

*Putting the landscape into the genomics of trees: approaches for understanding local adaptation and population responses to changing climate*

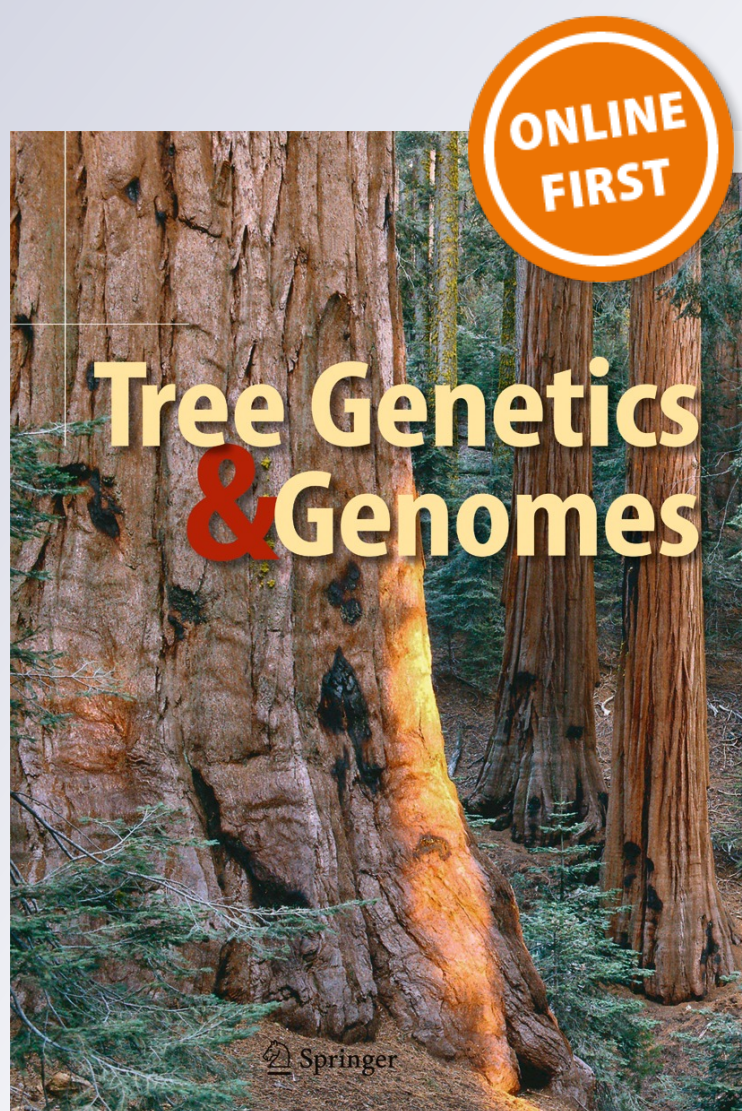
**V. L. Sork, S. N. Aitken, R. J. Dyer,  
A. J. Eckert, P. Legendre & D. B. Neale**

**Tree Genetics & Genomes**

ISSN 1614-2942

Tree Genetics & Genomes

DOI 10.1007/s11295-013-0596-x



**Your article is protected by copyright and all rights are held exclusively by Springer-Verlag Berlin Heidelberg. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your work, please use the accepted author's version for posting to your own website or your institution's repository. You may further deposit the accepted author's version on a funder's repository at a funder's request, provided it is not made publicly available until 12 months after publication.**

# Putting the landscape into the genomics of trees: approaches for understanding local adaptation and population responses to changing climate

V. L. Sork · S. N. Aitken · R. J. Dyer ·  
A. J. Eckert · P. Legendre · D. B. Neale

Received: 28 July 2012 / Revised: 6 January 2013 / Accepted: 14 January 2013  
© Springer-Verlag Berlin Heidelberg 2013

**Abstract** The Forest ecosystem genomics Research: supportTing Transatlantic Cooperation project (FoResTTraC, <http://www.foresttrac.eu/>) sponsored a workshop in August 2010 to evaluate the potential for using a landscape genomics approach for studying plant adaptation to the environment and the potential of local populations for coping with changing climate. This paper summarizes our discussions and articulates a vision of how we believe forest trees offer an unparalleled opportunity to address fundamental biological questions, as well as how the application of landscape genomic methods complement to traditional forest genetic approaches that provide critical information needed for natural resource management. In this paper, we will cover four topics. First, we begin by defining landscape genomics and briefly reviewing the

unique situation for tree species in the application of this approach toward understanding plant adaptation to the environment. Second, we review traditional approaches in forest genetics for studying local adaptation and identifying loci underlying locally adapted phenotypes. Third, we present existing and emerging methods available for landscape genomic analyses. Finally, we briefly touch on how these approaches can aid in understanding practical topics such as management of tree populations facing climate change.

**Keywords** Adaptive genetic variation · Climate change · Environmental association · Forest genetics · Genomics · Landscape genetics

Communicated by A. Abbott

V. L. Sork (✉)  
Department of Ecology & Evolutionary Biology,  
University of California, Los Angeles, Box 957239, Los Angeles,  
CA 90095-7239, USA  
e-mail: vlsork@ucla.edu

V. L. Sork  
Institute of the Environment and Sustainability, University  
of California, Los Angeles, Los Angeles, CA, USA

S. N. Aitken  
Centre for Forest Conservation Genetics, Department of Forest  
Sciences, Vancouver BC V6T 1Z4, Canada

R. J. Dyer · A. J. Eckert  
Department of Biology, Virginia Commonwealth University,  
Richmond, VA 23284, USA

P. Legendre  
Département de Sciences Biologiques, Université de Montréal,  
Montréal, Québec H3C 3J7, Canada

D. B. Neale  
Department of Plant Sciences, University of California, Davis,  
Davis, CA 95616, USA

## Introduction

Landscape genomics is the spatially explicit study of geographic patterns of genome-wide genetic variation (Holderegger et al. 2006; Manel et al. 2010a; Sork and Waits 2010). Traditionally, landscape genetics has used selectively neutral genetic markers and focused on the impact of landscape features in shaping movement of animals, gene movement of plants via seed and pollen, and geographic patterns of genetic variation, as discussed in numerous reviews on the topic (Balkenhol et al. 2009a, b; Cushman and Landguth 2010; Holderegger and Wagner 2008; Manel et al. 2003; Segelbacher et al. 2010; Storfer et al. 2007; Storfer et al. 2010). However, the availability of genome-wide sequences, even for non-model systems, creates the opportunity for a landscape “genomic” approach that can simultaneously examine the effects of demographic history, migration, and selection. To do so requires an important step of identifying genomic regions that underlie phenotypes involved in local adaptation. The power of landscape genomics is rooted in the ability to combine information



on the phenotype, the genotype, and the local environment of large numbers of spatially referenced samples collected across landscape scales (Fig. 1). Below, we briefly describe features of forest trees that make landscape genomic approaches tractable, and then outline complementary experimental and statistical methods available for assessing these elements of the adaptive landscape of forest trees.

### Trees as a tractable system

Tree species are of particular interest for management and conservation because of their economic and ecological importance. The effects of climate change on forest health, including the extant patterns and future potential of local adaptation, have the potential to directly impact the global carbon cycle and the rate of climate change due to the huge amount of carbon sequestered and stored by trees in forested ecosystems. A landscape genomic approach thus provides valuable insight for forest conservation and management in these times of human-induced rapid environmental change.

