

Chapter 17

Landscape Genetics

Kevin S. McKelvey, Samuel A. Cushman, and Michael K. Schwartz

In reading this book, you have observed that the spatial data used in landscape ecology come from many sources and in many forms. For many organisms, these data take the form of presence or absence at a location, or numbers of individuals at that same location. For species such as trees, where huge size differences exist between individuals, indices such as basal area, metric tons per hectare, or canopy cover are more useful than counts. For any measured species that is handled (or sampled noninvasively; Taberlet et al. 1999; Kendall and McKelvey 2008; Schwartz and Monfort 2008), an additional data source is available: the genetic data stored in the organism's tissue. If genetic samples are taken, then these data become another type of spatial data associated with the location where the organism was sampled. As such, genetic data can be analyzed with many of the same approaches used to analyze data of other types that vary spatially.

Genetic data, however, also have certain valuable properties that allow different sorts of analyses than are possible using many other data sources. Most data are representative of the present: where an organism is at the time of sampling, and what attributes are correlated with that presence. Genetic data allow us to look into the animal's past and ask where it was born, and where its parents were born. They also allow us to determine the individual's likely response to stressors such as heat, drought, or disease.

Landscape genetics has been viewed as a hybrid between population genetics and landscape ecology (Manel et al. 2003). Another way to view landscape genetics is as landscape ecology that uses spatial genetic patterns (e.g. genetic variation within or between individuals or groups). This is because landscape genetics treats the genetic patterns themselves as multivariate spatial data and seeks to infer ecological understandings by evaluating these patterns either in isolation or in conjunction with other spatial data. Because genetic data are very different from other data sources, we need to briefly discuss genetic data before we can explore their uses.

S.A. Cushman (✉), K.S. McKelvey and M.K. Schwartz
US Forest Service, Rocky Mountain Research Station, 800 E Beckwith, Missoula,
MT 59801, USA
e-mail: scushman@fs.fed.us

17.1 A Brief Primer on Genetics as they Relate to Landscape Analysis

The first thing to note about genetic data is that they represent an enormously rich data source. Even a primitive bacterium will generally have over one million base pairs in its genome. For example, the *Escherichia coli* K-12 genome contains 4,639,221-base pairs (Blattner et al. 1997). The human genome contains about 3 billion base pairs (Venter et al. 2001). Understanding the different patterns in the sequence of these base pairs is the business of molecular biology. How these patterns are arranged in geographic space is the business of landscape genetics or, more broadly, molecular ecology (Beebe and Rowe 2007).

Geneticists refer to genes, loci, and alleles. Genes are areas of genetic code associated with some function, loci are simply locations of interest within the code (although they can be genes), and alleles are different base pair patterns at a specific locus. Genetic variation is generated by random mutations at the level of the individual base pair of nucleotides (e.g. from adenine [A] to thymine [T], cytosine [C], or guanine [G]). Mutations lead to new alleles at loci which may or may not lead to different behavior of genes. If a mutation occurs in an area that codes for a protein (dubbed an “exon”) and causes a different amino acid to be produced, most likely this mutation will prove detrimental, will therefore not persist, and will not be observed in samples from live individuals in a population. The removal of these mutations by natural selection conserves the original sequence, and therefore all individuals in a population will frequently have the same sequence (no variation at that locus). However, once in awhile a mutation in an exon will be beneficial. These mutations may persist transiently, with the new mutation being superior and the original allele will vanish or at least diminish in frequency (called directional selection). When either selection (or a process called genetic drift and discussed below) causes one allele to persist, this is called fixation. In some cases, both states will be maintained by selection (called diversifying selection or balanced polymorphisms). For example, it has been shown that the Human Leukocyte Antigen locus that recognizes antigens and thus allows the immune system to respond is under balancing selection, where there is an advantage for an individual to be heterozygous, in that this allows greater protection from different pathogen species or strains. Mutations that are either removed or maintained by selection are not “neutral” but rather influence fitness and are therefore “adaptive”.

On the other extreme are areas of the genome which, as far as we can tell, do nothing at all, dubbed “introns” (within genes) or intergenic DNA (between genes). Because mutations in these areas generally have no known effect on the organism’s fitness, these areas are not conserved and therefore have the highest rates of observed mutations. In non-coding, or neutral areas, mutations build up at a predictable rate and therefore the number of mutations is related to time. In between these two extremes are non-coding sequences which are closely associated (physically linked by close proximity on a chromosome) with coding areas. Thus, there is likely some conservation within introns but they are variable enough to produce useful patterns. Further, their physical linkage to exons provides the potential to use intron patterns to track adaptive change.

While mutation is the ultimate source for genetic variation, the buildup of these mutations is a slow process. The movement of alleles across space, however, can be a rapid process. It occurs through reproduction and migration. Thus, spatial patterns of neutral genetic variation tend to be dominated by gene flow and drift and be little influenced by mutation. Gene flow is defined as exchange of genetic information (alleles) between populations through migration (Allendorf and Luikart 2007). Gene flow across landscapes makes populations or geographic areas more similar genetically. On the contrary, genetic drift causes populations or geographic areas across landscapes to become genetically differentiated. Genetic drift is defined as random changes in allele frequencies in populations between generations due to binomial sampling of genes during meiosis (gamete production). Genetic drift is more pronounced in small populations. For example, imagine a population of two individuals with two different alleles; by pure chance we may lose alleles not due to selection but by drift if one allele happens to not get passed down in a single generation.

