

## REVIEW

## Putting the 'landscape' in landscape genetics

A Storfer<sup>1</sup>, MA Murphy<sup>1</sup>, JS Evans<sup>2</sup>, CS Goldberg<sup>3</sup>, S Robinson<sup>3</sup>, SF Spear<sup>1</sup>, R Dezzani<sup>4</sup>, E Delmelle<sup>4</sup>, L Vierling<sup>5</sup> and LP Waits<sup>3</sup>

<sup>1</sup>School of Biological Sciences, Washington State University, Pullman, WA, USA; <sup>2</sup>USDA Forest Service – Rocky Mt. Research Station, Moscow, ID, USA; <sup>3</sup>Department of Fish and Wildlife Resources, University of Idaho, Moscow, ID, USA; <sup>4</sup>Department of Geography, University of Idaho, Moscow, ID, USA and <sup>5</sup>Department of Rangeland Ecology and Management, University of Idaho, Moscow, ID, USA

Landscape genetics has emerged as a new research area that integrates population genetics, landscape ecology and spatial statistics. Researchers in this field can combine the high resolution of genetic markers with spatial data and a variety of statistical methods to evaluate the role that landscape variables play in shaping genetic diversity and population structure. While interest in this research area is growing rapidly, our ability to fully utilize landscape data, test explicit hypotheses and truly integrate these diverse disciplines has lagged behind. Part of the current challenge in the development of the field of landscape genetics is bridging the communication and knowledge gap between these highly specific and technical disciplines. The goal of this review is to

help bridge this gap by exposing geneticists to terminology, sampling methods and analysis techniques widely used in landscape ecology and spatial statistics but rarely addressed in the genetics literature. We offer a definition for the term 'landscape genetics', provide an overview of the landscape genetics literature, give guidelines for appropriate sampling design and useful analysis techniques, and discuss future directions in the field. We hope, this review will stimulate increased dialog and enhance interdisciplinary collaborations advancing this exciting new field.

*Heredity* (2007) **98**, 128–142. doi:10.1038/sj.hdy.6800917; published online 1 November 2006

**Keywords:** landscape genetics; spatial statistics; spatial analysis; landscape ecology; spatial sampling; population genetics

## Introduction

Technological innovations in spatial analyses coupled with increased availability of spatial data and hypervariable genetic markers have resulted in great advances in our ability to study the influence of landscape variables, such as altitude, topography and ground cover, on genetic variation and structure. As a result, landscape genetics (Manel *et al.*, 2003) has emerged as a new research area that integrates landscape ecology, spatial statistics and population genetics. In contrast to traditional population genetics studies that were limited in spatial inference to tests of isolation-by-distance, landscape genetics provides a framework for testing the relative influence of landscape and environmental features on gene flow, genetic discontinuities (Guillot *et al.*, 2005a) and genetic population structure (Manel *et al.*, 2003; Holderegger and Wagner, 2006).

Understanding landscape effects on genetic connectivity provides insight into fundamental biological processes such as: metapopulation dynamics, speciation, and ultimately the formation of species' distributions. Landscape genetic analyses can also have great applied scientific value, such as identifying specific anthropogenic barriers that reduce gene flow or genetic diversity,

predicting the effects of proposed management alternatives on genetic variation and population connectivity, and identifying potential biological corridors to assist with reserve design.

Given these diverse research opportunities, landscape genetics is both challenging and exciting, as it brings together scientists from the broad disciplines of landscape ecology, spatial statistics, geography and population genetics. While several types of spatial statistical analyses have been used in geographical genetics (for review, see Epperson, 2003), there are many well-developed methods in landscape ecology and spatial statistics that have yet to be utilized. The vast array of spatial analysis techniques that can be applied to population genetic data make options for designing and conducting a landscape genetics study extremely diverse and potentially confusing. Better communication among landscape ecologists, spatial statisticians, remote-sensing scientists, geographers and population geneticists is key to integrating analysis methods and empirical data. To help bridge communication gaps, we have included a glossary of terminology used in spatial statistics and landscape ecology (Table 1), with the terms denoted in italics when first used in the text.

Our goals are to: (a) offer a definition of the term 'landscape genetics'; (b) review questions commonly addressed in the landscape genetics literature, (c) provide guidelines for sampling design, (d) highlight potentially useful analysis techniques; and (e) discuss future directions for the field. The most commonly used molecular tools for landscape genetic studies are neutral,

Correspondence: Dr LP Waits, Department of Fish and Wildlife Resources, University of Idaho, PO Box 441136, Moscow, ID 8344-1136, USA.

E-mail: lwaits@uidaho.edu

Received 13 April 2006; revised 20 September 2006; accepted 22 September 2006; published online 1 November 2006

**Table 1** Glossary of spatial statistics/landscape ecology/geographical genetics terms used in the text

<i>Term</i>	<i>Definition</i>	<i>Reference(s)</i>
Anisotropic (anisotropy)	Intensity and range of spatial dependence ( <i>autocorrelation</i> ) varies with direction and/or orientation. Directionality in environmental gradients, wind or water flow may lead to anisotropic patterns	Haining (2003), Fortin and Dale (2005)
Correlogram	A graph of the correlation coefficient calculated at various distance classes	Isaaks and Srivastava (1989)
Correspondence analysis	An ordination technique used to identify associations between variables by reciprocal averaging. There are several types of correspondence analysis	terBraak (1995), Angers <i>et al.</i> (1999)
CCA	An extension of reciprocal averaging techniques which combines regression and ordination techniques requiring that the ordination axes be linear combinations of the explanatory variables	
Gradient (analysis)	A continuous change in a variable over distance such as moisture or elevation. Gradient analysis has been popular in relating species abundance to continuous changes in environmental variables	Jongman <i>et al.</i> (1995)
Interpolation	Estimating data at unsampled locations using a mathematical model of the spatial pattern of sampled values. There are many types of interpolation models (e.g. inverse distance weighting, kriging models, tessellation models)	Isaaks and Srivastava (1989)
Kriging	A method of <i>interpolation</i> which uses known values at sampled locations and parameters from the <i>semivariogram</i> to estimate values at unsampled locations	Isaaks and Srivastava (1989)
Least-cost path	A line (vector) of 'least resistance' is generated between two observation points. The resistance refers to a resistance surface, or cost grid, generated according to some understanding of mobility (in landscape genetics, species mobility). The least-cost distance can be expressed in terms of geographic distance or total resistance along this path	Singleton <i>et al.</i> (2002), Adriaensen <i>et al.</i> (2003)
Moran's I	A weighted correlation coefficient used to detect departures from complete spatial randomness (positive or negative <i>spatial autocorrelation</i> ). The weighting scheme may be defined by distance between points or contiguity of polygons	Moran (1950)
Network analysis	Graph theory is applied to model and analyze linkages and flows in a network – a system of directional lines (vectors) connecting all of the observation points (nodes). A classic application of network analysis is measuring level and direction of traffic flows through road systems	Taaffe <i>et al.</i> (1996)
Neutral landscape models	A randomized landscape following some statistical quality of the original data. Used for permutation testing of hypotheses concerning landscape pattern or process	Gardner <i>et al.</i> (1987), Saura and Martínez-Millán (2001)
Point pattern statistics	A group of statistics used for analysis of observations collected at specific X, Y coordinates. The statistics can be used to quantify the pattern (e.g. clustered, uniform, random) of the points themselves or some response variable associated with the points	Diggle (2003)
Semivariogram	A graph of the <i>spatial dependence</i> between observations calculated at various distance classes. <i>Semivariogram</i> values become larger as observations become less similar, thereby offering an estimate of the distance needed between points before they become uncorrelated	Haining (2003), Fortin and Dale (2005)
Spatial autocorrelation	Observations are not independent identically distributed (iid), but are correlated over some distance in space. Positive autocorrelation indicates that nearby values are similar; negative autocorrelation indicates that nearby values are dissimilar	Haining (2003)
Spatial dependence	The influence of spatial relationships on observation attributes	Haining (2003)
Spatial generating process	Process that leads to the variation among observations over space	Turner <i>et al.</i> (2001)
Stationarity	Indicates that the variable of interest is influenced by a single process whose properties do not change with location or direction – it has a constant mean and variance across the study area	Bailey and Gatrell (1995), Fortin and Dale (2005)
Tessellation	A pattern of non-overlapping polygons is used to partition the area leaving no gaps. In landscape genetics, tessellation is generally used to create polygons around each point where genotypes are collected	Bailey and Gatrell (1995)

**Table 2** Recent works in landscape genetics, highlighting sampling design, molecular markers used, analytical techniques, and basic conclusions

