Landscape genetics: combining landscape ecology and population genetics

Stéphanie Manel[1], Michael K. Schwartz[2], Gordon Luikart[1] and Pierre Taberlet[1]

[1]Laboratoire d'Ecologie Alpine, Equipe Génomique des Populations et Biodiversité, UMR CNRS 5553, BP 53, Université Joseph Fourier, 38041 Grenoble Cedex 9, France
[2]Rocky Mountain Research Station, US Forest Service, 800 E. Beckwith, Missoula, MT 59801, USA

Understanding the processes and patterns of gene flow and local adaptation requires a detailed knowledge of how landscape characteristics structure populations. This understanding is crucial, not only for improving ecological knowledge, but also for managing properly the genetic diversity of threatened and endangered populations. For nearly 80 years, population geneticists have investigated how physiognomy and other landscape features have influenced genetic variation within and between populations. They have relied on sampling populations that have been identified beforehand because most population genetics methods have required discrete populations. However, a new approach has emerged for analyzing spatial genetic data without requiring that discrete populations be identified in advance. This approach, landscape genetics, promises to facilitate our understanding of how geographical and environmental features structure genetic variation at both the population and individual levels, and has implications for ecology, evolution and conservation biology. It differs from other genetic approaches, such as phylogeography, in that it tends to focus on processes at finer spatial and temporal scales. Here, we discuss, from a population genetic perspective, the current tools available for conducting studies of landscape genetics.

The recent improvements in molecular genetic tools, combined with existing or new statistical tools (e.g. geostatistics, maximum likelihood and Bayesian approaches) and powerful computers has led to the emergence of the field of landscape genetics, which is an amalgamation of molecular population genetics and LANDSCAPE ECOLOGY [1] (Box 1) (see Glossary). This discipline aims to provide information about the interaction between landscape features and microevolutionary processes, such as gene flow, genetic drift and selection. It will also aid in identifying cryptic boundaries, which are either breaks in the gene flow across populations without any obvious cause, or secondary contact among previously isolated populations. Landscape genetics can resolve population substructure across different geographical scales at fine taxonomic levels [2]; thus, it is different from the existing disciplines of biogeography, which focuses mainly on species diversity patterns at broad temporal and spatial scales [3], and PHYLOGEOGRAPHY, which combines phylogenetics and biogeography (Box 1). Landscape genetics data, collected at a finer scale than that typical of phylogenetics, will help our

Glossary

Assignment test: statistical approach that assigns an individual to the population from which its multilocus genotypes is most likely to be derived.
Cline: a character gradient; continuous variation in a character through a series of contiguous or adjacent populations. In population genetics, the character could be multi-locus genotypes (at the individual level) or single locus allelic frequencies.
Conservation unit: refers to either evolutionary significant unit or management unit, or any geographical units that managers feel important to conserve.
Environmental feature (or variable): biotic, climatic, edaphic and other conditions that comprise the immediate habitat of an organism.
Evolutionary significant unit: a collection of individuals or populations of a given species with a unique evolutionary history and genetic discreteness that makes them worthwhile to protect. The uniqueness is often inferred from phylogenies ideally combined with ecological data.
Genetic discontinuity: geographic zone of sharp genetic change.
F_{st}: measure of the level of population genetic differentiation at single loci, which reflects the proportion of allelic variation that occurs between populations.
Interpolation: (in a GIS sense) a process to estimate values at a geographical location for which no measurement data exists.
Isolation by distance: when genetic differentiation between individuals (or populations) increases with their geographical distance (because gene flow declines at larger distances).
Landscape ecology: study of interaction between spatial patterns and ecological processes; concerned with spatial extent and configuration at which ecological processes and patterns occur.
Management unit: populations with significant divergence of allele frequencies at nuclear or mitochondrial loci regardless of the phylogenetic distinctness of the alleles, (i.e. demographically distinct populations that should be managed to ensure the viability of the larger evolutionary significant unit, subspecies, or regional populations.
Mantel's test: measures association between elements of two matrices; can assess significance of this association by random permutations of the matrices.
Metapopulation: a group of subpopulations that exchange occasional migrants and that might be subject to local extinction and recolonization. In population genetics, a metapopulation is commonly modeled using the island model (all subpopulations exchange migrants equally) or the stepping stone model (only adjacent subpopulations exchange migrants).
Phylogeography: the examination of the geographic distributions of evolutionary lineages to understand the evolutionary history of a taxon.
Random pattern: spatial distribution pattern in which the presence of one individual has no influence on the distribution of other individuals.
Synthesis map: the distribution of synthetic variables obtained by multivariate analysis (e.g. PCA) with the objective to synthesize the information for a large number of loci and alleles, and display it across a map to facilitate visual interpretation of patterns.

Corresponding author: Stéphanie Manel (Stephanie.manel@ujf-grenoble.fr).

http://tree.trends.com 0169-5347/03/$ - see front matter © 2003 Elsevier Science Ltd. All rights reserved. doi:10.1016/S0169-5347(03)00008-9

Box 1. Historical perspectives of landscape genetics

The roots of landscape genetics can be traced back to the botanist and early biogeographer Augustin-Pyramus de Candolle (1778–1841) and Alfred Russel Wallace (1823–1913). De Candolle noted that there are differences in the pattern of distribution of organisms (taxa) across the landscape, which depend upon both 'physical causes' that are currently in action and other forces that no longer exist [63]. These forces, which operate on different timescales, are now studied in the fields of ecological and historical biogeography [3,64]. Similarly, during Wallace's journey throughout the Malay Archipelago, he noted a break ('boundary') in the fauna between the Australian Region and the Oriental Region [65].