For landscape genetics, patterns in neutral DNA, introns, and exons, each have different meanings, and can be used to ask different questions. Because changes in an organism's neutral DNA has no biological implications, these patterns are free to drift due to stochastic processes, and are free to flow due to organism movements. Patterns in neutral DNA, being free from natural selection, can be more easily interpreted, and a great deal of population and landscape genetics is based on analyzing these regions (Manel et al. 2003; Avise 2004; Allendorf and Luikart 2007). However, the evolution and movement of adaptive traits may be more pertinent for evaluating population status (e.g. long-term adaptive potential or current extinction risk), with the observed neutral patterns providing a potential index of adaptive differentiation between populations. Populations that have been separated for long periods of time, and therefore have many differences in neutral DNA, will also tend to develop local adaptations (e.g., Storfer et al. 1999). While, in a broad sense, large inter-population differences in neutral markers indicate potential for large differences in adaptive traits, the relationship is by no means constant. For example, Bromham and Woolfit (2004) found no consistent patterns in rates of molecular evolution based on neutral markers and rates of observed evolution of phenotypes. Using the degree of differentiation in neutral markers to infer the degree of adaptive change or divergence is therefore somewhat speculative.

In the future we will almost certainly evaluate both adaptive and neutral patterns (Luikart et al. 2003) with large amounts of information contained not only in the patterns themselves, but in the covariance patterns as well. This approach is the focus of the Landscape Genomics Chapter (Chapter 9, this volume; see also Gebremedhin et al. 2009).

17.1.1 Maternal and Paternal Inheritance

Almost all of the DNA in higher eukaryotes is nuclear, and therefore inherited from both parents. However, there are sequences that are inherited exclusively either

maternally or paternally. Exclusive maternally inherited DNA is associated with organelles (in animals mitochondria and in plants both mitochondria and chloroplasts) that contain their own DNA. They are largely not recombined with paternal DNA. In animals, the Y chromosome provides pure paternal inheritance in males; in plants, pollen contains extra-nuclear DNA providing richer data for tracking paternal lineages.

Genes inherited exclusively from maternal or paternal sources produce useful patterns because, lacking recombination (which shuffles DNA sequences), these patterns are easier to analyze, because differences in sex-related behavior can be inferred, and because there is only one “type” per individual (as opposed to nuclear DNA where there are two copies of each locus – one from each of the paired autosomal chromosomes). For example, in a famous case, analyses of mitochondrial DNA (mtDNA) for bears on the Admiralty, Baranof, and Chichagof (ABC) islands off the coast of Alaska indicated that the bears were highly divergent from mainland bears. However, nuclear DNA indicated that they were similar (Paetkau et al. 1998). Male bears crossed the passage to the mainland frequently, but apparently females never did. By comparing the patterns associated with the mixing of both sexes (nuclear) with those exclusively associated with females (mitochondrial), the movement patterns and relative isolation of the islands was elucidated.

The case of the ABC bears is an application of traditional population genetics, but the issues, movement and isolation, are commonly addressed through landscape genetics (especially when working at fine spatial scales). Indeed, the boundary between population and landscape genetics is blurry and in many cases a problem can be investigated using either approach. For the remainder of the chapter we will discuss the branch of landscape genetics that evaluates differences between individuals (e.g., Cushman et al. 2006) and does not seek to assign individuals to known populations (Holderegger and Wagner 2008 and Balkenhol et al. 2009). These approaches treat genetic differences as spatial gradients rather than as discrete clusters with relatively little internal pattern. In this simplified view, the difference between population and landscape genetics is analogous to the difference between ANOVA and regression. As with these statistical methods, the choice of approach should be dictated by the questions asked and the nature of the data; neither is uniformly superior. Similarly, just as regression using continuous variables and ANOVA represent two ends of a spectrum of models involving mixtures of categorical and continuous variables, landscape genetics can contain mixtures of gradients and groups.

17.2 The Nature of Genetic Data

As indicated above, genetics in general and landscape genetics in particular are huge fields; to avoid writing an entire book on it, we need to narrow down. For the remainder of the chapter, we will therefore discuss analyses of neutral nuclear DNA. For over a decade, the most common analyses of neutral DNA patterns have utilized “microsatellites”, thus we will only consider this type of genetic data

(Luikart and England 1999). Microsatellites are composed of simple sequence repeats of 1–6 base pairs. For example cytosine-adenine (CA) is a common (“dinucleotide”) repeat sequence found in genomes. In microsatellites, the same sequence of base pairs commonly will repeat 10–100 times (CA,CA,CA...CA) and mutations lead to variations in the number (n) of repeats (CA _{n}). Therefore genetic variation in microsatellites is scored in terms of the number of base pairs in the repeat (e.g. CA₁₀).

Microsatellite mutation occurs more frequently than single base pair mutations in sequences, and therefore microsatellite loci are often extremely variable. It is not uncommon to have ten or more lengths (alleles) of a particular microsatellite (locus) in a population (although some fish species have more than 45 alleles at a locus; Van Oppen et al. [1997]). They are also, as far as we can tell, truly neutral and different loci are often considered independent, although this can be tested with recent genome maps that are being produced. Because they are nuclear (and thus two alleles are contained in each individual at autosomal chromosomes), each locus can be either heterozygous (two different alleles) or homozygous (the same allele twice) increasing the between-individual variability. Because of the high number of potential combinations at any locus, we can often differentiate individuals (called DNA fingerprinting) in wild populations using 6–7 loci (Paetkau and Strobeck 1994).

17.2.1 Genetic Distance

Genetic distance is simply the degree to which genetic data between 2 individuals or populations vary. For individuals, it can be treated simply as the proportion of alleles different (e.g. Bray-Curtis, used in Cushman et al. 2006) or can be based on deviation from genetic expectations (e.g. a_r , Rousset 2000). Various distance metrics tend to be highly correlated; Cushman et al. (2006) found that the correlation between distance measured using the Bray-Curtis percentage dissimilarity measure (Legendre and Legendre 1998) and a_r (Rousset 2000) was 0.97 for their bear data.