Citation	Species	Sampling design	Marker(s)	Analysis methods	Inference
Arnaud (2003)	Landsnail ( <i>Helix aspersa</i> )	Unspecified	4 polymorphic enzymes 5 microsatellites	Mantel tests, correlograms, Moran's I	Migration occurs along hedge rows
Banks <i>et al.</i> (2005)	Marsupial ( <i>Antechinus agilis</i> )	Systematic	15 microsatellites mtDNA control region	Simulated annealing approach, partial Mantel, spatial autocorrelation, correlograms	Eucalyptus habitat corridors significantly increased gene flow relative to pine plantations
Cegelski <i>et al.</i> (2003)	Wolverine ( <i>Gulo gulo</i> )	Opportunistic (harvest)	10 microsatellites	F-statistics, Bayesian assignment	Wolverine populations are more genetically structured in fragmented habitat
Funk <i>et al.</i> (2005)	Spotted frog ( <i>Rana luteiventris</i> )	Stratified	6 microsatellites	AMOVA; (partial) Mantel test; Bayesian assignment test	Ridges and elevation are major barriers to gene flow
Guillot <i>et al.</i> (2005a)	Simulation & wolverine ( <i>Gulo gulo</i> )	Opportunistic (harvest)	10 microsatellites	Bayesian assignment, Voronoi tessellation	Identification of multiple populations, and cryptic barriers, with improved power over aspatial methods
Hirao and Kudo (2004)	Snowbed herbs ( <i>Peucedanum multivittatum</i> , <i>Veronica stelleri</i> , <i>Gentiana nipponica</i> )	Clusters	11 allozymes	(partial) Mantel test	Flowering time is more influential on genetic variation than isolation-by-distance
Hitchings and Beebe (1997)	common frog ( <i>Rana temporaria</i> )	Opportunistic (within urban and rural)	19 allozymes	F-statistics, Mantel test	Genetic diversity, gene flow, and fitness were lower in urban versus rural populations
Jacquemyn (2004)	Forest herb ( <i>Primula elatior</i> )	Random stratified	3 AFLP	Mantel test; AMOVA	Herbs found in younger forests had lower genetic diversity and differentiation than those in older forests
Keyghobadi <i>et al.</i> (1999, 2005a,b)	Alpine butterfly ( <i>Parnassius smintheus</i> )	Opportunistic (within meadows)	7 microsatellites	(partial) Mantel tests	Genetic distances between butterfly populations were increased through forest relative to meadows
Liepelt <i>et al.</i> (2002)	Silver fir ( <i>Abies alba</i> )	Opportunistic	mtDNA cpDNA	Haplotype maps, variogram model	Difference in semivariograms suggested wind pollinated influence on gene flow between refugia
Manni <i>et al.</i> (2004)	Simulation & human ( <i>Homo sapiens</i> )	Opportunistic	Y chromosome	Monmonier's algorithm	Identified significant genetic barriers between northwestern Africa and Iberian Peninsula
Michels <i>et al.</i> (2001)	Zooplankton ( <i>Daphnia ambigua</i> )	Connected ponds- clustered	12 allozymes	Mantel test	Gene flow was best correlated with two models – flow rate between ponds and dispersal rate
Pfenninger (2002)	Terrestrial snail ( <i>Pomatias elegans</i> )	Opportunistic (within different habitats)	1 allozyme	(partial) Mantel test; spatial autocorrelation	Gene flow correlated with ridge distance in two populations and, with habitat distance in one population
Piertney <i>et al.</i> (1998)	Red grouse ( <i>Lagopus lagopus scoticus</i> )	Opportunistic (harvest)	7 microsatellites	Mantel test, PCA, kriging (interpolation)	Interpolation of PCA scores identified genetic discontinuity, which corresponded to unsuitable habitat
Poissant <i>et al.</i> (2005)	Brook charr ( <i>Salvelinus fontinalis</i> )	Opportunistic (within both highland and lowland areas)	9 microsatellites	Neighbor-joining tree; AMOVA; <i>t</i> -test; (partial) Mantel	Gene flow correlated best with a path based on historical connectivity, but contemporary barriers limited gene flow
Roach <i>et al.</i> (2001)	Black-tailed prairie dogs ( <i>Cynomys ludovicianus</i> )	Opportunistic	7 microsatellites	Assignment tests, Mantel tests, information theoretic approach	Dispersal between colonies occurs regularly and likely along drainages
Scribner <i>et al.</i> (2005)	White tailed deer ( <i>Odocoileus virginianus</i> )	Opportunistic by quadrat	3 microsatellites	Mantel test	Greater thermal cover was significantly correlated with greater gene flow
Spear <i>et al.</i> (2005)	Tiger salamander ( <i>Ambystoma tigrinum</i> )	Opportunistic	8 microsatellites	Partial Mantel test; Spearman rank correlation coefficient	More of the genetic variability explained by including landscape variables than with distance alone
Vignieri (2005)	Pacific jumping mouse ( <i>Zapus trinotatus</i> )	Opportunistic (within three different river areas)	8 microsatellites	Mantel test; Moran's I	Least-cost habitat path that maximized riparian corridors and minimized elevation gain had the greatest correlation with genetic distance

hypervariable markers (e.g. amplified fragment length polymorphisms and microsatellites), and we generally assume their use throughout this review.

## Defining landscape genetics

In general, landscape genetics seeks to understand the influence of ecological processes (Turner *et al.*, 2001) on genetic variation by quantifying the relationship between landscape variables, population genetic structure and genetic variation (the latter two hereafter collectively referred to as 'genetic variation'). Since Manel *et al.* (2003) coined the term 'landscape genetics,' a diversity of published articles have been labeled landscape genetics because they incorporated geographic coordinates or landscape features when evaluating the spatial distribution of genetic variation (Table 2). These studies have varied extensively in their approach to evaluating relationships between landscape variables and genetic variation. We suggest that landscape genetics studies could benefit substantially by including explicit tests of the relative influence of landscape variables on genetic variation by incorporating robust, spatially informed study designs and spatial analyses. Thus, in this review, we define landscape genetics as research that explicitly quantifies the effects of landscape composition, configuration and matrix quality on gene flow and spatial genetic variation. This definition expands on the description of landscape genetics in Holderegger and Wagner (2006). Phylogeography (Avice, 2000) can also be used to quantify genetic variation in relation to ecological processes, but at a larger spatio-temporal scale than landscape genetics, making it more comparable to biogeography (Manel *et al.*, 2003).

## Major research categories in landscape genetics

There are a wide variety of basic and applied research questions that can be addressed using a landscape genetics approach (see Tables 2 and 3). We group these questions under five major research categories: (1) quantifying influence of landscape variables and configuration on genetic variation; (2) identifying barriers to gene flow; (3) identifying source-sink dynamics and movement corridors; (4) understanding the spatial and temporal scale of an ecological process; and (5) testing species-specific ecological hypotheses.

### Influence of landscape variables and configuration on genetic variation

Quantifying the effect of landscape configuration on gene flow has been a major focus of published landscape genetics studies (Manel *et al.*, 2003; Scribner *et al.*, 2005; see Table 2). Statistical analyses of genetic data have been used to identify the effects of matrix resistance on gene flow and genetic structure, including: cover type (Keyghobadi *et al.*, 1999; Spear *et al.*, 2005), stream distance (Roach *et al.*, 2001; Antolin *et al.*, 2006), historic landscape configuration (Holzhauer *et al.*, 2006), water flow rates (Michels *et al.*, 2001), ridge distances (Pfenninger, 2002; Funk *et al.*, 2005), thermal cover (Scribner *et al.*, 2005) and the effect of landscape configuration on allelic fixation time (Ezard and Travis, 2006). Recently, Sezen *et al.* (2005)

revealed the impact of landscape change and patch type on genetic variation in Costa Rican canopy palms (*Iriartea deltoidea*) by documenting a decrease in genetic diversity and an increase in the patch diameter of similar genotypes in second growth forests compared to old growth patches. This demonstrates that intrinsic scale, the area encompassed by a population as estimated by a genetic neighborhood, may change across a landscape due to the landscape composition and configuration. As another example, wolverine populations in intact habitats have a larger intrinsic scale than populations in fragmented habitats (Cegelski *et al.*, 2003).

### Identifying barriers

Identifying potential gene flow barriers is a major focus of landscape genetics research. While all landscape features affect gene flow, particular structures such as roads (Riley *et al.*, 2006), waterways (Antolin *et al.*, 2006) or mountain ridges (Funk *et al.*, 2005) are potentially impenetrable barriers. Genetic data have been used to identify abrupt breaks in gene flow (Dupanloup *et al.*, 2002; Manni *et al.*, 2004) as well as more gradual transitions (Geffen *et al.*, 2004). Barriers may also consist of microhabitats that prevent gene flow because they exceed a threshold for moisture, temperature or chemical tolerance for particular species (Palo *et al.*, 2004). Therefore, barrier identification has important implications for ecological (Walker *et al.*, 2003; Kreyer *et al.*, 2004; Funk *et al.*, 2005), conservation (Bhattacharya *et al.*, 2003; Miller and Waits, 2003; Dodd *et al.*, 2004) and evolutionary (Castella *et al.*, 2000; Broderick *et al.*, 2003; Cicero, 2004; Gee, 2004) investigations.

One distinct benefit of a landscape genetics approach is that spatially explicit techniques can allow researchers to identify barriers not detectable by traditional population genetic methods (Guillot *et al.*, 2005a; Coulon *et al.*, 2006). For example, Coulon *et al.* (2006) found genetic structuring in roe deer due to highways and rivers using spatial assignment test methods (in the program GENELAND, Guillot *et al.*, 2005b), whereas a non-spatial assignment test (STRUCTURE; Pritchard *et al.*, 2000) was not able to identify any genetic discontinuities. Landscape genetics can also be used to quantify the cumulative impact of a particular barrier type distributed across the landscape. For example, Epps *et al.* (2005) evaluated genetic diversity and structure in 27 desert bighorn sheep populations (*Ovis canadensis nelsoni*) and showed that diversity was negatively correlated with fenced highways, canals and human development. They also estimated a 'barrier effect distance' and suggested that any one of these barriers would create the same decrease in gene flow as 40 km of contiguous habitat.

### Source-sink dynamics

Understanding source-sink dynamics (Pulliam, 1988; Dias *et al.*, 1996) and variation in habitat quality can be useful for identifying corridors and guiding reserve design. Genetic data have been used to identify source and sink habitats for populations by identifying asymmetric gene flow using private alleles (Kennington *et al.*, 2003), and estimating the number of migrants into a population using either a coalescent approach (Beerli and Felsenstein, 2001) or assignment tests (Paetkau *et al.*, 1995; Wilson and Rannala, 2003). Theoretical population

**Table 3** Models that can explicitly incorporate landscape data organized by research questions presented in text

Question type	Appropriate model types	Example models	Sampling
Relative influence of landscape variables/ecological hypotheses	Matrix correlation	Mantel test, Partial Mantel test	All
	Autoregressive and geostatistical models	CAR, SAR, GWR	B, C, D, E, F, G, H
	Dispersal route analysis	Least-cost paths, network analysis, flow models	All
	Ordination	RDA, CCA, multidimensional scaling	B, C, D, E, F, G, H, Model based
Identify barriers	Landscape metrics	Many metrics calculated at landscape, class, patch level, and within different window sizes	E, G, H, Model based
	Point pattern	LISA, Cross-K	B, C, D, F, Model based
	Global autocorrelation	Moran's I, Geary's C	B, C, D, Model based
	Spatial interpolation	Inverse distance weighting, kriging	B, C, D, E, F, G, H, Model based
Source-sink	Assignment tests	Monmonier's algorithm, simulating annealing	B, C
	Algorithmic models	CART, MARS, RANDOM FORESTS	All
Reserve design/corridors	Spatially explicit metapopulations	RAMAS	All
	Network analysis	Flow models	All
Spatial scale	Spatial autocorrelation	Moran's I, Geary's C, LISA	B, C, D, Model based
	Path analysis	Least-cost paths, network analysis, flow models	All
	Spatially explicit metapopulations	RAMAS	All
	Matrix correlation	Mantel tests	All
	Landscape simulation	LANDIS, RAMAS GIS	All
Temporal scale	Regression	Geographically Weighted Regression	E, G, H
	Spatial autocorrelation	Moran's I, Geary's C, LISA	B, C, D, Model based
Model validation	Variance decomposition	Hierarchical analysis of variance, delineation of semivariance	All
	Simulation/permutation	Neutral landscapes	All
	Interpolation	Conditional simulation, randomization, Monte Carlo	A, B, C

Letters refer to sampling methods (Figure 1) that are most suited to the model type under most conditions.

models suggest evaluations of linkage disequilibrium can be used to detect sink habitats because disequilibrium is predicted to be higher in individuals from sinks due to immigrants from different sources (Nei and Le, 1973). Dias *et al.* (1996) empirically tested this theory and found higher linkage disequilibrium in blue tits (*Parus caeruleus*) living in a known sink habitat (evergreen forest) than that of birds sampled from a known source habitat (deciduous forest).

