CHAPTER EIGHT

Molecular Signatures of Adaptation and Selection in Forest Trees

Juan P. Jaramillo-Correa[*,¶,1], **Julien Prunier**[§], **Alejandra Vázquez-Lobo**[*,†], **Stephen R. Keller**[¶], **Alejandra Moreno-Letelier**[*,||]

[*]Department of Evolutionary Ecology, Institute of Ecology, Universidad Nacional Autónoma de México, México City, Mexico
[§]Department of Forestry, Centre for Forest Research and Institute for Systems and Integrative Biology, Université Laval, Québec, QC, Canada
[†]Research Centre for Biodiversity and Conservation, Universidad Autónoma del Estado de Morelos, Cuernavaca, Mexico
[¶]Department of Plant Biology, University of Vermont, Burlington, VT, USA
[||]Botanical Garden, Institute of Biology, Universidad Nacional Autónoma de México, México City, Mexico
[1]Corresponding author: E-mail: jaramillo@ecologia.unam.mx

Contents

Abstract

Uncovering the genes and molecular basis of phenotypic variation and adaptation is a major goal in conservation and evolutionary genetics; it also sets the basis for future operational breeding in commercial species, like forest trees. These taxa are characterized by their large size, growth habit and longevity, which hampers the use of reverse-genetic approaches (i.e. from gene function to phenotype) to pinpoint adaptive molecular variants. In this chapter, we summarize the basis of the forward-genetic approaches (i.e. from phenotype to gene function) currently used in forest trees. For each strategy, we provide a brief overview of the statistical approaches employed to identify candidate genes, and then highlight the main findings of landmark studies that provide

Advances in Botanical Research, Volume 74
ISSN 0065-2296
http://dx.doi.org/10.1016/bs.abr.2015.04.003

evidence for adaptation in forest trees. Adaptive and commercial traits are generally well inherited in trees, although they are mostly affected by the variation of multiple genes, each one accounting for a small part of the phenotypic variance of each character. However, some individual and important genes involved in growth, phenology, drought resistance and cold hardiness have been identified; many of them showing evidence of selection across multiple taxa (sometimes including angiosperms and gymnosperms). Future challenges for detecting the signatures and understanding the molecular basis of adaptation in trees include more adequate and precise phenotype assessment in natural populations, and the inclusion of gene interactions and epigenetic variations in current models. The implications of these findings in conservation and breeding of forest trees are finally discussed.

1. INTRODUCTION

Trees are woody vascular plants, angiosperms or gymnosperms that share attributes like large size, growth habit, longevity and high reproductive outcome (reviewed by Petit and Hampe (2006)). Although they do not belong to a unique phylogenetic group, they do conform to a particular life form that has independently arisen at different geological times from non-tree ancestors by convergent evolution across unrelated clades (reviewed by Niklas (1997), Groover (2005)). Individuals sharing this life form can achieve dominance in a variety of terrestrial environments, forming the base of entire ecosystems, which are often characterized by a tremendous biomass accumulation and productivity (Petit & Hampe, 2006; Whiteman, 2014). For such reason, trees have become the base of a whole industry relying on wood extraction, fruit production or the exploitation of other nonwood products such as latex for rubber manufacturing, metabolites with medical purposes (e.g. taxol) or oils for biofuel, among many others (reviewed by Whiteman (2014)). While often seen as undomesticated species, forest trees have been submitted to genetic improvement through selective breeding for over 50 years (Neale & Savolainen, 2004). More recently, they have been the subject of genomic-based studies aiming to understand the architecture of traits with commercial and adaptive interest, such as wood quality, cold hardiness or pest resistance (Neale & Kremer, 2011). Some of these studies have focused on the capacity of forest trees for rapid physiological and molecular adaptation at different temporal and spatial scales, and on their ability to undergo rapid range shifts in response to environmental changes (Savolainen, Pyhäjärvi, & Knurr, 2007).