Recently, there has been growing interest in combining the tools of molecular genetics with the principles of ecological biogeography and landscape ecology. This approach enables the spatial mapping of allele frequencies from one or more species (or populations) and, subsequently the correlation of such patterns with the current landscape. This represents a 'landscape genetic approach', which has been limited by the unavailability of enough molecular markers (or highly polymorphic markers) to examine biogeography at a fine spatial and temporal scale. The recent combination of spatial statistics, geographic information systems and numerous informative molecular markers will enable the amalgamation of ecological biogeography with molecular ecology to better understand population biology and evolution, especially at fine spatial scales. Most exciting might be our newfound ability to test some of the basic tenets of biogeography, and the chance to discover new, finer scale 'Wallace lines', or undetected breaks in distributions of subpopulations, individual groups or even family lineages.

understanding of the microevolutionary processes that generate genetic structure across space.

The two key steps of landscape genetics are the detection of GENETIC DISCONTINUITIES and the correlation of these discontinuities with landscape and ENVIRONMENTAL FEATURES, such as barriers (e.g. mountains, gradient of humidity). This approach provides information that is useful across many disciplines. For instance, detecting genetic discontinuities is necessary for evolutionary biologists and ecologists to understand how the movement of individuals or even gametes influences the genetic structure of a population. Understanding gene flow is also fundamental for ascertaining factors that enable or prevent local adaptation, and for describing dynamics that facilitate the spread of new, beneficial mutations [4,5]. However, the aim of many conservation biologists and managers is to determine what constitutes a natural break in, or between, populations. For example, the ability to delineate EVOLUTIONARY SIGNIFICANT UNITS, MANAGEMENT UNITS, or CONSERVATION UNITS is dependent on detecting population subdivision [6,7]. Thus, both academics and natural-resource managers can use landscape genetics widely.

The landscape genetics approach

Landscape genetics has emerged as the result of researchers explaining observed spatial genetic patterns by using landscape variables (Table 1). The most common spatial patterns described in the literature are: CLINES [8], ISOLATION BY DISTANCE [9], genetic boundaries (discontinuities) to gene flow [10,11], METAPOPULATIONS [12] and RANDOM PATTERNS [13]. The identification of these spatial genetic patterns requires the collection of genetic data from many individuals (or populations) whose exact geographical location is known. Ideally, in a landscape genetics approach, the individual is the operational unit of study. However, this can be extended to populations (using allele frequencies) if enough populations can be sampled. The advantages of using individuals as the operational unit are to avoid potential bias in identifying populations in advance and to conduct studies at a finer scale. After sampling, genetic and statistical tools are used to determine the spatial genetic pattern and to correlate it with landscape or environmental features (Table 1).

Genetic tools facilitating landscape genetics

Technological advances have enabled researchers to use markers with varying temporal or spatial resolution (e.g. mitochondrial DNA, microsatellites, amplified fragment length polymorphisms and Y chromosomes). Dozens of markers are available for numerous taxa (e.g. ungulates [14], mammals carnivores [15], birds [16] and fish [17]). The next major advance involves increased numbers of markers [including single nucleotide polymorphisms (SNPs), mapped loci and candidate genes] combined with a change to a population genomics framework [18]. Population genomics involves sampling many loci across the genome, in addition to the sampling of many individuals from each population. It also uses many loci from functional genes with known map locations. An important reason to sample many loci is to increase power to identify 'outlier' loci that might be under selection [19]. From a landscape genetic perspective, it is important to identify these selected loci because: (1) selected loci are not optimal for inferring population demographic history or for estimating population parameters, such as migration rates, genetic distance or substructure, because selection biases parameter estimation; (2) comparing distributions of neutral and selected markers [20], and/or quantitative characters [21] could help us to understand the mechanisms generating spatial patterns; and (3) the identification of loci under selection can help us to understand the genetic basis of local adaptation, adaptive differentiation and speciation [22].

Examining data in a landscape context will assist in detecting loci under selection by focusing sampling across potential selection gradients (e.g. temperature, humidity, soil quality, altitude etc.) and identifying the potential causes and consequences of selection (e.g. the geographical location of selection gradients or barriers, and the environmental or spatial variables correlated with the selected loci). For example, Lenormand *et al.* [23] studied two loci of pesticide resistance in the mosquito *Culex pipiens* along a north–south transect that crossed organophosphate-treated and nontreated areas in southern France. They found a decrease in pesticide resistance with increasing distance from the treated zone. This cline can be interpreted as a consequence of local adaptation when migration and selection act as antagonistic forces.

Table 1. Examples of studies assessing spatial patterns and landscape boundaries using genetic data

Spatial pattern	Species	Location	Landscape feature/ barrier	Genetic marker	No of pop. or individ.	Methods[a]	Results	Refs
Random	Lily *Ornithogalum montanum*	Italy	None	11 allozymes (36 alleles)	13 pop.	Spatial autocorrelation; Mantel tests (between genetic and geographic distance); Monmonier algorithm	No association between environmental parameters and allele frequencies	[13]
Boundary to gene flow	Brook char *Salvenus fontinalis*	Maine (USA)	Waterfalls	6 microsats	30 pop.	Linear regression between genetic (F_{st}) and environmental (chord) distance; Mantel test	Genetically more isolated at higher altitudes (i.e. above impassable waterfalls); unidirectional gene flow downstream	[41]
	Red grouse *Lagopus lagopus*	Scotland	Matrix of poor habitat	7 microsats	14 pop.	GIS, PCA	Interpolation of PC1 scores implied matrix of poor-habitat restricted gene flow between populations	[10]
	Alpine butterfly *Parnassus smintheus*	Kanasaskis region (Canada)	Forests			Mantel and partial Mantel test (between genetic and environmental distance)	Forests are barriers to gene flow	[11]
Clinal	Humans *Homo sapiens*	Japan	Maritime, montane, dialect	15 protein systems (24 alleles)	1125 pop.	Spatial autocorrelation; Interpolation of allele frequencies; Wombling	Some loci demonstrated clinal patterns along major axis of Japanese archipelago; 7 genetic boundaries coincide with landscape boundaries	[8]
Isolation by distance	Prickly forest skink *Gnypetoscincus queenslandiae*	Atherton Tabeland forest (Australia)	Forest	10 microsats	144 ind.	Regression of the genetic and geographical distances between pairs of individuals	Significant isolation by distance pattern	[28]
Meta-population	Black-tailed prairie dog *Cynomys ludovicianus*	Northern Colorado (USA)	None	7 microsats	13 pop., 155 ind.	Assignment tests	Populations undergo regular extinctions, which are recolonized by individuals from multiple source colonies	[56]

[a]Abbreviations: GIS, geographical information system; PCA, principle components analysis; Monmonier algorithm and Wombling: see Box 2.