Forest trees have life history and genomic characteristics that facilitate the application of a landscape genomic approach to be jointly informative about gene movement and local adaptation in natural environments. Trees live for long periods of time in one location, which contrasts dramatically with the situation of highly mobile animal species or herbaceous plants that can disperse annually (Bradshaw 1972; Petit and Hampe 2006). Unlike most agricultural crop plants, forest trees have not been domesticated, with populations possessing their full complement of genetic diversity

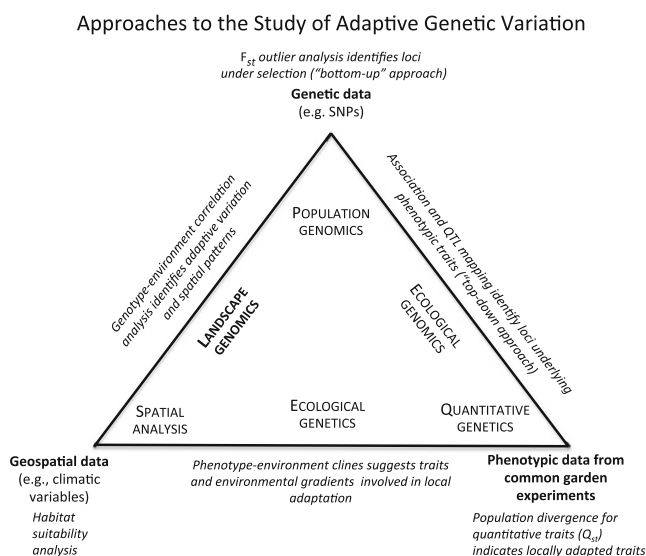
resulting from natural demographic and selective processes. Moreover, forest tree populations are often very large and frequently occupy highly heterogeneous environments, which results in large amounts of genetic variation within as well as among populations. High gene flow primarily via pollen can spread adaptive variants across a landscape, but selection for local adaptation appears to be strong, as indicated by strong population differentiation for key adaptive traits ( $Q_{st}$ ; Savolainen et al. 2007).

Some forest tree species, conifers in particular, have very large and complex genomes (e.g., conifers have genomes of 20+ GB, Morgante and De Paoli 2011), while others have more moderately sized genomes, such as oaks (~650–900 Mb in size, Favre and Brown 1996; Kremer et al. 2007; Zoldos et al. 1998) and poplars (about 550–650 Mb, Brunner et al. 2004; Tuskan et al. 2006). An analysis of DNA C-values across many plant taxa available through the Kew Garden website (<http://data.kew.org/cvalues/>) also indicates that trees species vary in genome size. With the advent of next-generation sequencing technologies, even the larger and more complex genomes of conifers are no longer intractable for study and the evolutionary and adaptive capacity of complex genomes of all sizes present exciting opportunities for discovery (Mackay et al. 2012).

A low level of linkage disequilibrium (LD) is another unique characteristic of most tree genomes (Moritsuka et al. 2012; Neale and Savolainen 2004). While this quality has sometimes been viewed as a liability in the discovery of genes underlying complex traits, it can now also be seen as an asset. The low levels of LD are the consequence of high outcrossing rates, long-distance propagule dispersal, and large effective population sizes. For landscape genomics, this property of tree populations makes it easier to identify sequences associated with specific phenotypes because the causal variants responsible for phenotypic traits can be located specifically within a sequence, especially once genome-wide marker sets become available (e.g., Parchman et al. 2012).

### Landscape genomic approaches to studying local adaptation

Landscape genomics can utilize any type of genome-wide marker (e.g., AFLPs or single-nucleotide polymorphisms (SNPs)). Increasingly, SNPs are preferable for analysis because they open the possibility of identifying functional genes and regulatory regions that underlie phenotypes. By analyzing a large number of SNPs for individuals sampled across a landscape, associations among genotypes, phenotypes and environments that reflect population response to the environment can be identified. Genome-wide association studies (GWAS) can be used to identify adaptive genetic variation by associating SNPs with a particular trait (Hancock and Di Rienzo 2008). This method is particularly effective for model systems such as *Arabidopsis* where traits have been mapped onto the genome (Atwell et al. 2010; Bergelson and Roux 2010; Fournier-Level



**Fig. 1** Schematic diagram of how phenotypic, genomic, and geospatial data are combined and analyzed to identify adaptive genetic variation. Disciplinary research areas are in CAPS. Analytical methods named along sides of triangles combine two data types; methods named at triangle vertices involve analysis of just one data type

et al. 2011). For non-model species, such as most tree species, a landscape approach would be to use the association between environmental gradients and genotypes to identify potential loci underlying selection. Such associations using high-resolution genome-wide scans are now feasible for non-model species, due to the emergence of reduced-representation DNA library sequencing using restriction-enzymes (e.g., Davey et al. 2011; Elshire et al. 2011; Miller et al. 2007). These methods can generate tens of thousands of SNPs across the entire genome at relatively low cost. While both GWAS and reduced-representation sequencing methods may find only a small number of causal variants of main effect, it is possible to identify multiple regions that explain a large part of the variation (e.g., Parchman et al. 2012; see Eckert and Dyer 2012). Moreover, they can be combined with targeted sequencing for candidate genes (e.g., Gnirke et al. 2009) to increase efficiency of testing whether allele frequencies associated with functional genes show environmental associations. In short, with large sample sizes of individuals and large numbers of genome-wide SNPs, landscape genomic analysis of genetic variation can provide the first level of evidence for the impact of selection in natural populations.

Traditional approaches to studying local adaptation in forest trees

Forest genetics has a long tradition of provenance (common garden) experiments yielding high quality phenotypic data for among and within-population variation (Langlet 1971). These experiments have yielded phenotypic data for quantitative genetic and ecological genetic analyses (Fig. 1). The provenance is the location or climate of origin of genetic materials from natural populations. Historically, these experiments were planted on multiple field sites and maintained for decades. Provenance trials are often partial reciprocal transplant experiments established on several sites. The original purposes of these trials were to quantify phenotypic differences among populations and to identify optimal seed sources for reforestation within regions. However, these experiments are now being repurposed for landscape genomic and climate change-related objectives. They may provide the opportunity to (1) quantify genetic clines reflecting local adaptation and identify the environmental gradients in which they occur; (2) quantify the degree of population differentiation and identify traits under divergent selection among environments; and (3) provide opportunities for linking genotypes to phenotypes using association mapping approaches described below.

Advantages of long-term field provenance trials for studying local adaptation in forest trees include the ability to assess survival and growth over the long term in real environments (Aitken et al. 2008; Chen et al. 2012). Disadvantages include: the extensive resources needed for establishment and maintenance of forest populations; the long-term nature of data

collection; and the difficulty of assessing traits other than survival and growth. Many provenance trials sample populations from within relatively small geopolitical regions, thus range-wide patterns of variation cannot be studied and the geographic signal to noise ratio is low. This sampling strategy often includes fewer provenances than ideal for landscape genomics. Field test sites often sample only productive area within these regions, and do not push genotypes to the limits of their climatic tolerances. Despite these limitations, provenance trials have generated a wealth of knowledge about local adaptation within tree species, particularly for widespread, economically important species used in reforestation (Morgenstern 1996). They also provide invaluable information on the genetic basis of clinal patterns of phenotypic traits important to local adaptation. Knowledge of population differentiation and potentially important phenotype–environment correlations in quantitative traits, such as timing of bud burst, leaf morphology, and many others, can be used to design landscape genomic studies and to guide sampling strategies for forest management and restoration. Well-designed short-term seedling provenance trials in controlled climate growth chambers can supplement information from field trials by exposing populations to climatic extremes, and to sets of climatic treatments that differentiate temperature, moisture and photoperiodic effects.

Both long- and short-term seedling common gardens can have population and landscape-genomic applications that were never envisioned or were not feasible at the time they were established. For example, a range-wide short-term common garden experiment was established in 2002 for quantifying phenotypic variation among populations and for assessing neutral population structure with microsatellite markers in *Picea sitchensis* (Mimura and Aitken 2007; 2010). As technologies evolved, these genetic materials and phenotypes were used to assess population differences in gene expression to identify candidate genes for bud phenology and cold hardiness (Holliday et al. 2008), to test for signatures of selection and identify SNPs in those candidate genes (Holliday et al. 2010b), and then to associate SNPs with adaptive phenotypic variation (Holliday et al. 2010a). This common garden also illustrates a common sampling issue with many existing trials: the correlations among geographic distance, climatic distance, and shared population recolonization history made it difficult to adjust for population structure while retaining signatures of adaptation. In this situation, adjustments for population structure designed to avoid false positives for testing SNPs for signatures of local adaptation can lead to false-negatives and underestimate selection. However, in many cases, these approaches are yielding meaningful results (see Eckert et al. 2012).

