MOLECULAR ECOLOGY

Molecular Ecology (2010) doi: 10.1111/j.1365-294X.2010.04717.x

Perspectives on the use of landscape genetics to detect genetic adaptive variation in the field

STÉPHANIE MANEL,*† STÉPHANE JOOST,‡ BRYAN K. EPPERSON,§ ROLF HOLDEREGGER,¶ ANDREW STORFER,** MICHAEL S. ROSENBERG,†† KIM T. SCRIBNER, ‡‡ AURÉLIE BONIN§§ and MARIE-JOSÉE FORTIN¶¶

*Laboratoire Population Environnement Développement, UMR 151 UP/IRD, Université de Provence, 3 place Victor Hugo, 13331 Marseille Cedex 03, France, †Laboratoire d'Ecologie Alpine, UMR-CNRS 5553, Université Joseph Fourier, BP53 38041 Grenoble Cedex 9, France, ‡Laboratoire de Systèmes d'Information Géographique (LASIG), Ecole Polytechnique Fédérale de Lausanne (EPFL), Bâtiment GC, Station 18, 1015 Lausanne, Switzerland, §Department of Forestry, Michigan State University, East Lansing, MI 48824, USA, ¶WSL Swiss Federal Research Institute, Zürcherstrasse 111, CH-8903 Birmensdorf, Switzerland, **School of Biological Sciences, Washington State University, Pullman, WA 99164-4236, USA, ††Center for Evolutionary Medicine and Informatics, The Biodesign Institute, and School of Life Sciences, Arizona State University, Tempe, AZ 85287-4501, USA, ‡‡Department of Fisheries & Wildlife and Dept. of Zoology, Michigan State University, East Lansing, MI 48824, §§Department of Botany University of British Columbia 3529-6270 University Boulevard Vancouver, BC, V6T 1Z4, Canada, ¶¶Department of & Evolutionary Biology, University of Toronto, Toronto, ON, Canada M5S 3G5*

Abstract

Understanding the genetic basis of species adaptation in the context of global change poses one of the greatest challenges of this century. Although we have begun to understand the molecular basis of adaptation in those species for which whole genome sequences are available, the molecular basis of adaptation is still poorly understood for most non-model species. In this paper, we outline major challenges and future research directions for correlating environmental factors with molecular markers to identify adaptive genetic variation, and point to research gaps in the application of landscape genetics to real-world problems arising from global change, such as the ability of organisms to adapt over rapid time scales. High throughput sequencing generates vast quantities of molecular data to address the challenge of studying adaptive genetic variation in non-model species. Here, we suggest that improvements in the sampling design should consider spatial dependence among sampled individuals. Then, we describe available statistical approaches for integrating spatial dependence into landscape analyses of adaptive genetic variation.

Keywords: computational approach, genome scan, local adaptation, landscape genomics, molecular techniques, regression analysis

Received 30 November 2009; revision received 14 May 2010; accepted 14 May 2010

Introduction

Can species adapt to global change? Environmental change at all spatial scales is rapidly altering selection regimes for global flora and fauna (Reusch & Wood 2007). One of the most challenging questions of our time is whether adaptive evolution can keep pace with the rate and direction of selection that is imposed by humans (Hendry *et al.* 2008). It is critical to assess how genetic diversity may change and at what cost to the maintainance of population viability in the long-term (Lynch & Lande 1993). Adaptive genetic diversity dictates narrower tolerance limits to changing environmental conditions for specific populations than for a species as a whole (Etterson 2008). It is predicted that many species are able to shift their geographic ranges to track global change (Parmesan 2001), but the general potential of species to adapt to rapid change is still debated (Davis *et al.* 2005; Reusch & Wood 2007).

Correspondence: Stéphanie Manel,
E-mail: stephanie.manel@ujf-grenoble.fr

Local adaptation results from the balance between gene flow and many natural selection factors, including the climate (Savolainen *et al.* 2007; Hoffmann & Willi 2008). Adaptive differentiation among populations within a species has been documented through the study of clinal variation in physiological, phenological and fitness traits in relation to gradients in climate (Davis *et al.* 2005; Gienapp *et al.* 2008). Broadly speaking, however, we are only just beginning to understand the genomic basis of phenotypic traits associated with local adaptation for species whose whole genomes have been sequenced (e.g. Begun *et al.* 2007; Turner *et al.* 2008). Additionally, for a limited number of species such as forest trees, studies have been able to build upon a long history of common garden experimentation (Bradshaw *et al.* 1995; Neale & Savolainen 2004; Neale 2007; Neale & Ingvarsson 2008; Grattapaglia *et al.* 2009), allowing the characterization of the geographic pattern of neutral and adaptive genetic variation in relation to geography and climate (Savolainen *et al.* 2007; Aitken *et al.* 2008; Eckert *et al.* 2009a,b). This knowledge may aid the understanding of these species' responses to rapid climate change in the future (Sork *et al.* 2010).

Nonetheless, recent and upcoming advances in high throughput DNA sequencing leads to ever increasing availability of genomic sequences, facilitating an enhanced understanding of the genetic basis of current and future adaptation in a broad variety of species (Segelbacher *et al.* 2010). As a result, the limiting factor in future studies will no longer be the molecular laboratory workload, but rather the development of statistical, bioinformatics and modelling tools for identifying both genes or gene networks under selection (McCarthy *et al.* 2008), as well as the environmental factors acting as selective pressures. We refer to such a framework for understanding the spatial distribution of adaptive genetic variation using genomic tools (Box 1) as landscape genomics (Joost *et al.* 2007), and in Box 2 we clarify the use of this and other related terms.

Our focus is on studies and methods that will assess spatial correlations of particular molecular markers with environmental variables (Hamilton *et al.* 2002; Manel *et al.* 2009, 2010; Schwartz *et al.* 2010; Poncet *et al.* 2010). The pattern of genetic variation observed in such loci along environmental gradients has usually been interpreted as being caused by natural selection (Endler 1986; Schmidt *et al.* 2008). Further, we focus on the effects of evolutionary processes that can operate over much smaller spatial and temporal scales than those typically employed in phylogeographic studies (Manel *et al.* 2003). However, it is important to point out that larger time scale effects such as selective sweeps, a form of genetic hitchhiking where neutral

Box 1. Overview of advantages and drawbacks of the main genomic resources available for landscape genomics studies

AFLPs (Amplified Fragment Length Polymorphisms) and related markers

Until recently, the AFLP technique was the method of choice to obtain large numbers of molecular markers for non-model organism genomic studies, since it does not require prior sequenced-based information (Meudt & Clarke 2007). For example, in one of the most comprehensive AFLP-based genome scans, 1300 AFLP markers were surveyed to investigate the genetic basis of host specialization in the larch budmoth (Emelianov *et al.* 2004). AFLP markers are biallelic, dominant and they usually cover the entire genome although they sometimes tend to cluster around centromeres. A recurring issue associated with the AFLP technique is fragment size homoplasy (Vekemans *et al.* 2002), which occurs when non-homologous AFLP fragments co-migrate. In a variant of the AFLP protocol, the Diversity Array Technology (DArT), up to several thousands of DNA polymorphisms can be detected in a single hybridization assay on a microarray slide (Jaccoud *et al.* 2001). The major advantage of DArTs over AFLPs is that their sequences are easily accessible.