While angiosperm and coniferous trees share many attributes that affect their evolutionary trajectories (Petit & Hampe, 2006), they also differ in

several fundamentally important ways that are relevant to the search for the genetic basis of adaptively important traits. Several trait differences separate angiosperm from conifer trees, including their reproductive biology (e.g. flowering and fruiting vs cone production), pollination syndromes, their perennial or deciduous habits, as well as wood architectural traits such as the composition of lignin and the ratio of cellulose/hemicellulose, as well as the use of tracheids and vessels for water transport (Niklas, 1997). Some of these traits have begun to be addressed by association genetics studies in both gymnosperm and angiosperm trees by using different strategies (e.g. whole genome vs candidate gene studies), while others still await comparative genomic data. From a strictly genomic point of view, the most striking difference between gymnosperm and angiosperm trees is the size and content of their genomes, which has great relevance for performing studies aimed at discovering the genomic signatures of adaptation. For instance, conifers have huge and highly complex genomes (>20 Gb) characterized by the proliferation of repetitive sequences from transposable elements (e.g. Birol et al., 2013; Neale et al., 2014; Nystedt et al., 2013), while angiosperm trees generally have more compact and less redundant genomes (e.g. Myburg et al., 2014; Tuskan et al., 2006). Indeed, while conifer genomics has recently received a boost from published draft genome assemblies of Norway spruce (*Picea abies*; Nystedt et al., 2013), white spruce (*Picea glauca*; Birol et al., 2013), and loblolly pine (*Pinus taeda*; Neale et al., 2014), angiosperm tree genomics has benefited for nearly a decade from the first forest tree to have its whole genome sequenced and annotated – the black cottonwood, *Populus trichocarpa* (Salicaceae) (Tuskan et al., 2006).

Progress in forest tree genomics appears nevertheless limited when compared to that in model species such as *Arabidopsis* or maize. Advances in forest tree genomics are hampered by the long generation times of species and the lack of well-defined mutants for specific genes, which constrains the use of reverse-genetic approaches. Such strategies rely on the *prior* knowledge of gene functions and the study of their effects on the phenotype of mutants for those particular genes (but see Séguin, Lachance, Désjardins, Leplé, and Pilate (2014)). Instead, dissecting the genomic architecture of complex phenotypic traits in forest trees has principally relied on forward-genetic approaches (Figure 1), and it is following that logic that this chapter is organized. Forward genetics is based on the observation, selection and intensive study of individuals exhibiting contrasting phenotypes (e.g. height, wood quality, drought or pest resistance, etc.) to uncover the genes

