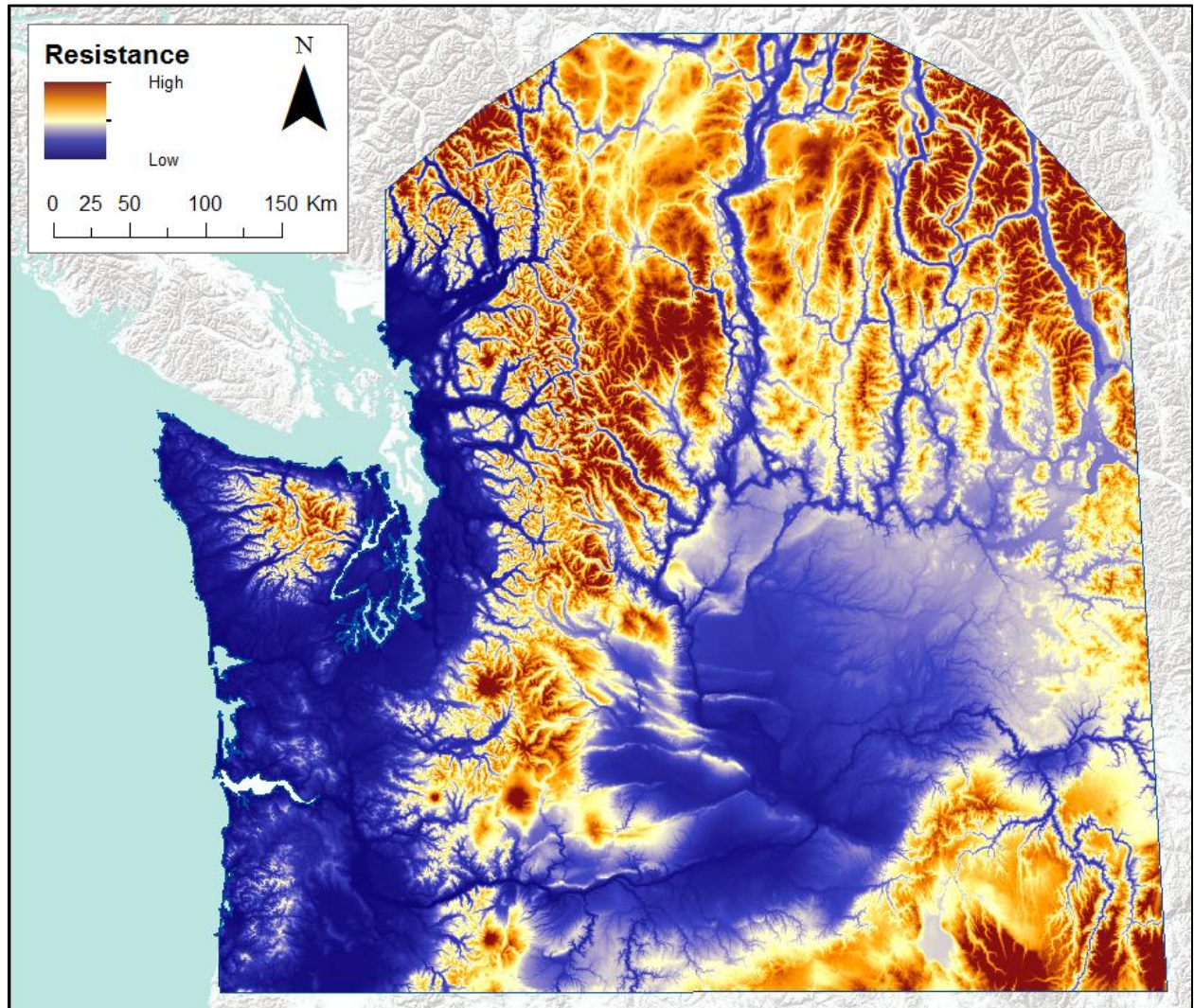
Figure 2. Landscape resistance surface for elevation. Low elevations are expected to provide the least resistance to cougar gene flow, and high elevations the greatest.
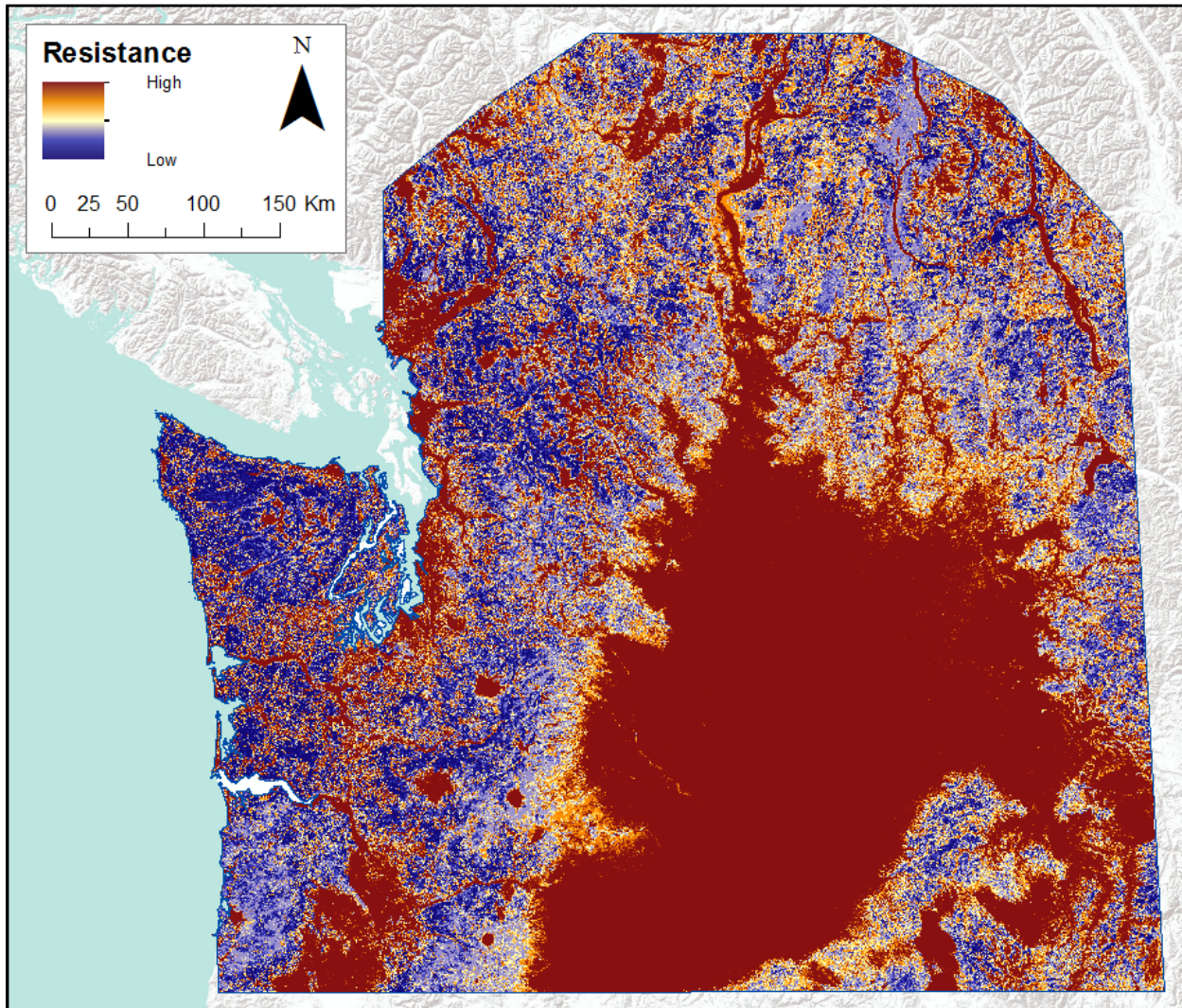
Figure 3. Landscape resistance surface for forest canopy cover. Resistance to cougar gene