Microsatellites

Microsatellites are codominant and generally multiallelic (Zane *et al.* 2002). This makes them useful to monitor decreases in intrapopulation genetic variability observed in the vicinity of adaptive genes (Schlötterer 2002) or to identify particular alleles specifically associated with environmental variables (Joost *et al.* 2007), for example. However, microsatellites have a high mutation rate and a complex mutation pattern, characteristics which can be difficult to accommodate when searching for selection of signatures using traditional population genomics models (Vitalis *et al.* 2001). Moreover, microsatellites can be sparse in the genome of some species and thus difficult to find (Schlötterer 2004). Up to now, the development of hundreds of microsatellites was time-consuming and expensive (Zane *et al.* 2002), and these markers were also not particularly amenable to massively parallel genotyping. As a result, only in model species were microsatellite resources sufficient to be exploited in a population genomics context (Luikart *et al.* 2003). Fortunately, the increased availability of high-throughput

sequencing data will greatly facilitate microsatellite discovery and typing in non-model species (Hudson 2008). Microsatellites are featured in several studies reported in this issue. For instance, based on microsatellite data, Sork *et al.* (2010) detected climatically-associated genetic variation in populations of valley oak in California, suggesting that the potential for future adaptation in the face of climate change is limited in this long-lived species.

SNPs (Single Nucleotide Polymorphisms)

SNPs are the most abundant type of polymorphism in genomes (Schlötterer 2004). For example, on average there is one SNP every Kb in the 3-billion-base human genome (Zhang & Hewitt 2003). They are usually biallelic and evolve according to a simple infinite sites mutation model (Schlötterer 2004). One of the major drawbacks of SNPs is their susceptibility to ascertainment bias, i.e. the bias introduced by using a subset of the studied individuals or populations for marker discovery purposes and which can lead to a skew in the distribution of allelic frequencies (Morin *et al.* 2004). Detecting SNPs also requires a priori information on the studied genome sequence (Morin *et al.* 2004), but once this task is completed, SNPs present a high potential for an automated high-throughput analysis at a moderate cost (Schlötterer 2004). The most impressive SNP datasets have long been restricted to model species: for instance, more than 10 000 SNPs were surveyed to examine the effects of differentiation and selection in the human (Akey *et al.* 2004) and mouse (Harr 2006) genomes. Fortunately, next-generation sequencing technologies are expected to give a substantial boost to the use of SNPs for both model and non-model organisms. For example, Turner *et al.* (2010) investigated the genetic basis of adaptation to serpentine soils in *Arabidopsis lyrata* using about 8 millions polymorphims (mostly SNPs) identified in Solexa sequencing data. The markers showing the highest genetic differentiation between soil types were preferentially situated in genes involved in heavy metal detoxification and calcium/magnesium transport. These genes thus constitute good candidate for serpentine adaptation.

EST (Expressed Sequence Tag)-based molecular markers and other markers derived from next-generation sequencing data.

ESTs are short (~200–700 nucleotides) subsequences of transcribed and spliced DNA, generated by partially sequencing a pool of mRNAs (Bouck & Vision 2007). One of the most exciting prospects offered by next-generation sequencing technologies is the development of EST libraries for a wider range of species (Hudson 2008). These libraries can be astutely exploited to identify EST-based markers. These markers (classical microsatellites or SNPs) are usually located within a coding or a transcribed but untranslated region of a gene (Bouck & Vision 2007); but, they can also be assayed in non-transcribed sequences flanking genes by using a primer anchoring within the EST and another primer complementary to an adaptor-ligated restriction site (Bouck & Vision 2007). EST-based markers are thus tightly associated to gene-rich regions, which is particularly useful when searching for signatures of selection (Bonin 2008). Other types of promising marker systems building on next-generation sequencing data include the CRoPS (Complexity Reduction of Polymorphic Sequences; van Orsouw *et al.* 2007) and the RAD (Restriction-site associated DNA; Baird *et al.* 2008) methods. The practical and analytical shortcomings of all these new markers are nonetheless poorly understood. For example, the use of normalized EST libraries can theoretically bias the estimation of alleles frequencies at EST-based markers by favoring the sequencing of low-frequencies alleles. Similarly, the impact of sequencing errors on marker discovery remains to be explored.

alleles closely linked to a selectively favoured allele can increase in frequency (Hedrick 2005), can also affect the contemporary spatial distributions of genetic variation, even at fine spatial scales (Schonswetter *et al.* 2005; Knowles 2009). Large scale spatial effects from the distant past, such re-immigration after the last recent glacial epoch and subsequent refugia can also affect current spatial genetic, again even at fine scales (e.g. Boys *et al.* 2005). The potentially confounding effects of past events must be carefully considered using population genetic theory and by determining the appropriate spatial and temporal scales (discussed by Anderson *et al.* 2010).

Here, we identify some of the major challenges and future research directions in the study of the effects of environment on the adaptive genetic response of non-model organisms. We also identify gaps in the acquisition of molecular-genetic and environmental data that currently limit the application of landscape genomics to real-world problems. We discuss the importance of sampling design, which is strongly influenced by spatial dependencies among sampling points (Muirhead *et al.* 2008; Schwartz & McKelvey 2009; Anderson *et al.* 2010).

Finally, we suggest statistical approaches for integrating spatial dependence in analyses of genomic data.

Molecular data in landscape genomics

The main goal of landscape genomics is to identify loci having adaptive significance in the genome by combining genomic and environmental data (Box 2) (Joost *et al.* 2007). Landscape genomics has the remarkable characteristic of not requiring phenotypic data on the adaptive trait(s) of interest, which can be laborious to collect especially for wild and/or endangered species. In that respect, it differs from other classical strategies aimed at unraveling the genetic basis of adaptation, such as Quantitative Trait Loci (QTL) analysis or association mapping or quantitative genetics studies (Stinchcombe & Hoekstra 2008).

A prerequisite of the landscape genomics approach is to survey many genetic loci (typically several hundred or more) scattered in the genome of many individuals in order to discover genomic regions under selection, either directly or more likely though physical linkage (Box 3) (Luikart *et al.* 2003; Storz 2005). Several genomic resources can advantageously be exploited to this end (Box 1). Yet until recently, the amplified fragment length polymorphism (AFLP) technique has often been the most efficient option in terms of effort and costs to screen the genome of non-model species (Luikart *et al.* 2003). Hundreds of AFLP markers spanning the whole genome can be obtained relatively easily for any organism, without a priori sequence knowledge (Meudt & Clarke 2007). However, it is very laborious to link markers showing a signature of selection with the actual gene or mutation under selection (Bonin 2008). Moreover, obtaining sufficient AFLP markers to adequately saturate the genome is difficult, especially in species where linkage disequilibrium decays rapidly (Bonin 2008). As a result, AFLP-based genome scans have largely failed to pinpoint potential adaptive gene(s) or mutation(s) (but see Wood *et al.* 2008; Manel *et al.* 2010; Poncet *et al.* 2010).

Soon such technical limitations will disappear, owing to recent advances of next-generation sequencing technologies and its increasing affordability (Box 1) (Hudson 2008). The phrase 'next-generation sequencing' refers to the series of recent technologies capable of producing up to millions of relatively short sequence reads

Box 2. Clarification of terms

A number of recent terms, including landscape genetics (Manel *et al.* 2003), landscape genomics (Luikart *et al.* 2003; Joost *et al.* 2007), molecular genecology (Hamilton *et al.* 2002; Skot *et al.* 2002), and ecological genomics (Ungerer *et al.* 2008), have recently been introduced to describe studies aimed at understanding the impact of the environment/landscape on genetic response. These are in fact not new research fields, but rather involve the interdisciplinary integration of multiple pre-existing research disciplines, including spatial statistics, landscape ecology, population genetics and molecular biology. These terms were initially introduced to facilitate the discussion of researchers across disciplines; however, the multiplication of similar terms has led to the need for clarification.