#### Identifying loci under divergent selection

Barrett and Hoekstra (2011) describe two general approaches that can be used to find genes underlying complex adaptive

traits. The “top-down” approach analyzes genotypic and phenotypic variation to find marker loci that co-segregate with phenotypic variation for putatively locally adaptive traits, either using quantitative trait loci (QTL) mapping or association mapping of individuals in common gardens (ecological genomics; Fig. 1). This can be done using relatively few markers for specific candidate gene loci, or using a whole-genome scan with many markers. QTL mapping was a more commonly used approach before whole-genome scan technology was widely available.

QTL mapping uses a pedigreed mapping population usually comprising full-sib progeny arrays from known parents to identify genomic regions containing loci associated with phenotypic variation. It relies on linkage disequilibrium between markers and causal loci, and thus a given QTL trait may be population or pedigree specific. This approach has been quite widely used in forest trees to search for genes controlling complex traits such as growth, phenology, disease resistance, cold hardiness, drought tolerance, and wood properties (Table 6-5, Nichols and Neale 2010), and extensive population resources have been developed for some species. In some cases, these resources have been put into clonal archives where researchers can access these materials for future investigations (<http://dendrome.ucdavis.edu/ftgsc>). For example, QTLs for bud phenology and cold hardiness have been found in *Quercus* spp. (Derory et al. 2010; Gailing et al. 2009), *Pinus sylvestris* (Hurme et al. 1997; 2000) and *Pseudotsuga menziesii* (Jermstad et al. 2001a, b; 2003).

The other top-down approach is genetic association mapping, where the genomic mapping resolution is much greater than QTL mapping due to low linkage disequilibrium in natural population samples compared to pedigreed progeny arrays (e.g., Eveno et al. 2008; Ingvarsson et al. 2008; Grivet et al. 2011; Pelgas et al. 2011; Prunier et al. 2011; Chen et al. 2012). Association mapping can be conducted using a single, un-pedigreed mapping population, or using individuals sampled from multiple populations and adjusted for population structure. It can focus on SNPs within candidate genes, or can use a whole-genome scan approach. This ecological genomic approach (see Fig. 1) has been used to identify candidate gene associations with bud phenology and cold hardiness in *P. menziesii* (Eckert et al. 2009a) and *P. sitchensis* (Holliday et al. 2010a).

The alternative to the “top-down” approach is of course the “bottom-up” approach, requiring genotypes but not phenotypes (e.g., Gonzalez-Martinez et al. 2006; Ingvarsson et al. 2008; Eveno et al. 2008; Grivet et al. 2011; Pelgas et al. 2011; Prunier et al. 2011; Chen et al. 2012). Here, marker loci (usually SNPs) are genotyped in individuals from different populations or environments, and tested for patterns of variation that indicate they are under selection (population genomics; Fig. 1). The loci tested in a bottom-up approach can either originate from candidate genes or from a whole-genome scan. Multiple methods

have been developed for this type of analysis and have all been applied to tree populations. Examples include methods based on the site-frequency spectrum (e.g., Eckert et al. 2009b), methods based on correlations with environmental variables (e.g., Eckert et al. 2010b), and those based on allele frequency differences among populations (e.g., Prunier et al. 2011). The most popular of these for the study of local adaptation in forest trees are methods that identify outliers with respect to population structure (i.e.,  $F_{st}$  outlier analysis).  $F_{st}$  outlier analyses to identify loci showing strong differentiation across populations do not require spatially explicit data because they do not associate allele frequencies with environmental gradients, but they can identify loci under selection when compared either to the entire genome or to a set of non-candidate markers. Significant outliers may indicate divergent selection and a role in local adaptation as well as purifying selection for those loci that are less differentiated than expected (e.g., Beaumont and Nichols 1996; Bonin et al. 2007; Joost et al. 2007). However, significant outliers are not always reliable indicators of selection and demographic history can generate a large number of false positives. A landscape genomic approach based on association between SNPs and environmental gradients can be another source of evidence for the impact of selection. This environmental association analysis is particularly powerful when a large sample of SNPs included in the design so that it is possible to separate out the genome-wide impact of demographic history from local selection on individual loci.

The advantage of the bottom-up approach is that costly and time-consuming common garden experiments are not needed, and no prior knowledge of phenotypic traits involved in local adaptation is required. One disadvantage is that environmental association analyses do not necessarily relate a variant under selection to a phenotype. A second disadvantage is that most phenotypic traits involved in local adaptation are highly polygenic, with phenotypic variation resulting from several to many genes. Kremer and Le Corre (2012) recently showed through simulations that if phenotypic traits involve more than a few genes, the underlying loci may show only modest levels of population differentiation ( $F_{st}$ ), similar to that of selectively neutral loci, even when there is strong population differentiation for the phenotypic trait ( $Q_{st}$ ). Nonetheless, environmental association analysis does not necessarily require large allele frequency difference among populations when numerous populations are sampled across a gradient (see Eckert et al. 2010b).

Combined approaches that utilize information on population phenotypes, genotypes, and environments will be the most informative for understanding the genetic and genomic basis of local adaptation to forest trees, the environmental characteristics that drive divergent selection, and the phenotypic traits that confer fitness to those environments (Anderson et al. 2011, Barrett and Hoekstra 2011). The various methods used in top-down and bottom-up approaches mentioned above have different sensitivities to the detection of loci under selection partly due to differences in the genomic

architecture of adaptive traits (Kremer and Le Corre 2012) and also due to differences across loci in the timing and strength of selection and overall differences in the background level of differentiation across loci (Le Corre and Kremer 2012). One way to approach these discrepancies is to use the bottom-up approach first and the top-down approach second, as illustrated in Eckert et al. (Eckert et al. 2009a, b; see Box 1, Fig. 2).

**Box 1: Loblolly pine (*Pinus taeda* L.) as a case study in determining patterns of adaptive variation across a landscape**

A series of studies in *Pinus taeda* illustrate how a multi-faceted approach can be used to discover the genes underlying complex adaptive traits and determine patterns of adaptive variation in these genes across the natural range of loblolly pine (Fig. 2). These studies were done over a 5-year period (2006–2011) and all employed first-generation Sanger sequencing technology. To begin, the “bottom-up” approach was used to discover SNPs and estimate departures from selective neutrality in a small number of candidate genes thought to be involved in water-use efficiency (Gonzalez-Martinez et al. 2006). Statistical tests for departure from selective neutrality suggested that two genes might be under selection, one of which was annotated as the early-response-drought gene (*erd3*). Next, the “top-down” association approach was used to test for associations between alleles of these genes and water-use efficiency, as measured by carbon isotope ratios (Gonzalez-Martinez et al. 2008). In this case, five genes were associated with water-use efficiency, including two dehydrin genes, but not the *erd3* gene. These results affirm the recommendation of Barrett and Hoekstra (2011) that multiple approaches are really essential. These two studies, which were essentially repeated a few years later using a very large set of candidate genes (Cumbie et al. 2011), illustrate the emerging genomic studies to link specific genes to phenotypes under selection.

Following these “top-down” and “bottom-up” studies, two landscape genomic studies were conducted to search for patterns of adaptation in genes across the landscape. Two different approaches were used and in both methods correction for association due solely to demographic effects was implemented. In the first, an association approach was used where an index for environmental aridity was associated with genotype (Eckert et al. 2010b). Five loci associated with aridity were found in this study. In the second approach, a Bayesian approach was used to search for association among candidate gene genotypes and a suite of geographic and environmental factors, including aridity (Eckert et al. 2010a). This approach revealed yet another set of genes potentially responding to selection and underlying local adaptation.