flow is expected to decrease with increasing forest canopy cover. In Circuitscape forest canopy cover was treated as a conductance surface, where conductance is the reciprocal of resistance, however it is displayed here as resistance for ease of comparison with other resistance surfaces.
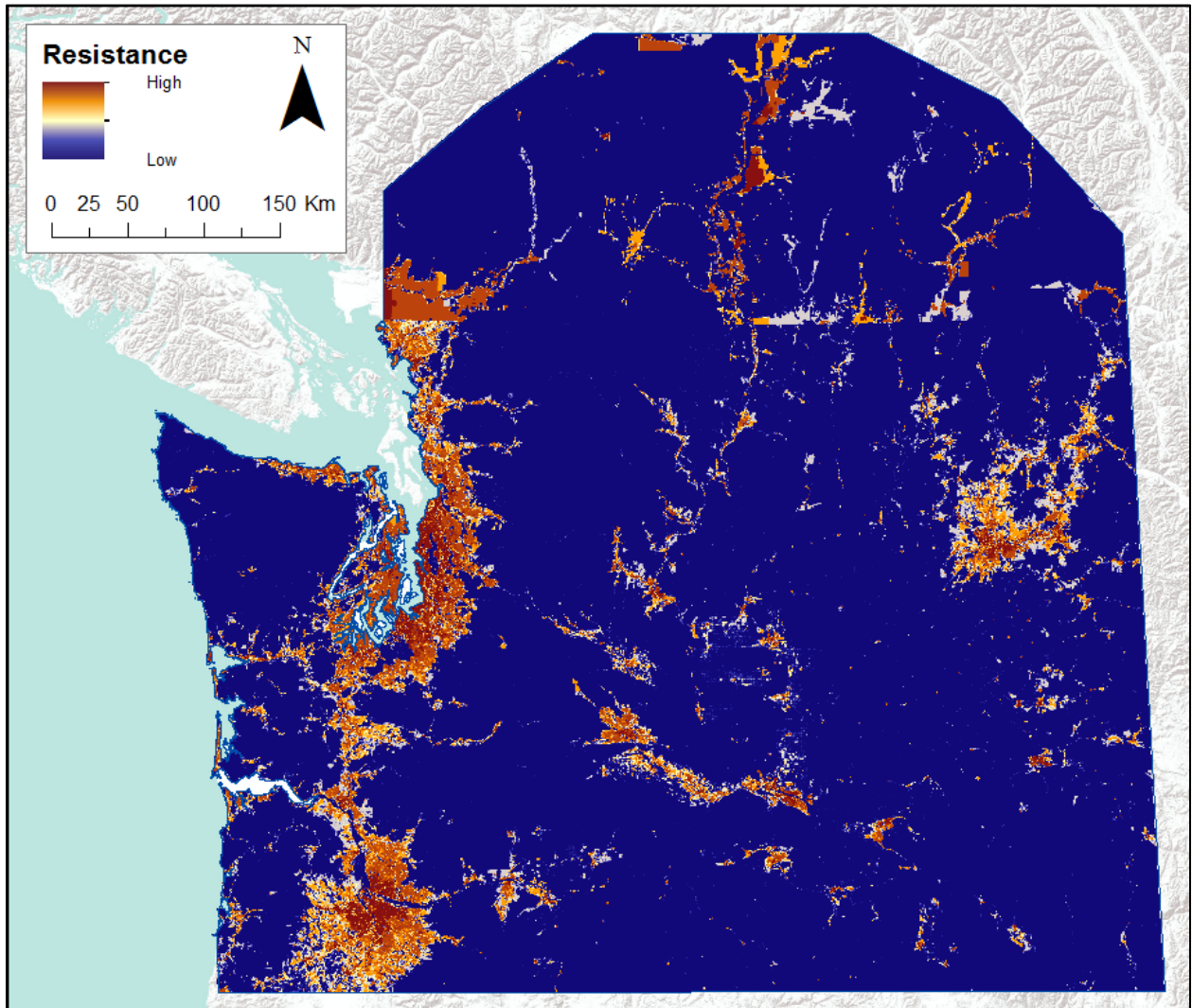