Landscape genetics (Manel *et al.* 2003) aims to provide information about the interaction between landscape features and microevolutionary processes, such as gene flow, genetic drift or selection. Most current applications of landscape genetics focus on gene flow and migration (processes that can either facilitate or constrain local adaptation), i.e. the effect of the environment on the selectively neutral component of genetic diversity (Storfer *et al.* 2007; Anderson *et al.* 2010). However, landscape genetics also aims to correlate allele frequencies with the environment in order to understand the effect of the environment on the adaptive component of genetic diversity (Holderegger *et al.* 2006).

Landscape genomics (Luikart *et al.* 2003; Joost *et al.* 2007) uses correlation studies between the genomic data and the environment to identify genes either potentially linked to candidate genes or the genes themselves under selection. Landscape genomics is included in landscape genetics, but refers more specifically to the use of the future large amount of genetic data due to high-throughput sequencing. Landscape genomics is thus at the interface of bioinformatics, genomics, spatial statistics and landscape ecology.

Molecular genecology (Hamilton *et al.* 2002) is the study of geographical clines in the frequencies of alleles and their relationship to ecological clines in environmental conditions. Its objectives are largely the same as for the other research fields listed above.

Ecological genomics (Ungerer *et al.* 2008) integrates over several disciplines and seeks to understand the genetic mechanisms underlying responses of organisms to their natural environment. It is broader than landscape genetics and genomics, since it further includes experimental and laboratory approaches.

(35–1000 bases) at once thanks to the high parallelization of the sequencing process. Data throughput will continue to scale up in the near future with the ongoing development of real-time single-molecule sequencing technologies targeting longer reads (Hudson 2008). As a result, companies like VisiGen are aiming for a $1000 (human) genome, and we expect that studies of non-model organisms too will necessarily benefit from this 'genomic revolution'. Currently, and for a wide range of species, it is financially feasible to sequence Expressed Sequence Tag (EST; see Box 1) libraries and develop EST-associated molecular markers (e.g. Vera *et al.* 2008). Interestingly, unlike AFLPs or 'classical' microsatellites or Single Nucleotide Polymorphysms (SNPs), these markers occur in gene-rich regions of the genome, i.e. those most likely to be under selection. EST-based genome scans have already been used to identify promising candidate genes for adaptation in various species such as white spruce (Namroud *et al.* 2008), salmon (Vasemagi *et al.* 2005) and seagrass (Oetjen & Reusch 2007). However, the use of high-throughput sequencing techniques in a landscape genomics context is still in its infancy (but see Eckert *et al.* 2009a; 2010; for recent applications). Additionally, increasing the number of analysed loci will inevitably raise concerns about linkage disequilibrium, as is discussed in Box 3 (Segelbacher *et al.* 2010). Furthermore, current landscape genomic studies, currently for the most part at the exploratory stage, need to move forward to the confirmatory stage of proving the adaptive significance of identified loci linked to genes under selection (Reusch & Wood 2007).

Box 3. A cautionary note on linkage disequilibrium and multilocus genetics

Multilocus genetic processes are likely to figure prominently in the future of landscape genetics and genomics. Note, however, that spatial genetic structure and admixture could create linkage disequilibrium (LD) between physically unlinked markers as well as between unlinked markers and adaptive candidate genes. This can result in a two locus version of the Wahlund effect and hence a bias in ascertainment of genetic variability or population structure (Prout 1973; Christiansen & Feldman 1975). Primary among forces creating LD that is useful for gene discovery may be genetic hitchhiking effects (Thomson 1977; Asmussen & Clegg 1981; Ewens 2004; Hedrick 2005), which can take a number of forms, most importantly the accumulation of neutral mutations near alleles of loci that have undergone long term natural selection. Admixture can be a problem in an existing study system, or it could become a problem as populations go extinct or are founded and colonized. As an example, for human disease gene discovery, admixture-caused population level LD between markers and known genes is a major confounding problem. The solution often is to add genetic transmission tests and to analyse data using the Transmission Disequilibrium Test TDT (Spielman *et al.* 1993) or similar methods.

In general, little is known about multilocus genetics in a spatially explicit framework. One computer simulation study with selectively neutral genes and low amounts of dispersal in an isolation by distance process for a large population showed that LD was very small at the population-wide level, whether or not the two loci considered were physically linked. However, LD was large at smaller spatial scales, again irrespective of physical linkage, suggesting LD changes across different spatial scales (Epperson 1995). Moreover the relationship of LD with recombination rates is also scale dependent. If the complexity of the environment or the landscape are added, appropriate analytical models quickly become intractable, making computer simulations necessary (Epperson *et al.* 2010).

Environmental data

Landscape genomic studies either use environmental data collected in the field or take advantage of existing GIS databases. Recent increases in the availability of digital environmental data from remote sensors and weather stations have now made many global environmental data sets freely available (Box 4). In the absence of local environmental data, global environmental data can serve as valuable surrogates in landscape genomics studies. Yet, depending on the spatial (and temporal) scale of study question, detailed local measurements with high precision (e.g. spatial resolution $\leq 1\ m^2$) may often be needed to understand local microevolutionary processes (Anderson *et al.* 2010). Micro-environmental data may be gathered using special sensor networks installed in the field, providing high-resolution eco-climatic data. The US National Ecological Observatory Network (NEON) will likely encompass sensor networks throughout the USA to gather long-term data on ecological responses to changes in land use and climate at a cost of $400–450 million USD (Keller *et al.* 2008). Such high resolution data form the very foundation for future research to investigate the current local adaptation of organisms, which has been shaped by past selection.

Future environmental data acquisition for use in landscape genomics should: (1) use measures of environmental conditions within the home range of mobile organisms (Moorcroft & Lewis 2006); (2) complement coarser environmental data sets acquired over several decades (e.g. LANDSAT data) with local high resolution environmental data (e.g. fine scale IKONOS data); (3) make use of performance increases in data from new satellites or sensors-networks (4) make use of underexploited Digital Elevation Models (DEMs); and (5) use spatio-temporal three-dimensional data (Gugerli *et al.* 2008) instead of point environmental data, as is especially important in studies of vagile animals. We could then precisely match genetic data to environment at adequate spatial and temporal scales. For example Sork *et al.* (2010) used fine scale climate data at a scale appropriate for the genetic data to understand how climate change shapes the evolutionary response of Californian valley oak (*Quercus lobata*).

Spatial aspects specific to landscape genomics

To predict the future geographical range of a species, it is crucial to understand how species biologically respond to spatial heterogeneity of the environment or landscape at multiple spatial and temporal scales (Fortin & Dale 2005). Current species distributions are the result of many confounding processes, including population demography and history, phylogeographic history, behavior, physiological tolerances, competition, response to human land use change and adaptation to the environment (Gaston 2003). The interplay between selection and gene flow strongly influences biotic processes linked to adaptation (Savolainen *et al.* 2007; Holderegger & Wagner 2008).