Physical locations of corridors have also been identified using landscape genetic approaches. For example, Banks *et al.* (2005) used spatial autocorrelation analysis of genetic structure to demonstrate that riparian strips of native eucalyptus forest facilitate dispersal in a marsupial carnivore (*Antechinus agilis*) compared to the matrix habitat of pine plantations. This study also helped detect a dispersal threshold for male *A. agilis* by showing that fragmented pine plantation habitat significantly reduced dispersal at distances greater than 750 m. *Least-cost analysis* also has been valuable in identifying landscape variables that facilitate gene flow and may function as corridors (Spear *et al.*, 2005; Vignieri, 2005). For example, Vignieri (2005) found that least-cost paths that minimized elevation gain and maximized riparian forest cover were strongly correlated with gene flow in the Pacific jumping mouse (*Zapus trinotatus*).

### Spatial and temporal scales

Genetic variation may respond differently over varying spatial or temporal scales, which is a critical issue in defining research questions and subsequent study design in landscape ecology and spatial statistics (for review, see Gardner, 2001). The scale at which particular landscape variables have the greatest influence on gene flow (i.e. process scale) may give insight into species' biology. For example, Trapnell and Hamrick (2004) showed that the contributions of pollen and seed movement to overall gene flow in the Central American epiphytic orchid, *Laelia rubescens*, were scale-dependent. Primary factors governing gene flow were seed gravity (seed dispersal) at the finest spatial scale, hummingbird behavior (pollen dispersal) at the intermediate scale, and wind (occasional seed dispersal) at the broadest scale.

Temporal scale may also have a significant impact on landscape genetics. For example, Ramstad *et al.* (2004) detected significant genetic structure among sockeye salmon (*Oncorhynchus nerka*) populations in Alaska and used simple and partial Mantel tests to evaluate the relative influence of different ecological and evolutionary factors on genetic differentiation. They found that temporal isolation based on spawning time and founder effects associated with ongoing glacial retreat and colonization of new spawning habitats contributed

significantly to genetic population structure, while geographic distance and spawning habitat differences did not have significant influence.

### Species-specific hypothesis testing

Landscape genetics offers new approaches for testing hypotheses specifically related to how the ecology of the study species shapes patterns of genetic variation, such as identification of bioregions (Sacks *et al.*, 2004), potential response to climate change (Rehfeldt *et al.*, 1999), and ecological variables (Jørgensen *et al.*, 2005). For example, Sacks *et al.* (2004) tested the hypothesis that coyotes would exhibit natal-biased dispersal by evaluating population genetic structure across four contiguous habitat bioregions in Northern California and found genetic groupings could best be explained by habitat bioregions and not habitat barriers. Rehfeldt *et al.* (1999) concluded that the ability of western US conifer species to adapt to climate change may be constrained due to significant geographic structure in quantitative genetic traits associated with elevation and latitude. In a study of herring (*Clupea harengus*) in the Baltic Sea, Jørgensen *et al.* (2005) showed that the ecological variables that best explained genetic variation were salinity, surface temperature and spawning time.

## Study design

### Importance of study design

Historically, population genetic studies often relied on opportunistically collected samples from known localities or in easily accessible areas. Opportunistic sampling may fail to capture the spatial variation or *spatial dependency* of the system, resulting in difficulty detecting spatial relationships or erroneous model inferences (Legendre, 2002; Fortin and Dale, 2005). Thus, studies should be designed to sample the variable(s) of interest within the scale of spatial dependency (Coulon *et al.*, 2004). In addition, population genetics studies have been traditionally designed to collect samples from a minimum of 20–30 individuals per *a priori* delineated 'population' (Nei, 1978). However, *a priori* delineation is no longer necessary due to the development of genetic clustering algorithms, such as assignment tests (Pritchard *et al.*, 2000; Wilson and Rannala, 2003; Manel *et al.*, 2005). In addition, models of landscape influence on genetic variation often require more continuously distributed sampling, which can be analyzed with a wide variety of spatial analysis techniques reviewed herein (see Table 3).

Careful consideration of scale in study design is also critical because arbitrarily defined scales may lead to erroneous conclusions or fail to capture appropriate variability in the landscape (King, 1990; Gardner, 2001). We refer to scale as the appropriate spatial or temporal dimensions at which processes can be observed and quantified (for review, see Dungan *et al.*, 2002). Under this definition, scale has two relevant components: 'grain' and 'extent' (O'Neill *et al.*, 1988). Grain is the minimum resolution of the data and extent is the total area of interest. Accounting for grain and extent in both the dependent and independent variables ensures that the appropriate scale for a specific question(s) is captured (for review, see Marceau, 1999). When design-

ing a landscape genetics study, researchers can determine appropriate sampling scale with pilot data (e.g. movement data of the study organism), or data from similar organisms in other areas. Ideally, a spatial sampling scheme should capture the range of spatial variation in both landscape variables and organismal genetic variability by collecting observations across a range of variances (e.g. from low to high local *spatial autocorrelation* and from low to high genetic relatedness, respectively) (Cressie, 1993; Goovaerts, 1997; Fortin and Dale, 2005).

If pilot data are available, exploration of intrinsic scale and *stationarity* can also be executed before sampling or modeling. Most spatial statistical models assume stationarity in the response variable that is implicitly a function of a *spatial generating process* (Fortin and Dale, 2005). In other words, the underlying assumption of most spatial statistical models is that the same ecological processes (e.g. slope, moisture and topography) influence the response variable (e.g. genetic variation) in the same way across the whole study area. However, this assumption is likely to be violated in complex landscapes. Violations of stationarity can be tested with local indicators of spatial association (LISA) (Anselin, 1995). Non-stationarity can then be accounted for by incorporating locally adjusted methods such as geographically weighted regression or locally coherent spatial regression (Haining, 2003; Fotheringham *et al.*, 2004).

In landscape genetic studies, it is important to consider how the effects of temporal scale influence genetic variation. The theoretical genetics literature has addressed how temporal heterogeneities in demographic parameters such as dispersal and density can have large effects on spatial genetic variation (Whitlock, 1992; Leblois *et al.*, 2004). However, this complexity is often overlooked in landscape genetic studies. For species with overlapping generations and cyclic changes in density, allele frequency shifts among cohorts and sampling years are expected (Jorde and Ryman, 1995). If not accounted for, temporal genetic heterogeneity can be incorrectly interpreted as true population differentiation, particularly in cases when genetic differentiation is weak (Waples, 1998).

### Incorporating landscape data

Landscape genetic studies require data from two distinct sources: landscape data (e.g. remotely sensed data, digital elevation models, field collections) and multilocus genetic data. Landscape data are gathered in a number of ways, including: field surveys, aerial remote sensing and/or satellite remote sensing. As these methods contain several potential sources of error that may obscure relationships with genetic variation (e.g. temporal sampling frequency, data generalization; Burroughs 1986), use of data sets with error documentation and associated metadata is highly desirable (please see the US Federal Geographic Data Committee website, [http://www.fgdc.gov/standards/standards\\_publications](http://www.fgdc.gov/standards/standards_publications), for lists of metadata standards).

Recent advances in fine scale data resolution (<4 m) and analysis methods have greatly improved the spatial accuracy and precision of detecting, classifying and delineating landscape habitat characteristics in both two (Wulder *et al.*, 2004; Greenberg *et al.*, 2005) and

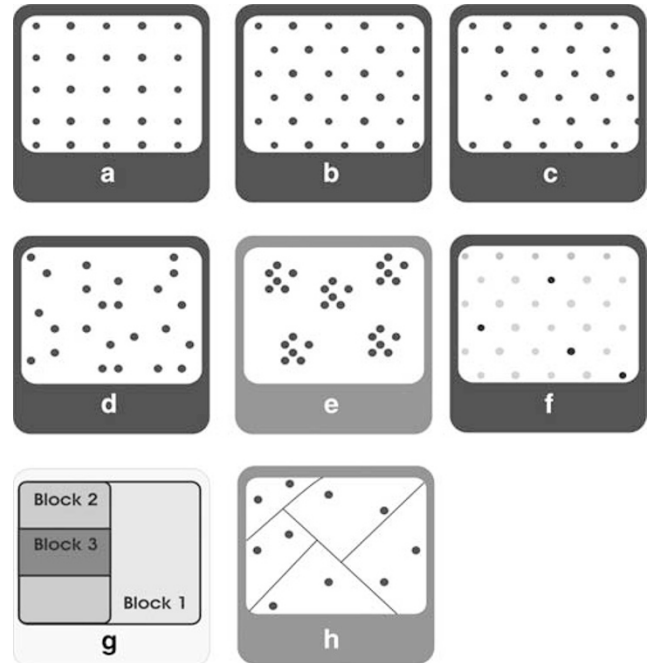
three dimensions (Lefsky *et al.*, 2002). Such fine scale data can help create detailed digital elevation models and spatially explicit vegetation canopy structure products. At coarser spatial scales (resolution 250 m–1.1 km), image data are acquired for the entire globe twice per day, enabling analyses at unprecedented temporal resolution (Rahman *et al.*, 2004; Running *et al.*, 2004) to compare vegetation phenology with genetic variation in plants and animals. However, it is important to note that the scale at which data are collected should match the scale of the study questions and hypotheses.

### Sampling

When selecting a sampling design, there are several important considerations. First, what is the research or management question? For instance, identifying specific barriers to gene flow will require a different sampling scheme than identifying source and sink habitats. Second, is the species continuously distributed across the landscape or does it have a spatially random or aggregated distribution? To make inferences about landscape influence on genetic variation, the sampling scheme should capture the variability in the independent landscape variables of interest (Bueso and Angulo, 1999). Third, based on the sampling design and data collected, which types of statistical analysis models are appropriate? (see Table 3; Forman, 1997).

Simple sampling designs discussed in this paper include systematic (uniform, unaligned, random start) and random (simple random or clustered) (Figure 1). Systematic sampling designs (Figure 1a–c) are more efficient in covering the extent of the study area than random samples (Harrison and Dunn, 1993). However, a uniform systematic design (Figure 1a) can be biased when it coincides in frequency with a regular pattern in the landscape due to a fixed sampling interval, and may miss *anisotropic* effects (Berry and Baker, 1968; Fortin and Dale, 2005). Random sampling designs have the advantage of producing a spatially unbiased sample and can be applied to continuously distributed species or a subset of the landscape based on required habitat conditions (Figure 1d–e). A random sample can be beneficial because it generates a wide range of distances among points, helping describe the relationships between observations across space if sampling is dense enough such that sample observations cluster appropriately to optimize *semivariogram* estimation. However, the distribution of samples may not be representative of the underlying geographic surface, because for most samples drawn, some areas will be oversampled while others will remain undersampled (Griffith and Amrhein, 1997).