The studies describe a pluralistic approach that might be used to discover genes underlying complex adaptive traits in forest trees and identify patterns of local adaptation across heterogeneous landscapes. Next-generation sequencing technologies will allow such studies to be conducted at a genome-wide level with 1,000s or even 10,000s of geo-referenced samples. With careful sampling designs, studies in forest trees are well positioned to shed new light on the molecular basis of local adaptation in plants (Eckert and Dyer 2012).

with environmental factors than do the purely genotypic and phenotypic methods described above (see landscape genomics; Fig. 1). One approach is to genotype individuals from populations across a heterogeneous landscape for SNPs within loci that are identified to functional gene sequences, followed by an outlier analysis of correlation probabilities between these individual SNPs and a specific environmental factor (e.g., Coop et al. 2010; Hancock et al. 2008), referred to as an environmental association analysis or environmental outlier analysis. Ideally, an independent set of hundreds or thousands of non-candidate random or putatively neutral SNPs can be used to control for the background demographic history experienced by the entire genome. The method of Coop et al. (2010) called *Bayenv* uses a Bayesian Markov chain Monte Carlo analysis to generate an initial null model of how SNP frequencies vary across sites and then tests whether correlations of candidate gene SNPs associated with particular environmental variables are stronger than expected under a null model. Eckert et al. (2010a) successfully used *Bayenv* to identify specific individual SNPs and their association with individual climate variables, which requires adjustment for the numerous multiple correlation tests. Several studies introduce techniques to conduct environmental association analyses (e.g., Eckert et al. 2010a; Foll and Gaggiotti 2006; Fumagalli et al. 2011; Joost et al. 2007), but the Kang et al. (2010) method is probably the most robust and versatile for a variety of sample designs. However, all of them share the same limitation of testing a large number of SNPs and climate variables one at a time for phenotypic traits that are often polygenic and for which epistatic interactions may be more important (Holliday et al. 2012; Le Corre and Kremer 2012).

We propose that landscape genomic studies should consider applying multivariate statistical approaches (e.g., Kremer et al. 1997, Grivet et al. 2008) to genome-wide sampled loci, following an earlier tradition of forest genetics that deployed multivariate statistics to identify seed zones (Westfall and Conkle 1992). Community ecologists have also used such methods to analyze species diversity patterns because the data are highly multidimensional or multivariate and their variance is expected to display spatial structures across landscapes as species respond to both neutral and adaptive processes.

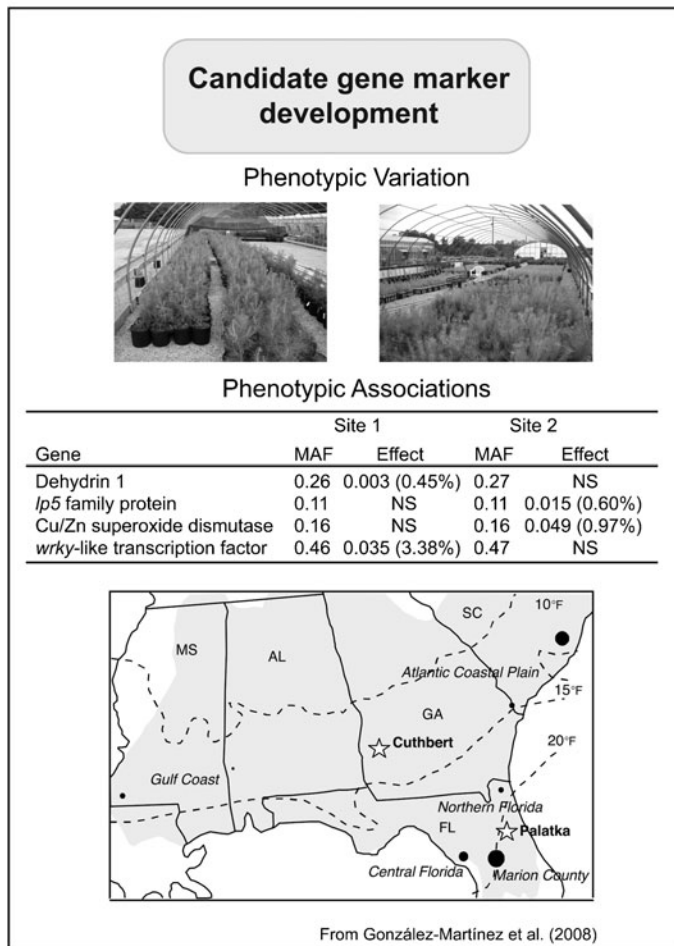
Multivariate statistical models offer several advantages. First, multivariate statistics can simultaneously examine multiple loci without blurring the effects of individual genes (Lasky et al. 2012; Mosca et al. 2012). These multivariate tests can be used without assuming multivariate normality. Second, it is possible to apply a global test of significance to multivariate data without the need for any correction due to multiple tests, which is required when conducting multiple tests of SNPs to get correct experiment-wise error rate. If the

## Statistical methods in landscape genomics

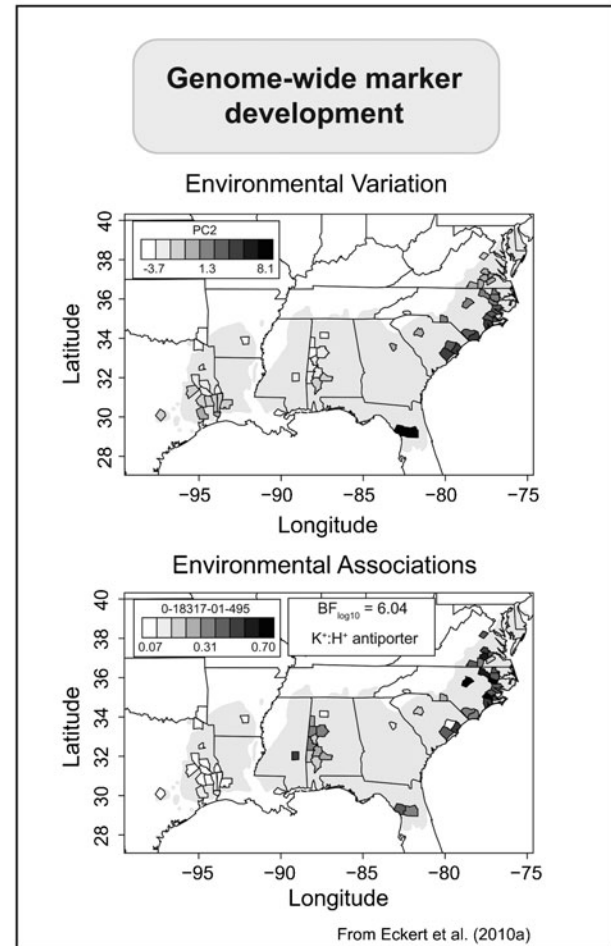
Landscape genomic methods offer powerful approaches for identifying the geographic mosaic of adaptation correlated



## Bottom-Up Approach



## Top-Down Approach



**Fig. 2** The bottom-up (*left*) and top-down (*right*) approaches for identifying adaptive genetic variation. The bottom-up approach starts with the definition of putatively adaptive genes using model plant species and then progresses to link variant alleles of these genes with phenotypes or the environment. Here an example from loblolly pine (*Pinus taeda* L.) is used to illustrate the point for carbon isotope ratios measured within two common gardens (e.g., Gonzalez-Martinez et al. 2008; see also Cumbie et al. 2011). The top-down approach starts with genome-wide marker sets and aims to identify adaptive genetic variation from a large

sample of random markers. An example from loblolly pine is used to illustrate linking environmental variation (e.g., a principal component (PC) based on climate variables) with genetic variation for thousands of single-nucleotide polymorphisms (SNPs) in an effort to identify those overly associated with the environment (e.g., Eckert et al. 2010a). The difference between these approaches, however, lies in the a priori specification of putatively adaptive genes and not with the type of association analysis (i.e., phenotypes versus environment)

overall test is significant and a sufficient number of adaptive SNPs are included, one can proceed to assessing the effects of individual genes (with or without a formal testing procedure). Finally, multivariate methods provide useful summary graphs that allow the visualization of results for multiple genes and multiple environmental variables and help in their interpretation. We anticipate that future landscape genomic studies will utilize multivariate tools, but these studies will also need to design their data collection and models so that they control for spatial autocorrelation of genotypes due to isolation by distance and shared demographic history (Vasemagi and Primmer 2005).