Figure 4. Landscape resistance surface for human population density. Resistance to cougar gene flow is expected to increase with increasing human population density.
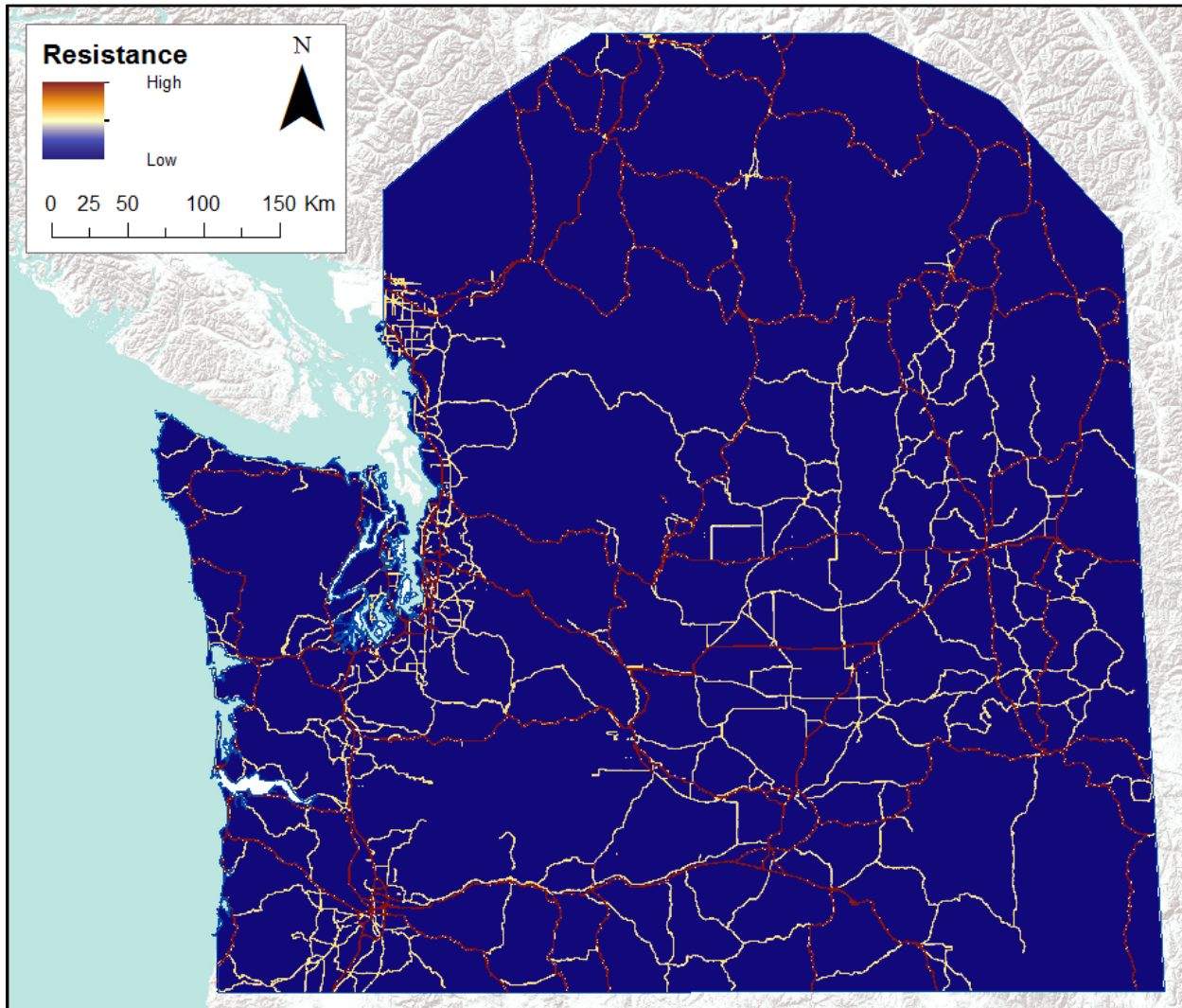
Figure 5. Landscape resistance surface for highways. Resistance to cougar gene flow is expected to increase with increasing highway traffic volume.