Statistical tools to identify spatial genetic patterns

In some situations, it is easy to define subpopulations or demes on the basis of spatial clustering of individuals (e.g. ponds of fish in isolated lakes or birds nesting on archipelago islands). When this is the case, methods such as Wright's F_{ST} or methods derived from ASSIGNMENT TESTS [24] might be preferred. However, individuals are often not arranged in a clustered distribution, but are uniformly distributed across space. This is when the landscape approach is most valuable.

From the perspective of landscape genetics, spatial genetic patterns would be assessed at an individual level without defining populations in advance. The idea of a method to assess spatial genetic patterns at an individual level dates back to Wright's concept of the neighbourhood in a continuously distributed population [25]. He considered that the basic unit of population structure is the 'neighborhood', its size defined as $4\pi D\sigma^2$, where D is the population density and σ^2 is the mean axial square of parent–offspring dispersal rate.

Table 2. Examples of methods used to identify genetic spatial patterns and to test for correlation between genetic and environmental variables

Method	Spatial pattern identified	Conditions of application[e]	Software and Web address[f]	Refs
Linear regression between genetic differentiation[a] and logarithm of distance	Isolation by distance	Continuous populations	Genpop (http://www.cefe.cnrs-mop.fr/)	[57]
Bayesian clustering assignment	Population units	Hardy-Weinberg and linkage equilibrium in each population; independence among loci; each population has same probability of being origin; population of origin is sampled	Structure (http://pritch.bsd.uchicago.edu/) Partition (http://www.univ-montp2.fr/~genetix/partition/partition.htm)	[31] [32]
Spatial autocorrelation and correlogram[b]	Cline, isolation by distance, Random	Defining arbitrary distances classes	AIDA (http://www.unife.it/genetica/Giorgio/giorgio_soft.html#AIDA) SGS (http://kourou.cirad.fr/genetique/software.html) R (http://www.fas.umontreal.ca/biol/casgrain/fr/labo/R/index.html) SPAGeDi (http://www.ulb.ac.be/sciences/lagev/software.html) GenAlEx (http://www.anu.edu.au/BoZo/GenAlEx/)	[58] [59][g] [60] [61]
Interpolation	Genetic boundary		No free software known	
Wombling or Monmonier algorithm	Genetic boundary		No developed software known	
Synthesis map	Visual insight on pattern concerned	Use of multivariate analysis (PCA) and interpolation before drawing map	No free software known	
Mantel test	Test correlations[c]		Genpop (http://www.cefe.cnrs-mop.fr/) R (http://www.fas.umontreal.ca/biol/casgrain/fr/labo/R/index.html) GenAlEx (http://www.anu.edu.au/BoZo/GenAlEx/)	[57][g] [61][g]
Partial mantel test	Test partial correlation[d]		R (http://www.fas.umontreal.ca/biol/casgrain/fr/labo/R/index.html) FSTAT (http://www.unil.ch/izea/softwares/fstat.html)	 [62]

[a]Genetic differentiation is a multilocus estimator of an $F_{st}/(1 - F_{st})$ analogue between pairs of individuals.
[b]Only the scale of the pattern is identified rather than the exact location of the pattern.
[c]Between pairwise genetic distances and one geographical distance or one 'ecological' distance along waterways, valleys, and so on.
[d]Between pairwise genetic distances and geographical distances and/or 'ecological' distances along water ways, valleys, and so on.
[e]Data required are geographical and genetic data from a single sampling.
[f]Websites are subject to change; this list is not inclusive.
[g]Developed by P. Casgrain and P. Legendre.

This size enables the estimation of dispersal distances if density is known [25].

Here, we describe six methods that provide the statistical cornerstone of the landscape genetics approach (Table 2).

Mantel's test and regression analysis

For continuously distributed populations, Rousset [26] derived an individual-based method to detect an isolation-by-distance pattern and to estimate neighbourhood size using regression analyses. A MANTEL'S TEST [27] between the genetic differentiation and geographical distance tests for the presence of an isolation-by-distance pattern between individuals (Table 2). A regression of the genetic differentiation (e.g. relatedness) between pairs of individuals and their geographical distance enables an estimation of the neighbourhood size to be made. For example, in the case of the prickly forest skink *Gnypetoscincus queenslandiae*, a significant isolation-by-distance was found using the Mantel's test between the individual values of differentiation (based on relatedness) and the geographical distance of individuals [28] (Table 1). Subsequently, the authors used the neighbourhood size and demographic density estimates to calculate skink average dispersal distance.