These general sampling schemes can be implemented within more complex designs such as spatial hierarchical sampling (Figure 1f), nested sampling (Figure 1g) and stratified sampling (Figure 1h). In a hierarchical sampling design, one measurement (e.g. cover type) is collected more intensively and at a finer spatial scale than a second variable (e.g. genotypes) (Haining, 1990). Landscape variables may be less expensive to measure and change more rapidly in space than genetic variation, making spatial hierarchical sampling a viable option in landscape genetic studies. Spatially nested sampling designs generally work well for populations that are



**Figure 1** Visual representation of discussed sampling approaches. Diagrams with black backgrounds represent sampling approaches appropriate for continuously distributed populations, gray backgrounds represent sampling approaches appropriate for continuously distributed or clustered populations, and the white background represents a sampling approach most appropriate for clustered populations. Systematic sampling (a–c) designs assume a relatively even distribution of potential observations across the landscape and therefore can be used with continuously distributed populations. Black dots represent sampling locations for (a) uniform sampling, (b) unaligned sampling and (c) random start systematic sampling. Random sampling designs (d–e) ensure independence between samples because each location (or cluster) in a study area has the same probability of being chosen (Berry and Baker, 1968). Black dots represent sampling locations for (d) simple random sampling and (e) clustered random sampling. More complex sampling designs can be applied with either simple or random designs. Hierarchical sampling (f) is generally appropriate for collection of data at two scales: black dots indicate sample locations for the broad scale variable, a second variable is sampled at finer scale at the black and gray dots. Nested sampling (g) requires the landscape to be partitioned into systematic or irregular sample units (i.e. 'blocks'), which are then further subdivided into sample units nested within this first level, and so forth (Haining, 1990; Thompson and Seber, 1996; Haining, 2003). Block 1 represents the largest area, which is then partitioned into two Block 2 units. The Block 2 units are then further subdivided into the smallest units (Block 3). In stratified sampling (h) the landscape is partitioned based on some variable(s) that are unequally represented in the landscape and observation(s) are taken from each stratum (Hudak *et al.*, 2004). Lines represent strata designed to partition variation across the landscape, black dots represent random sample location within each strata. Shown is equal sampling by strata, but sampling can also be unequal.

naturally clustered or for questions of spatial scale. For example, at a broad scale (block 1), limits to gene flow across the landscape may be explained by ridgelines and canyons (Figure 1g). However, at a finer scale (block 3), limits to gene flow may be driven by microhabitat conditions.

A stratified sampling design can be implemented to capture the range of variability across landscape variable(s) of interest and to account for unbalanced

representation of landscape variables (e.g. 80% of land cover is of one class) (Figure 1h). Any of the previously mentioned sampling designs can be applied within each stratum. Stratified random sampling is particularly appropriate for *gradient analysis* and the impact of landscape configuration when variables (e.g. rainfall) are unequally represented across the study area, when the sampling area is large compared to the number of feasible observations, or when the process (e.g. gene flow) is non-stationary. Stratification guarantees a spread of observations across defined conditions in the study area, while randomization leads to a wide range of distances among pairs of points.

## Analysis methods

Analysis techniques in landscape genetic studies have utilized several statistical approaches including: assignment tests, tests of matrix correlations, dispersal route analysis and autocorrelation. Other well-developed methods in landscape ecology, spatial statistics and geographical genetics that have received less attention are: ordination, landscape metrics and *spatial interpolation*. These methods can be applied using a wide variety of statistical software (see Appendix A) and are discussed in detail below.

### Assignment tests

While traditional population genetic statistics (e.g.  $F_{ST}$ ,  $R_{ST}$ ,  $\theta$ , Nei's D) often require *a priori* population delineation, assignment test approaches (Pritchard *et al.*, 2000; Manel *et al.*, 2005) can be used to gain insight into landscape genetic patterns without *a priori* delineation of populations. For example, Proctor *et al.* (2005) sampled brown bears (*Ursus arctos*) on both sides of highways in southern British Columbia and Alberta, Canada, and an ecologically similar 'control' valley not impacted by a transportation corridor. Using assignment tests, they detected lower levels of gene flow across the highway than in the control valley.

Another approach is to identify spatial genetic discontinuities. Manni *et al.* (2004) used Voronoi polygons to build geometric relationships between populations and then applied a modified version of Monmonier's algorithm (1973) to identify genetic discontinuities to infer a barrier to gene flow. Monmonier's algorithm creates a break line through the vectors between observations (i.e. the lines of the Voronoi polygons) based on the greatest slope, which reflects higher genetic distances.

Spatial assignment tests try to identify genetic discontinuities in populations that are spatially contiguous (Guillot *et al.*, 2005a,b; Corander *et al.*, 2006; Francois *et al.*, 2006; Manel *et al.*, submitted). GENELAND is a spatial-assignment test that combines Voronoi tessellation with a Bayesian assignment approach to identify genetic spatial discontinuities (Guillot *et al.*, 2005b). Populations are specified conditionally on the spatially weighted variance structures identified within each Voronoi polygon (Guillot *et al.*, 2005b). Manel *et al.* (submitted) used a moving window approach to apply a local assignment test (Rannala and Mountain, 1997) and generate a spatially referenced probability map for finding an individual's genotype across the landscape. Genetic discontinuities are identified by calculating the mean

slope for all individual probability maps and identifying areas of high mean slope as putative population boundaries (Manel *et al.*, submitted). Note, however, that all the above spatial assignment methods perform best with a relatively continuous distribution of sample locations, such as generated by unaligned/random start systematic sampling or dense random sampling (Figure 1).

### Matrix correlations

Mantel tests (Mantel, 1967) and partial Mantel tests (Smouse *et al.*, 1986; Legendre *et al.*, 2002) have been widely used in landscape genetic studies. For example, Keyghobadi *et al.* (1999) used Mantel tests to demonstrate that meadows facilitated higher gene flow among alpine butterfly (*Parnassius smintheus*) populations than forests. Mantel tests have also been used to explain difference in flowering time of three species of alpine meadow plants based on approximate date of snowmelt across a range of elevations (Hirao and Kudo, 2004).

Partial Mantel tests have been used to expand analyses to incorporate multiple landscape variables. For example, the relative importance of partial Mantel variables was evaluated using Akaike's Information Criterion (AIC) and showed that distance measured along drainages and age of the colony were the most important predictors of genetic distance among black-tailed prairie dog (*Cynomys ludovicianus*) colonies (Roach *et al.*, 2001). In addition, Spear *et al.* (2005) included landscape variables (including 10 cover types) in their analysis of tiger salamander gene flow and showed that including topographic distance, rivers and shrub habitat improved the overall model fit by 50% compared to using straight-line Euclidean distance. It is noteworthy that there has been recent controversy about the statistical validity of partial Mantel tests (Raufaste and Rousset, 2001; Castellano and Balletto, 2002; Rousset, 2002), particularly given their widespread use in landscape genetic studies. However, variations of the partial Mantel test may circumvent certain bias and autocorrelation problems (Legendre, 1993, 2002).

### Dispersal route analysis

Dispersal route analyses (including alternate paths, least-cost paths and some *network analyses*) test the correlations of alternative paths with genetic distance to infer the most likely path of gene flow based on either pre-defined paths or *a priori* weights along a vector or a friction surface (e.g. Michels *et al.*, 2001; Spear *et al.*, 2005). Alternate path approaches have included distances along rivers for aquatic or riparian associated species (Michels *et al.*, 2001; Funk *et al.*, 2005; Poissant *et al.*, 2005; Vignieri, 2005) or the shortest straight-line path through suitable habitat (Pfenninger, 2002; Arnaud, 2003; Coulon *et al.*, 2004). Least-cost models have been based on water flow for passive dispersal (Michels *et al.*, 2001), topography and moisture (Spear *et al.*, 2005; Vignieri, 2005) and thermal cover (Scribner *et al.*, 2005). For example, Arnaud (2003) calculated a least-cost path between snail sampling sites that followed canal embankments considered good habitat for snail dispersal. Using Mantel *correlograms*, least-cost distance had a stronger correlation with genetic distance than did straight-line distance, and thus, reserve design may focus on conservation of canal



embankments. These studies demonstrated that predefined paths and least-cost paths usually explain more variation in gene flow than straight-line Euclidean distance as estimated in traditional isolation-by-distance studies.

### Autocorrelation

Typical spatial analyses of genetic variation involve quantifying genetic variation over distance (conceptually similar to isolation-by-distance) based on spatial autocorrelation such as a *Moran's I* statistic (Epperson and Li, 1996; Arnaud, 2003) or *semivariograms* (Liepelt *et al.*, 2002; Wagner *et al.*, 2005). A semivariogram provides a graphical description of the variance between binned observations at different distance classes (e.g. <5, 5–10 m, etc) across the landscape (Fortin and Dale, 2005). Thus, *semivariogram* values (in the presence of autocorrelation) are generally small when distance between observations is close, and become larger as distances increase (Fortin and Dale, 2005). For example, Wagner *et al.* (2005) used directional semivariograms to show distinct *anisotropy* (i.e. directionality) in genetic variation as influenced by mode of reproduction (clonal or sexual) and wind direction in the epiphytic lichen, *Lobaria pulmonaria*. Semivariogram modeling mathematically describes spatial autocorrelation, and by fitting a variogram model to the empirical semivariogram, the population variance can be estimated accounting for spatial autocorrelation. Thus, semivariogram modeling provides parameter estimates that are comparable between studies (Wagner *et al.*, 2005).

Spatial autocorrelation approaches are also useful for defining the scale of genetic neighborhoods. Peakall *et al.* (2003) used spatial autocorrelation incorporating multiple microsatellite loci to identify the scale of genetic variation over a 1 km transect in bush rats (*Rattus fuscipes*), categorized by age and sex. They found that subadult bush rats had a larger extent of genetic similarity than adults (400 versus 200 m) and that females had higher similarity over a larger area than males (400 versus 300 m). Additional methods exist for examining effects of multiple scales, such as hierarchical analysis of variance and delineation of semivariance (Cressie, 1993; Liburne *et al.*, 2004). These methods can be coupled with simulation techniques to examine the expected genetic neighborhood size, or testing how variability or pattern changes at different spatial scales (Gardner *et al.*, 1987; Saura and Martínez-Millán, 2001; Wu, 2004).

### Ordination

Ordination is another powerful tool that can be applied to identify the influence of continuous variables and gradients on genetic variation. Ordination is used to elucidate the pattern of the spatial arrangement of individuals using canonical correlation, principal components analysis or multidimensional scaling (Jongman *et al.*, 1995). *Canonical correspondence analysis* (CCA) is a type of ordination especially suited for investigating relationships between species and their environment to effectively identify gradients in environmental variables (ter Braak and Verdonschot, 1995); space can be explicitly incorporated as a covariate in CCA (Borcard *et al.*, 1992). CCA was applied to genetic data to determine that

genetic variation of the brook charr (*Salvelinus fontinalis*) was influenced by drainage pattern, lake altitude, stocking history and number of fish species (Angers *et al.*, 1999). Genetic data were incorporated in the model by constructing a matrix of allele frequencies at each locus in each lake in the study area. In a recent landscape genetic analysis of cougars (*Puma concolor*) (McRae *et al.*, 2005), non-metric multidimensional scaling of shared allele distances (Bowcock *et al.*, 1994) was used to detect a spatial discontinuity associated with desert and grassland habitats in the southwestern USA.