Species distributional response to environmental conditions is a phenomenon that is often referred to as spatial dependence (Legendre 1993; Fortin & Dale 2005; Wagner & Fortin 2005). Species spatial aggregation occurs as well due to biotic processes such as dispersal and species interactions. These spatial structures create spatial autocorrelated genetic data. The degree of spatial autocorrelation in genetic data can be measured though various spatial autocorrelation coefficients (Fortin & Dale 2005). For animal species, a hypothetical example of the effects of habitat (Fig. 1) on spatial dependence of genetic associations among individuals for a neutral genetic locus versus a locus under selection is represented by Fig. 1. Spatial distributions of genotypes at the two loci are characterized by different spatial autocorrelation patterns. Measures of spatial structure for genotypes at the neutral locus often exhibit an isolation by distance pattern (Wright 1943) reflecting localized breeding and gene flow (Fig. 1b). In the example, environmental variables intrinsic to forest habitats

Box 4. Initiatives to map the environment

Large scale database measures. The Global Map project (http://www.globalmap.org/) exemplifies the trend toward constructing freely available, large-scale environmental data sets. It will include elevation, land cover (including vegetation) and land use data, as well as transportation infrastructure and political boundaries. The project is supervised by the International Steering Committee for Global Mapping (Secretariat of ISCGM 1998) with over 90 participating countries (Verdin & Jenson 1996). The main international global environmental geodata sources are included into the Global Map project and are available over the Internet from the Secretariat of ISCGM housed within the Geographical Survey Institute of Japan.

Several important international or national agencies have made efforts to freely distribute geo-environmental data describing the earth at different resolutions and for different periods. Primary among these are the European Environment Agency (EEA; http://www.eea.europa.eu/), American agencies such as USGS and NASA, and LANDSAT satellite images (http://www.landsat.org), which have offered global orthorectified data free of charge. Moreover, the Global Biodiversity Information Facility (GBIF; http://www.gbif.org/) is an international organization which aims to make the world's biodiversity geodata digitally available (including data on livestock species). Finally, UNEP documents the Global Environment Outlook (http://www.unep.org/geo). This UN report presents the challenges facing the Earth in safeguarding the environment and moving towards a more sustainable future, and it proposes a data compendium with a list of all key data providers (http://geocompendium.grid.unep.ch/).

Local scale sensor measures. With regard to local scale, research in landscape genomics will benefit from an ongoing major technological revolution in the acquisition of high spatial and temporal resolution environmental data. Sensor networks can be used for survey of the environment at many different scales, from continental systems designed to measure global change to recent advances allowing high resolution monitoring of specific habitats. They can be combined with computational tools including high-performance communication networks, data storage systems, GIS and visualization environments (Rundel *et al.* 2009). Moreover, resulting data can be easily integrated with remote sensing or other types of standard sets of eco-climatic parameters. The main quality of sensor net-

works lies in their capacity to extend spatial and temporal scales of observation, affording opportunities to obtain unexpected results and to develop new research paradigms (Porter *et al.* 2009).

Multiscale measures and Digital Elevation Models (DEM). DEM, using elevation measures (from database or direct measures) and numeric models, can provide a diversity of morphometric (slope, aspect, curvature), hydro-morphometric (e.g. wetness), and also climatic indicators (e.g. solar radiation). The USGS Earth Resources Observation and Science Center (http://eros.usgs.gov) distributes global digital raster data sets with spatial resolutions ranging from 1 km (GTOPO30) to 90 m resolution (SRTM), and even 30 m for the United States and territorial islands. These data sets can be completed with increasingly available Very High Resolution DEMs (1 m for XY coordinates, and ~0.5 m for Z) acquired with LIDAR (Light Detection And Ranging) technology, and are able to generate high-resolution habitat predictors (Andrew & Ustin 2009). This underexploited tool can provide multiscale data (Lassueur *et al.* 2006) to be used in landscape genomics studies.

confer a selective advantage to a certain genotype that is selected against in intervening grassland habitat. Spatial patterns in genotypic variation in the locus under selection is therefore the end result of convergence, isolation by distance and environmental effects intrinsic to forest as well as to grassland habitats.

In addition, current species distributions may have resulted from adaptations to environmental conditions that no longer exist (i.e. ancestral vs. current niche) (Wiens & Graham 2005; James *et al.* 2007; Roe *et al.* 2009). Indeed, a species optimal habitat may have already been lost or changed due to either natural or human influence. In such circumstances, current environmental–genotypic relationships would not be reliable as indicators of a species' genetic responses to environmental changes (Cushman *et al.* 2009).

Sampling design

Sampling of populations adapted to different habitats, climates, land uses or management systems (e.g. for livestock) must be carefully designed and statistically analysed over appropriate geographic scales (Lohr 1999; Fortin & Dale 2005; Muirhead *et al.* 2008; Schwartz & McKelvey 2009; Anderson *et al.* 2010). Sampling effort that is too low (Muirhead *et al.* 2008; Schwartz & McKelvey 2009), in light of the multiplicity of landscape and environmental factors acting at multiple spatial and temporal scales, limits the signal to noise ratio of spatial genetic structure and could lead to misinterpretation of spatial statistics. To interpret detection of significant relationships between genetic and spatial environmental data, one should distinguish among the effects of sampling effort (Fortin & Dale 2005; Muirhead *et al.* 2008), sampling design (Fortin *et al.* 1989; Legendre *et al.* 2002), the power of statistical methods employed (Fortin & Dale 2005) and the confounding effects of multiple spatio-temporal scales (Dungan *et al.* 2002; Geffen *et al.* 2004; Boulet *et al.* 2007). Accordingly, sampling designs should be stratified across environmental variables of interest using current landscapes features and environmental conditions as a quasi-experimental design to test specific hypotheses. In fact, landscape heterogeneity itself can be used as a quasi-experimental design to test specific hypotheses. For example, samples taken along an altitudinal gradient could be used to determine local adaptation to climatic conditions (e.g. Bonin *et al.* 2006). Also, it is important to assess the effective distance and pathway that an organism would use to move in a heterogeneous landscape (Spear *et al.* 2010), in order to relate the genetic diversity to the appropriate landscape features. For example, Vignieri (2005) tested whether individuals of the Pacific jumping mouse were using riparian zones or mountains to move between areas.

In quantifying the spatial structure of genetic data, it is difficult to tease apart the relative proportion of spatial dependence versus spatial autocorrelation described in the previous section, which are always confounded in both plants and animals. A potential solution to this problem is spatially-nested sampling designs, whereby for animals the distance between sampling locations varies from less than that of the species daily movement (i.e. to capture the degree of spatial autocorrelation) to beyond natal dispersal (i.e. to determine the environmental-species relationship) (Fortin & Dale 2005). Another way would be to perform model-based sampling to account for known environmental structure or gradients (de Gruijter & ter Braak 2004).

Challenges in spatial analysis of adaptive loci

The null hypothesis is that there is no correlation between a particular allele and environmental factors such as temperature or moisture apart from that which may be caused by limited dispersal and genetic drift. Relating specific alleles to an environmental variable is similar in some regards to association studies (Gupta *et al.* 2005; Balding 2006) that link alleles to phenotypes or to studies correlating species occurrence to environmental variables, as in ecological niche models. Methods used to tackle this problem have ranged from simple approaches such as linear regression to more sophisticated approaches such as generalized additive