Spatial autocorrelation methods

Spatial autocorrelation methods have been widely used in genetics since the late 1970s [29] to assess associations between the genetic relatedness of pairs of individuals and the geographical distance (Table 1). These methods test whether the observed genotype of an individual at one locality (quantified by allele frequencies, genetic-distance, morphometric traits, etc.) is dependent on the genotype of

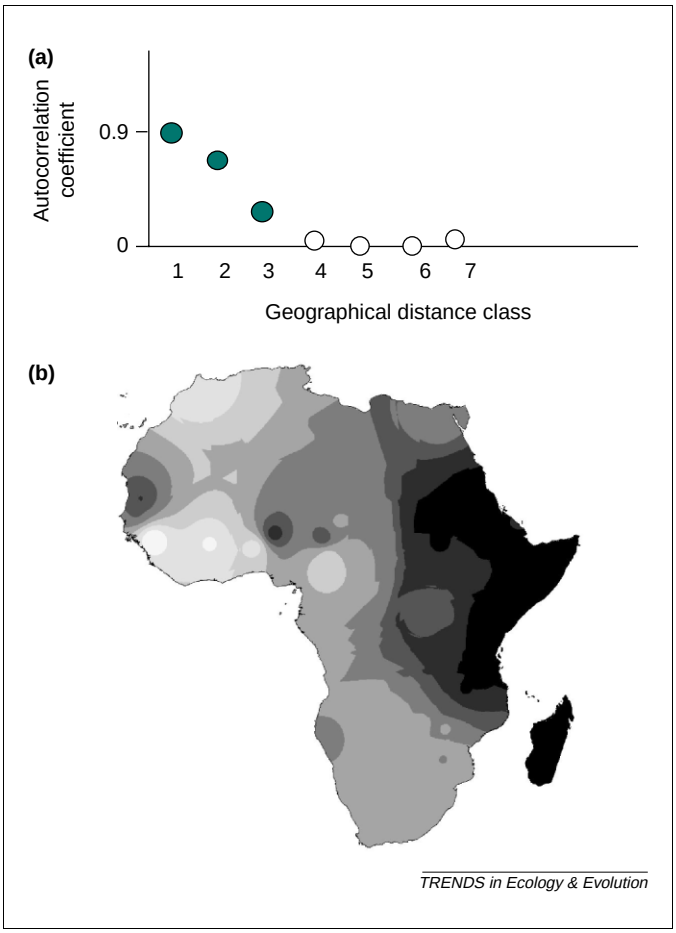

Fig. 1. Statistical tools to identify spatial genetic patterns. (a) A correlogram interpreted as a pattern of isolation by distance. Shaded circles represent autocorrelation coefficient and open circles are nonsignificant coefficients. (b) A synthetic map illustrating the geographical variation of the first principal component (PC1), which accounts for 38% of the total allelic variation (183 microsatellite allele frequencies) in cattle across Africa [37]. The PC1 component follows a gradient that peaks (darkest shading) in central East Africa and correlates with known introgression of alleles from recently introduced Indian *Bos indicus* cattle into African *B. taurus* cattle. Reproduced, with permission, from [37].

an individual at a neighbouring locality. Then a spatial correlogram evaluates the behavior of autocorrelation as a function of distance (Fig. 1a). For example, Cassens *et al.* [9] used spatial autocorrelation to analyze mitochondrial haplotypes of 129 individual of European otters *Lutra lutra* across Europe. They showed that geographical distance is the main determinant of spatial genetic structure in the eastern German otter populations. This result is important for the conservation of the otter by enabling conservationists to determine the proper spatial scale for management and demographic studies. Spatial autocorrelations can also help identify clines [8] (Table 1). Smouse and Peakall [2] extended the classic spatial autocorrelation measures that used one allele at a time to a multiallelic and multilocus approach. They applied their method to the orchid *Caladenia tentaculata,* and showed that the multiallelic treatment (to assess the spatial pattern) reduced the variation among alleles in comparison to the single allelic analysis. Although spatial autocorrelation can determine the scale of the spatial pattern, it cannot identify the specific location of a genetic discontinuity (such as a river, mountain, etc) [30].

Bayesian clustering approaches

Bayesian clustering approaches are derived from traditional assignment tests that use individual multilocus genotypes to assign individuals to populations [24]. These approaches identify a population (i.e. random mating groups) with minimal assumptions about the location of population boundaries [31,32]. These methods first attempt to group individuals into sets (populations) of random mating individuals that minimize Hardy–Weinberg (HW) and gametic disequilibrium (GD) across the data set. Departure from HW and GD leads to the splitting of the population into subpopulations. Subsequently, all individuals of unknown origin can then be assigned to potential populations of origin taking into consideration the probability of them belonging to each population sampled, as determined by the genetic data.

Knowing the spatial coordinates of the sampling location of the assigned individual enables the dispersal distance for the individual to be estimated. Using this approach, Stow *et al.* [33] showed that adult Cunningham's skinks *Egernia cunninghami* disperse shorter distances in a cleared site than they do in a forested site, which suggests that deforestation would result in the fragmentation and isolation of such skink populations.

Although these Bayesian approaches have enormous potential, they also have limits; the assumption of random mating excludes the use of them on species that are partially or wholly selfing or asexual. Furthermore, there are events other than population subdivision that lead to gametic disequilibrium and departures from HW proportions, such as small populations sizes, bottlenecks, inbreeding and admixture. Lastly, using this method to assign individuals of unknown origin to populations requires that the true population of origin be sampled [34].

Multivariate analyses and synthesis maps

Another approach to the identification of spatial patterns is the use of multivariate analyses, such as principal component analysis (PCA). PCA summarizes all the variation for many loci in the study area. It is usually conducted with populations as the study unit, but can accommodate individuals as the operational unit. The INTERPOLATION of the major principal components derived from PCA leads to a SYNTHESIS MAP. This approach does not test the presence of a spatial pattern, but gives insight into the concerned spatial pattern. Piertney *et al.* [10] used PCA to cluster allele frequencies of spatially referenced populations of red grouse *Lagopus lagopus scoticus*. The first two principal components derived from PCA were then interpolated to obtain a synthesis map (Table 1). Synthesis maps have been widely used to visualize clines from allozyme data in human populations [35,36] and, more recently, for microsatellite data from live stock in Africa [37] (Fig. 1b).