The expected genetic distance between individuals in a well mixed population will related to the frequency within the population of alleles at each sampled locus. However, if individuals within a geographic area are related, the frequency of genotypes and alleles will differ from their global expectations. For example, if we were to sample from a family where all offspring had the same parents, no more than four alleles would exist at each locus across the entire family group regardless of the global allelic frequencies (this could also be as little as one allele if both parents were homozygous for the same allele at this locus). Similarly, this family would have, overall, more alleles in common than random expectations. Because this group is closer genetically than expectation, by definition it has to be further apart genetically from other family groups than random expectation. Similarly, if two populations are separated for a long time, allele frequencies will diverge due to the random processes of drift and mutation, and samples from individuals between the two populations will be genetically more dissimilar than would be expected

if they were freely mixing. In landscape genetics, patterns of genetic distance are the spatial information of interest: mating behavior and movement patterns can be inferred by relating these patterns to the landscapes in which they occur.

A common spatial pattern of genetic distance in terrestrial organisms is that organisms that are spatially proximal will be more genetically similar than organisms that are distant. This is because, for animals, an organism is more likely to find a proximal organism than one that is distant, and the resulting offspring are more likely to settle nearby. For plants, pollen is more likely to be transferred to proximal plants (in some cases to other parts of the same plant), and seed densities likely are highest close to the parent plant. These types of patterns are referred to as “isolation by distance” (Wright 1943) and for many organisms can be considered the expectation distribution: on a completely featureless landscape, we would expect patterns to be associated with Euclidian distance. When population processes lead to isolation by distance, genetic distances will form gradients, and landscape genetic approaches are therefore appropriate. We can next consider what happens when landscape features interact with levels of gene flow and enhance or retard genetic exchange across the landscape.

17.3 Common Approaches in Landscape Genetics

As might be expected, early landscape genetic approaches sought to use spatial genetic patterns to define populations. Once genetic discontinuities (barriers) and populations were defined, standard population genetic approaches as well as individual based approaches could be applied (Manel et al. 2003 and references therein). The fundamental idea was that genetic differences formed a gradient surface with areas of rapid change and “flat” areas. Steep slopes would be related to barriers to gene flow, and flat areas with freely mixed or “panmitic” populations. Rather than assuming prior knowledge of where the steep areas were, the genetic patterns would allow determination of population boundaries (Barbujani et al. 1989; Barbujani and Sokal 1990; Stenico et al. 1998; Simoni et al. 1999). While an excellent idea in theory, in practice genetic data are very noisy when used in this manner and are likely highly influenced by sampling scheme (Novembre and Stephens 2008; Schwartz and McKelvey 2009; Frantz et al. 2009). Thus early methods such as the Monmonier algorithm (Monmonier 1973) tended to get stuck in local minima and failed to reliably reproduce simulated populations. Of the various approaches to determining minima simulated annealing is the most robust and Dupanloup et al. (2002) were able to improve results using these methods. However, this approach has not become widespread. Instead, a posteriori grouping of genetic data into populations is largely done using non spatial (e.g. STRUCTURE; Prichard et al. (2000)) or semi-spatial (e. g. Geneland; Guillot et al. [2005]) clustering algorithms.

Coulon et al. (2004) correlated pair-wise genetic distance between samples with landscape features viewed as having different resistances to gene flow. Least cost

resistance paths were created between all individuals and correlations between path resistances and genetic distance higher than those between Euclidian and genetic distance were viewed as supportive of the putative resistance model. More recently McRae (2006) conceived landscape resistance as a series of parallel circuits connected by resistors. Gene flow was viewed as analogous to electron flow, and, as this approach had been used in engineering to determine electron flow through metal plates, numerical approaches were available. The advantage of the circuit analogy is that gene flow would be altered not only by the minimum resistance between points, but also by the width and therefore number of alternative routes available. Recently, Garroway et al. (2008) utilized graph theoretic approaches similar to those that have been used to analyze social networks to evaluate genetic data. This approach allows the macro-scale characteristics of the network to be determined, potentially allowing deeper insights into the biology driving genetic interchange than are possible through evaluations of resistance to travel.

Once factors controlling (or at least correlating with) gene flow have been identified, the model can be applied to a landscape to determine likely areas for both corridors and barriers to flow. For many animals, the advent of GPS tracking should, in turn, allow the direct confirmation of the genetically derived connectivity rules, although genetics may reflect more historical gene flow integrated across many individuals, whereas GPS tracking could reflect more recent gene flow for a few select individuals as well as movements not related to mating, reproductive success, and gene flow.

17.4 Examples of Connectivity Modeling

In the remainder of this chapter we illustrate some of the potentials associated with landscape genetics with examples taken from our own work. The organisms involved are black bears (*Ursus americanus*) and wolverines (*Gulo gulo*) and in both cases our research group has used resistance modeling to infer corridors and barriers, and subsequently utilized these models to develop putative corridor structures. In the case of the wolverine, we had a coherent a priori habitat use model test. In the case of the black bear, we explored multiple models, but used “causal modeling” (Legendre and Troussellier 1988; Legendre 1993) to separate spurious correlations and identify biologically relevant factors.

17.4.1 Wolverine Example

Wolverine populations, at their southern extent are limited to high elevation areas. Not only do they need to den in snow (Magoun and Copeland 1998), but year-around habitat use is limited to high elevation areas, possibly due to aversion to summer heat (Hornocker and Hash 1981). Wolverines have large home ranges, on the order

of 100 km² for females (Copeland et al. 2007). Hence, the mountainous areas where they live contain relatively few individuals (Squires et al. 2007) and continued presence therefore requires movement between the mountainous islands.