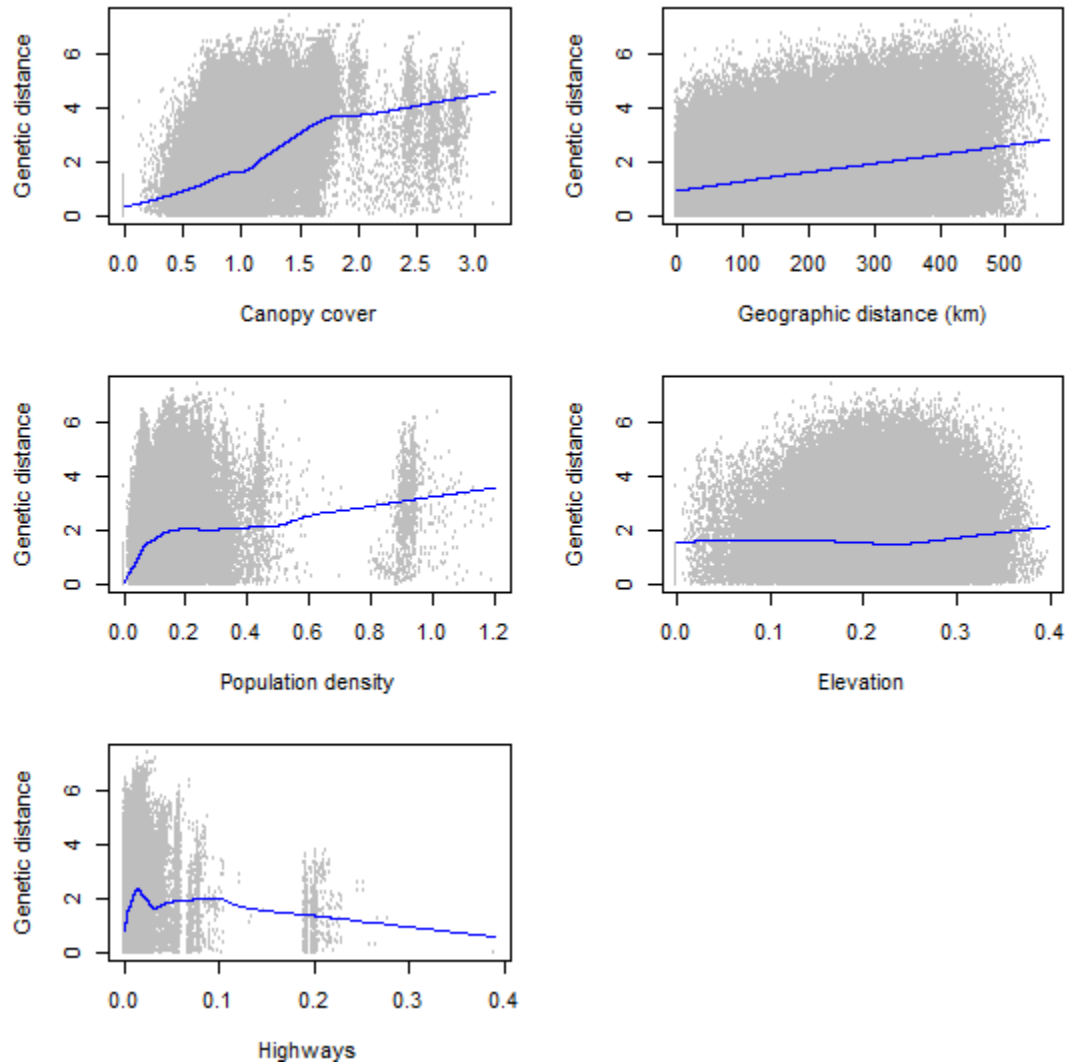**Figure 6.** Scatterplots of PCA-based genetic distance and raw resistance estimates. The X-axis for the geographic distance plot is shown in units of km, while all other X-axes are shown in terms of Circuitscape resistance (unitless). Locally-weighted scatterplot smoothing (LOESS) lines (in blue) have been fitted to illustrate trends, but should not be interpreted as evidence of a significant relationship.
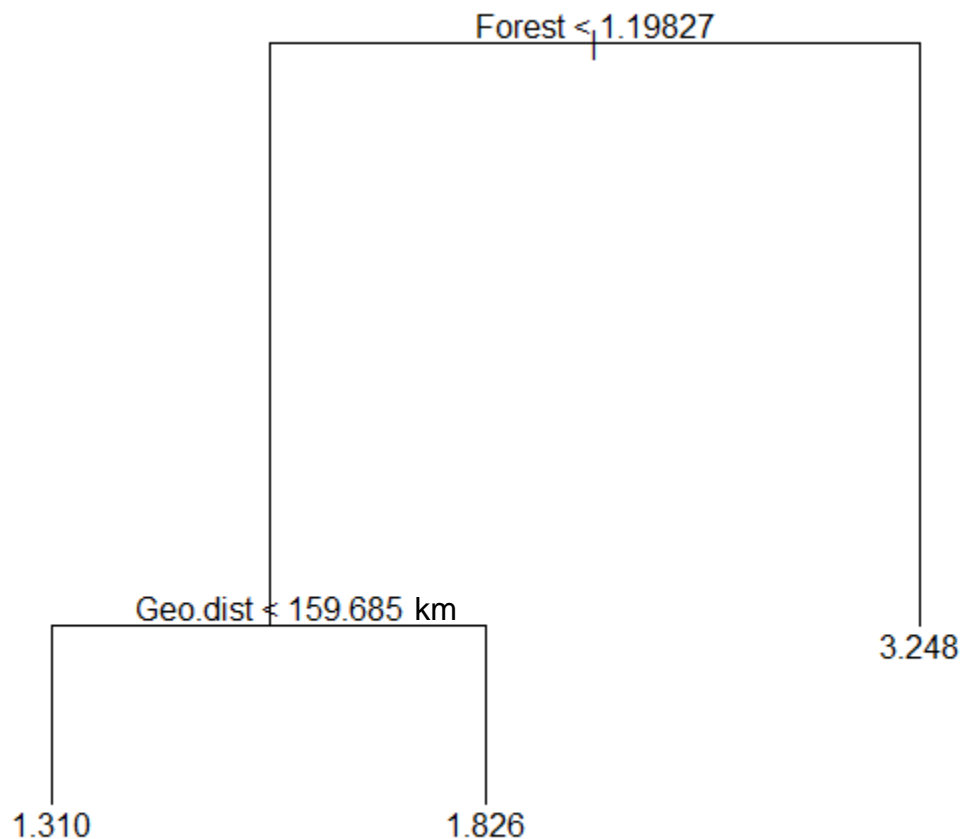
Figure 7. Single regression tree explaining variation in PCA-based genetic distance. Numbers at the terminal nodes represent the mean genetic distance for that group, where genetic distance ranges from 0 to 7.5, with a mean of 1.7. The top node splits the data based on resistance due to the reciprocal of forest canopy cover, which ranges from 0 to 3.2, with a mean of 0.9. The second node splits the data based on geographic distance, shown in km. The reduction in deviance from each split is proportional to branch length.