Monmonier's algorithm and wombling

Barbujani [30] reviewed two specific methods to infer genetic boundaries from allele frequency spatial distributions: Monmonier's algorithm [38] and wombling [39,40]. Principles of both methods are given in Box 2. Both have great potential to delineate genetic boundaries.

Box 2. Principle of the Womble and Monmonier methods to localize genetic boundaries in the landscape

The maximum difference Monmonier algorithm

The objective of Monmonier algorithm is to visualize data contained in a genetic distance matrix on a geographical map to identify boundaries [38,66,67].

(1) Samples (i.e. individuals) are located on a map according to their relative geographical position.

(2) The Delaunay triangulation approach [68] is used to connect the adjacent geographical positions of the samples on the map (Fig. I), resulting in a network that connects all the samples.

(3) Genetic distances between neighbouring samples are calculated and associated to each edge (corresponding to the neighbouring sample) of the network (Fig. I).

(4) Monmonier's [38] maximum – difference algorithm is then used to identify boundaries:

- The edge of the network with the largest associated distance is selected and used to begin extending a boundary (arrows in Fig. I) perpendicular to the edges of the network.
- The edge directly adjacent to the growing boundary with the largest genetic distance is selected to extend the boundary.
- The two previous steps are repeated until the growing boundary meets another boundary or reaches the edge of the area under study.

The Womble approach

The Womble approach locates boundaries across a surface for an interpolated variable (i.e. allele frequency surface) by searching for regions in which the absolute value of the surface slope is large [39]. Barbujani and Sokal [69] extended the concept to assess both the magnitude and the direction of slope surfaces.

(1) The values of a continuous variable (i.e. allelic frequencies) are transformed into a continuous distribution by interpolation and/or construction of a regular lattice (where samples are on the node of the lattice). A surface or a lattice is then obtained for each allele.

(2) The partial derivative of the considered variable (i.e. allele frequencies) is computed at each lattice point (or surface), recording both the magnitude and the direction of the maximum slope. The magnitude and slope angles at each lattice point are averaged for all allele frequency surfaces, resulting in a 'systemic function' in which each lattice point is represented by a vector characterized by magnitude and direction (Fig. II).

(3) Testing the significance of the slope magnitude is analogous to hypothesis testing in statistics. Barbujani and Sokal [69] chose to consider only lattice points whose magnitude fell in the highest 5% of all simulated slopes. In addition, they imposed a connectedness criterion among high magnitude lattice points: their direction does not differ by $>30°$.

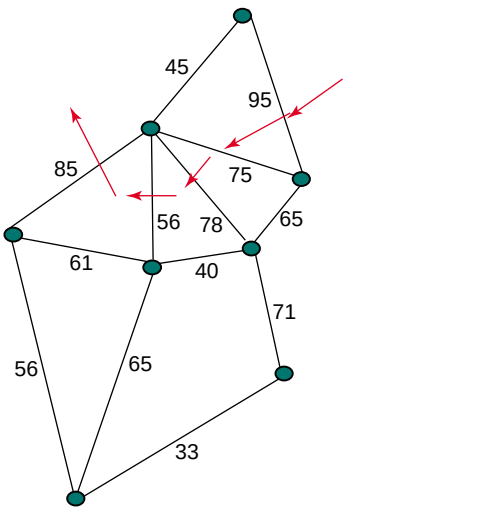

Fig. I. Example of a hypothetic Delaunay triangulation. The tops of the triangles correspond to the geographical position of the samples and are connected by the Delaunay triangulation approach. Genetic distance between two samples is given by the number indicated on each edge of the triangles. The first boundary is drawn with arrows, following Monmonier algorithm described in the Box 2 [38].

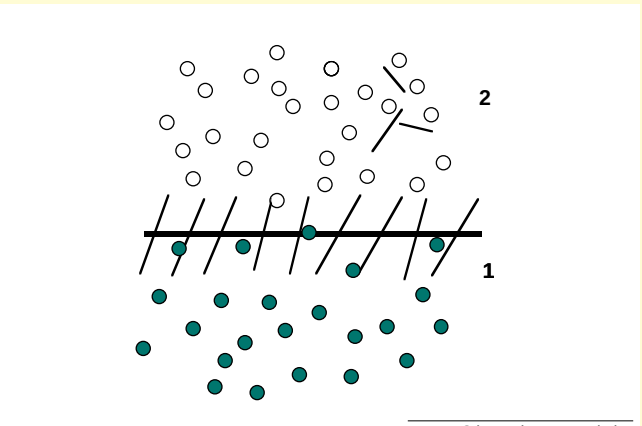

Fig. II. Example of a hypothetical Womble approach. Data (white and green circles) consist of allele frequencies (e.g. at one locus) calculated in different localities (with at least 20 individuals in each). Allele frequencies were transformed by interpolation (not shown) and the interpolated surface was then subject to the Womble method described in Box 2. The partial derivative of the allele frequencies is computed (at each grid point) and the magnitude of the resulting slope is proportional to the length of the rods on the figure. Only magnitudes significant at 5% are shown (Box 2). The direction of the slopes is indicated by the direction of the rods. The 1 indicates rods that could be connected by a genetic discontinuity between green and white localities (large rod) because their direction does not differ by $>30°$. The 2 corresponds to noise.

Correlating genetic patterns with landscape and environmental features

Once a spatial genetic pattern is identified, it is possible to test for correlations with environmental or landscape variables. Here, we describe statistical tests that correlate genetic patterns with environmental variables, and a visual approach that provides insight into the correlations.