Aubry et al. (2007) suggest that wolverine range is limited by areas covered by persistent snow through the end of the denning period in mid May. In 8 radio-telemetry studies in the conterminous US and in British Columbia, Canada, areas characterized by persistent spring snow through the denning period (hereafter “snow layer”) contained all recorded dens and 89% of all telemetry relocations (J. Copeland pers. Comm). Therefore, from a practical standpoint the snow layer defined the areas where wolverines both reproduced and utilized on a daily basis.

Based on their size, these islands are, from the standpoint of population stability, extremely small. Many likely contain a single family consisting of a mature male, several mature females and their progeny. For population stability, likely the single most important attribute is the ability to provide demographic and genetic rescue, and to rapidly recolonize areas when local extinctions occur (see Squires et al. 2007). We therefore needed to determine the dispersal rules and likely travel routes used by wolverines.

A spatially distributed sample of tissues from 210 wolverines existed for Montana, Idaho, and northern Wyoming. These data were derived from various studies and from trapping in Montana (Schwartz et al. in press). These samples have been analyzed in several papers (Kyle and Strobeck 2002; Cegelski et al. 2003, 2006; Guillot et al. 2005) and clustering algorithms have been used to define sub-populations. Additionally, autocorrelation analysis (Smouse and Peakall 1999) indicates a positive gradient of correlation extending approximately 180 km (Schwartz et al. in press). While both the clustering analyses and autocorrelation demonstrate that genetic data are spatially structured, the presence of strong gradients coupled with irregular sampling procedures likely lead to artificial clusters not indicative of actual movement and isolation patterns (Schwartz and McKelvey 2009). Gradient approaches, however, are appropriate for these types of genetic patterns and samples.

Using the approach of comparing pair-wise genetic distance with putative landscape resistance maps (Coulon et al. 2004, 2006; Cushman et al. 2006), Schwartz et al. tested the hypothesis that wolverines dispersed by keeping as much as possible within the mountainous areas where they both were born and to which they limited their daily movements. Because of the concordance between the snow layer, denning and daily movement, Schwartz et al. constructed resistance surfaces in which movements exterior to the areas of snow had higher resistance than movements staying within snow areas. Following Coulon et al. (2004, 2006) and Cushman et al. (2006) Schwartz et al. used least cost paths to construct the minimum resistance paths between each pair of individuals and used Mantel tests to measure the correlation between least cost path resistances and genetic distance. These correlations were tested against a Mantel test comparing genetic distance to Euclidian distance. Lastly, partial Mantel tests were used to separate the effects of the snow layer from the effects of Euclidian distance, and determine which was likely the primary cause of the observed correlations.

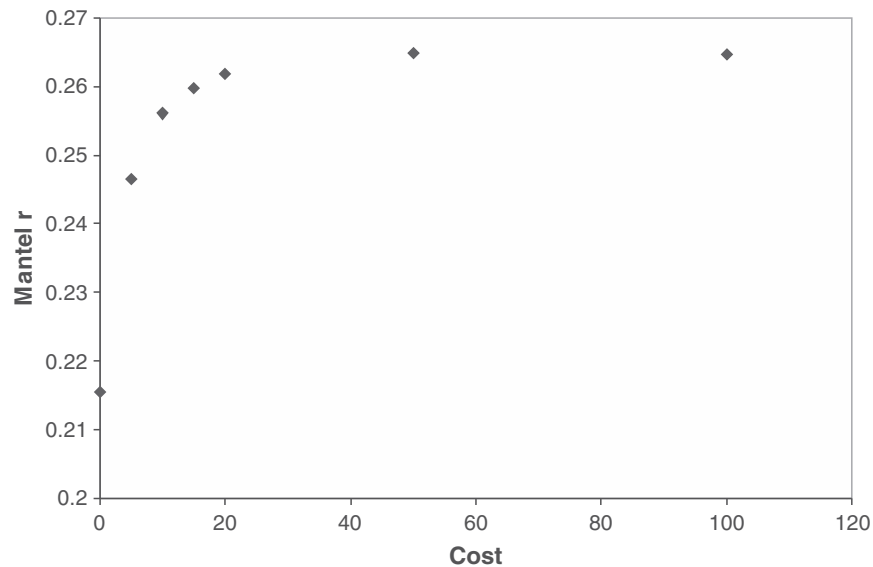


Fig. 17.1 Correlations against cost for various models testing whether areas characterized by persistent spring snow facilitated wolverine dispersal. Costs are the relative cost of moving through non-snow rasters (snow rasters always have Cost = 1)

For wolverines, rather extreme models, in which least cost resistance paths were maximally within the snow layer performed best, and all models with less resistance through snow areas were significantly better than the Euclidian distance model (Fig. 17.1).

The major goal of the analysis was to quantify the degree of population connectivity across a large portion of the Northern Rocky Mountains and identify important potential corridors or barriers to movement. As the identified resistance map only provides point-wise information and connectivity is mediated through movement, Schwartz et al. utilized the source–destination least path approach introduced by Cushman et al. (2008) to map broad scale landscape connectivity. Because wolverines were well distributed within the mountain islands and because all dens were within these areas, Schwartz et al. placed theoretical start locations in all snow areas large enough to hold a female wolverine home range (and hence a den). Using the best resistance model identified above, they built pairwise least cost paths between all theoretical start locations to determine the optimal routes from any area where a wolverine might be born to any area where a wolverine might breed and leave its genes behind. Kernels were draped across each path and the kernels added to produce a visualization of theoretical path densities across the landscape (Fig. 17.2).