### Landscape metrics

Landscape metrics can be applied to landscape genetics studies to quantify landscape structure in multivariate models that explore genetic variation. Landscape metrics quantify the spatial arrangement of landscape patches representing three levels: landscape, class and patch (for reviews, McGarigal and Marks, 1995; Haines-Young and Chopping, 1996). Landscape-level metrics quantify a cumulative structural measure (e.g. mean patch size) of all patches. Class metrics quantify a particular level (e.g. specific vegetation type), and patch metrics describe each individual patch (e.g. amount of edge relative to interior). Landscape metrics can be particularly powerful in metapopulation analysis where patch dynamics (e.g. connectivity and/or fragmentation) can greatly influence demography and gene flow. While methods for understanding influences of patch dynamics on ecological process have been well-developed in landscape ecology (O'Neill *et al.*, 1988; McGarigal and Marks, 1995; Gustafson, 1998), they have been used less often in landscape genetics. However, Keyghobadi *et al.* (2005b) used landscape metric analysis of mean patch size and distance to edge to demonstrate that habitat fragmentation was associated with higher genetic differentiation and lower gene flow in an alpine meadow-dwelling butterfly (*Parnassius smintheus*) than in more contiguous landscapes.

### Spatial interpolation

A relatively promising but underutilized methodology for landscape genetics is spatial interpolation, which can be particularly useful for continuously distributed species and derived data (as collected under systematic or dense random sampling designs). Interpolation techniques can be used to represent allele frequency data across a landscape surface. Interpolation provides a way to predict values and corresponding levels of uncertainty for the variable of interest between points where observations have been made. Previous studies that applied interpolation typically used mitochondrial DNA markers (Barbujani *et al.*, 1989; Dupanloup *et al.*, 2002), which generally lack the resolution to reveal fine scale-structure necessary for landscape genetic analysis. However, Pierny *et al.* (1998) interpolated a map of allelic composition by *kriging* the scores from the first axis of a principle component analysis of allele frequencies to examine spatial trends in a microsatellite study of genetic variation of the red grouse (*Lagopus lagopus scoticus*). Kriging uses a modeled *semivariogram* to create an interpolated continuous surface from the observed data (Isaaks and Srivastava, 1989). When a kriged principal component surface was overlaid with habitat

in a GIS, barriers to gene flow were visually found to coincide with areas of unsuitable habitat.

## Future directions

Future advances in this field will depend on the development of new methods for translating genetic data into a form that can be analyzed with well-developed techniques from spatial analysis and landscape ecology. There are currently several future directions that could improve our ability to better integrate landscape level analyses with population genetic data. These include: improvements in the representation of genetic data for spatial analysis, improving power for analysis of noisy ecological data, expanding available techniques to include multivariate approaches, model validation, and simulation of landscape and species distribution changes.

### Representation of genetic data

It is critical to recognize that genetic data are different from ecological data typically collected for spatial analysis. In population genetics, data most often take the form of a multilocus genotype, which is not a direct measurement of a result of a spatial process (e.g. microclimate). These multilocus genotypes are only meaningful in relation to other individuals or populations, while most ecological field measurements are meaningful as independent data points (e.g. soil moisture at point *y*). However, spatial analysis methods typically require data to be collected as comparable point values that are geographically referenced to sampling locations on the landscape (Epperson, 2003).

If multilocus genetic data can be represented as point data, *point pattern statistics* are a powerful set of spatial statistics for identifying local autocorrelation structure (Anselin, 1995), clustering and spatial dependence (Diggle, 2003). For example, point pattern statistics have been used to map risk of sudden oak death according to landscape position, topography and solar insolation (Kelly and Meentemeyer, 2002). Two recent studies have demonstrated the potential for using point pattern statistics with spatially referenced genotypic data (Shimatani, 2002; Shimatani and Takahashi, 2003). For example, Shimatani (2002) used point pattern statistics to show that a secondary growth beech (*Fagus crenata*) forest patch contained trees from both the seed source population and the harvested population, suggesting that reproduction before harvest had provided important recruitment and genetic diversity to the population. A promising future direction would be to use point pattern statistics to identify clusters of closely related individuals and then map potential dispersal corridors based on associated landscape variables.

A second major consideration of the use of genetic data with landscape or ecological variables is that genetic variation results from both historic and contemporary processes. Therefore, current genetic variation may be more representative of processes that occurred as a result of a previous landscape configuration. If sufficient data are available, genetic data should be tested with both past and present landscapes to determine the strongest association. For example, Keyghobadi *et al.* (2005a) found that while genetic differentiation had the strongest correlation with contemporary habitat fragmentation,

heterozygosity was better explained by fragmentation that existed 40 years in the past. These results are supported by recent work in the bush-cricket, *Metrioptera roeseli*, where landscape changes have occurred on a faster timescale that which resulting genetic changes can be detected (Holzhauer *et al.*, 2006).

### Noisy data

Both ecological and genetic data can have a high level of noise relative to statistical signal. One possible approach to addressing this challenge is the application of algorithmic models, powerful analysis tools for complex data sets that can incorporate numerous independent and dependent variables. Algorithmic models have been used to describe the complex relationship between environmental conditions, spatial arrangement and species abundance (Déath and Fabricius, 2000; Aitkenhead *et al.*, 2004). Classification and Regression Trees (CART; Breiman *et al.*, 1984) and randomforests (Breiman, 2001) are non-parametric models that partition independent variables into levels that span the range of the dependent variable, and are particularly robust with variables that are categorical or ranked (e.g. land cover type, soil type). Finally, classification tree type models are robust to overfitting (Breiman, 2001) and are capable of dealing with multiple non-linear relationships. These types of models could be applied in landscape genetics to model multiple response variables including genetic discontinuities, migration rates and relatedness.

### Improving multivariate analyses

One of the difficulties with many of the currently applied techniques, such as the Mantel test or Moran's *I*, is the inability to employ multiple variables and interaction terms in a single model. Spatial regression models, such as autoregression, have been widely used in landscape ecology to evaluate the association between abundance and multiple environmental variables. Keitt *et al.* (2002) tested a variety of conditional (CAR) and simultaneous (SAR) autoregressive models, as well as exponential (ER) and Gaussian (GA) geostatistical models to account for spatial autocorrelation in species-habitat relationships. CAR and SAR models build on traditional linear models by incorporating a local spatial autocorrelation term, while ER and GA model the correlation between errors as a function of distance, thus accounting for spatial pattern. In one of their three case studies of bank vole (*Clethrionomys glareolus*) habitat selection among five environmental variables, vole abundance increased with vegetation height (which provided shelter from predators) and decreased with lichen and moss cover (indicative of dry habitats) (Keitt *et al.*, 2002). Autoregressive analyses are preferable over standard regression analyses when there is significant spatial autocorrelation (Fortin and Dale, 2005), which is often the case with genetic data. Autoregressive and geostatistical models could be applied in a landscape genetic context after converting multilocus genotypes into a measure of a spatial process.

### Model validation

Once sampling has occurred, genotypic data have been generated, and a model has been constructed, it is important to test the significance of the selected land-

scape genetic model(s). Ideally, researchers could replicate study areas and analyses when evaluating models; however, logistics may prevent replication due to the nature of ecological systems and financial and/or time constraints. One way to assess generalization of results is to implement permutation or simulation tests to discriminate whether observed relationships between landscape variables and genetic variation are not due to random chance alone (Gardner *et al.*, 1987; Turner, 2005). Neutral landscapes are random realizations of the landscape and can be used to rule out correlations due to chance or extraneous influences (Gardner *et al.*, 1987). Simulated neutral landscapes that follow a function of statistical similarity represented in the empirical landscape (e.g. autocorrelation) (Gardner *et al.*, 1987; Saura, 2002) can be used to generate a null distribution and *p*-values can be obtained via permutation (Li and Wu, 2004). By using this approach, researchers can determine if observed relationships fall outside the 95% confidence interval of the distribution, thus suggesting statistical significance of model variable(s). Monte Carlo simulation can also be used to quantify uncertainty in a random field when using an interpolation model such as kriging or when testing the uncertainty of a particular landscape variable used in a model. The null hypothesis that all genetic observations are spatially independent of one another can be empirically rejected by bootstrapping or other randomization tests.

#### Predicting landscape and species distribution changes

Predicting the effects of landscape change on genetic variation is an important future application of landscape genetics. For example, one may simulate the effect of different proposed road widths or alternative forest harvest plans on spatial patterns of genetic variation for a species of concern. This type of study may allow managers to choose a management alternative that minimizes impacts on the focal species while allowing some development to take place. Using landscape change simulation models, such as LANDIS (Mladenoff, 2004), in conjunction with habitat suitability and metapopulation models, such as RAMAS (Akçakaya, 2002), researchers can model population viability across different hypothetical (or alternative proposed) landscapes. While these models require demographic data for the species of interest, genetic data could greatly inform dispersal and habitat suitability parameters of the models.

Landscape genetic studies also hold promise for application to predicting and controlling spread of invasive species (Sakai *et al.*, 2001) or escape of transgenes (Pilson and Prendeville, 2004; Marvier and Van Acker, 2005). For example, identification of landscape barriers may allow researchers to predict locations of barriers to invasive species spread and thus the potential extent of the damage. At a broader scale, researchers could use spatial data to investigate the landscape-level distribution of invasive species to simulate possible spread. Similarly, simulation modeling may allow researchers to determine the type and size of barrier that could be implemented to contain invasive species or transgenes. For example, when planting transgenic crops, simulations could be conducted to determine the appropriate buffer size such that trans-

genic pollen is unlikely to escape and thus avoid potential hybridization with native species (Rieger *et al.*, 2002).

## Conclusions

The goals of this review are to: discuss landscape genetics as a discipline, suggest appropriate methods for the sampling design, and show clear benefits of utilizing well-developed methods in landscape ecology and spatial statistics for analysis. The diversity of vocabulary (e.g. Table 1), methods (e.g. Tables 2 and 3) and ideas presented in this review clearly indicate the challenges for collaborations among experts in population genetics, landscape ecology and spatial statistics. Experts working in these disciplines do not communicate regularly, and such communication is critical for advancement of landscape genetics as a discipline. The field of landscape genetics has been moving forward rapidly, with much recent literature devoted to using landscape genetics as a tool (Table 3). To advance the field, we recommend that workshops and courses be developed to foster cross-disciplinary communication and collaborations that will advance the methodology and approaches used in landscape genetics. In 2005, the International Association of Landscape Ecologists contributed greatly to this effort by holding a landscape genetics symposium at their annual meeting, and the ideas from this conference are now shared with the broader community in a special issue of *Landscape Ecology* (Holderegger and Wagner, 2006). We look forward to a future with increased discussion and collaboration among geneticists, landscape ecologists and spatial statisticians.