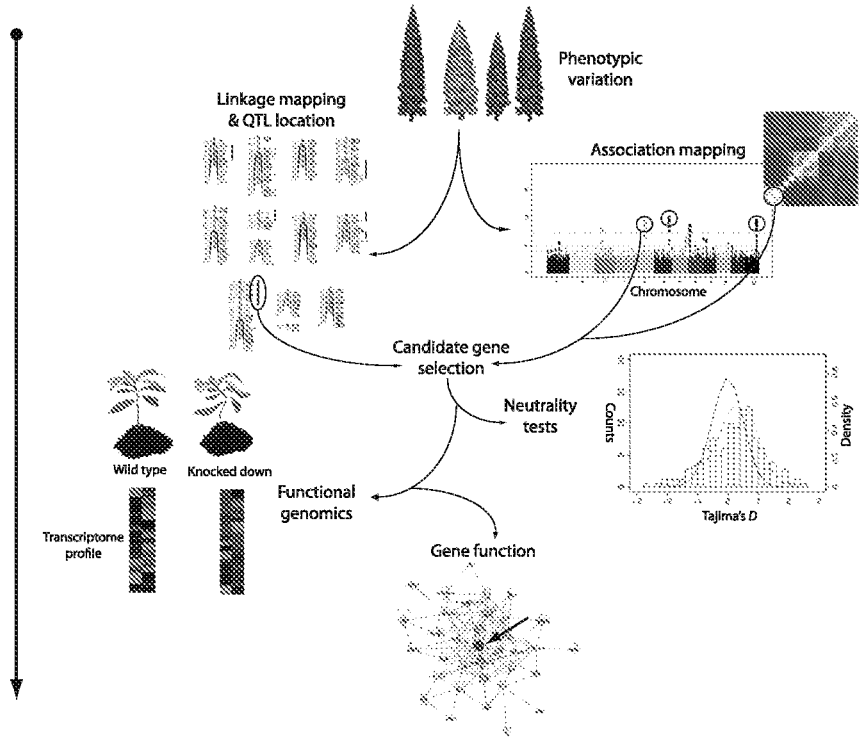

Figure 1 Schematic representation and illustrations of some the forward-genetics approaches used to detect candidate genes involved in adaptation in forest trees that are discussed throughout this chapter. (See colour plate)

underlying such phenotype variation. Forward-genetic approaches can vary from quantitative trait loci (QTL) detection and association mapping to comparative transcriptome profiling (Figure 1). In this chapter, we summarize the rationale behind each one of these methods, together with an overview of the statistical approaches that are currently used to identify candidate genes/alleles underlying trait variation. Then, we highlight the main findings of landmark studies that provide evidence for adaptation in forest trees.

2. QUANTITATIVE TRAITS VARIATION

Adaptation mostly arises from selection on phenotypes and their underlying genetic variants (i.e. alleles) that increase fitness in a given set of environmental conditions (Kremer & Le Corre, 2012; Le Corre & Kremer, 2003).

Traditionally, genetic variation in adaptively important phenotypes in forest trees has been studied by using reciprocal transplant experiments or common garden tests (also called provenance trials in the forestry literature) using a quantitative genetics framework (e.g. Falconer & MacKay, 1996) that allows estimation of trait heritability within populations and trait genetic differentiation among populations (Q_{ST}). These approaches allow explicit quantification of genetic versus environmental effects on the phenotypes, and can indirectly test if trait evolution has been driven by local adaptation (Latta, 1998).

Most quantitative genetics studies in forest trees have long focused in traits of economic interest (Gonzalez-Martinez, Ersoz, Brown, Wheeler, & Neale, 2006), although height, diameter, pest resistance, cold hardiness, wood quality and bud phenology (mostly bud flush and bud set timings) are the most commonly surveyed characters (e.g. Holliday, Ralph, White, Bohlmann, & Aitken, 2008; Jaramillo-Correa, Beaulieu, & Bousquet, 2001; Rehfeldt, Ying, Spittlehouse, & Hamilton, 1999). Height and diameter can be affected by many factors, such as local climate or soils, and reflect the overall performance of a tree in a test site. Because of this, heritability estimates for these traits are usually low to moderate (i.e. $h^2 < 0.2$) and vary across sites, populations and tree age (e.g. Benowicz & El-Kassaby, 1999; Li, Beaulieu, Corriveau, & Bousquet, 1993; Li, Beaulieu, Daoust, & Plourde, 1997). On the other hand, characters such as pest resistance, cold hardiness, wood density and bud phenology usually exhibit moderate to high heritabilities (i.e. h^2 between 0.3 and 0.8; Howe et al., 2003; Porth, Klápště, Skyba, Hamelinnan et al., 2013), which suggests that abundant genetic variation exists for these traits, and therefore can be shaped by selection.

Quantitative genetic theory shows that trait differentiation among populations (Q_{ST}) is expected to be equivalent to allele frequency differentiation at molecular markers (F_{ST} or G_{ST}; Nei, 1977; Wright, 1951) if selection is not acting; thus, the comparison of these two estimates is thought to test for the occurrence of local adaptation in phenotypic traits (Leinonen, McCairns, O'Hara, & Merilä, 2013; Leinonen, O'Hara, Cano, & Merilä, 2008; McKay & Latta, 2002). As with heritability, most traits exhibiting high Q_{ST} estimates (i.e. above 0.3) are those related to phenology and defence, while traits associated with the general form of an individual (i.e. diameter, branch number) mostly have lower Q_{ST} values (below 0.2; reviewed by Leinonen et al. (2008), Savolainen, Lascoux, and Merilä (2013)). It must be noted, however, that several authors argue that Q_{ST} versus F_{ST} comparisons may not be valid under all situations (McKay & Latta, 2002). The discrepancy between these two measures may reflect differences in their genetic architecture. That is,