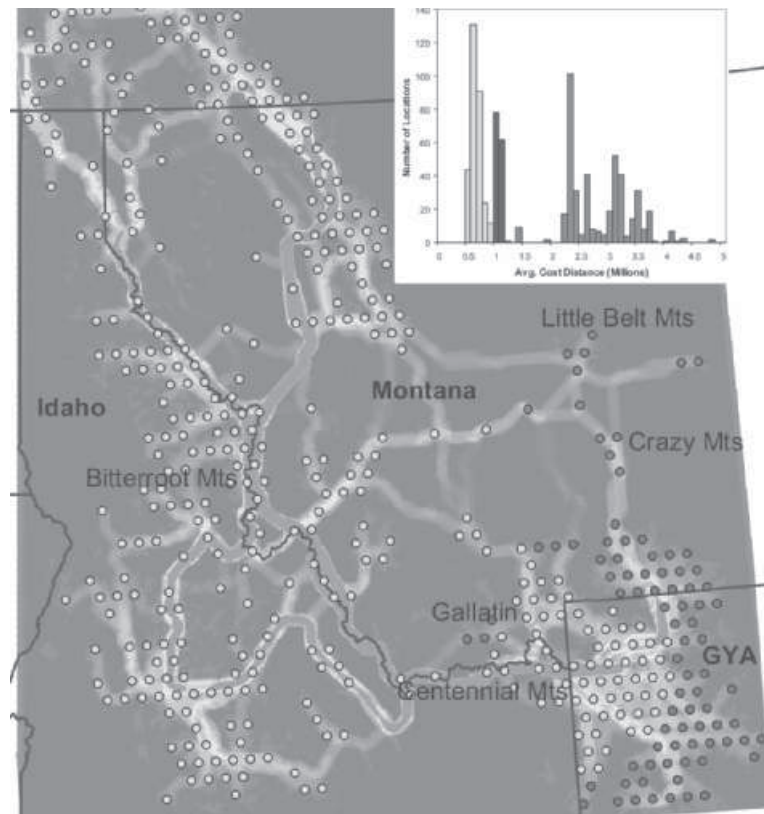


Fig. 17.2 (a) Map of the northern US Rockies with cumulative least-cost paths between systematically placed locations (circles) in spring snow cover cells. Areas in red are predicted to be used more often than those in cooler colors. The color of the circle corresponds to the average cost distance between that location and all other locations, based on modeling. (b) Cumulative plot of the average cost distance (in millions of cost units) between each systematically placed location and all other locations. The graph was divided into four modes (three within the northern US Rockies, and 1 between the Greater Yellowstone Area and Colorado). The yellow mode has the lowest average cost distances (within the Northern US Rockies), the blue bars the next lowest, the pink bars (Crazy and Little Belt Mountains) have the greatest average cost distances in the Northern US Rockies, and the green bars show the distances between all points from Colorado to the Greater Yellowstone Area

17.4.2 Bear Example

The second example is an effort to predict landscape resistance as experienced by American black bear as a function of elevation, slope, roads, land-use and forest cover. The study area consisted of an approximately 3,000-km² area of the Selkirk and Purcell mountain ranges in the extreme northern tip of Idaho State, USA (Cushman et al. 2006) and to apply this model to predict movement corridors and

identify potential barriers for gene flow between Yellowstone National Park and the Canadian Border in the Montana, USA. (Cushman et al. 2008).

The topography in the Cushman et al. (2006) study area is mountainous, with elevation ranging from approximately 700–2,400 m. The Kootenai River trench bisects the study area, separating the Selkirk Mountains on the west from the Purcell Mountains on the east with a 5–8-km-wide unforested, agricultural valley and a broad, deep river. Mountains outside the valley are heavily forested, with *Abies lasiocarpa* (subalpine fir) and *Picea engelmannii* (Engelmann spruce) codominant above 1,300 m and a diverse mixed conifer forest dominating below 1,300 m.

We used noninvasive hair snaring to obtain genetic samples, following the protocols of Proctor et al. (2005). Sampling stations were set at 266 plots distributed on a 1.6-km grid. We detected no deviations from Hardy–Weinberg proportions, found no evidence of gametic disequilibrium, and expected heterozygosity in the Purcell and Selkirk mountain ranges were 0.78 and 0.80, respectively, while observed heterozygosity was 0.76 and 0.80 for each range, respectively (Schwartz et al. 2006).

A priori, we identified three potential drivers of genetic structure in this black bear population, including isolation by the Kootenai River valley as a barrier, isolation by geographic distance, and isolation by landscape-resistance gradients. Our goal was to determine the relative support for isolation by distance and barriers in comparison to isolation by gradients of landscape resistance.

In addition, there are many possible alternative models for landscape effects on population connectivity of this species. A priori, we proposed elevation, roads, topographical slope, land-use and land-cover as factors hypothesized to affect gene flow in this species. In addition, each of these factors could influence bear gene flow in a number of potential ways. Therefore, we constructed a factorial of multiple levels of each of these factors, totaling 108 alternative landscape resistance hypotheses (Cushman et al. 2006).

We used the Bray–Curtis percentage dissimilarity measure to calculate genetic distance among individuals (Legendre and Legendre 1998), producing a matrix containing the genetic distances among all pairs of sampled bears. We used the ArcGIS to produce cost distance matrices that correspond to each of the 110 alternative models (108 landscape models, isolation by distance, and isolation by the Kootenai River barrier). The cost distance matrices corresponding to each landscape resistance hypotheses were created using the COSTDISTANCE function (ESRI 2003). This entailed creating cost matrices reflecting the least-cost distance from the location at which each bear's DNA sample was recorded to every other bear's location across each of the 108 resistance surfaces.