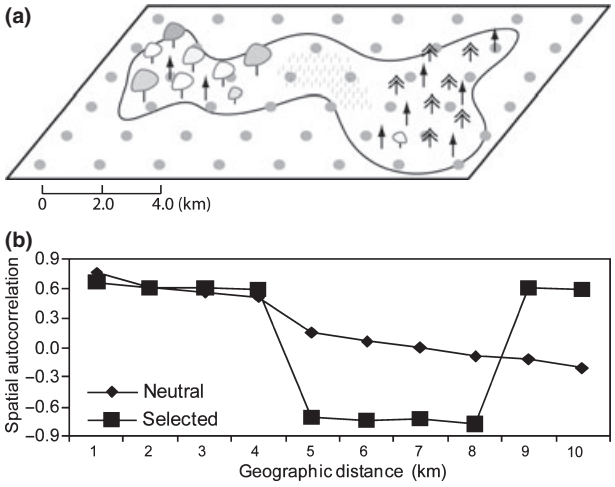

Fig. 1 Hypothetical example, for an animal species, of the effects of habitat on the spatial structure of genetic associations among individuals for nominal data (i.e. like and unlike genotypes; Sokal & Oden 1978) for a neutral genetic locus and locus under selection. (a) This spatially heterogeneous landscape has three land cover types (deciduous forest, grassland and coniferous forest) where an animal species is present in all three types. The gray dots indicate sampling locations. (b) Spatial distributions of genotypes at the two loci are characterized by different autocorrelation patterns obtained at the sampling locations. Measures of spatial dependence for genotypes at the neutral locus exhibit an isolation by distance (Wright 1943) pattern reflecting localized breeding and gene flow. Environmental variables intrisic to forest habitats confer a selective advantage to a certain genotype that is selected against in intervening grassland habitat. Spatial patterns in genotypic variation in the locus under selection appear to be due to convergence, isolation by distance and environmental effects intrinsic to forest as well as to grassland habitats.

models (Guisan *et al.* 2002; Pearman *et al.* 2008). Potential solutions have been proposed to consider explicitly the spatially dependent nature of the data using spatial regression methods (Dormann *et al.* 2007; Diniz *et al.* 2009; Dormann 2009). However, applications to detect loci potentially under selection are still lacking (but see Manel *et al.* 2010). Ideally, statistical methods in landscape genomics should consider both (1) spatial autocorrelation in allele frequencies generated by biotic processes (i.e. gene flow) which are distance related; and (2) unaccounted spatially structured environmental variables resulting in a spatial structuring of allele frequency distribution (Manel *et al.* 2010). Spatial regression methods have been put forth (e.g. conditional autoregressive models and simultaneous autoregressive models) to consider spatial dependence between individuals/loci and biotic processes by incorporating geographic space in the model structure. A promising spatial regression approach is the method of Moran's eigenvector maps (MEM) (Borcard & Legendre 2002; Dray *et al.* 2006; Diniz-Filho *et al.* 2009). MEM variables are the eigenvectors of a spatial weighting matrix calculated from the sampling locations' geographic coordinates. MEM analysis produces uncorrelated spatial eigenfunctions used to dissect the spatial patterns of the studied variation (allele frequencies in the present context) into separate scales to be used as predictors in regression. To detect loci potentially linked to genes under selection, Manel *et al.* (2010) used multiple linear regressions to correlate single AFLP allele frequencies from a large genome scan of *Arabis alpina* with environmental variables. To consider unmeasured variables in the analysis which potentially create spatial structure in allele distribution, they used only broad-scale principal coordinates of neighbour matrices (MEMs) as explanatory variables.

When sample size is small, spatial regression methods may not be appropriate given that the signal to noise ratio is generally low; geographically weighted regression has been proposed as one promising alternative (Fotheringham *et al.* 2002). In non-stationary circumstances, i.e. when spatial autocorrelation and effects of environmental correlates are not constant across the region, regression tree methods (e.g. CART, random forest, boosted regression; Elith & Graham 2009) offer alternatives to spatial regression (Dormann *et al.* 2007; Fortin & Melles 2009). Regression tree methods are based on an iterative procedure that splits the observations (samples) into a series of two groups in a hierarchical 'tree' (dendrogram-like) structure where the values of dependent variable are similar within each group based on a specific value of one of the independent variables (quantitative or qualitative independent values). Usually the first deeper splits reflect mostly large spatial scales processes while the last shallower splits in the tree structure correspond to localize spatial effects.

Consideration of spatial autocorrelation (i.e. biotic and abiotic processes) in the models allows the determination of processes governing allele frequency variation, but results may be strongly affected by sampling and stochastic variation (Slatkin & Arter 1991). Population and family structures have also been highlighted as a confounding issue in the inference of natural selection (Balding 2006; Excoffier *et al.* 2009). Bayesian geographical analysis approaches have been recently introduced to address this problem by testing for correlations between allele frequencies and environmental variables after correcting for background levels of population structure and differences in sample size (Felsenstein 2002; Yu *et al.* 2006; Hancock *et al.* 2008). Using this approach, Hancock *et al.* (2008) found evidence of a selective effect of the climate on metabolism genes in humans from the analysis of the association between 973 SNPs and climatic variables. This approach requires

that populations are known or defined in advanced (i.e. from genetic structure) to be able to estimate allele frequency, which is not always possible depending on the species and the sampling (Manel *et al.* 2005). Studies of adaptive genetic variation can benefit by genotyping many populations, in a broad range of conditions (Turner *et al.* 2010). It appears that in some cases, it is more effective to sample a large number of locations with fewer individuals than to sample many individuals in only a few locations (Poncet *et al.* 2010).

Once a model has been chosen, it is necessary to choose among multiple, ideally uncorrelated explanatory variables. Model selection procedures are commonly used for this purpose by giving a weight (score of importance) to each explanatory variable (Burnham & Anderson 2002). Such analyses result in choosing the factors that explain the highest proportion of variation in the dependent variable (usually allele frequency variation in landscape genomics studies).

Models will likely become increasingly complex in addressing landscape genomics issues, able to account for longer term effects, various modes of adaptive selection, linkage disequilibrium (Servin & Stephens 2007), pleiotropy and epistasis, structural versus regulatory genetic effects, as well as being able to compare multiple null and alternate hypotheses, and tailored to the characteritics of the study species. In light of the great complexity of landscape genetic processes, the goal of predicting the population genetic effects following projected global changes for a given species will require that appropriate models be constructed carefully, taking into account as many details of organismal biology as possible. The modes of selection responsible for current adaptation must be determined and implemented, and projection models should be spatially explicit and include both stochastic and uncertainty components. Due to such complexity, most models will be based on computer simulations (see Epperson *et al.* 2010). We are currently working on programs that ultimately will be able to model multiple distributions of a very wide range of patterns of environmental variables (and how these impose selection), include complex patterns of dispersal and genetic transmission, are multilocus, and allow environmental patterns to change over time. Again, such approaches to projection will not necessarily be simple, and careful attention should be paid to model assumptions and sources of error.

Conclusions

Forthcoming whole genome data sets will propel molecular ecology into a new dimension of genetic and evolutionary analysis. Landscape genomics, via studying the spatial distribution of loci of adaptive or ecological significance in natural populations, will contribute to the better understanding of plant and animal adaptation to their environment and inform management of genetic resources in response to adaptation to global change. Recent studies investigated the geographic and environmental pattern in SNP's associated with candidate genes, opening new insights in the understanding of the potential of populations to adapt to climate change (Eckert *et al.* 2009a,b; Eckert *et al.* 2010). Such studies provide an opportunity to resolve unanswered questions such as: does adaptation to local environments involve new mutations or standing genetic variation? How many genes influence ecologically important traits (Orr 2005; Stinchcombe & Hoekstra 2008)? The next step is to model spatially explicit forecasts of population genetic responses to climate shifts. What is needed are spatially explicit metapopulation and continous space models (Wade & McCauley 1988; Harding *et al.* 1998) with directional spatial shifting of the environment. Models for genes that are differentially selected in direct response to climate change will necessarily add other layers, as well as the potentially complex interactions between dispersal and selection.

Acknowledgements

This work was conducted as a part of the Landscape Genetics Working Group supported by the National Center for Ecological Analysis and Synthesis, a Center funded by NSF (Grant #DEB-0553768), the University of California, Santa Barbara, and the State of California. SM was also supported by the Institut Universitaire de France. RH was also supported by the CCES-BIOCHANGE project of the ETH-domain. MSR was also supported by the National Evolutionary Synthesis Center (NESCent: NSF #EF-0905606) and NSF # DBI-0542599. We also thank Victoria Sork and anonymous reviewers for helpful comments on an earlier version of the manuscript.