Mantel's test

Mantel's test can also measure the association between genetic distance and an environmental variable, such as forest cover or temperature [11,41,42] (Table 1). Although the standard Mantel test enables only a comparison between two variables to be made, a partial Mantel test can be used to compare three or more variables. For example, Riginos and Nachman [43] analyzed the mitochondrial DNA haplotype differentiation in relation to habitat variables for 105 blennoid fish *Axoclinus nigricaudus*. After conducting a standard multiple regression between genetic distance and predictors of landscape and environmental variables (i.e. biogeography, geographical distance, habitat discontinuities because of sandy shores or deep open water), they assessed the significance of the predictive variables by comparing each

coefficient of the regression against those generated by permuting the predictive variables. The partial coefficients of three of the predictive variables (biogeography, geographical distance and water discontinuities) were significantly greater than that expected if they have been randomly associated. The authors concluded that the level of genetic differentiation among populations of *A. nigricaudus* was due to the combined influences of biogeography, geographical distance and habitat discontinuities.

Canonical correspondence analysis

In spite of their usefulness, the Mantel or partial Mantel tests cannot be used to quantify the amount of variation explained by environmental factors. Canonical correspondence analysis (CCA) has been used as an alternative because it enables genetic diversity to be related to environmental factors and can test for environmental factors that contribute significantly to the explanation of the variations in genetic diversity [44,45]. Angers *et al.* [44] applied a CCA to analyze simultaneously the contribution of drainage pattern, altitude and human-induced factors (e.g. sport fishing and stocking of domestic brook charr *Salvelinus fontinalus*) on the genetic diversity of brook charr using five microsatellites. They detected an influence of these three factors on the pattern of genetic diversity (population divergence) among and within populations, which is helpful for the conservation of this species and the planning of further experiments.

Geographical information systems

In parallel to statistical tests, geographical information systems (GIS) can be used to visualize spatial genetic patterns (e.g. boundaries) and also to generate hypotheses about the cause of genetic boundaries, because GIS enables landscape variables to be overlaid onto genetic data. GIS leads to the creation of synthesis maps providing various geostatistical tools for interpolation (e.g. [46,47]). For example, Piertney *et al.* [10] analyzed microsatellite allele frequencies of 14 populations of red grouse from northeast Scotland through the use of PCA. Using PC1 scores, the authors found zones where the slope of the interpolated genetic surface was significantly higher than other areas and suggested that these zones were potential barriers. These putative barriers were overlain onto a variety of land maps that corresponded to areas of unsuitable grouse habitat (Table 1).

Testing correlation between two maps

Piazza *et al.* [47] found a significant correlation between gene frequency gradients in humans and archaeological dates of the first Neolithic European farmers by using a Pearson's correlation coefficient modified for spatial data. This study suggests that the greatest influence on European genetic variation corresponds to the migration of Neolithic farmers from the area of origin of agriculture in the Middle East. A statistical test has also been developed to test whether two boundaries occur in exactly the same place [48].

Prospects

We have reviewed some advances in the analyses of genetic spatial patterns and their correlation with landscape and environmental features. The largest advance is the movement towards methods that do not require assumptions of population boundaries beforehand. The development of the fast-moving field of landscape genetics has benefited from the recent development of molecular tools and the new and existing statistical tools developed in landscape ecology. Landscape genetics should advance greatly our understanding of how landscape and environmental features influence gene flow, population structure and local adaptation. But landscape genetics is not an end in itself. The techniques used in this field will describe spatial genetic patterns, but more importantly will lead to exploration of the processes that caused the patterns.

In the laboratory, SNPs are likely to become widely used in the future, because of their favourable attributes: they are the most abundant class of polymorphisms in genomes [49,50], they can be genotyped in automated systems more easily than can microsatellites, and their mutation dynamics are more easily modeled and less variable among loci. Finally, they can often be identified in candidate genes with potentially important functions [51].

In statistics, the near future will see the development of statistical tests to detect selection across space [52] and to identify spatial patterns using linked loci instead of using only independent loci [53]. Most of the assumptions required in many of these statistical methods are likely to be violated in natural populations, and the impact of these violations on the inference reliability is uncertain. It is essential to validate the performance of statistical methods (e.g. by simulations). For example, Epperson and Li [54] characterized the statistical properties of spatial autocorrelation statistics for a wide range of sampling methods and sampling size, with careful considerations of appropriate spatial scale of sampling. It is important that studies such as this [54] are conducted and compared with other statistical approaches, because the success of studies of landscape genetics will depend on a change of sampling strategies in population genetics.

Landscape genetics implies random sampling across the entire study area and not just sampling some individuals in each of several populations identified beforehand. Unfortunately, the establishment of large-scale sampling programs to assess the genetic variability is usually difficult. Here, landscape management activities can help greatly: existing inventory and monitoring programs can provide environmental data and large numbers of genetic samples across large landscapes. These data can then be explored to find genetic discontinuities and identify areas of conservation concern. In general, more detailed investigations (i.e. simulations) are necessary to establish the precise relationships between properties of the statistical methods used and sampling strategy in terms of number of samples and loci, distance among samples and overall economic cost [55]. All these improvements will help landscape genetics fulfil its enormous potential for facilitating studies in ecology, evolution and conservation biology.

Acknowledgements

M.K.S. was supported by the Northern Region of the US Forest Service, and the Rocky Mountain Research Station. K. McKelvey, F. Samson and

J. Hoog provided helpful ideas for this review. We thank M. Arroyo, D. Tallmon, P. England and A. Peja-Pereira for comments on the paper and O. Hanotte and co-authors for providing Fig. 1b. We apologize to the many authors whose interesting and relevant work could not be cited because of space limitations.