We next identified which of the seven models had the strongest support. All 110 resistance hypotheses were significantly related to the pattern of genetic difference among bears. The barrier hypothesis was ranked 102 out of 110 in terms of the magnitude of the Mantel correlation coefficient. The isolation-by-distance hypothesis was more strongly supported and was ranked 35 out of the 110 total models. The most supported model predicts that gene flow in this population is influenced predominantly by landscape-resistance gradients with no significant independent relationships with the Kootenai River barrier or geographical distance.

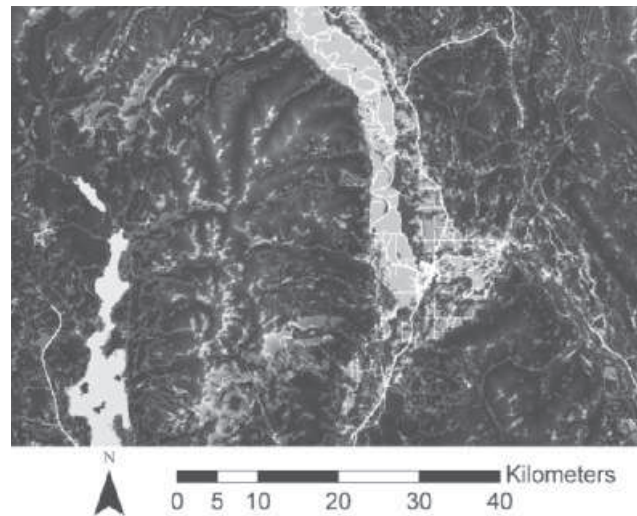


Fig. 17.3 Predicted landscape resistance to black bear gene flow in Idaho Study Area, inferred from causal modeling, Cushman et al. (2006). Color ramp ranges from low resistance (relative resistance 1) in black to high resistance (relative resistance maximum 62) in white

After the effects of distance are partialled out, ranking the ten significant partial models by the Monte Carlo permuted *p*-value value provided a means to determine which hypotheses have the greatest support and to identify the combination of landscape factors most related to the genetic structure of this population. The ten significant partial models were all concentrated in one small area of a four-dimensional factorial space (roads, elevation, forest cover, slope), indicating a unimodal peak of support. Within this small area of concentration, the best-supported models were associated with minimum resistance to movement at middle elevations, high resistance of nonforested habitat to movement, no relationship with slope, and equivocal support for the different levels of road resistance. A map extrapolating the understanding from the most highly supported model across the entire study area is shown in Fig. 17.3 .

Similar to the approach used to model wolverine corridors (above), we modeled movement from source locations distributed at 2-km intervals along the forested portion of the Canadian border between the Idaho–Washington state border and the eastern edge of Glacier National Park (160 locations) to 160 destination locations along the northern boundary of Yellowstone National Park (Cushman et al. 2008). We then computed the least-cost paths across the best-supported resistance surface developed by Cushman et al. (2006) for all combinations of the 160 northern and 160 southern points (25,600 paths).

Despite the large number of combinations, all least-cost paths followed only three major routes (Fig. 17.4). The most important route, in terms of the number of source–destination paths it contained, followed the eastern edge of the mountainous terrain along the Bridger and Belt mountains and crossed into the Bob Marshall Wilderness

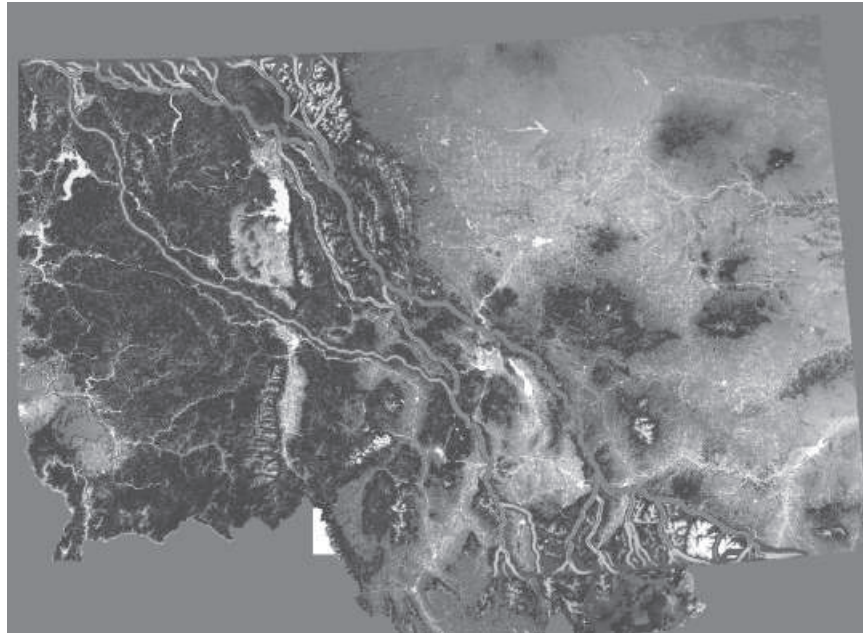


Fig. 17.4 Map of expected importance of least cost routes from Yellowstone National Park to the Canadian border based on the best resistance map identified in Cushman et al. 2006. The relative strength of the predicted route is shown in a color ramp from red (very highly important) to blue (relatively less important) based on the number of source–destination paths crossing that location

and then into Glacier National Park. The second major route proceeded north through the Gravelly and Tobacco Root ranges and then proceeded north roughly following the foothills along the east side of the Swan and North Fork Flathead rivers. The final route was much smaller, in terms of the number of source–destination paths it contained, but was important for connectivity between the northwestern part of the study area and the GYA. This route followed the Clark Fork River and continued northward through the Cabinet and Selkirk mountains to Canada. The majority of the length of the major paths fell within federal ownership (US Forest Service [USFS] and National Park Service [NPS]). Nevertheless, only 24% of the major paths length fell within designated wilderness or national park, and less than half were in wilderness, park, or designated roadless areas (Cushman et al. 2008).