References

Aitken SN, Yeaman S, Holliday JA, Wang TL, Curtis-McLane S (2008) Adaptation, migration or extirpation: climate change outcomes for tree populations. *Evolutionary Applications*, **1**, 95–111.

Akey JM, Eberle MA, Rieder MJ *et al.* (2004) Population history and natural selection shape patterns of genetic variation in 132 genes. *Plos Biology*, **2**, 1591–1599.

Anderson C, Epperson B, Fortin M-J *et al.* (2010) Considering spatial and temporal scales in landscape genetics studies of gene flow. *Molecular Ecology*, doi: 10.1111/j.1365-294X.2010.04757.x.

Andrew ME, Ustin SL (2009) Habitat suitability modelling of an invasive plant with advanced remote sensing data. *Diversity and Distributions*, **15**, 627–640.

Asmussen MA, Clegg MT (1981) Dynamics of the linkage disequilibrium function under models of gene-frequency hitchhiking. *Genetics*, **99**, 337–356.

Baird NA, Etter PD, Atwood TS *et al.* (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE*, **3**, e3376.

Balding D (2006) A tutorial on statistical methods for population association studies. *Nature Reviews Genetics*, **7**, 781–791.

Begun DJ, Holloway AK, Stevens K *et al.* (2007) Population genomics: whole-genome analysis of polymorphism and divergence in *Drosophila simulans*. *Plos Biology*, **5**, 2534–2559.

Bonin A (2008) Population genomics: a new generation of genome scans to bridge the gap with functional genomics. *Molecular Ecology*, **17**, 3583–3584.

Bonin A, Miaud C, Taberlet P, Pompanon F (2006) Explorative genome scan to detect candidate loci for adaptation along a gradient of altitude in the common frog (*Rana temporaria*). *Molecular Biology and Evolution*, **23**, 773–783.

Borcard D, Legendre P (2002) All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. *Ecological Modelling*, **153**, 51–68.

Bouck A, Vision T (2007) The molecular ecologist's guide to expressed sequence tags. *Molecular Ecology*, **16**, 907–924.

Boulet M, Couturier S, Cote SD, Otto RD, Bernatchez L (2007) Integrative use of spatial, genetic, and demographic analyses for investigating genetic connectivity between migratory, montane, and sedentary caribou herds. *Molecular Ecology*, **16**, 4223–4240.

Boys J, Cherry M, Dayanandan S (2005) Microsatellite analysis reveals genetically distinct populations of red pine (*Pinus resinosa*, Pinaceae). *American Journal of Botany*, **92**, 833–841.

Bradshaw HD, Wilbert SM, Otto KG, Schemske DW (1995) Genetic-Mapping of floral traits associated with reproductive isolation in monkeyflowers (Mimulus). *Nature*, **376**, 762–765.

Burnham K, Anderson D (2002) *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, 2nd edn. Springer-Verlag, Heidelberg.

Christiansen F, Feldman M (1975) Subdivided populations: a review of one- and two-locus deterministic theory. *Theoretical Population Biology*, **7**, 13–38.

Cushman SA, McKelvey KS, Schwartz MK (2009) Use of empirically derived source-destination models to map regional Conservation Corridors. *Conservation Biology*, **23**, 368–376.

Davis MB, Shaw RG, Etterson JR (2005) Evolutionary responses to changing climate. *Ecology*, **86**, 1704–1714.

Diniz-Filho JAF, Nabout JC, Telles MPD, Soares TN, Rangel T (2009) A review of techniques for spatial modeling in geographical, conservation and landscape genetics. *Genetics and Molecular Biology*, **32**, 203–211.

Dormann CF (2009) Response to Comment on ''Methods to account for spatial autocorrelation in the analysis of species distributional data: a review''. *Ecography*, **32**, 379–381.

Dormann CF, McPherson JM, Araujo MB *et al.* (2007) Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography*, **30**, 609–628.

Dray S, Legendre P, Peres-Neto PR (2006) Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). *Ecological Modelling*, **196**, 483–493.

Dungan JL, Perry JN, Dale MRT *et al.* (2002) A balanced view of scale in spatial statistical analysis. *Ecography*, **25**, 626–640.

Eckert AJ, Bower AD, Wegrzyn JL *et al.* (2009a) Asssociation genetics of coastal douglas fir (*Pseudotsuga menziesu var. menziesii*, Pinaceae). I. Cold-hardiness related traits. *Genetics*, **182**, 1289–1302.

Eckert AJ, Pande B, Ersoz ES *et al.* (2009b) High-throughput genotyping and mapping of single nucleotide polymorphisms in loblolly pine (*Pinus taeda L.*). *Tree Genetics & Genomes*, **5**, 225–234.

Eckert AJ, Bower AD, Gonzalez-Martinez SC, Wegrzyn JL, Neale DB (2010) Back to nature: genetic association in a landscape context. *Molecular Ecology*, doi: 10.1111/j.1365-294X.2010.04698.x.

Elith J, Graham CH (2009) Do they? How do they? WHY do they differ? On finding reasons for differing performances of species distribution models. *Ecography*, **32**, 66–77.

Emelianov I, Marec F, Mallet J (2004) Genomic evidence for divergence with gene flow in host races of the larch budmoth. *Proceedings of the Royal Society of London, Series B*, **271**, 97–105.

Endler J (1986) *Natural Selection in the Wild*. Princeton University Press, New Jersey, USA.

Epperson BK (1995) Spatial structure of 2-locus genotypes under isolation by distance. *Genetics*, **140**, 365–375.

Epperson BK, McRae B, Scribner K *et al.* (2010) Utility of computer simulation in landscape genetics. *Molecular Ecology*, doi: 10.1111/j.1365-294X.2010.04678.x.

Etterson (2008) Evolution in response to climate change. In: *Conservation Biology: Evolution in Action* (eds Fox C, Carroll SB), pp. 145–163. Oxford University Press, New York.

Ewens WJ (2004) *Mathematical Population Genetics*. Springer-Verlag, Berlin.

Excoffier L, Hofer T, Foll M (2009) Detecting loci under selection in a hierarchically structured population. *Heredity*, **103**, 285–298.

Felsenstein J (2002) *Contrasts for a Within-Species Comparative Method*. Oxford University Press, Oxford.

Fortin M-J, Dale M (2005) *Spatial Analysis. A Guide for Ecologists*. Cambridge University Press, Cambridge.

Fortin MJ, Melles S (2009) *Avian Spatial Responses to Forest Spatial Heterogeneity at the Landscape Level: Conceptual and Statistical Challenges*. Springer, New York.

Fortin MJ, Drapeau P, Legendre P (1989) Spatial autocorrelation and sampling design in plant ecology. *Vegetatio*, **83**, 209–222.

Fotheringham AS, Brunsdon C, Charlton ME (2002) *Geographically Weighted Regression: the Analysis of Spatially Varying Relationships*. Wiley, Chichester.

Gaston K (2003) *The Structure and Dynamics of Geographic Ranges*. Oxford, UK.

Geffen E, Anderson MJ, Wayne RK (2004) Climate and habitat barriers to dispersal in the highly mobile grey wolf. *Molecular Ecology*, **13**, 2481–2490.

Gienapp P, Teplitsky C, Alho JS, Mills JA, Merila J (2008) Climate change and evolution: disentangling environmental and genetic responses. *Molecular Ecology*, **17**, 167–178.