Figure 8. Relative influence of landscape variables on explained deviance of boosted regression tree model. Total explained deviance = 19.2%.
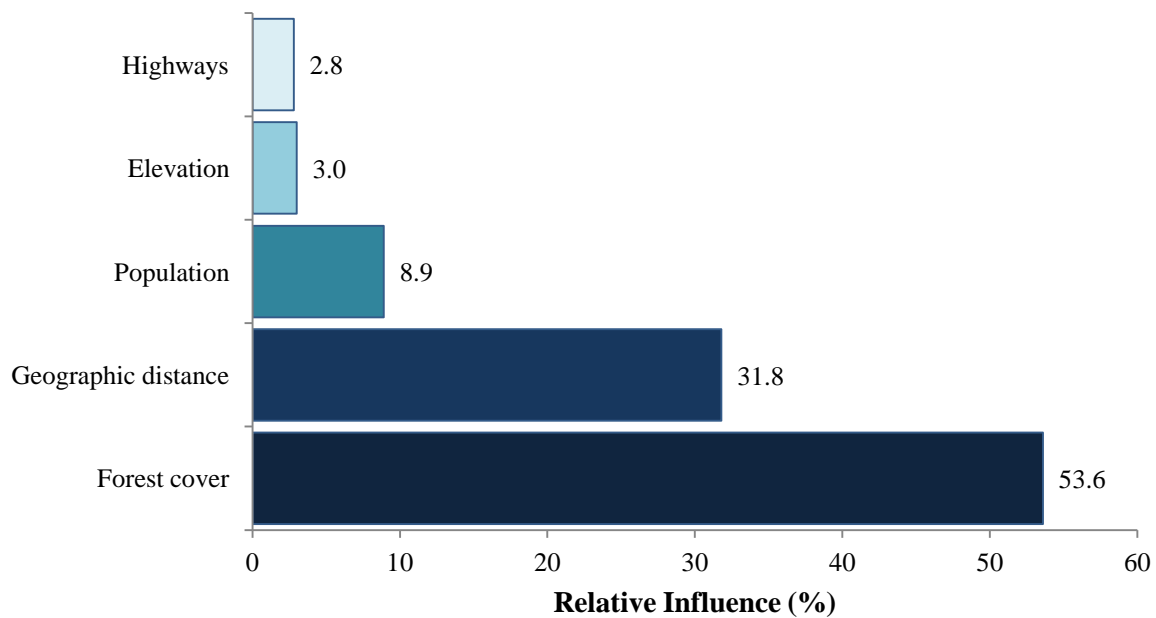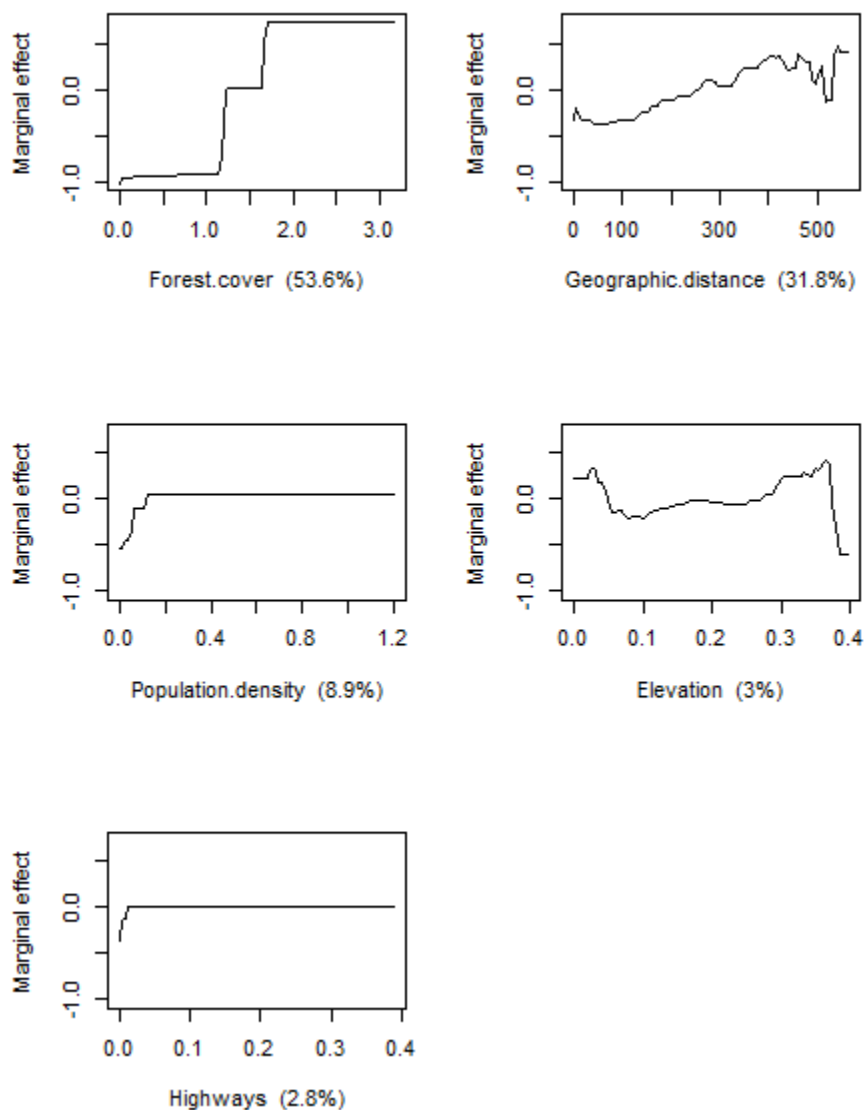
Figure 9. Partial dependence plots from boosted regression tree modeling, in order of decreasing relative influence. The Y-axes shows the marginal effect of resistance on PCA-based genetic distance. Negative marginal effects correspond to a decrease in genetic distance between individuals, and vice versa. The X-axis for the geographic distance plot is shown in units of km, while all other X-axes are shown in terms of Circuitscape resistance (unitless). In order to model landscape resistance, the reciprocal of forest cover was used, where resistance was greatest for unforested areas and least for densely forested areas. The relative influence of each variable on the explained deviance is shown in parentheses; total deviance explained by the model = 19.2%.