17.5 Conclusions

Landscape genetics is still a nascent science. While we continue to be amazed both by the power and flexibility of these data, we are just beginning to scratch the surface. The strength and consistency of the results associated with the wolverine

and bear examples in this chapter greatly exceeded our expectations. As new approaches emerge and, with the cost of genetic analyses halving every 2 years (Hauser and Seeb 2008), these approaches will likely become pervasive in the next decade (see Chapter 9, this volume). There are already strong movements to augment ongoing traditional monitoring efforts with genetic sampling. In many cases, the additional cost of collecting genetic material is trivial. This trend will lead to spatial genetic datasets that are much more regular and extensive than are currently available. In short, we expect one of the most commonly analyzed types of spatial datasets in the near future will be genetic datasets, as they provide powerful questions to long standing ecological and conservation questions. However, this rapid advance in availability of large genomic data sets will present major computational and database challenges. The size of genomic datasets is vast, which poses a major challenge for storage, organization and file access. In addition, analyzing these large data sets to extract meaningful information in the form of neutral markers, introns and exons is a massive datamining challenge. For the potential of landscape genetics to be realized there must be simultaneous advances in ecological informatics (Chapter 8, this volume), database architecture (Chapter 12, this volume), and major improvements in statistical data mining approaches to associate genetic patterns with landscape and environmental features (Chapter 16, this volume).

References

- Allendorf FW, Luikart G (2007) Conservation and the genetics of populations. Blackwell, Hoboken, NJ
- Aubry KB, McKelvey KS, Copeland JP (2007) Distribution and broadscale habitat relations of the wolverine in the contiguous United States. *J Wildl Manag* 71:2147–2158
- Awise JC (2004) The hope, hype, and reality of genetic engineering. Oxford University Press, Oxford
- Balkenhol N, Gugerli F, Cushman SA, Waits L, Coulon A, Arntzen J, Holderegger R, Wagner H (2007) Identifying future research needs in landscape genetics: where to from here? *Landscape Ecol* 24:455–463
- Barbujani G, Oden NL, Sokal RR (1989) Detecting regions of abrupt change in maps of biological variables. *System Zool* 38:376–389
- Barbujani G, Sokal RR (1990) Zones of sharp genetic change in Europe are also linguistic boundaries. *Proc Nat Acad Sci* 87:1816–1819
- Blattner FR, Plunkett III G, Bloch CA, et al (1997) The complete genome sequence of *Escherichia coli* K-12. *Science* 277:1453–1462
- Beebee, Rowe (2007) An introduction to molecular ecology. 2nd edition. Wiley, New York.
- Bromham L, Woolfit M (2004) Explosive radiations and the reliability of molecular clocks: island endemic radiations as a test case. *System Biol* 53:758–766
- Cegelski C, Waits L, Anderson N (2003) Assessing population substructure and gene flow in Montana wolverines (*Gulo gulo*) using assignment-based approaches. *Mol Ecol* 12:2907–2918
- Cegelski CC, Waits LP, Anderson NJ, Flagstad O, Strobeck C, Kyle CJ (2006) Genetic diversity and population structure of wolverine (*Gulo gulo*) populations at the southern edge of their current distribution in North America with implications for genetic viability. *Conserv Genet* 7:197–211