Among the methods of multivariate analysis, a first category comprises descriptive methods, like ordination (e.g., principal component analysis, PCA) and clustering. Ordination methods plot the observations along the main axes of variation (gradients) of the multivariate data, allowing researchers to see how the populations sampled are distributed along these gradients and how the different genes are related to these gradients. Clustering or partitioning methods produce coarser models of the data that are divided into groups under the hypothesis that discontinuities exist among the groups. These statistical methods are often applied to survey data as a first contact to learn something



about them, like the main ordination trends and the loadings in PCA. With a large number of loci, PCA of genetic data can provide detailed maps that correspond to geographic and environmental features.

A second important category of statistical methods, which should also prove important in landscape genomics, contains methods to analyze the spatial structure of multivariate response data. New methods have recently been proposed to learn about the multi-scale spatial structure of the data, and derived methods such as multi-scale ordination (Wagner 2003; Wagner 2004) and multi-scale variation partitioning (Peres-Neto et al. 2006). These methods that include spatial eigenfunction analyses, such as principle coordinates of neighbor matrices, Moran's eigenvector maps (MEM) and asymmetric eigenvector maps (Blanchet et al. 2008; Borcard and Legendre 2002; Borcard et al. 2004; Dray et al. 2006; Griffith and Peres-Neto 2006) have recently been reviewed by Legendre and Legendre (2012, Chapter 14). One example of this approach is a landscape genetic study of *Arabis alpina* that associated AFLP allele frequencies of populations with climate variables controlling for geographic structure (Manel et al. 2010b; 2012). By incorporating broad-scale MEM as explanatory variables that serve as proxies for unmeasured environmental variables, the authors found that the same climate factors, temperature and precipitation, were associated with AFLP allele distributions across 13 alpine species. These spatial ecology methods hold promise for future modeling with SNPs.

A third category contains methods to analyze the relationships between multivariate genetic data and multivariate environmental variables, such as canonical analysis (or constrained ordination). Sork et al. (2010) used canonical correlation analysis of *Quercus lobata* data in California to show that gradients in climate variables were associated with multivariate genetic data even after controlling for spatial autocorrelation due to location, thus providing evidence that selection is shaping the genetic structure of populations and identifying which climate variables are most correlated. The CCA method, which assumes no cause and effect between genetic and environmental variables, can control for demographic effects, if spatial autocorrelation is included in the model design. Redundancy analysis (RDA) and partial RDA analyses are alternative and robust approaches that can control for spatial effects while analyzing others. RDA which is analogous to a multivariate regression analysis, has been applied to the analysis of gene expression data in mussels (Boutet et al. 2009), to association analysis of SNPs in populations of *Arabidopsis thaliana* throughout Europe with climate variation (Lasky et al. 2012), and to multivariate genomic and transcriptomic data in trees (Sork et al. unpublished MS, Gugger, Cokus, and Sork unpublished data). It is possible to carry out correct tests of significance of the response to environmental data in the presence of spatial correlation (Peres-Neto and Legendre 2010).

One statistical challenge in clinal analysis of genetic variation is how to best sample the landscape in order to separate the effects of selectively neutral processes, including gene flow and demographic history, from selection. This challenge is present whether variation is in the form of phenotypic data from common gardens or molecular marker data, as mentioned above. These effects are dependent on: (1) physical distance, (2) differential selection, which is dependent on environmental distance, (3) patterns of phenotypic variation among populations, and (4) on quantitative trait variation within populations (Savolainen et al. 2007; Yeaman and Jarvis 2006). For landscape clinal analyses, the number of localities sampled is essential. Including samples from many locations with varying geographic and climate distances among them facilitates the identification of genes under positive selection and specific SNP alleles associated with phenotypes (St. Clair et al. 2005). In contrast to some traditional forest genetic approaches that involve sampling tens of individuals per location and few locations, allowing for quantification of within- as well as among-population variation based on allele frequencies (e.g., Coop et al. 2010), some landscape genomic and clinal analyses are powerful when using seeds from just one or two parents per location and sampling a larger number of localities (e.g., Sork et al. 2010, St. Clair et al. 2005). This sampling approach has been used effectively in forest tree genecological studies (see, Campbell 1979, Langlet 1971). In sum, any sampling strategies that reduce correlations between neutral genetic distance resulting from demographic history, and environmental distance will be less subject to false positive signals of adaptation.

#### Tree phenotypic response to climate change

Globally, tree populations are facing an enormous threat from rapid climate change, as their long generation time limits their capacity to adapt readily putting them at particular risk (Aitken et al. 2008). Landscape genomics has the potential to quantify the capacity for populations to adapt to new climatic conditions, to identify geographic regions where tree populations may not have sufficient diversity for rapid adaptation, and to guide seed choices for reforestation in areas likely to experience warmer climates than currently exist. Studies are needed where population differentiation, clinal analyses of environmental data, and marker-phenotype associations can be analyzed using phenotypic data from multiple sites or environments.

Quantifying phenotypic population response functions, also known as norms of reaction, to an environmental factor such as temperature or moisture requires multiple common garden test environments, including some that exceed population tolerances (Aitken et al. 2008). Where adequate field-based provenance trials are not already established,

short-term common garden experiments in controlled climate environments testing responses to varying temperature and moisture regimes can rapidly provide the necessary data. Such data allow a robust assessment of genotype-by-environment interaction, phenotypic plasticity, and population responses. This approach is particularly important for predicting the effects of climate change on populations because response functions predict the performance of a single population across a range of climatic conditions. The recently developed universal response function (URF) combines data from multiple environments for multiple provenances, and quantifies both the genetic (provenance, reflecting past selection) and environmental (site, reflecting phenotypic plasticity) components on phenotypes (Wang et al. 2010). By combining observations across provenances and sites, or even between experiments, the URF approach requires smaller sample sizes than traditional response function analyses. Future work will be needed to integrate provenance analyses with genomic data to identify causal variants to genotype–environment interactions (e.g., Crossa et al. 2010). Results from *A. thaliana* suggest that some loci may be adaptive in some environments and selectively neutral in others (Fournier-Level et al. 2011). In the end, a combination of landscape genomic analyses of adaptive genetic and phenotypic variation will be necessary to understand the ability of natural tree populations to adapt to rapid climate change. This integration will allow for a rigorous assessment of the need for human interventions such as assisted migration to relocate populations within existing species ranges, or to assist the range expansion of species.

## Summary and conclusions

Landscape genomic approaches provide valuable tools for understanding how selection in natural forest tree populations is shaping local adaptation, while controlling for the potentially confounding historical impacts of gene flow and genetic drift. By using the latest high throughput genome sequencing tools to genotype a large number of SNPs across the genome, and associating those markers with climatic and other environmental variables, it is possible to begin to sort out adaptive from neutral variation. Moreover, the growing availability of genomic resources developed for tree species allows us to identify specific loci associated with traits that underlie plant fitness. These resources for some trees include a completely sequenced and annotated genome, while other species have a growing number of EST libraries for functional genes. Thus, landscape genomic methods allow us to test for correlations between climate and SNP frequencies across landscapes, and describe geographic patterns of adaptive genetic variation using existing or new genomic resources. For many forest trees, provenance common gardens at multiple sites will allow complementary information on how genes identified through landscape genomic analysis contribute to phenotypes, while

for others, new short- or long-term common gardens containing more strategically sampled material will be needed. All this genetic and environmental information provides a valuable foundation for analyzing how tree populations might respond to climate change. Thus, landscape genomics is a useful tool for understanding the extent to which climate has shaped geographic genetic structure in the past, and for identifying forest populations that may be at risk with the predicted rapid future climate change.