LITERATURE CITED

Anderson, A.E., D.C. Bowden, and D.M. Kattner. 1992. The puma on Uncompahgre Platueau, Colorado. Technical Publication No. 40. Denver: Colorado Division of Wildlife.

Anderson, Jr., C.R., F.G. Lindzey, and D.B. McDonald. 2004. Genetic structure of cougar populations across the Wyoming Basin: Metapopulation or megapopulation? Journal of Mammalogy 85: 1207-1214.

Andreasen, A.M., K.M. Stewart, W.S. Longland, J.P. Beckmann, and M.L. Forister. 2012. Identification of source-sink dynamics in mountain lions of the Great Basin. Molecular Ecology 21: 5689-5701.

Balkenhol, N. 2009. Evaluating and improving analytical approaches in landscape genetics through simulations and wildlife case studies. Ph.D. dissertation, University of Idaho.

Balkenhol, N., and L.P. Waits. 2009. Molecular road ecology: exploring the potential of genetics for investigating transportation impacts on wildlife. Molecular Ecology 18: 4151-4164.

Balkenhol, N., L.P. Waits, and R.J. Dezzani. 2009. Statistical approaches in landscape genetics: an evaluation of methods for linking landscape and genetic data. Ecography 32: 818-830.

Beier, P. 1995. Dispersal of juvenile cougars in fragmented habitat. The Journal of Wildlife Management 59: 228-237.

Beier, P. 2010. A focal species for conservation planning. Pages 177-189 in M.G. Hornocker and S. Negri, editors. Cougar: Ecology and conservation. University of Chicago Press, Chicago IL.

Cooley, H.S., R.B. Wielgus, G. Koehler, and B. Maletzke. 2009. Source populations in carnivore management: cougar demography and emigration in a lightly hunted population. Animal Conservation 12: 321-328.

Crown Registry and Geographic Base. 2012. Products – Gridded DEMs. http://archive.ilmb.gov.bc.ca/crgb/products/imagery/gridded.htm. Accessed 28 March 2012.

Culver, M. 1999. Molecular genetic variation, population structure, and natural history of free-ranging pumas (*Puma concolor*). Ph.D. dissertation, University of Maryland, College Park, Maryland.

Culver, M., W.E. Johnson, J. Pecon-Slattery, and S.J. O'Brien. 2000. Genomic ancestry of the American puma (*Puma concolor*). Journal of Heredity 91: 186-197.

Cushman, S.A., K.S. McKelvey, J. Hayden, and M.K. Schwartz. 2006. Gene flow in complex landscapes: Testing multiple hypotheses with causal modeling. The American Naturalist 168: 486-499.

De' Ath, G. 2007. Boosted trees for ecological modeling and prediction. Ecology 88: 243-251.

De' Ath, G., and K.E. Fabricius. 2000. Classification and regression trees: a powerful yet simple technique for ecological data analysis. Ecology 81: 3178-3192.

Dickson, B.G., and P. Beier. 2002. Home-range and habitat selection by adult cougars in southern California. The Journal of Wildlife Management 66: 1235-1245.

Dickson, B.G., and P. Beier. 2007. Quantifying the influence of topographic position on cougar (*Puma concolor*) movement in southern California, USA. Journal of Zoology 271: 270-277.

Dickson, B.G., J.S. Jenness, and P. Beier. 2005. Influence of vegetation, topography, and roads on cougar movement in southern California. Journal of Wildlife Management 69: 264-276.

Dratch, P.A. 1983. Genetic subdivision in Scottish red deer and North American wapiti. Ph.D. dissertation, University of Edinburgh, Scotland.

Elith, J., J.R. Leathwick, and T. Hastie. 2008. A working guide to boosted regression trees. Journal of Animal Ecology 77: 802-813.

Ernest, H.B., W.M. Boyce, V.C. Bleich, B. May, S.J. Stiver, and S.G. Torres. 2003. Genetic structure of mountain lion (*Puma concolor*) populations in California. Conservation Genetics 4: 353-366.

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

Fox, J., and S. Weisberg. 2011. An {R} Companion to Applied Regression, Second Edition. Sage: Thousand Oaks CA. URL: http://socserv.socsci.mcmaster.ca/jfox/Books/Companion

Freeland, J.R. 2006. Molecular Ecology. Hoboken, NJ: John Wiley and Sons.

Garroway, C.J., J. Bowman, and P.J. Wilson. 2011. Using a genetic network to parameterize a landscape resistance surface for fishers, *Martes pennanti*. Molecular Ecology 20: 3978-3988.

Goslee, S.C., and D.L. Urban. 2007. The ecodist package for dissimilarity-based analysis of ecological data. Journal of Statistical Software 22: 1-19.

Goudet, J. 2001. FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). Available from http://www.unil.ch/izea/softwares/fstat.html.

Goudet, J., N. Perrin, and P. Waser. 2002. Tests for sex-biased dispersal using bi-parentally inherited genetic markers. Molecular Ecology 11: 1103-1114.