Q_{ST} for most quantitative traits results from the combined action (and interaction) of many genetic variants, each of which may contribute a small amount to the total variance of a trait. These many small effect variants may experience strong patterns of covariance or statistical linkage disequilibrium (LD) across populations that shift trait means more than would be predicted from the combined frequency changes at individual loci (Kremer & Le Corre, 2012; McKay & Latta, 2002). On the other hand, F_{ST} is typically estimated from physically unlinked molecular markers and does not incorporate covariances or statistical LD among loci.

Comparisons of Q_{ST} estimates across studies must be interpreted with caution, since they will necessarily depend on the breadth of population sampling and the age and ecological conditions at which quantitative traits were evaluated (Hamilton, Lexer, & Aitken, 2013). For instance, after sampling *Picea mariana* populations in a latitudinal transect in northeastern Canada, a Q_{ST} of 0.91 was estimated for bud set (Morgenstern, 1969), while a much lower estimate ($Q_{ST} = 0.22$) was obtained for this trait when surveying more southern populations (Prunier, Laroche, Beaulieu, & Bousquet, 2011). Similarly, low Q_{ST} values were determined for cold hardiness (0.02–0.32) in *Picea sitchensis* populations from an island (Campbell, Pawuk, & Harris, 1989), while a much higher estimate (0.89) was obtained when the whole species distribution was surveyed (Mimura & Aitken, 2007). Variation in Q_{ST} estimates has also been observed among test sites when evaluating the same provenances. For instance, phenotypic differentiation for susceptibility to *Mycosphaerella* leaf disease varied between 0.05 and 0.25 in five different trials of *Eucalyptus globulus* in Tasmania; these values were partially correlated to the severity of the infection at each test site (Hamilton et al., 2013). Nevertheless, in spite of these caveats, Q_{ST} is still regarded as a useful mean to distinguish whether population differentiation in complex polygenic traits has been driven by selective or stochastic factors, particularly in genome-wide contexts that allow establishing a more accurate link between molecular and phenotypic variation (see Leinonen et al. (2013) for a review).

3. LINKAGE MAPPING AND QTL DETECTION

While a quantitative genetics framework allows studying genetic variation and inferring selection on phenotypes, a more mechanistic understanding of adaptive evolution requires knowing the genomic architecture of adaptive traits and linking molecular and phenotypic variation. From

this point of view, identifying QTL in a linkage map is often the first step. This relies on identifying genetic markers that co-segregate with morphological variation in a linkage genetic map built from the progeny of individuals exhibiting contrasting phenotypes (usually an F2 or a backcross; Falconer & MacKay, 1996). However, this method can be troublesome for long-lived species like forest trees, for which the establishment of experimental mapping populations can be expensive and attaining the number of generations can take from years to decades.

The first linkage maps for forest trees (and other taxa) were built based on dominant markers, like RAPDs or AFLPs (e.g. Bradshaw & Stettler, 1995; Scalfi, Troggio, Piovani, et al., 2004). Later, they incorporated codominant and gene-derived markers such as SSRs, ESTPs and SNPs, which facilitated the co-location of QTLs and the identification of the putative coding genes involved; they also allowed comparisons across closely related taxa in cases where markers were transferable among them (e.g. Brondani, Williams, Brondani, & Grattapaglia, 2006; Paolucci et al., 2010; Pelgas et al., 2006). These studies have mainly aimed to decipher the genetic architecture and location of genome regions underlying quantitative traits with economical interest, including wood yield and quality or pest resistance (e.g. Bradshaw & Stettler, 1995; Scalfi et al., 2004). However, QTLs for adaptation to abiotic conditions like temperature or water availability have been also investigated. For instance, in *Picea glauca*, a genetic map that has been steadily developed since 2006 (Pelgas et al., 2006; Pelgas, Bousquet, Meirmans, Rutland, & Isabel, 2011) has allowed the identification of several regions involved in cold adaptation; these regions contain from 8 to 38 loci, and are spread across the genome (Pelgas et al., 2011). Similarly, between 11 and 15 QTLs involved in autumn and spring cold hardiness were located in a *Pseudotsuga menziensii* linkage map (Jermstad et al., 2001). Many other examples of adaptive QTL studies exist for conifers (e.g. Prunier et al., 2013; reviewed by Rutland et al. (2011)); however, most of them have the unfortunate drawback that each individual QTL identified usually accounts only for a very small part of the total adaptive variation, and altogether they rarely explain more than 10% of the phenotypic variance (PVE; see Figure 2; see also Marguerit et al. (2014) for a remarkable exception of a QTL for water-use efficiency accounting for 60% of the phenotypic variation). This last point should be particularly highlighted given that the rather low population size used in these studies (<300 trees) have certainly led to overestimated QTL effects (Xu, 2003).