## Recommendations for future research directions and priorities

1. Successful application of landscape genomics research in forest trees is dependent on easy access to well-curated and comprehensive data resources. While published genomic data are usually curated and publically available, the same tradition does not exist for the vast phenotypic data collected from provenance trials or for the corresponding environmental data. Thus, a priority area must be funding for and required use of databases that capture and curate genomic, phenotypic and environmental data for a large number of species.
2. Scientific collaborations between those focused on phenotypic studies and those involved in population and landscape genomics should be increased. The former tend to be conducted by applied resource management agencies, while the latter are concentrated in academic institutions. The archiving of tissue samples (e.g., desiccated leaf samples stored at room temperature) from individuals being phenotyped in short- and long-term provenance common gardens would greatly facilitate future landscape genomic studies at relatively low cost. Data from a broader range of species with different life histories is also needed as the existing literature is dominated by commercially important, wide-ranging, abundant wind-pollinated and wind-dispersed temperate and boreal species (e.g., widespread conifers).
3. Increased support for development of bioinformatic tools and interfaces to facilitate easy access to and foster broad usage of the database resources. Greater attention to appropriate spatial scales between genetic (sampling) data and downscaled climatic data is also needed, which may require methodological innovations in down-scaling the climate data to facilitate the appropriate sampling design.
4. Future research in landscape genomics needs to integrate geospatial genetic analysis with traditional population genetic theory. Statistical methods are available to control for background genetic structure, but better sampling design and improved analytical methods will make it easier to separate effects of demographic history and migration from the effects of selection on genetic variation in continuously distributed forest populations,

which typically have high gene flow and weak population structure across large environmental gradients.

5. Future landscape genomic analysis and clinal analyses should take advantage of existing spatial ecological analytical methods. Multivariate statistical models will offer a useful tool for understanding the role of multiple loci and factors, and they do not require multivariate normality. One advantage is that they can provide a global test of significance prior to tests of individual genes that will indicate whether a structure or relationship is present in the multivariate data without the need for any correction. If the overall test is significant, one can proceed to assessing the effect of individual genes (with or without a formal testing procedure).
6. We recommend that future work on response to climate change translate existing geospatial genetic analyses into predictions of adaptive responses that are applicable to forest genetic resource conservation and management for future climates.

## References

- Aitken SN, Yeaman S, Holliday JA, Wang TL, Curtis-McLane S (2008) Adaptation, migration or extirpation: climate change outcomes for tree populations. *Evol Appl* 1(1):95–111
- Anderson JT, Willis JH, Mitchell-Olds T (2011) Evolutionary genetics of plant adaptation. *Trends in Genetics* 27(7):258–266
- Atwell S, Huang YS, Vilhjalmsen BJ et al (2010) Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. *Nature* 465(7298):627–631. doi:10.1038/nature08800
- Balkenhol N, Gugerli F, Cushman SA, Waits LP, Coulon A, Arntzen JW, Holderegger R, Wagner HH (2009a) Identifying future research needs in landscape genetics: where to from here? *Landsc Ecol* 24(4):455–463
- Balkenhol N, Waits LP, Dezzani RJ (2009b) Statistical approaches in landscape genetics: an evaluation of methods for linking landscape and genetic data. *Ecography* 32(5):818–830
- Barrett RDH, Hoekstra HE (2011) Molecular spandrels: tests of adaptation at the genetic level. *Nat Rev Genet* 12(1):767–780
- Beaumont MA, Nichols RA (1996) Evaluating loci for use in the genetic analysis of population structure. *Proc R Soc London, Ser B* 263(1377):1619–1626
- Bergelson J, Roux F (2010) Towards identifying genes underlying ecologically relevant traits in *Arabidopsis thaliana*. *Nat Rev Genet* 11(12):867–879. doi:10.1038/nrg2896
- Blanchet FG, Legendre P, Borcard D (2008) Forward selection of explanatory variables. *Ecology* 89(9):2623–2632. doi:10.1890/07-0986.1
- Bonin A, Nicole F, Pompanon F, Miaud C, Taberlet P (2007) Population adaptive index: a new method to help measure intraspecific genetic diversity and prioritize populations for conservation. *Conserv Biol* 21(3):697–708. doi:10.1111/j.1523-1739.2007.00685
- Borcard D, Legendre P (2002) All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. *Ecol Model* 153(1–2):51–68. doi:10.1016/S0304-3800(01)00501-4
- Borcard D, Legendre P, Avois-Jacquet C, Tuomisto H (2004) Dissecting the spatial structure of ecological data at multiple scales. *Ecology* 85(7):1826–1832. doi:10.1890/03-3111
- Boutet I, Tanguy A, Guen DL, Piccino P, Hourdez S, Legendre P, Jollivet D (2009) Global depression in gene expression as a response to rapid thermal changes in vent mussels. *Proceedings of the Royal Society B-Biological Sciences* 276:3071–3079
- Bradshaw AD (1972) Some of the evolutionary consequences of being a plant. *Evol Biol* 5:25–47
- Brunner AM, Busov VB, Strauss SH (2004) Poplar genome sequence: functional genomics in an ecologically dominant plant species. *Trends Plant Sci* 9(1):49–56. doi:10.1016/j.plant.2003.11.006
- Campbell RK (1979) Genecology of Douglas-fir in a watershed in the Oregon Cascades. *Ecology* 60(5):1036–1050
- Chen J, Kallman XM, Gyllenstrand N, Zaina G, Morgante M, Bousquet J, Eckert A, Wegrzyn J, Neale D, Lagercrantz U, Lascoux M (2012) Disentangling the roles of history and local selection in shaping clinal variation of allele frequencies and gene expression in Norway spruce (*Picea abies*). *Genetics* 191:865–881
- Coop G, Witonsky D, Di Rienzo A, Pritchard JK (2010) Using environmental correlations to identify loci underlying local adaptation. *Genetics* 185(4):1411–1423. doi:10.1534/genetics.110.114819
- Crossa J, de los Campos G, Perez P et al (2010) Prediction of genetic values of quantitative traits in plant breeding using pedigree and molecular markers. *Genetics* 186(2):713–U406. doi:10.1534/Genetics.110.118521
- Cumbie WP, Eckert A, Wegrzyn J, Whetten R, Neale D, Goldfarb B (2011) Association genetics of carbon isotope discrimination, height and foliar nitrogen in a natural population of *Pinus taeda* L. *Heredity* 107(2):105–114. doi:10.1038/hdy.2010.168
- Cushman SA, Landguth EL (2010) Spurious correlations and inference in landscape genetics. *Mol Ecol* 19(17):3596–3602
- Davey JW, Hohenlohe PA, Etter PD, Boone JQ, Catchen JM, Blaxter ML (2011) Genomewide genetic marker discovery and genotyping using next-generation sequencing. *Nat Rev Genet* 12:499–510
- Derory J, Scotti-Saintagne C, Bertocchi E, Le Dantec L, Gaignic N, Jauffres A, Casasoli M, Chancerel E, Bodenes C, Alberto F, Kremer A (2010) Contrasting relations between diversity of candidate genes and variation of bud burst in natural and segregating populations of European oaks. *Heredity* 105(4):401–411. doi:10.1038/hdy.2009.170
- Dray S, Legendre P, Peres-Neto PR (2006) Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). *Ecological Modeling* 196:483–493
- Eckert AJ, Dyer RJ (2012) Defining the landscape of adaptive genetic diversity. *Mol Ecol* 21(12):2836–2838. doi:10.1111/j.1365-294X.2012.05615.x
- Eckert AJ, Bower AD, Wegrzyn JL, Pande B, Jermstad KD, Krutovsky KV, Clair JBS, Neale DB (2009a) Association genetics of coastal Douglas fir (*Pseudotsuga menziesii* var. *menziesii*, Pinaceae). I. Cold-hardiness related traits. *Genetics* 182(4):1289–1302. doi:10.1534/Genetics.108.102350
- Eckert AJ, Wegrzyn JL, Pande B, Jermstad KD, Lee JM, Liechty JD, Tarse BR, Krutovsky KV, Neale DB (2009b) Multilocus patterns of nucleotide diversity and divergence reveal positive selection at candidate genes related to cold hardiness in coastal Douglas fir (*Pseudotsuga menziesii* var. *menziesii*). *Genetics* 183(1):289–298
- Eckert AJ, Bower AD, Gonzalez-Martinez SC, Wegrzyn JL, Coop G, Neale DB (2010a) Back to nature: ecological genomics of loblolly pine (*Pinus taeda*, Pinaceae). *Mol Ecol* 19(17):3789–3805
- Eckert AJ, van Heerwaarden J, Wegrzyn JL, Nelson CD, Ross-Ibarra J, Gonzalez-Martinez SC, Neale DB (2010b) Patterns of population structure and environmental associations to aridity across the range of loblolly pine (*Pinus taeda* L., Pinaceae). *Genetics* 185(3):969–982
- Eckert AJ, Shahi H, Datwyler SL, Neale DB (2012) Spatially variable natural selection and the divergence between parapatric subspecies