Graves, T.A., T.N. Wasserman, M.C. Ribeiro, E.L. Landguth, S.F. Spear, N. Balkenhol, C.B. Higgins, M.J. Fortin, S.A. Cushman, and L.P. Waits. 2012. The influence of landscape characteristics and home-range size on the quantification of landscape-genetics relationships. Landscape Ecology 27: 253-266.

Guillot, G., and F. Rousset. 2013. Dismantling the Mantel tests. Methods in Ecology and Evolution (in press).

Guillot, G., A. Estoup, F. Mortier, and J.F. Cosson. 2005. A spatial statistical model for landscape genetics. Genetics 170: 1261-1280.

Jombart, T. 2012. An introduction to adegenet 1.3-4. 67 p. Available at: http://cran.r-project.org/web/packages/adegenet/vignettes/adegenet-basics.pdf.

Jombart, T., S. Devillard, A.B. Dufour, and D. Pontier. 2008. Revealing cryptic spatial patterns in genetic variability by a new multivariate method. Heredity 101: 92-103.

Kertson, B.N., R.D. Spencer, J.M. Marzluff, J. Hepinstall-Cymerman, and C.E. Grue. 2011. Cougar space use and movements in the wildland-urban landscape of western Washington. Ecological Applications 21: 2866-2881.

Koen, E.L., C.J. Garroway, P.J. Wilson, and J. Bowman. 2010. The effect of map boundary on estimates of landscape resistance to animal movement. PLoS ONE 5: e11785.

Landguth, E.L, and S.A. Cushman. 2010. CDPOP: A spatially explicit cost distance population genetics program. Molecular Ecology Resources 10: 156-161.

Legendre, P., and M.J. Fortin. 2010. Comparison of the Mantel test and alternative approaches for detecting complex multivariate relationships in the spatial analysis of genetic data. Molecular Ecology Resources 10: 831-844.

Legendre, P., and L. Legendre. 2012. Numerical Ecology. Third English Edition. Elsevier, Oxford.

Legendre, P., F.J. Lapointe, and P. Casgrain. 1994. Modeling brain evolution from behavior: a permutational regression approach. Evolution 48: 1487-1499.

Logan, K.A., and L.L. Irwin. 1985. Mountain lion habitats in the Big Horn Mountains, Wyoming. Wildlife Society Bulletin 13: 257-262.

Logan, K.A., and L.L. Sweanor. 2000. Puma. Pages 347-377 *in* S. Demarais and P.R. Krausman, editors. Ecology and management of large mammals in North America. Englewood Cliffs, NJ: Prentice Hall.

Logan, K.A., and L.L. Sweanor. 2010. Behavior and social organization of a solitary carnivore. Pages 105-117 *in* M. Hornocker and S. Negri, editors. Cougar ecology and conservation. Chicago: University of Chicago Press.

Loxterman, J.L. 2010. Fine scale population genetic structure of pumas in the Intermountain West. Conservation Genetics 12: 1049-1059.

Maehr, D.S., E.D. Land, D.B. Shindle, O.L. Bass, and T.S. Hoctor. 2002. Florida panther dispersal and conservation. Biological Conservation 106: 187-197.

Maletzke, B.T. 2010. Effects of anthropogenic disturbance on landscape ecology of cougars. Ph.D. dissertation, Washington State University.

Manel, S., M.K. Schwartz, G. Luikart, and P. Taberlet. 2003. Landscape genetics: combining landscape ecology and population genetics. Trends in Ecology and Evolution 18: 189-197.

Marquardt, D.W . 1970. Generalized inverses, ridge regression and biased linear estimation. Technometrics 12: 591-612.

McRae, B.H. 2006. Isolation by resistance. Evolution 60: 1551-1561.

McRae, B.H., and Shah, V.B. 2011. Circuitscape User Guide. The University of California, Santa Barbara. Available at: http://www.circuitscape.org.

McRae, B.H., P. Beier, L.E. Dewald, L.Y. Huynh, and P. Keim. 2005. Habitat barriers limit gene flow and illuminate historical events in a wide-ranging carnivore, the American puma. Molecular Ecology 14: 1965-77.

McRae, B.H., B.G. Dickson, T.H. Keitt, and V.B. Shah. 2008. Using circuit theory to model connectivity in ecology, evolution, and conservation. Ecology 89: 2712-2724.

Menotti-Raymond, M.A., and S.J. O'Brien. 1995. Evolutionary conservation of ten microsatellite loci in four species of Felidae. Journal of Heredity 86: 319-322.

Menotti-Raymond, M.A., V.A. David, L.A. Lyons, A.A. Schaffer, J.F. Tomlin, M.K. Hutton, and S.J. O'Brien. 1999. A genetic linkage map of microsatellites in the domestic cat (*Felis catus*). Genomics 57: 9-23.

Murphy, M.A., J.S. Evans, and A. Storfer. 2010. Quantifying Bufo boreas connectivity in Yellowstone National Park with landscape genetics. Ecology 91: 252-261.

Musial, T.S. 2009. Evolutionary and landscape genetic structure of *Puma concolor* in Oregon. MS thesis, Portland State University, Portland, Oregon.

Nei, M. 1987. Molecular evolutionary genetics. New York, NY: Columbia University Press.

Newby, J. 2011. Puma Dispersal Ecology in the Central Rocky Mountains. MS thesis, University of Montana.

Park, S.D.E. 2001. Trypanotolerance in west African cattle and the population genetic effects of selection. The University of Dublin, Dublin, Ireland.

Patterson, N., A.L. Price, and D. Reich. 2006. Population structure and eigenanalysis. PLOS Genetics 2: 2074-2093.

Peakall, R. and P.E. Smouse. 2006. Genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes 6: 288–295.

Pimm, S.L., L. Dollar, and O.L. Bass Jr. 2006. The genetic rescue of the Florida panther. Animal Conservation 9: 115-122.