## References

- Adriaensen F, Chardon JP, de Blust G, Swinnen E, Villalba S, Gulinck H *et al.* (2003). The application of 'least-cost' modeling as a functional landscape model. *Landscape Urban Plan* **64**: 233–247.
- Aitkenhead MJ, Mustard MJ, McDonald AJS (2004). Using neural networks to predict spatial structure in ecological systems. *Ecol Model* **179**: 393–403.
- Akçakaya HR (2002). *RAMAS Metapop: viability analysis for stage-structured metapopulations (version 4.0)*. Applied biomathematics: Setauket, New York.
- Angers B, Magnan P, Plante M, Bernatchez L (1999). Canonical correspondence analysis for estimating spatial and environmental effects on microsatellite gene diversity in brook charr (*Salvelinus fontinalis*). *Mol Ecol* **8**: 1043–1053.
- Anselin L (1995). Local indicators of spatial association. *Geogr Anal* **27**: 93–115.
- Antolin M, Savage L, Eisen R (2006). Landscape features influence genetic structure of black-tailed prairie dogs (*Cynomys ludovicianus*). *Landscape Ecol* **21**: 867–875.
- Arnaud J-F (2003). Metapopulation genetic structure and migration pathways in the land snail *Helix aspersa*: influence of landscape heterogeneity. *Landscape Ecol* **18**: 333–346.
- Avise J (2000). *Phylogeography*. Harvard University Press: Cambridge, MA.
- Bailey TC, Gatrell AC (1995). *Interactive spatial data analysis*. Longman: Harlow, UK.
- Banks SC, Lindenmayer DB, Ward SJ, Taylor AC (2005). The effects of habitat fragmentation via forestry plantation establishment on spatial genotypic structure in the small marsupial carnivore, *Antechinus agilis*. *Mol Ecol* **14**: 1667–1680.

- Barbujani G, Oden NL, Sokal RR (1989). Detecting regions of abrupt change in maps of biological variables. *Syst Zool* **38**: 376–389.
- Beerli P, Felsenstein J (2001). Maximum likelihood estimation of a migration matrix and effective population sizes in *n* subpopulations by using a coalescent approach. *Proc Natl Acad Sci USA* **98**: 4563–4568.
- Berry B JL, Baker AM (1968). *Geographic Sampling*. Prentice-Hall: Englewood Cliffs, NJ.
- Bhattacharya M, Primack RB, Gerwein J (2003). Are roads and railroads barriers to bumblebee movement in a temperate suburban conservation area? *Biol Conserv* **109**: 37–45.
- Borcard D, Legendre P, Drapeau P (1992). Partialling out the spatial component of ecological variation. *Ecology* **73**: 1045–1055.
- Bowcock AM, Ruiz-Linares A, Tomfohrde J, Minch E, Kidd JR, Cavalli-Sforza LL (1994). High resolution of human evolutionary trees with polymorphic microsatellites. *Nature* **368**: 455–457.
- Breiman L (2001). *Random Forests*. University of California: Berkeley, CA.
- Breiman L, Friedman JH, Olshen RA, Stone CJ (1984). *Classification and Regression Trees*. The Wadsworth Statistics/probability series. CRC Press: Boca Raton, FL.
- Broderick D, Idaghdour Y, Korrida A, Hellmich J (2003). Gene flow in great bustard populations across the Strait of Gibraltar as elucidated from excremental PCR and mtDNA sequencing. *Conserv Genet* **4**: 793–800.
- Bueso M, Angulo J (1999). Criteria for multivariate spatial sampling design based on covariance matrix perturbation. In: Soares A, Gomex-Hernandez J, Froidevaux R (eds). *Geostatistics for Environmental Applications*. Springer: NY. pp 491–502.
- Burrough PA (1986). *Principles of Geographic Information Systems*. Oxford University Press: Oxford, UK.
- Castella V, Ruedi M, Excoffier L, Ibanez C, Arlettaz R, Hausser J (2000). Is the Gibraltar Strait a barrier to gene flow for the bat *Myotis myotis* (Chiroptera Vespertilionidae)? *Mol Ecol* **9**: 1761–1772.
- Castellano S, Balleto E (2002). Is the partial mantel test inadequate? *Evolution* **56**: 1871–1873.
- Cegelski C, Waits LP, Anderson NJ (2003). Assessing population structure and gene flow in Montana wolverines (*Gulo gulo*) using assignment-based approaches. *Mol Ecol* **12**: 2907–2918.
- Cicero C (2004). Barriers to sympatry between avian sibling species (Paridae: *Baeolophus*) in local secondary contact. *Evolution* **58**: 1573–1587.
- Corander J, Marttinen P, Sirén J, Tang J (2006). BAPS: Bayesian Analysis of Population Structure, Manual v. 4.1. Available at: <http://www.mni.helsinki.fi/jic/bapspage.html>.
- Coulon A, Cosson JF, Angibault JM, Cargnelutti B, Galan M, Morellet N 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 JF, Angibault JMA, Aulagnier S, Cargnelutti B et al. (2006). Genetic structure is influenced by landscape features: empirical evidence from a roe deer population. *Mol Ecol* **15**: 1669–1679.
- Cressie NAC (1993). *Statistics for Spatial Data*. Wiley: New York.
- Deáth G, Fabricius KE (2000). Classification and regression trees: a powerful yet simple technique for ecological data analysis. *Ecology* **81**: 3178–3192.
- Dias PC, Verheyen GR, Raymond M (1996). Source-sink populations in Mediterranean blue tits: evidence using single-locus minisatellite probes. *J Evol Biol* **9**: 965–978.
- Diggle PJ (2003). *Statistical Analysis of Spatial Point Patterns*. Academic Press: New York.
- Dodd CK, Barichivich WJ, Smith LL (2004). Effectiveness of a barrier wall and culverts in reducing wildlife mortality on a heavily traveled highway in Florida. *Biol Conserv* **118**: 619–631.
- Dungan JL, Perry JN, Dalehgb MRT, Legendre P, Citron-Pousty S, Fortin MJ et al. (2002). A balanced view of scale in statistical analysis. *Ecography* **25**: 626–640.
- Dupanloup I, Schneider S, Excoffier L (2002). A simulation annealing approach to define the genetic structure of populations. *Mol Ecol* **11**: 2571–2581.
- Epperson B (2003). *Geographical Genetics*. Princeton University Press: Princeton and Oxford.
- Epperson BK, Li T (1996). Measurement of genetic structure within populations using Moran's spatial autocorrelation statistics. *Proc Natl Acad Sci USA* **93**: 10528–10532.
- Epps CW, Palsbøll PJ, Wehausen JD, Roderick GK, Ramey II RR, McCullough DR (2005). Highways block gene flow and cause a rapid decline in genetic diversity of desert bighorn sheep. *Ecol Lett* **8**: 1029–1038.
- Ezard THG, Travis JMJ (2006). The impact of habitat loss and fragmentation on genetic drift and fixation time. *Oikos* **114**: 367–375.
- Forman RTT (1997). *Land Mosaics: The Ecology of Landscapes and Regions*. Cambridge University Press: New York.
- Fortin M-J, Dale M (2005). *Spatial Analysis: A Guide for Ecologists*. Cambridge University Press: Cambridge, UK.
- Fotheringham AS, Brunsdon C, Charlton M (2004). *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. J Wiley & Sons, Inc.: New York.
- Francois O, Ancelet S, Guillot G (2006). Bayesian clustering using hidden markov random fields in spatial population genetics. *Genetics*, 3 August [E-pub ahead of print].
- Funk WC, Blouin MS, Corn PS, Maxell BA, Pilliod DS, Amish S et al. (2005). Population structure of Columbia spotted frogs (*Rana luteiventris*) is strongly affected by the landscape. *Mol Ecol* **14**: 483–496.
- Gardner RH (2001). *Scaling Relations in Experimental Ecology*. Columbia University Press: New York.
- Gardner RH, Milne BT, Turner MG, O'Neill RV (1987). Neutral models for the analysis of broad-scale landscape pattern. *Landscape Ecol* **1**: 19–28.
- Gee JM (2004). Gene flow across a climatic barrier between hybridizing avian species, California and Gambel's quail. *Evolution* **58**: 1108–1121.
- Geffen E, Anderson MJ, Wayne RK (2004). Climate and habitat barriers to dispersal in the highly mobile grey wolf. *Mol Ecol* **13**: 2481–2490.
- Goovaerts P (1997). *Geostatistics for Natural Resources Evaluation*. Oxford University Press: Oxford.
- Greenberg JA, Dobrowski SZ, Ustin SL (2005). Shadow allometry: estimating tree structure parameters using hyper-spatial image analysis. *Remote Sens Environ* **97**: 15–25.
- Griffith D, Amrhein C (1997). *Multivariate Statistical Analysis for Geographers*. New Jersey: Prentice Hall. 345p.
- Guillot G, Estoup A, Mortier F, Cosson JF (2005a). A spatial statistical model for landscape genetics. *Genetics* **170**: 1261–1280.
- Guillot G, Mortier F, Estoup A (2005b). Geneland: a computer package for landscape genetics. *Mol Ecol Notes* **5**: 712–715.
- Gustafson EJ (1998). Quantifying landscape spatial pattern: what is the state of the art? *Ecosystems* **1**: 143–156.
- Haines-Young R, Chopping M (1996). Quantifying landscape structure: a review of landscape indices and their application to forested landscapes. *Prog Phys Geogr* **20**: 418–445.
- Haining R (1990). *Spatial Data Analysis in the Social and Environmental Sciences*. Cambridge University Press: Cambridge.
- Haining R (2003). *Spatial Data Analysis: Theory and Practice*. Cambridge University Press: Cambridge, UK.
- Harrison A, Dunn R (1993). Problems of sampling the landscape. In: Haines-Young R, Green DR, Cousins S (eds). *Landscape Ecology and GIS*. Taylor and Francis: London. pp 101–109.