References

1 Turner, M. *et al.* (2001) *Landscape Ecology in Theory and Practice*, Springer-Verlag
2 Smouse, P.E. and Peakall, R. (1999) Spatial autocorrelation analysis of individual multiallele and multilocus genetic structure. *Heredity* 82, 561–573
3 Brown, J.H. and Lomolino, M.V. (1998) *Biogeography* (2nd edn), Sinauer Associates
4 Sork, V.L. *et al.* (1999) Landscape approaches to historical and contemporary gene flow in plants. *Trends Ecol. Evol.* 14, 219–224
5 Reed, D.H. and Frankham, R. (2001) How closely correlated are molecular and quantitative measures of genetic variation? A meta-analysis. *Evolution* 55, 1095–1103
6 Crandall, K.A. *et al.* (2000) Considering evolutionary processes in conservation biology. *Trends Ecol. Evol.* 15, 290–295
7 Moritz, C. (2002) Strategies to protect biological diversity and the evolutionary processes that sustain it. *Syst. Biol.* 5, 238–254
8 Sokal, R.R. and Thomson, B.A. (1998) Spatial genetic structure of human populations in Japan. *Hum. Biol.* 70, 1–22
9 Cassens, I. *et al.* (2000) Mitochondrial DNA variation in the European otter (*Lutra lutra*) and the use of spatial autocorrelation analysis in conservation. *J. Hered.* 91, 31–35
10 Piertney, S. *et al.* (1998) Local genetic structure in red grouse (*Lagopus lagopus scoticus*): evidence from microsatellite DNA markers. *Mol. Ecol.* 7, 1645–1654
11 Keyghobadi, N. *et al.* (1999) Influence of landscape on the population genetic structure of the alpine butterfly *Parnassius smintheus* (Papilionidae). *Mol. Ecol.* 8, 1481–1495
12 Hanski, I. (1998) Metapopulation dynamics. *Nature* 396, 41–49
13 Piglucci, M. and Barbujani, G. (1991) Geographical patterns of gene frequencies in Italian populations of *Ornithogalum montanum* (Liliaceae). *Genet. Res.* 58, 95–104
14 Slate, J. *et al.* (1998) Bovine microsatellite loci are highly conserved in red deer (*Cervus elaphus*), sika deer (*Cervus nippon*) and Soay sheep (*Ovis aries*). *Anim. Genet.* 29, 307–315
15 Menotti-Raymond, M.A. and O'Brien, S.J. (1995) Evolutionary conservation of ten microsatellite loci in four species of Felidae. *J. Hered.* 86, 319–322
16 Atzmon, G. *et al.* (2002) DNA markers and crossbreeding scheme as means to select sires for heterosis in egg production of chickens. *Anim. Genet.* 34, 132–139
17 Hogan, Z.S. and May, B.P. (2002) Twenty-seven new microsatellites for the migratory Asian catfish family Pangasiidae. *Mol. Ecol. Notes* 2, 38–41
18 Black, W.C. *et al.* (2001) Population genomics: genome-wide sampling of insect populations. *Annu. Rev. Entomol.* 46, 441–469
19 Vitalis, R. *et al.* (2001) Interpretation of variation across marker loci as evidence of selection. *Genetics* 158, 1811–1823
20 Hardy, O.J. *et al.* (2000) Spatial autocorrelation of allozyme and quantitative markers within a natural population of *Centaurea jacea* (Asteraceae). *J. Evol. Biol.* 13, 656–667
21 McKay, J.K. and Latta, R.G. (2002) Adaptative population divergence: markers, QTL and traits. *Trends Ecol. Evol.* 17, 285–291
22 Wu, C. (2001) The genic view of the process of speciation. *J. Evol. Biol.* 14, 851–865
23 Lenormand, T. *et al.* (1999) Tracking the evolution of insecticide resistance in the mosquito *Culex pipiens*. *Nature* 400, 861–864
24 Paetkau, D. *et al.* (1995) Microsatellite analysis of population structure in Canadian polar bears. *Mol. Ecol.* 4, 347–354
25 Wright, S. (1943) Isolation by distance. *Genetics* 28, 114–138
26 Rousset, F. (2000) Genetic differentiation between individuals. *J. Evol. Biol.* 13, 58–62
27 Mantel, N. (1967) The detection of disease clustering and a generalised regression approach. *Cancer Res.* 27, 209–220
28 Sumner, J. *et al.* (2001) Neighbourhood size, dispersal and density estimates in the prickly forest skink (*Gnypetoscincus queenslandiae*) using individual genetic and demographic methods. *Mol. Ecol.* 10, 1917–1927
29 Sokal, R. and Neal, O. (1978) Spatial autocorrelation in biology, 2. Some implications and four applications of evolutionary interest. *Biol. J. Linn. Soc.* 10, 229–249
30 Barbujani, G. (2000) Geographic patterns: how to identify them and why. *Hum. Biol.* 72, 133–153
31 Pritchard, J.K. *et al.* (2000) Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959
32 Dawson, K. and Belkhir, K. (2001) A Bayesian approach to the identification of panmictic populations and the assignment of individuals. *Genet. Res.* 78, 59–77
33 Stow, A.J. *et al.* (2001) The impact of habitat fragmentation on dispersal of Cunningham's skink (*Egernia cunninghami*): evidence from allelic and genotypic analysis of microsatellites. *Mol. Ecol.* 10, 867–878
34 Cornuet, J.-M. *et al.* (1999) New methods employing multilocus genotypes to select or exclude populations as origins of individuals. *Genetics* 153, 1989–2000
35 Piazza, A. *et al.* (1981) Synthetic gene frequency maps of man and selective effects of climate. *Proc. Natl. Acad. Sci. U. S. A.* 78, 2638–2642
36 Cavalli-Sforza, L. *et al.* (1996) *The History and Geography of Human Genes*, Princeton University Press
37 Hanotte, O. *et al.* (2002) African pastoralism: genetic imprints of origins and migrations. *Science* 296, 336–339
38 Monmonier, M. (1973) Maximum-difference barriers: an alternative numerical regionalization method. *Geog. Anal.* 3, 245–261
39 Womble, W. (1951) Differential systematics. *Science* 28, 315–322
40 Barbujani, G. *et al.* (1989) Detecting regions of abrupt change in maps of biological variables. *Syst. Zool.* 38, 389–1989
41 Castric, V. *et al.* (2001) Landscape structure and hierarchical genetic diversity in the brook charr. *Salvelinus fontinalus*. *Evolution* 55, 1016–1028
42 Spong, G. and Creel, S. (2001) Deriving dispersal distances from genetic data. *Proc. R. Soc. Lond. Ser. B* 268, 2571–2574
43 Riginos, C. and Nachman, M.W. (2001) Population subdivision in marine environments: the contributions of biogeography, geographical distance and discontinuous habitat to genetic differentiation in a blennioid fish, *Axoclinus nigricaudus*. *Mol. Ecol.* 10, 1439–1453
44 Angers, B. *et al.* (1999) Canonical correspondence analysis for estimating spatial and environmental effects on microsatellite gene diversity in brook charr (*Salvelinus fontinalus*). *Mol. Ecol.* 8, 1043–1053
45 Gram, W.K. and Sork, V.L. (2001) Association between environmental and genetic heterogeneity in forest tree populations. *Ecology* 82, 2012–2021
46 Piazza, A. *et al.* (1981) The making and testing of geographic gene-frequency maps. *Biometrics* 37, 635–659
47 Piazza, A. *et al.* (1995) Genetics and the origin of European languages. *Proc. Natl. Acad. Sci. U. S. A.* 92, 5836–5840
48 Jacquez, G. (1995) The map comparison problem: tests for the overlap of geographical boundaries. *Stat. Med.* 14, 2343–2361
49 Reich, D.E. *et al.* (2001) Linkage disequilibrium in the human genome. *Nature* 411, 199–204
50 Primmer, C.R. *et al.* (2002) Single-nucleotide polymorphism characterization in species with limited available sequence information: high nucleotide diversity revealed in the avian genome. *Mol. Ecol.* 11, 603–612
51 Tabor, H.K. *et al.* (2002) Candidate-gene approaches for studying complex genetic traits: practical considerations. *Nat. Rev. Genet.* 3, 391–396
52 Nielsen, R. (2001) Statistical tests of selective neutrality in the age of genomics. *Heredity* 86, 641–647
53 Estoup, A. *et al.* (2000) Juxtaposed microsatellite systems as diagnostic markers for admixture: an empirical evaluation with brown trout (*Salmo trutta*) as model organism. *Mol. Ecol.* 9, 1873–1886
54 Epperson, B.K. and Li, T. (1996) Measurement of genetic structure within population using Moran's spatial autocorrelation statistics. *Proc. Natl. Acad. Sci. U. S. A.* 93, 10528–10532
55 Diniz-Filho, J.A.F. and Telles, M.P.C. (2002) Spatial autocorrelation