- of lodgepole pine (*Pinus contorta*, Pinaceae). *Am J Bot* 99(8):1323–1334
- Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One* 6(5):e19379
- Eveno E, Collada C, Guevara MA, Leger V, Soto A, Diaz L, Leger P, Gonzalez-Martinez SC, Cervera MT, Plomion C, Garnier-Gere PH (2008) Contrasting patterns of selection at *Pinus pinaster* Ait. drought stress candidate genes as revealed by genetic differentiation analyses. *Mol Biol Evol* 25:417–437
- Favre J, Brown S (1996) A flow cytometric evaluation of the nuclear DNA content and GC percent in genomes of European oak species. *Ann For Sci* 53:915–917
- Foll M, Gaggiotti OE (2006) Identifying the environmental factors that determine the genetic structure of populations. *Genetics* 174:875–891
- Fournier-Level A, Korte A, Cooper MD, Nordborg M, Schmitt J, Wilczek AM (2011) A map of local adaptation in *Arabidopsis thaliana*. *Science* 333(6052):86–89
- Fumagalli M, Sironi M, Pozzoli U, Ferrer-Admetlla A, Pattini L, Nielsen R (2011) Signatures of environmental genetic adaptation pinpoint pathogens as the main selective pressure through human evolution. *PLoS Genetics* 7(11):e1002355
- Gailing O, Vornam B, Leinemann L, Finkeldey R (2009) Genetic and genomic approaches to assess adaptive genetic variation in plants: forest trees as a model. *Physiol Plant* 137(4):509–519. doi:10.1111/j.1399-3054.2009.01263.x
- Gnirke A, Melnikov A, Maguire J et al (2009) Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. *Nat Biotechnol* 27:182–189
- Gonzalez-Martinez SC, Ersoz E, Brown GR, Wheeler NC, Neale DB (2006) DNA sequence variation and selection of tag single-nucleotide polymorphisms at candidate genes for drought-stress response in *Pinus taeda* L. *Genetics* 172(3):1915–1926. doi:10.1534/Genetics.105.047126
- Gonzalez-Martinez SC, Huber D, Ersoz E, Davis JM, Neale DB (2008) Association genetics in *Pinus taeda* L. II. Carbon isotope discrimination. *Heredity* 101(1):19–26. doi:10.1038/hdy.2008.21
- Griffith DA, Peres-Neto PR (2006) Spatial modelling in ecology: the flexibility of eigenfunction spatial analyses. *Ecology* 87:2603–2613
- Grivet D, Sork VL, Westfall RD, Davis FW (2008) Conserving the evolutionary potential of California valley oak (*Quercus lobata* Née): a multivariate genetic approach to conservation planning. *Mol Ecol* 17(1):139–156. doi:10.1111/j.1365-294x.2007.03498.x
- Grivet D, Sebastiani F, Alia R, Bataillon T, Torre S, Zabal-Aguirre M, Vendramin GG, Gonzalez-Martinez SC (2011) Molecular footprints of local adaptation in two Mediterranean conifers. *Mol Biol Evol* 28:101–116
- Hancock AM, Di Rienzo A (2008) Detecting the genetic signature of natural selection in human populations: models, methods, and data. *Annu Rev Anthropol* 37:197–217
- Hancock AM, Witonsky DB, Gordon AS, Eshel G, Pritchard JK, Coop G, Di Rienzo A (2008) Adaptations to climate in candidate genes for common metabolic disorders. *PLoS Genetics* 4(2):e32. doi:10.1371/journal.pgen.0040032
- Holderegger R, Wagner HH (2008) Landscape genetics. *Bioscience* 58(3):199–207
- Holderegger R, Kamm U, Gugerli F (2006) Adaptive vs. neutral genetic diversity: implications for landscape genetics. *Landsc Ecol* 21(6):797–807
- Holliday JA, Ralph SG, White R, Bohlmann J, Aitken SN (2008) Global monitoring of autumn gene expression within and among phenotypically divergent populations of Sitka spruce (*Picea sitchensis*). *New Phytol* 178(1):103–122
- Holliday JA, Ritland K, Aitken SN (2010a) Widespread, ecologically relevant genetic markers developed from association mapping of climate-related traits in Sitka spruce (*Picea sitchensis*). *New Phytol* 188(2):501–514. doi:10.1111/j.1469-8137.2010.03380
- Holliday JA, Yuen M, Ritland K, Aitken SN (2010b) Postglacial history of a widespread conifer produces inverse clines in selective neutrality tests. *Mol Ecol* 19(18):3857–3864. doi:10.1111/j.1365-294X.2010.04767
- Holliday JA, Wang T, Aitken S (2012) Predicting adaptive phenotypes from multilocus genotypes in Sitka spruce (*Picea sitchensis*) using Random Forest. *G3: Genes| Genomics| Genetics* 2:1085–1093. doi:10.1534/g3.112.002733
- Hurme P, Repo T, Savolainen O, Paakkonen T (1997) Climatic adaptation of bud set and frost hardiness in Scots pine (*Pinus sylvestris*). *Can J Forest Res* 27(5):716–723
- Hurme P, Sillanpää MJ, Arjas E, Repo T, Savolainen O (2000) Genetic basis of climatic adaptation in Scots pine by Bayesian quantitative trait locus analysis. *Genetics* 156(3):1309–1322
- Ingvarsson PK, Garcia MV, Luquez V, Hall D, Jansson S (2008) Nucleotide polymorphism and phenotypic associations with and around the phytochrome B2 locus in European aspen (*Populus tremula*, Salicaceae). *Genetics* 178:2217–2226
- Jermstad KD, Bassoni DL, Jech KS, Wheeler NC, Neale DB (2001a) Mapping of quantitative trait loci controlling adaptive traits in coastal Douglas-fir. timing of vegetative bud flush. *Theor Appl Genet* 102(8):1142–1151
- Jermstad KD, Bassoni DL, Wheeler NC, Anekonda TS, Aitken SN, Adams WT, Neale DB (2001b) Mapping of quantitative trait loci controlling adaptive traits in coastal Douglas-fir. II. Spring and fall cold-hardiness. *Theor Appl Genet* 102(8):1152–1158
- Jermstad KD, Bassoni DL, Jech KS, Ritchie GA, Wheeler NC, Neale DB (2003) Mapping of quantitative trait loci controlling adaptive traits in coastal Douglas fir. III. Quantitative trait loci-by-environment interactions. *Genetics* 165(3):1489–1506
- Joost S, Bonin A, Bruford MW, Despres L, Conord C, Erhardt G, Taberlet P (2007) A spatial analysis method (SAM) to detect candidate loci for selection: towards a landscape genomics approach to adaptation. *Mol Ecol* 16(18):3955–3969. doi:10.1111/j.1365-294X.2007.03442.x
- Kang HM, Sul JH, Service SK, Zaitlen NA, Kong S-y, Freimer NB, Sabatti C, Eskin E (2010) Variance component model to account for sample structure in genome-wide association studies. *Nat Genet* 42(4):348–354
- Kremer A, Le Corre V (2012) Decoupling of differentiation between traits and their underlying genes in response to divergent selection. *Heredity* 108(4):375–385. doi:10.1038/hdy.2011.81
- Kremer A, Zanetto A, Ducoussou A (1997) Multilocus and multitrait measures of differentiation for gene markers and phenotypic traits. *Genetics* 145(4):1229–1241
- Kremer A, Casasoli M, Barreneche T et al (2007) Fagaceae trees. In: Kole C (ed) Genome mapping and molecular breeding in plants, vol 7th edn, Forest trees. Springer, New York, p pp 161
- Langlet O (1971) Two hundred years of geneecology. *Taxon* 20:653–722
- Lasky JR, Des Marais DL, McKay JK, Richards JH, Juenger TE, Keitt TH (2012) Characterizing genomic variation of *Arabidopsis thaliana*: the roles of geography and climate. *Mol Ecol* 21:5512–5529. doi:10.1111/j.1365-294X.2012.05709.x
- Le Corre V, Kremer A (2012) The genetic differentiation at quantitative trait loci under local adaption. *Mol Ecol* 21:1548–1566
- Legendre P, Legendre L (2012) Numerical ecology, 3rd English edition developments in environmental modelling, vol 24. Elsevier, Amsterdam
- Mackay J, Dean JF, Plomion C, Peterson DG, Canovas FM, Pavy N, Ingvarsson PK, Savolainen O, Guevara MA, Fluch S, Vinceti B, Abarca D, Diaz-Sala C, Cervera MT (2012) Towards decoding