- Hirao AS, Kudo G (2004). Landscape genetics of alpine-snowbed plants: comparisons along geographic and snow-melt gradients. *Heredity* **93**: 290–298.
- Hitchings SP, Beebee TJC (1997). Genetic substructuring as a result of barriers to gene flow in urban *Rana temporaria* (common frog) populations: implications for biodiversity conservation. *Heredity* **79**: 117–127.
- Holderegger R, Wagner HH (2006). A brief guide to landscape genetics. *Land Ecol* **21**: 793–796.
- Holzhauser S, Ekschmitt K, Sander A-C, Dauber J, Wolters V (2006). Effect of historic landscape change on the genetic structure of the bush-cricket (*Metrioptera roeseli*). *Landscape Ecol* **21**: 891–899.
- Hudak AT, Crookston NL, Evans JS, Falkowski MJ, Smith AMS, Morgan P et al. (2004). Regression modeling and mapping of coniferous forest basal area and tree density from discrete-return lidar and multispectral satellite data. *Can J Remote Sens* **32**: 126–138.
- Isaaks EH, Srivastava RM (1989). *An Introduction to Applied Geostatistics*. Oxford University Press: New York.
- Jacquemyn H (2004). Genetic structure of the forest herb *Primula elatior* in a changing landscape. *Mol Ecol* **13**: 211–219.
- Jongman RHG, Braak CJFT, Tongeren OFRV (eds) (1995). *Data Analysis in Community and Landscape Ecology*. Cambridge University Press: Cambridge.
- Jorde PE, Ryman N (1995). Temporal allele frequency change and estimation of effective size in populations with overlapping generations. *Genetics* **139**: 1077–1090.
- Jørgensen HBH, Hansen MM, Bekkevold D, Ruzzante DE, Loeschcke V (2005). Marine landscapes and population genetic structure of herring (*Clupea harengus* L.) in the Baltic Sea. *Mol Ecol* **14**: 3219–3234.
- Keitt TH, Bjornstad ON, Dixon PM, Citron-Pousty S (2002). Accounting for spatial pattern when modeling organism-environment interactions. *Ecography* **25**: 616–625.
- Kelly M, Meentemeyer RK (2002). Landscape dynamics of the spread of sudden oak death. *Photogram Eng Remote Sens* **68**: 1001–1009.
- Kennington WJ, Gockel J, Partridge L (2003). Testing for asymmetrical gene flow in a *Drosophila melanogaster* body-size cline. *Genetics* **165**: 667–673.
- Keyghobadi N, Roland J, Matter SF, Strobeck C (2005a). Among- and within-patch components of genetic diversity respond at different rates to habitat fragmentation: an empirical demonstration. *Proc R Soc Lond B Biol Sci* **272**: 553–560.
- Keyghobadi N, Roland J, Strobeck C (2005b). Genetic differentiation and gene flow among populations of the alpine butterfly, *Parnassius smintheus*, vary with landscape connectivity. *Mol Ecol* **14**: 1897–1909.
- Keyghobadi N, Roland J, Strobeck S (1999). Influence of landscape on the population genetic structure of the alpine butterfly *Parnassius smintheus* (Papilionidae). *Mol Ecol* **8**: 1481–1495.
- King AW (1990). Translating models across scales in the landscape. *Quantitative Methods of Landscape Ecology*. Springer: Berlin, Allemagne. pp 479–517.
- Kreyer D, Oed A, Ealther-Hallwig K, Frank R (2004). Are forests potential landscape barriers for foraging bumblebees? Landscape scale experiments with *Bombus terrestris* agg. and *Bombus pascuorum* (Hymenoptera, Apidae). *Biol Conserv* **116**: 111–118.
- Leblois R, Rousset F, Estoup A (2004). Influence of spatial and temporal heterogeneities on the estimation of demographic parameters in a continuous population using individual microsatellite data. *Genetics* **166**: 1081–1092.
- Lefsky MA, Cohen WB, Parker GG, Harding DJ (2002). Lidar remote sensing for ecosystem studies. *BioScience* **52**: 19–30.
- Legendre P (1993). Spatial autocorrelation: trouble or new paradigm? *Ecology* **74**: 1659–1673.
- Legendre P (2002). The consequences of spatial structure for the design and analysis of ecological field surveys. *Ecography* **25**: 601–615.
- Legendre P, Dale MRT, Fortin M-J, Gurevitch J, Hohn M, Myers D (2002). The consequences of spatial structure for the design and analysis of ecological field surveys. *Ecography* **25**: 601–615.
- Li H, Wu J (2004). Use and misuse of landscape indices. *Landscape Ecol* **19**: 155–169.
- Liburne LR, Webb TH, Francis GS (2004). The scale matcher: a procedure for assessing scale compatibility of spatial data and models. *Int J Geogr Inf Sci* **18**: 257–280.
- Liepert S, Bialozyt R, Ziegenhage B (2002). Wind-dispersed pollen mediates postglacial gene flow among refugia. *Proc Natl Acad Sci USA* **99**: 14590–14594.
- Manel S, Berthoud F, Bellemain E, Gaudeul M, Luikart G, Swenson JE et al. (submitted). A new individual-based spatial approach for identifying genetic discontinuities in natural populations. *Mol Ecol*.
- Manel S, Gaggiotti OE, Waples RS (2005). Assignment methods: matching biological questions with appropriate techniques. *Trends Ecol Evol* **20**: 136–142.
- Manel S, Schwartz MK, Luikart G, Taberlet P (2003). Landscape genetics: combining landscape ecology and population genetics. *Trends Ecol Evol* **18**: 189–197.
- Manni F, Guerard E, Heyer E (2004). Geographic patterns of (genetic, morphologic, linguistic) variation: how barriers can be detected by using Monmonier's algorithm. *Hum Biol* **76**: 173–190.
- Mantel N (1967). The detection of disease clustering and a generalized regression approach. *J Cancer Res* **27**: 209–220.
- Marceau DJ (1999). The scale issue in social and natural sciences. *Can J Remote Sens* **25**: 347–356.
- Marvier M, Van Acker RC (2005). Can crop transgenes be kept on a leash? *Frontiers Ecol Environ* **3**: 99–106.
- McGarigal K, Marks BJ (1995). FRAGSTATS. Spatial analysis program for quantifying landscape structure. *USDA Forest Service General Technical Report PNW-GTR-351*.
- McRae BH, Beier P, Dewald LE, Huynh LY, Keim P (2005). Habitat barriers limit gene flow and illuminate historical events in a wide-ranging carnivore, the American puma. *Mol Ecol* **14**: 1965–1977.
- Michels E, Cottenie K, Neys L, DeGalas K, Coppin P, DeMeester L (2001). Geographical and genetic distances among zooplankton populations in a set of interconnected ponds: a plea for using GIS modeling of the effective geographical distance. *Mol Ecol* **10**: 1929–1938.
- Miller CR, Waits LP (2003). The history of effective population size and genetic diversity in the Yellowstone grizzly (*Ursus arctos*): implications for conservation. *Proc Natl Acad Sci USA* **100**: 4334–4339.
- Mladenoff DJ (2004). LANDIS and forest landscape models. *Ecol Model* **180**: 7–19.
- Monmonier MS (1973). Maximum-difference barriers: an alternative numerical regionalization method. *Geogr Anal* **5**: 245–261.
- Moran C (1950). Notes on continuous stochastic phenomena. *Biometrika* **37**: 17–23.
- Nei M (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* **89**: 583–590.
- Nei M, Le WH (1973). Linkage disequilibrium in subdivided populations. *Genetics* **75**: 213–219.
- O'Neill RV, Krummel JR, Gardner RH, Sugihara G, Jackson B, DeAngelis DL et al. (1988). Indices of landscape pattern. *Landscape Ecol* **1**: 153–162.
- Paetkau D, Calvert W, Stirling I, Strobeck C (1995). Microsatellite analysis of a population structure in Canadian polar bears. *Mol Ecol* **3**: 489–495.
- Palo JU, Lesbarreres D, Schmeller DS, Primmer CR, Marilea J (2004). Microsatellite marker data suggest sex-biased