- Coulon A, Cosson J-F, Angibault JM, et al (2004) Landscape connectivity influences gene flow in a roe deer population inhabiting a fragmented landscape: an individual-based approach. *Mol Ecol* 13:2841–2850
- Coulon A, Guillot G, Cosson J, Angibault J, Aulagnier S, et al (2006) Genetics structure is influenced by landscape features. Empirical evidence from a roe deer population. *Mol Ecol* 15:1669–1679
- Copeland JP, Peek JM, Groves CR, et al (2007) Seasonal habitat associations of the wolverine in central Idaho. *J Wildl Manag* 71:2201–2212
- Copeland JP, McKelvey KS, Aubry KB, et al (In review) The bioclimatic envelope of the wolverine: Do environmental constraints limit their geographic distribution?
- Cushman SA, McKelvey KS, Hayden J, Schwartz MK (2006) Gene-flow in complex landscapes: testing multiple hypotheses with causal modeling. *Am Nat* 168:486–499
- Cushman SA, McKelvey KS, Schwartz MK (2009) Identifying corridors for the movement of black bear (*Ursus americanus*) between Yellowstone National Park and the Canadian Border. *Conserv Biol* 23:368–376
- Dupanloup I, Schneider S, Excoffier L (2002) A simulated annealing approach to define the genetic structure of populations. *Mol Ecol* 11:2571–2581
- ESRI (2003) ArcGIS. Environmental systems research incorporated, Redlands, CA
- Frantz AC, Cellina S, Krier A, Schley L, Burke T (2009) Using spatial Bayesian methods to determine the genetic structure of a continuously distributed population: clusters or isolation by distance? *J App Ecol*. doi: 10.1111/j.1365-2664.2008.01606.x
- Garroway CJ, Bowman J, Carr D, Wilson PJ (2008) Applications of graph theory to landscape genetics. *Evolut Appl* 1:620–630
- Gebremedhin B, Ficetola GF, Naderi S, Rezaei H.-R, Maudet C, Rioux D, Luikart G, Flagstad Ø, Thuiller W, Taberlet P (2009) Frontiers in identifying conservation units: from neutral markers to adaptive genetic variation. *Anim Conserv* 12:107–109
- Guillot G, Mortier F, Estoup A (2005) Geneland: a computer package for landscape genetics. *Mol Ecol Notes* 5:712–715
- Hauser L, Seeb JE (2008) Advances in molecular technology and their impact on fisheries genetics. *Fish Fisheries* 9:473–486
- Hornocker MG, Hash HS (1981) Ecology of the wolverine in northwestern Montana. *Canad J Zool* 59:286–1301
- Holderegger R, Wagner HH (2008) Landscape genetics. *Biosci* 58:199–207
- Kendall, KC, McKelvey KS (2008) Hair Collection. In R. Long, P. MacKay, J. Ray, and W. Zielinski, editors. *Noninvasive survey methods for North American Carnivores*. Island. Washington, DC
- Kyle CJ, Strobeck C (2002) Connectivity of peripheral and core populations of North American wolverines. *J Mammal* 83:1141–1150
- Legendre P (1993) Spatial autocorrelation: trouble or new paradigm? *Ecology* 74:1659–1673
- Legendre P, Legendre L (1998) *Numerical ecology*, 2nd edition. Elsevier, Amsterdam
- Legendre P, Troussellier M (1988) Aquatic heterotrophic bacteria: modeling in the presence of spatial autocorrelation. *Limnol Oceanography* 33:1055–1067
- Luikart G, England PR, Tallmon D, Jordan S, Taberlet P. (2003) The power and promise of population genomics: from genotyping to genome typing. *Nat Rev Genet* 4:981–994.
- Luikart G, England PR (1999) Statistical analysis of microsatellite DNA data. *Trends Ecol Evol* 14:253–256.
- Magoun AJ, Copeland JP (1998) Characteristics of wolverine reproductive den sites. *J Wildl Manag* 62:1313–1320
- Manel S, Schwartz MK, Luikart G, Taberlet P (2003) Landscape genetics: the combination of landscape ecology and population genetics. *Trends Ecol Evol* 18:189–197
- McRae BH (2006) Isolation by resistance. *Evolution* 60:1551–1561
- Monmonier MS (1973) Maximum-difference barriers: an alternative numerical regionalization method. *Geograph Anal* 3:245–261

- Novembre J, Stephens M (2008) Interpreting principal component analyses of spatial population genetic variation. *Nat Genet* 40:646–649
- Paetkau D, Shields GF, Strobeck C (1998) Gene flow between insular, coastal and interior populations of brown bears in Alaska. *Mol Ecol* 7:1283–1292
- Paetkau D, Strobeck C (1994) Microsatellite analysis of genetic variation in black bear populations. *Mol Ecol* 3:489–495
- Prichard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155:945–959
- Proctor MF, McLellan BN, Strobeck C, Barclay RMR (2005) Genetic analysis reveals demographic fragmentation of grizzly bears yielding vulnerability by small populations. *Proc Roy Soc B: Biol Sci* 272:2409–2416
- Rousset F (2000) Genetic differentiation between individuals. *J Evol Biol* 13:58–62
- Schwartz MK, Copeland JP, Anderson NJ, et al (In Press) Wolverine gene flow across a narrow climatic niche. *Ecology*
- Schwartz MK, Cushman SA, McKelvey KS, Hayden J, Engkjer C (2006) Detecting genotyping errors and describing black bear movement in northern Idaho. *Ursus* 17:138–148
- Schwartz MK, McKelvey KS (2009) Why sampling scheme matters: the effect of sampling scheme on landscape genetic results. *Conserv Genet* 10:441–452
- Schwartz MK, Monfort SL (2008) Genetic and endocrine tools for carnivore surveys. R. Long, P. MacKay, J. Ray, and W. Zielinski (eds.) *Noninvasive survey methods for North American Carnivores*. Island, Washington, DC
- Simoni L, Guerresi P, Pettener D, Barbujani G (1999) Patterns of gene flow inferred from genetic distances in the Mediterranean region. *Human Biol* 71:399–415
- Smouse PE, Peakall R (1999) Spatial autocorrelation analysis of individual multiallele and multilocus genetic structure. *Heredity* 82:561–573
- Squires JR, Copeland JP, Ulizio TJ, Schwartz MK, Ruggiero LF (2007) Sources and patterns of wolverine mortality in western Montana. *J Wildl Manag* 71:2213–2220
- Stenico M, Nigro L, Barbujani G (1998) Mitochondrial lineages in Ladin-speaking communities of the eastern Alps. *Proc Roy Soc London, Ser B, Biol Sci* 265:555–561
- Storfer A, Cross J, Rush V, Caruso J (1999) Adaptive coloration and gene flow as a constraint to local adaptation in the streamside salamander, *Ambystoma barbouri*. *Evolution* 53:889–898.
- Taberlet P, Waits LP, Luikart G (1999) Non-invasive sampling: look before you leap. *Trends Ecol Evol* 14:323–327.
- Van Oppen MJH, Rico C, Deutsch JC, Turner GF, Hewitt GM (1997) Isolation and characterization of microsatellite loci in the cichlid fish *Pseudotropheus zebra*. *Mol Ecol* 6:387–388
- Venter JC, Adams MD, Myers EW, et al (2001) The sequence of the human genome. *Science* 291:1304–1351
- Wright S (1943) Isolation by distance. *Genetics* 28:114–138