Raufaste, N., and F. Rousset. 2001. Are partial Mantel tests adequate? Evolution 55: 1703-1705.

Rice, W.R. 1989. Analyzing tables of statistical tests. Evolution 43: 223-225.

Ridgeway, G. 2013. gbm: Generalized Boosted Regression Models. R package version 2.0-8.

    http://CRAN.R-project.org/package=gbm

Riley, S.P.D., J.P. Pollinger, R.M. Sauvajot, E.C. York, C. Bromley, T.K. Fuller, and R.K.

    Wayne. 2006. FAST-TRACK: A southern California freeway is a physical and social

    barrier to gene flow in carnivores. Molecular Ecology 15: 1733-1741.

Ripley, B. 2012. tree: Classification and regression trees. R package version 1.0-33.

    http://CRAN.R-project.org/package=tree

Robinson, H.S., R.B. Wielgus, H.S. Cooley, and S.W. Cooley. 2008. Sink populations in

    carnivore management: Cougar demography and immigraton in a hunted population.

    Ecological Applications 18: 1028-1037.

Roelke, M.E., J. Martenson, and S.J. O'Brien. 1993. The consequences of demographic

    reduction and genetic depletion in endangered Florida panther. Current Biology 3:

    340-350.

Rousset, F., 2008. Genepop'007: a complete reimplementation of the Genepop software for

    Windows and Linux. Molecular Ecology Resources 8: 103-106.

Shirk, A.J., D.O. Wallin, S.A. Cushman, C.G. Rice, and K.I. Warheit. 2010. Inferring

    landscape effects on gene flow: a new model selection framework. Molecular

    Ecology 19: 3603-3619.

Shirk, A.J., S.A. Cushman, and E.L. Landguth. 2012. Simulating pattern-process

    relationships to validate landscape genetic models. International Journal of Ecology

    2012: 1-8.

Sinclair, E.A., E.L. Swenson, M.L. Wolfe, D.C. Choate, B. Bates, and K.A. Crandall. 2001. Gene flow estimates in Utah's cougars imply management beyond Utah. Animal Conservation 4: 257-264.

Spear, S.F., N. Balkenhol, M.-J. Fortin, B.H. McRae, and K.I.M. Scribner. 2010. Use of resistance surfaces for landscape genetic studies: considerations for parameterization and analysis. Molecular Ecology 19: 3576-3591.

Spear, S.F., J. Baumsteiger, and S. Storfer. 2011. Type N Experimental Buffer Treatment Study: Baseline Measures of Genetic Diversity and Gene Flow of Three Stream-Associated Amphibians. Cooperative Monitoring Evaluation and Research Report, CMER 06-605. Washington Department of Natural Resources, Olympia, WA.

Spreadbury, B.R., R.R. Kevin Musil, Jim Musil, Chris Kaisner, John Kovak. 1996. Cougar population characteristics in southeastern British Columbia. Journal of Wildlife Management 60: 962-969.

Stoner, D.C., and M.L. Wolfe. 2012. How does a widely distributed felid respond to natural and anthropogenic habitat fragmentation? Wild Felid Monitor 5: 1-27.

Storfer, A., M.A. Murphy, J.S. Evans, C.S. Goldberg, S. Robinson, S.F. Spear, R. Dezzani, E. Delmelle, L. Vierling, and L.P. Waits. 2007. Putting the 'landscape' in landscape genetics. Heredity 98: 128-142.

Sturges, H. 1926. The choice of a class-interval. Journal of the American Statistical Association 21: 65-66.

Sweanor, L.L., K.A. Logan, and M.G. Hornocker. 2000. Cougar dispersal patterns, metapopulation dynamics, and conservation. Conservation Biology 14: 798-808.

Taylor, S., C. Buergelt, M. Roelke-Parker, B. Homer, and D. Rotstein. 2002. Causes of mortality of free-ranging Florida panthers. Journal of Wildlife Diseases 38: 107-114.

[USGS] United States Geological Survey. 2012. Seamless Data Warehouse. http://seamless.usgs.gov/index.php. Accessed 28 March 2012.

van Oosterhout, C., W.F. Hutchinson, D.P.M.,Wills, and P. Shipley. 2004. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Resources 4: 535-538.

Verbeylen, G., L. De Bruyn, F. Adriaensen, and E. Matthysen. 2003. Does matrix resistance influence red squirrel (*Sciurus vulgaris*) distribution in an urban landscape? Landscape Ecology 18: 791-805.

[WDFW] Washington Department of Fish and Widlife. 2011. 2011 Game status and trend report. Wildlife Program, Washington Department of Fish and Wildlife, Olympia, WA.

[WHCWG] Washington Wildlife Habitat Connectivity Working Group. 2010. Washington Connected Landscapes Project: Statewide Analysis. Washington Departments of Fish and Wildlife, and Transportation, Olympia, WA.

[WSDOT] Washington State Department of Transportation. 2012. Washington State Department of Transportation geodata catalog. Available at: http://www.wsdot.wa.gov/Mapsdata/geodatacatalog/. Accessed 12 April 2012.

Wasserman, T. N., S. A. Cushman, M. K. Schwartz, and D. O. Wallin. 2010. Spatial scaling and multi-model inference in landscape genetics: Martes americana in northern Idaho. Landscape Ecology 25:1601-1612.

Weir, B.S., and C.C. Cockerham. 1984. Estimating F-statistics for the analysis of population structure. Evolution 38: 1358-1370.

Wiken, E., F. Jiménez Nava, and G. Griffith. 2011. North American Terrestrial Ecoregions—Level III. Commission for Environmental Cooperation, Montreal, Canada.

Wright, S. 1943. Isolation by distance. Genetics 28: 114-138.