- dispersal in the common frog *Rana temporaria*. *Mol Ecol* **13**: 2865–2869.
- Peakall R, Ruibal M, Linenmayer D (2003). Spatial autocorrelation analysis offers new insights into gene flow in the Australian bush rat, *Rattus fuscipes*. *Evolution* **57**: 1182–1195.
- Pfenninger M (2002). Relationship between microspatial population genetic structure and habitat heterogeneity in *Pomatias elegans* (O.F. Muller 1774) (Caenogastropoda, Pomatiidae). *Biol J Linn Soc Lond* **76**: 565–575.
- Piertney SB, MacColl ADC, Bacon PJ, Dallas JF (1998). Local genetic structure in red grouse (*Lagopus lagopus scoticus*): evidence from microsatellite DNA markers. *Mol Ecol* **7**: 1645–1654.
- Pilson D, Predeville HR (2004). Ecological effects of transgenic crops and the escape of transgenes into wild populations. *Annu Rev Ecol Syst* **35**: 149–174.
- Poissant J, Knight TW, Ferguson MM (2005). Nonequilibrium conditions following landscape rearrangement: the relative contribution of past and current hydrological landscapes on the genetic structure of a stream-dwelling fish. *Mol Ecol* **14**: 1321–1331.
- Pritchard JK, Stephens M, Donnelly P (2000). Inference of population structure using multilocus genotype data. *Genetics* **155**: 945–959.
- Proctor MF, McLellan BN, Barclay RMR (2005). Genetic analysis reveals demographic fragmentation of grizzly bears yielding vulnerably small populations. *Proc R Soc Lond B Biol Sci* **272**: 240–2416.
- Pulliam HR (1988). Sources, sinks and population regulation. *Am Nat* **132**: 652–661.
- Rahman AF, Cordova VD, Gamon JA (2004). Potential of MODIS ocean bands for estimating CO<sub>2</sub> flux from terrestrial vegetation: a novel approach. *Geophys Res Lett* **31**: Art. No. L10503.
- Ramstad KM, Woody CA, Stage GK, Allendorf FW (2004). Founding events influence genetic population structure of sockeye salmon (*Oncorhynchus nerka*) in Lake Clark, Alaska. *Mol Ecol* **13**: 277–290.
- Rannala B, Mountain JL (1997). Detecting immigration by using multilocus genotypes. *Proc Natl Acad Sci USA* **94**: 9197–9201.
- Raufaste N, Rousset F (2001). Are partial mantel tests adequate? *Evolution* **55**: 1703–1705.
- Rehfeldt GH, Ying CC, Spittlehouse DL, Hamilton Jr DA (1999). Genetic responses to climate in *Pinus contorta*: niche breadth, climate changes, and reforestation. *Ecol Monogr* **69**: 375–407.
- Rieger MA, Lamond M, Preston C, Powles SB, Rouch RT (2002). Pollen-mediated movement of herbicide resistance between commercial canola fields. *Science* **296**: 2386–2388.
- Riley SPD, Pollinger JP, Sauvajot RM, York EC, Bromley C, Fuller TK et al. (2006). A southern California freeway is a physical and social barrier to gene flow in carnivores. *Mol Ecol* **15**: 1733–1741.
- Roach JL, Stapp P, Horne BV, Antolin MF (2001). Genetic structure of a metapopulation of black-tailed prairie dogs. *J Mammal* **82**: 946–959.
- Rousset F (2002). Partial mantel tests: reply to Castellano and Balletto. *Evolution* **56**: 1874–1875.
- Running SW, Nemani RR, Heinsch FA (2004). A continuous satellite-derived measure of global terrestrial primary production. *BioScience* **54**: 547–560.
- Sacks BN, Brown SK, Ernest HB (2004). Population structure of California coyotes corresponds to habitat-specific breaks and illuminates species history. *Mol Ecol* **13**: 1265–1275.
- Sakai AK, Allendorf F, Holt JS, Lodge DM, Molofsky J, With KA et al. (2001). The population biology of invasive species. *Annu Rev Ecol Syst* **32**: 305–332.
- Saura S (2002). Effects of minimum mapping unit on land cover data spatial configuration and composition. *Int J Remote Sens* **23**: 4853–4880.
- Saura S, Martínez-Millán J (2001). Landscape patterns simulation with a modified random clusters method. *Landscape Ecol* **15**: 661–678.
- Scribner KT, Blanchong JA, Bruggeman DJ, Epperson BK, Lee C-Y, Pan Y-W et al. (2005). Geographical genetics: conceptual foundations and empirical applications of spatial genetic data in wildlife management. *J Wildlife Manage* **69**: 1434–1453.
- Sezen UU, Chazdon RL, Holsinger KE (2005). The palm trees in a second growth forest in Costa Rica are much less genetically diverse than trees in the adjacent old growth. *Science* **307**: 891.
- Shimatani K (2002). Point process for fine-scale spatial genetics and molecular ecology. *Biom J* **44**: 325–352.
- Shimatani K, Takahashi M (2003). On methods of spatial analysis for genotyped individuals. *Heredity* **91**: 173–180.
- Singleton PH, Gaines WL, Lehmkuhl JF (2002). Landscape permeability for large carnivores in Washington: a geographic information system weighted-distance and least-cost corridor assessment. U.S. Forest Service Research Paper PNW-RP-549.
- Smouse PE, Long JC, Sokal RR (1986). Multiple regression and correlation extensions of the Mantel test of matrix correspondence. *Syst Zool* **35**: 627–632.
- Spear SF, Peterson CR, Matacq M, Storfer A (2005). Landscape genetics of the blotched tiger salamander (*Ambystoma tigrinum melanostictum*). *Mol Ecol* **14**: 2553–2564.
- Taaffe EJ, Gauthier HL, O'Kelly ME (1996). *Geography of transportation*, 2nd edn, Simon and Schuster: Upper Saddle River, New Jersey.
- ter Braak CJF, Verdonschot PFM (1995). Cononical correspondence analysis and related multivariate methods in aquatic ecology. *Aquat Sci* **5/4**: 1–35.
- ter Braak CJF (1995). Canonical correspondence analysis and related multivariate methods in aquatic ecology. *Aquat Sci* **55**: 255–289.
- Thompson SK, Seber GAF (1996). *Adaptive Sampling*. Wiley: New York.
- Trapnell DW, Hamrick JL (2004). Partitioning nuclear and chloroplast variation at multiple spatial scale in the neotropical epiphytic orchid, *Laulia rubescens*. *Mol Ecol* **13**: 2655–2666.
- Turner M (2005). Landscape ecology: what is the state of the science. *Annu Rev Ecol Syst* **36**: 319–344.
- Turner MG, Gardner RH, O'Neill RV (eds) (2001). *Landscape Ecology in Theory and Practice: Pattern and Process*. Springer: New York.
- Vignieri SN (2005). Streams over mountains: influence of riparian connectivity on gene flow in the Pacific jumping mouse (*Zapus trinotatus*). *Mol Ecol* **14**: 1925–1937.
- Wagner H, Holderegger R, Werth S, Gugerli F, Hoesbee SE, Scheidegger C (2005). Variogram analysis of the spatial genetic structure of continuous populations using multilocus microsatellite data. *Genetics* **169**: 1739–1752.
- Walker RS, Novaro AJ, Branch LC (2003). Effects of patch attributes, barriers, and distance between patches on the distribution of a rock-dwelling rodent (*Lagidium viscacia*). *Landscape Ecol* **18**: 187–194.
- Waples RS (1998). Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. *J Hered* **81**: 267–276.
- Whitlock MC (1992). Temporal fluctuations in demographic parameters and the genetic variance among populations. *Evolution* **46**: 608–615.
- Wilson GA, Rannala B (2003). Bayesian inference of recent migration rates using multilocus genotypes. *Genetics* **163**: 1177–1191.
- Wu J (2004). Effects of changing scale on landscape pattern analysis: scaling relations. *Landscape Ecol* **19**: 125–138.
- Wulder MA, Hall RJ, Coops NC, Franklin SE (2004). High spatial resolution remotely sensed data for ecosystem characterization. *BioScience* **54**: 511–521.

## Appendix A

Resources for use in landscape genetics and spatial analysis. Italics indicate free data or software.

### Spatial software

**GRASS** – Free extensive GIS system. (Unix/Linux) (<http://www.opengrass.org>)

**SAGA** – Free open source GIS system with raster and vector capabilities. (Windows) (<http://www.saga-gis.uni-goettingen.de/html/index.php>)

**SPRING** – Free raster and vector GIS system with robust database and modeling environment. (Windows) (<http://www.dpi.inpe.br/spring/english/>)

**CHIPS** – Free remote sensing image processing. (Windows) (<http://www.geogr.ku.dk/chips/index.htm>)

**MultiSpec** – Free remote sensing image processing. (Windows/Macintosh) (<http://dynamo.ecn.purdue.edu/%7Ebiehl/MultiSpec/>)

**OpenSourceGIS** – Links to Free Open Source/GIS related software. (<http://opensourcegis.org/>)

**ArcScripts** – Free ESRI download site for scripts and utilities. (ArcInfo, ArcGIS, and ArcView) (<http://arcscripts.esri.com>)

**R** project for statistical computing – Free statistical analysis software with several packages for spatial modeling. (Windows/UNIX) (<http://www.r-project.org/>)

**Rgeo** – Web pages for analysis of spatial data using R. (Windows/UNIX) (<http://sal.uiuc.edu/csiss/Rgeo/>)

**ADE4** – Ecological data analysis using exploratory and Euclidean methods. (Windows) (<http://pbil.univ-lyon1.fr/ADE-4/ADE-4.html>)

**GeoDA** – ESDA, autocorrelation, spatial regression, and LISA. (Windows) (<http://www.csiss.org/clearing-house/GeoDa/>)

**Geographically Weighted Regression** – Inexpensive, software for GWR. (Windows) (<http://ncg.nuim.ie/ncg/GWR/>)

**OpenBUGS** – Bayesian analysis using MCMC. (Windows) (<http://mathstat.helsinki.fi/openbugs/>)

**Fragstats** – Landscape metrics (raster based). (Windows) (<http://www.umass.edu/landeco/research/fragstats/fragstats.html>)

**LEAP-II** – Landscape metrics (raster based). (Windows) ([http://www.ai-geostats.org/software/Geostats\\_software/leap.htm](http://www.ai-geostats.org/software/Geostats_software/leap.htm))

**V-LATE** – Landscape metrics extension for ESRI ArcGIS with (vector based). (Windows) (<http://www.geo.sbg.ac.at/larg/vlate.htm>)

**STARMA/IEAST** – STARMA modeling. (<http://fried-for.msu.edu/>)

**SIMMAP** – Landscape pattern simulation using modified random clusters. (Windows) (<http://web.udl.es/usuarios/saura/>)

**RMLANDS** – Simulation of natural and anthropogenic disturbances (<http://www.umass.edu/landeco/research/rmlands/rmlands.html>)

**LANDIS-II** – Simulation of forest succession, disturbance, and seed dispersal (<http://landis.forest.wisc.edu/>)

**RAMAS** – Landscape metapopulation models. Moderately priced. (Windows) (<http://www.ramas.com/software.htm>)

**CrimeStat** – Point pattern statistics. (Windows) (<http://www.icpsr.umich.edu/CRIMESTAT/>)

**Point Pattern Analysis (PPA)** – Web based point pattern analysis. (<http://www.nku.edu/~longa/cgi-bin/cgi-tcl-examples/generic/ppa/ppa.cgi>)

### Genetic software

**Alleles in Space** – Joint spatial and genetic analysis (<http://www.marksgeneticsoftware.net/>)

**BAPS** – Bayesian Assignment of Population Structure – Bayesian identification of population sub-division and population assignment, 4.1 includes a spatial model (<http://www.rni.helsinki.fi/~jic/bapspage.html>)

**Barrier** – Identifies genetic barriers using Monmonier's algorithm (<http://www.mnhn.fr/mnhn/ecoanthropologie/software/barrier.html>)

**BayesASS** – Bayesian assignment test, identifies migrants (<http://www.rannala.org/labpages/software.html>)

**Circuitscape** – Identifies genetic breaks based on circuit theory (<http://www2.for.nau.edu/SOFArchive/Graduate-Research/bhm2/circuitscape.htm>)

**GenALEX** – Genetic Analysis in Excel including relatedness, F-stats, Mantel tests, autocorrelation ([http://www.anu.edu.au/BoZo/GenALEX/genalex\\_download.php](http://www.anu.edu.au/BoZo/GenALEX/genalex_download.php))

**Geneland** – Population assignment incorporating georeferences data ([http://www.inapg.inra.fr/ens\\_rech/mathinfo/personnel/guillot/Geneland.html](http://www.inapg.inra.fr/ens_rech/mathinfo/personnel/guillot/Geneland.html))

**IBD** – Analyses genetic isolation by geographic distance (<http://phage.sdsu.edu/~jensen/>)

**IMMANC** – Detect immigration/dispersal (<http://www.rannala.org/labpages/software.html>)

**McMantel** – Mantel test between genetic and geographic distance matrixes dbmcd@uwyo.edu (Contact author)

**Migrate** – Detect immigration/dispersal using coalescent modeling (<http://evolution.genetics.washington.edu/lamarc/migrate.html>)

**Partition** – Identifies population sub-division and assigns individuals to populations (<http://www.univ-montp2.fr/~genetix/partition/partition.htm>)

**PCAGEN** – Principal components analysis (<http://www2.unil.ch/popgen/softwares/pcagen.htm>)

**SPAGeI** – Spatial genetic structure of individuals and/or populations (<http://www.ulb.ac.be/sciences/ecoevol/spagedi.html>)

**Structure** – Bayesian identification of population sub-division and population assignment (<http://pritch.bsd.uchicago.edu/software.html>)

**SGS** – Autocorrelation of microsatellite data (<ftp://ghd.dnsalias.net/degen/software.html>)

**TTPGA** – Mantel tests (<http://www.marksgeneticsoftware.net/>)

### Miscellaneous resources

**AI-GEOSTAT** – A geostatistics resource web site with a mail list server, FAQ, and software resources (<http://www.ai-geostats.org/>)

**Center for Spatially Integrated Social Sciences (CSISS)** – Extensive spatial literature search engine (<http://www.csiss.org/>)

**Brian Epperson's web site** – Resources for geographical genetics (<http://www.for.msu.edu/faculty/fmembers/epperson.htm>)

**Pierre Legendre's web site** – Resources for spatial ecology (<http://www.bio.umontreal.ca/legendre/index-English.ht>)