analysis and the identification of operational units for the conservation in continuous populations. *Conserv. Biol.* 16, 924–935
56 Roach, J.L. *et al.* (2001) Genetic structure of a metapopulation of black-tailed prairie dogs. *J. Mammal.* 82, 946–959
57 Raymond, M. and Rousset, F. (1995) Genpop (Version 1.2): population genetics software for exact tests and ecumenicism. *J. Hered.* 86, 248–249
58 Bertorelle, G. and Barbujani, G. (1995) Analysis of DNA diversity by spatial autocorrelation. *Genetics* 140, 811–819
59 Degen, B. *et al.* (2001) SGS-Spatial genetic software: a computer program for analysis of spatial genetic and phenotypic structures of individuals and populations. *J. Hered.* 92, 447–448
60 Hardy, O. and Vekemans, X. (2002) Spagedi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Mol. Ecol. Notes* 2, 618
61 Peakall, R. and Smouse, P.E. (2001) *GenAlEx V5: Genetic Analysis in Excel. Population Genetic Software for Teaching and Research*, Australian National University
62 Goudet, J. (1995) FSTAT version 1.2: a computer program to calculate F statistics. *J. Hered.* 86, 485–486
63 Crisci, J.V. (2001) The voice of historical biogeography. *J. Biogeogr.* 28, 157–168
64 Crisp, M. (2001) Historical biogeography and patterns of diversity in plants, algae and fungi: introduction. *J. Biogeogr.* 28, 153–156
65 Wallace, A.R. (1860) On the Zoological Geography of the Malay Archipelago. *J. Proc. Linn. Soc. Zool.* 4, 172–184
66 Petit, E. *et al.* (2001) No evidence of bottleneck in the postglacial recolonization of Europe by the noctule bat (*Nyctalus noctula*). *Evolution* 53, 1247–1258
67 Mani, F. and Barrai, I. (2001) Genetic structure and linguistic boundaries in Italy: a microregional approach. *Hum. Biol.* 73, 335–347
68 Brassel, K.E. and Reif, D. (1979) A procedure to generate Thiessen polygons. *Geog. Anal.* 325, 31–36
69 Barbujani, G. and Sokal, R. (1989) Zones of sharp genetic change in Europe are also linguistic boundaries. *Proc. Natl. Acad. Sci. U. S. A.* 87, 1816–1819

Could you name the most significant papers published in the life sciences this month?

Updated daily, **Research Update** on BioMedNet presents short, easy-to-read commentary about the latest hot papers, enabling you to keep abreast with advances across the life sciences.
Written by laboratory scientists with a keen understanding of their field, **Research Update** will clarify the significance and future impact of this research.

Articles will be freely available for a promotional period until 31st March 2003.
Our experienced in-house team is under the guidance of a panel of experts from across the life sciences, who offer suggestions and advice to ensure that we have high calibre authors and have spotted the 'hot' papers.

Visit the **Research Update** daily at http://update.bmn.com and sign up for email alerts to make sure you don't miss a thing.

This is your chance to have your opinion read by the life science community: if you would like to contribute contact us at research.update@elsevier.com