Grattapaglia D, Plomion C, Kirst M, Sederoff RR (2009) Genomics of growth traits in forest trees. *Current Opinion in Plant Biology*, **12**, 148–156.

de Gruijter JJ, ter Braak CJF (2004) Design-based versus model-based sampling strategies: comment on R.J. Barnes

''Bounding the required sample size for geologic site characterization''. *Mathematical Geology*, **24**, 859–864.

Gugerli F, Englisch T, Niklfeld H *et al.* (2008) Relationships among levels of biodiversity and the relevance of intraspecific diversity in conservation—a project synopsis. *Perspectives in Plant Ecology Evolution and Systematics*, **10**, 259–281.

Guisan A, Edwards TC, Hastie T (2002) Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecological Modelling*, **157**, 89–100.

Gupta PK, Rustgi S, Kulwal PL (2005) Linkage disequilibrium and association studies in higher plants: present status and future prospects. *Plant Molecular Biology*, **57**, 461–485.

Hamilton NRS, Skot L, Chorlton KH, Thomas ID, Mizen S (2002) Molecular genecology of temperature response in Lolium perenne: 1. *Preliminary analysis to reduce false positives. Molecular Ecology*, **11**, 1855–1863.

Hancock AM, Witonsky DB, Gordon AS *et al.* (2008) Adaptations to climate in candidate genes for common metabolic disorders. *Plos Genetics*, **4**, e32.

Harding JS, Benfield EF, Bolstad PV, Helfman GS, Jones EBD (1998) Stream biodiversity: The ghost of land use past. *Proceedings of the National Academy of Sciences, USA*, **95**, 14843–14847.

Harr B (2006) Genomic islands of differentiation between house mouse subspecies. *Genome Research*, **16**, 730–737.

Hedrick PW (2005) *Genetics of Populations*. Sudbury, Massachusetts.

Hendry A, Farrugia T, Kinnison M (2008) Human influences on rates of phenotypic change in wild animal populations. *Molecular Ecology*, **17**, 20–30.

Hoffmann A, Willi Y (2008) Detecting genetic response to environmental change. *Nature Reviews Genetics*, **9**, 421–432.

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

Holderegger R, Kamm U, Gugerli F (2006) Adaptive vs. neutral genetic diversity: implications for landscape genetics. *Landscape Ecology*, **21**, 797–807.

Hudson ME (2008) Sequencing breakthroughs for genomic ecology and evolutionary biology. *Molecular Ecology Resources*, **8**, 3–17.

Jaccoud D, Peng K, Feinstein D, Kilian A (2001) Diversity Arrays: a solid state technology for sequence information independent genotyping. *Nucleic Acids Research*, **29**, E25.

James PMA, Fortin MJ, Fall A, Kneeshaw D, Messier C (2007) The effects of spatial legacies following shifting management practices and fire on boreal forest age structure. *Ecosystems*, **10**, 1261–1277.

Joost S, Bonin A, Bruford MW *et al.* (2007) A spatial analysis method (SAM) to detect candidate loci for selection: towards a landscape genomics approach to adaptation. *Molecular Ecolgy*, **16**, 3955–3969.

Keller M, Schimel DS, Hargrove WW, Hoffman FM (2008) A continental strategy for the National Ecological Observatory Network. *Frontiers in Ecology and the Environment*, **6**, 282–284.

Knowles LL (2009) Statistical phylogeography. *Annual Review of Ecology Evolution and Systematics*, **40**, 593–612.

Lassueur T, Joost S, Randin CF (2006) Very high resolution digital elevation models: do they improve models of plant species distribution? *Ecological Modelling*, **198**, 139–153.

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

Legendre P, Dale MRT, Fortin MJ *et al.* (2002) The consequences of spatial structure for the design and analysis of ecological field surveys. *Ecography*, **25**, 601–615.

Lohr (1999) *Sampling: Design and Analysis*, Duxbury.

Luikart G, England PR, Tallmon D, Jordan S, Taberlet P (2003) The power and promise of population genomics: from genotyping to genome typing. *Nature Reviews Genetics*, **4**, 981–994.

Lynch M, Lande L (1993) *Evolution and Extinction in Response to Environmental Change*. Sinauer, Sunderland, MA.

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

Manel S, Gaggiotti OE, Waples RS (2005) Assignment methods: matching biological questions with appropriate techniques. *Trends in Ecology and Evolution*, **20**, 136–142.

Manel S, Conord C, Després L (2009) Genome scan to assess the respective roles of host-plant and environmental constraints on the adaptation of a widespread insect. *BMC Evolutionary Biology*, **9**, 288.

Manel S, Poncet B, Legendre P, Gugerli F, Holderegger R (2010) Common factors drive adaptive genetic variation at different spatial scales in *Arabis alpina*. *Molecular Ecology*, doi: 10.1111/j.1365-294X.2010.04716.x.

McCarthy MI, Abecasis GR, Cardon LR *et al.* (2008) Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nature Reviews Genetics*, **9**, 356–369.

Meudt HM, Clarke AC (2007) Almost forgotten or latest practice? AFLP applications, analyses and advances. *Trends in Plant Science*, **12**, 106–117.

Moorcroft P, Lewis M (2006) *Mechanistic Home Range Analysis*. Princeton University Press, Princeton.

Morin PA, Luikart G, Wayne RK (2004) SNPs in ecology, evolution and conservation. *Trends in Ecology & Evolution*, **19**, 208–216.

Muirhead JR, Gray DK, Kelly DW *et al.* (2008) Identifying the source of species invasions: sampling intensity vs. genetic diversity. *Molecular Ecology*, **17**, 1020–1035.

Namroud MC, Beaulieu J, Juge N, Laroche J, Bousquet J (2008) Scanning the genome for gene single nucleotide polymorphisms involved in adaptive population differentiation in white spruce. *Molecular Ecology*, **17**, 3599–3613.

Neale DB (2007) Genomics to tree breeding and forest health. *Current Opinion in Genetics & Development*, **17**, 539–544.

Neale DB, Ingvarsson PK (2008) Population, quantitative and comparative genomics of adaptation in forest trees. *Current Opinion in Plant Biology*, **11**, 149–155.

Neale DB, Savolainen O (2004) Association genetics of complex traits in conifers. *Trends in Plant Science*, **9**, 325–330.

Oetjen K, Reusch TBH (2007) Genome scans detect consistent divergent selection among subtidal vs. intertidal populations of the marine angiosperm *Zostera marina*. *Molecular Ecology*, **16**, 5156–5167.

Orr HA (2005) The genetic theory of adaptation: A brief history. *Nature Reviews Genetics*, **6**, 119–127.

van Orsouw NJ, Hogers RCJ, Janssen A *et al.* (2007) Complexity reduction of polymorphic sequences (CRoPS (TM)): a novel approach for Large-Scale Polymorphism discovery in complex Genomes. *PLoS ONE*, **2**, e1172.

Parmesan C (2001) Detection of range shifts: general methodological issues and case studies of butterflies. In:

Fingerprints of Climate Change – Adapted Behaviour and Shifting Species Ranges (eds Walther G-R, Burga CA, Edwards PJ). pp. 57–76. Kluwer Academic, New York.

Pearman PB, Randin CF, Broennimann O *et al.* (2008) Prediction of plant species distributions across six millennia. *Ecology Letters*, **11**, 357–369.

Poncet B, Herrmann D, Gugerli F *et al.* (2010) Tracking genes of ecological relevance using a genome scan in two independent regional population samples of *Arabis alpina*. *Molecular Ecology*, doi: 10.1111/j.1365-294X.2010.04696.x.