- the conifer Giga-genome. *Plant Molecular Biology* 80:555–569. doi:10.1007/s1103-012-9961-7
- Manel S, Schwartz MK, Luikart G, Taberlet P (2003) Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology & Evolution* 18(4):189–197
- Manel S, Joost S, Epperson BK, Holderegger R, Storer A, Rosenberg MS, Scribner KT, Bonin A, Fortin MJ (2010a) Perspectives on the use of landscape genetics to detect genetic adaptive variation in the field. *Mol Ecol* 19(17):3760–3772
- Manel S, Poncet BN, Legendre P, Gugerli F, Holderegger R (2010b) Common factors drive adaptive genetic variation at different spatial scales in *Arabidopsis thaliana*. *Mol Ecol* 19(17):3824–3835. doi:10.1111/j.1365-294X.2010.04716.x
- Manel S, Gugerli F, Thuiller W, Alvarez N, Legendre P, Holderegger R, Gielly L, Taberlet P (2012) Broad-scale adaptive genetic variation in Alpine plants is mainly driven by temperature and precipitation. *Mol Ecol* 19:3824–3835
- Miller MR, Dunham JP, Amores A, Cresko WA, Johnson EA (2007) Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome Res* 17(2):240–248. doi:10.1101/gr.5681207
- Mimura M, Aitken SN (2007) Adaptive gradients and isolation-by-distance with postglacial migration in *Picea sitchensis*. *Heredity* 99(2):224–232
- Mimura M, Aitken SN (2010) Local adaptation at the range peripheries of Sitka spruce. *J Evol Biol* 23(2):249–258
- Morgante M, De Paoli E (2011) Toward the conifer genome sequence. In: Plomion C, Bousquet J, Kole C (eds) *Genetics, Genomics and Breeding of Conifers*. Trees, pp 389–403. Edenbridge and CRC, New York
- Morgenstern EK (1996) Geographic variation in forest trees: genetic basis and application of knowledge in silviculture. Univ of British Columbia Press, Vancouver
- Moritsuka E, Hisataka Y, Tamura M, Uchiyama K, Watanabe A, Tsumura Y, Tachida H (2012) Extended linkage disequilibrium in noncoding regions in a conifer, *Cryptomeria japonica*. *Genetics* 190(3):1145. doi:10.1534/genetics.111.136697
- Mosca E, Eckert AJ, Di Pierro EA, Rocchini D, La Porta N, Belletti P, Neale DB (2012) The geographical and environmental determinants of genetic diversity for four alpine conifers of the European Alps. *Mol Ecol* 21:5530–5545. doi:10.1111/mec.12043
- Neale DB, Savolainen O (2004) Association genetics of complex traits in conifers. *Trends Plant Sci* 9(7):325–330. doi:10.1016/J.Plants.2004
- Nichols KM, Neale DB (2010) Association genetics, population genomics, and conservation: Revealing the genes underlying adaptation in natural populations of plants and animals. In: DeWoody JA, Bickham JW, Michler CH, Nichols KM, Olin E, Rhodes J, Woeste KE (eds) *Molecular approaches in natural resource conservation and management*. Cambridge University Press, New York, NY, pp 123–168
- Parchman TL, Gompert Z, Mudge J, Schilkey FD, Benkman CW, Buerkle CA (2012) Genome-wide association genetics of an adaptive trait in lodgepole pine. *Mol Ecol* 21(12):2991–3005. doi:10.1111/j.1365-294X.2012.05513.x
- Pelgas B, Bousquet J, Meirmans PG, Ritland K, Isabel N (2011) QTL mapping in white spruce: gene maps and genomic regions underlying adaptive traits across pedigrees, years and environments. *BMC Genomics* 12:145
- Peres-Neto PR, Legendre P (2010) Estimating and controlling for spatial structure in the study of ecological communities. *Glob Ecol Biogeogr* 19(2):174–184. doi:10.1111/j.1466-8238.2009.00506.x
- Peres-Neto PR, Legendre P, Dray S, Borcard D (2006) Variation partitioning of species data matrices: estimation and comparison of fractions. *Ecology* 87:2614–2625
- Petit RJ, Hampe A (2006) Some evolutionary consequences of being a tree. *Annu Rev Ecol Evol Syst* 37:187–214. doi:10.1146/annurev.ecolsys.37.091305.110215
- Prunier J, Laroche J, Beaulieu J, Bousquet J (2011) Scanning the genome for gene SNPs related to climate adaptation and estimating selection at the molecular level in boreal black spruce. *Mol Ecol* 20(8):1702–1716. doi:10.1111/j.1365-294X.2011.05045.x
- Savolainen O, Pyhajarvi T, Knurr T (2007) Gene flow and local adaptation in trees. *Annu Rev Ecol Evol Syst* 38:595–619
- Segelbacher G, Cushman SA, Epperson BK, Fortin M-J, Francois O, Hardy OJ, Holderegger R, Taberlet P, Waits LP, Manel S (2010) Applications of landscape genetics in conservation biology: concepts and challenges. *Conserv Genet* 11(2):375–385
- Sork VL, Waits L (2010) Contributions of landscape genetics—approaches, insights, and future potential. *Mol Ecol* 19(17, SI):3489–3495. doi:10.1111/j.1365-294X.2010.04786.x
- Sork VL, Davis FW, Westfall R, Flint A, Ikegami M, Wang HF, Grivet D (2010) Gene movement and genetic association with regional climate gradients in California valley oak (*Quercus lobata* Née) in the face of climate change. *Mol Ecol* 19(17):3806–3823. doi:10.1111/j.1365-294X.2010.04726
- St. Clair JB, Mandel NL, Vance-Borland KW (2005) Genecology of Douglas fir in western Oregon and Washington. *Ann Bot* 96:1199–1214
- Storer A, Murphy MA, Evans JS, Goldberg CS, Robinson S, Spear SF, Dezzani R, Delmelle E, Vierling L, Waits LP (2007) Putting the 'landscape' in landscape genetics. *Heredity* 98(3):128–142
- Storer A, Murphy MA, Spear SF, Holderegger R, Waits LP (2010) Landscape genetics: where are we now? *Mol Ecol* 19(17):3496–3514
- Tuskan GA, DiFazio S, Jansson S et al (2006) The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* 313(5793):1596–1604. doi:10.1126/science.1128691
- Vasemagi A, Primmer CR (2005) Challenges for identifying functionally important genetic variation: the promise of combining complementary research strategies. *Mol Ecol* 14(12):3623–3642
- Wagner HH (2003) Spatial covariance in plant communities: Integrating ordination, geostatistics, and variance testing. *Ecology* 84(4):1045–1057. doi:10.1890/0012-9658(2003)084[1045:scipci]2.0.co;2
- Wagner HH (2004) Direct multi-scale ordination with canonical correspondence analysis. *Ecology* 85(2):342–351. doi:10.1890/02-0738
- Wang TL, O'Neill GA, Aitken SN (2010) Integrating environmental and genetic effects to predict responses of tree populations to climate. *Ecol Appl* 20(1):153–163
- Westfall RD, Conkle MT (1992) Allozyme markers in breeding zone 1050 designation. *New Forests* 6:279–309
- Yeaman S, Jarvis A (2006) Regional heterogeneity and gene flow maintain variance in a quantitative trait within populations of lodgepole pine. *Proceedings of the Royal Society B-Biological Sciences* 273(1594):1587–1593. doi:10.1098/rspb.2006.3498
- Zoldos V, Papes D, Brown S, Panaud O, Siljak-Yakovlev S (1998) Genome size and base composition of seven *Quercus* species: inter- and intra-population variation. *Genome* 41:162–168