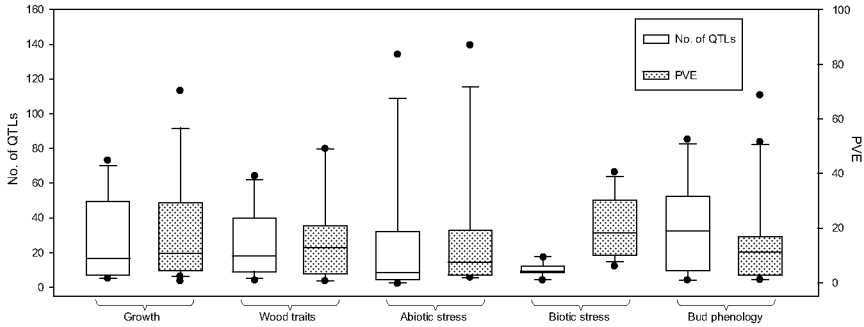

Figure 2 Number of quantitative trait loci (QTL), and percentage of the variance explained (PVE) by each individual QTL for various phenotypic and adaptive traits in a collection of 85 studies in forest trees reported herein and in Muranty et al. (2014; Table 2). Boxes denote interquartile ranges, whiskers 95% confidence intervals, and horizontal lines within boxes represent the mean.

In angiosperm trees, several studies have obtained similar results to those in conifers. For instance, 32 QTLs related to bud burst timing were placed on a *Quercus robur* linkage map, each one explaining 3–11% of the phenotypic variance (Derory et al., 2010), while 31 loci involved in bud set were identified in *Castanea sativa*, with weak to medium PVE (4–17%) (Casasoli et al., 2004). Other studies have found fewer QTLs associated with adaptive traits (Savolainen et al., 2007). For example, in *Q. robur*, 2–10 QTLs related to water-logging tolerance were identified (Parelle et al., 2007), while six regions associated with cold adaptation were mapped in a *Populus* hybrid cross, each one having a low to medium phenotypic effect (PVE = 5–18%; Frewen et al., 2000). Equally low QTL numbers but with higher PVE were observed for other traits, such as resistance to leaf blight in the tropical rubber tree *Hevea* spp. (5 QTLs with PVE = 4–44%; Lespinasse et al., 2000), or to pathogen moulds in *Theobroma cacao* (6 QTLs with PVE = 12–28%; Risterucci, Paulin, Ducamp, N'goran, & Lanaud, 2003). Thus, the number of QTLs involved, and their effect size, seems more related to variability among different traits than to differences in angiosperm/conifer evolution.

Altogether, QTL mapping studies have shown that the genes related to adaptation in forest trees are often spread over the entire genome, they are seldom co-located across closely related species and each one of them has a low to medium phenotypic effect. However, given the usually large genomes of many tree taxa, and the few generations of recombination possible in most linkage mapping studies in trees (due to long generation time), this

approach is best suited to identifying broad linkage blocks associated with an adaptive trait, and rarely allows pinpointing the specific polymorphisms, sequences or even genes involved in adaptation. They are thus mostly useful for delineating candidate genomic regions encompassing adaptive variation that can be targeted with other approaches, such as genotype–phenotype association mapping.