Porter JH, Nagy E, Kratz TK *et al.* (2009) New eyes on the world: advanced sensors for ecology. *BioScience*, **59**, 385–397.

Prout T (1973) Appendix to J. B. Mitton and R. K. Koehn. Population genetics of marine pelecypods. III. Epistasis between functionally related isozymes in Mytilas edulis. *Genetics*, **73**, 487–496.

Reusch TBH, Wood TE (2007) Molecular ecology of global change. *Molecular Ecology*, **16**, 3973–3992.

Roe JH, Brinton AC, Georges A (2009) Temporal and spatial variation in landscape connectivity for a freshwater turtle in a temporally dynamic wetland system. *Ecological Applications*, **19**, 1288–1299.

Rundel PW, Graham EA, Allen MF, Fisher JC, Harmon TC (2009) Environmental sensor networks in ecological research. *New Phytologist*, **182**, 589–607.

Savolainen O, Pyhajarvi T, Knurr T (2007) Gene flow and local adaptation in trees. *Annual Review of Ecology Evolution and Systematics*, **38**, 595–619.

Schlötterer C (2002) A microsatellite-based multilocus screen for the identification of local selective sweeps. *Genetics*, **160**, 753–763.

Schlötterer C (2004) The evolution of molecular markers—just a matter of fashion? *Nature Reviews Genetics*, **5**, 63–69.

Schmidt PS, Serrao EA, Pearson GA *et al.* (2008) Ecological genetics in the north Atlantic: environmental gradients and adaptation at specific loci. *Ecology*, **89**, S91–S107.

Schonswetter P, Stehlik I, Holderegger R, Tribsch A (2005) Molecular evidence for glacial refugia of mountain plants in the European Alps. *Molecular Ecology*, **14**, 3547–3555.

Schwartz MK, McKelvey KS (2009) Why sampling scheme matters: the effect of sampling scheme on landscape genetic results. *Conservation Genetics*, **10**, 441–452.

Schwartz M, Luikart G, McKelvey K, Cushman S (2010) Landscape genomics: a brief perspective. In: *Spatial Complexity, Informatics, and Wildlife Conservatio* (eds Huettmann F, Cushman S). pp. 165–174, Springer-Verlag, Berlin.

Segelbacher G, Cushman S, Epperson B *et al.* (2010) Landscape Genetics: concepts and Challenges in a Conservation Context. *Conservation Genetics.*, **11**, 375–385.

Servin B, Stephens M (2007) Imputation-based analysis of association studies: candidate regions and quantitative traits. *Plos Genetics*, **3**, 1296–1308.

Skot L, Hamilton NRS, Mizen S, Chorlton KH, Thomas ID (2002) Molecular genecology of temperature response in Lolium perenne: 2. association of AFLP markers with ecogeography. *Molecular Ecology*, **11**, 1865–1876.

Slatkin M, Arter HE (1991) Spatial autocorrelation methods in population-genetics. *American Naturalist*, **138**, 499–517.

Sokal R, Oden N (1978) Spatial autocorrelation in biology. 2. Some implications and four applications of evolutionnary interest. *Biological Journal of Limnean Society*, **10**, 229–249.

Sork V, Davis F, Wang HF *et al.* (2010) Gene movement and genetic association with regional climate gradients in California valley oak (Quercus lobata Née). *Molecular Ecology*, doi: 10.1111/j.1365-294X.2010.04726.x.

Spear S, Fortin M, McRae B, Scribner K (2010) Use of resistance surfaces for landscape genetic studies: consideration for parameterization and analysis. *Molecular Ecology*, doi: 10.1111/j.1365-294X.2010.04657.x.

Spielman RS, McGinnis RE, Ewens WJ (1993) Transmission test for linkage disequilibrium—The insulin gene region and insulin-dependent diabetes-mellitus (Iddm). *American Journal of Human Genetics*, **52**, 506–516.

Stinchcombe JR, Hoekstra HE (2008) Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits. *Heredity*, **100**, 158–170.

Storfer A, Murphy M, Evans J *et al.* (2007) Putting the landscape in the landscape genetics. *Heredity*, **98**, 1–15.

Storz JF (2005) Using genome scans of DNA polymorphism to infer adaptive population divergence. *Molecular Ecology*, **14**, 671–688.

Thomson G (1977) Effect of a selected locus on linked neutral loci. *Genetics*, **85**, 753–788.

Turner TL, Levine MT, Eckert ML, Begun DJ (2008) Genomic analysis of adaptive differentiation in Drosophila melanogaster. *Genetics*, **179**, 455–473.

Turner JRG, Bourne EC, Wettberg EJ *et al.* (2010) Population resequencing reveals local adaptation of *Arabidopsis lyrata* to serpentine soils. *Nature genetics*, doi: 10.38/ng.515.

Ungerer MC, Johnson LC, Herman MA (2008) Ecological genomics: understanding gene and genome function in the natural environment. *Heredity*, **100**, 178–183.

Vasemagi A, Nilsson J, Primmer CR (2005) Expressed sequence tag-linked microsatellites as a source of gene-associated polymorphisms for detecting signatures of divergent selection in Atlantic salmon (*Salmo salar L.*). *Molecular Biology and Evolution*, **22**, 1067–1076.

Vekemans X, Beauwens T, Lemaire M, Roldan-Ruiz I (2002) Data from Amplified Fragment Length Polymorphism (AFLP) markers show indication of size homoplasy and of a relationship between degree of homoplasy and fragment size. *Molecular Ecology*, **11**, 139–151.

Vera JC, Wheat CW, Fescemyer HW *et al.* (2008) Rapid transcriptome characterization for a nonmodel organism using 454 pyrosequencing. *Molecular Ecology*, **17**, 1636–1647.

Verdin KL, Jenson SK (1996) Development of continental scale digital elevation models and estraction of hydrographic features. In: *Third International Conference/Workshop on Integrating Geographic Information Systems and Environmental Modelin*, National Center for Geographic Information and Analysis, Santa Fe, New Mexico. Available from http://www1.gsi.go.jp/geowww/globalmap-gsi/gtopo30/papers/santafe3.html.

Vignieri SN (2005) Streams over mountains: influence of riparian connectivity on gene flow in the Pacific jumping mouse (*Zapus trinotatus*). *Molecular Ecology*, **14**, 1925–1937.

Vitalis R, Dawson K, Boursot P (2001) Interpretation of variation across marker loci as evidence of selection. *Genetics*, **158**, 1811–1823.

Wade MJ, McCauley DE (1988) Extinction and recolonization—their effects on the genetic differentiation of local-populations. *Evolution*, **42**, 995–1005.

Wagner HH, Fortin MJ (2005) Spatial analysis of landscapes: concepts and statistics. *Ecology*, **86**, 1975–1987.

Wiens JJ, Graham CH (2005) Niche conservatism: integrating evolution, ecology, and conservation biology. *Annual Review of Ecology Evolution and Systematics*, **36**, 519–539.

Wood HM, Grahame JW, Humphray S, Rogers J, Butlin RK (2008) Sequence differentiation in regions identified by a genome scan for local adaptation. *Molecular Ecology*, **17**, 3123–3135.

Wright S (1943) Isolation by distance. *Genetics*, **28**, 114–138.

Yu JM, Pressoir G, Briggs WH *et al*. (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nature Genetics*, **38**, 203–208.

Zane L, Bargelloni L, Patarnello T (2002) Strategies for microsatellite isolation: a review. *Molecular Ecology*, **11**, 1–16.

Zhang DX, Hewitt GM (2003) Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. *Molecular Ecology*, **12**, 563–584.