4. ASSOCIATION MAPPING

Also known as LD mapping, this method has proven to be a good alternative to QTL mapping in forest trees, as it can be readily used in natural populations without prior controlled crosses or experimental designs. Association mapping methods take advantage of the genetic diversity existing in natural stands to test for genotype–phenotype correlations in individuals, while controlling for background effects of population structure and kinship. The presence of historical recombination in natural populations allows for much greater precision at identifying genotype–phenotype associations, since markers must be in very tight LD with causal variants (Soto-Cerda & Cloutier, 2010). Such covariation is usually detected by surveying loci scattered all over the genome (Neale & Savolainen, 2004), although association studies performed with a few specific (and carefully selected) candidate genes or genomic regions have also provided good results for traits for which the potential causal genes are known (e.g. González-Martínez et al., 2006; Ingvarsson, García, Hall, Luquez, & Jansson, 2006).

When used in combination with dense high-throughput genotyping (e.g. SNPs identified from next-generation sequencing), genome-wide association studies (GWAS) can detect loci closely associated with causal variants, or even the quantitative trait nucleotides themselves, especially for lineages with high recombination rates like forest trees (Soto-Cerda & Cloutier, 2010). However, caution is needed, given that GWAS inferences can be biased by population structure and admixture, genetic drift and/or environmental variation through time (e.g. De Mita et al., 2013), which must all be taken into account through a series statistical corrections to avoid the detection of false positives. For example, population structure can produce a high rate of false positives in association studies when there is a correspondence between trait variation and the axes of ancestry across populations (e.g. Ingvarsson & Street, 2011). Consequently, the use of generalized linear models (GLM), unified mixed models (UMM; Yu

et al., 2006) or Bayesian approaches (Günther & Coop, 2013), that allow controlling for population and family structure, has become regular practice in association studies (e.g. Cappa et al., 2013; Kang et al., 2008; Uchiyama et al., 2013). For instance, Cappa et al. (2013) greatly decreased the number of spurious SNP-associations with growth and wood properties in *E. globulus* after performing such corrections with an UMM.

Another way to increase statistical power for identifying genotype–phenotype associations is to consider haplotypes rather than single-SNP variation (Akey, Jin, & Xiong, 2001). For example, only six marker–trait pairs explaining between 0.1 and 15% of the phenotypic variation were detected in *Populus nigra* when analyzing associations between chemical wood properties and individual SNPs, while the number of significant associations and PVE largely increased when these variants were concatenated in haplotypes (Guerra et al., 2013). Similar results were obtained in *P. trichocarpa* when using the same strategy (Wegrzyn et al., 2010).

In forest trees, GWAS employing thousands to millions of markers spread across the genome have allowed the identification of multiple loci related to traits with ecological and economic importance. Therefore, the discovery of associations is not dictated by the choice of candidate genes, but instead by the density of the markers employed. The most powerful GWAS study performed to date in a tree was based on whole genome re-sequencing (c. 18 million SNPs) in *P. trichocarpa* (Evans et al., 2014). In this study, the strongest candidate selected regions associations were the *FT2* paralog of *FLOWERING LOCUS T*, which correlated with continued height growth and bud set, and the *PHYTOCHROME AND FLOWERING TIME 1* gene (*PFT1*), which was associated with bud flush timing. This study further showed that the genetic architecture of most trait associations involved SNPs in intronic or intergenic regions, suggesting that regulatory regions make up a large fraction of the standing genetic variance influencing adaptive phenotypes in natural populations.

Further surveys in *Populus* include extensive phenotyping efforts followed by GWAS, which made it possible to infer the genetic architecture underlying complex traits in provenance trials and common gardens (McKown, Guy, et al., 2014; McKown, Klápště, et al., 2014; Porth, Klápště, Skyba, Hannemann et al., 2013; Porth, Klápště, Skyba, Lay et al., 2013). For example, a series of association tests for phenological traits revealed a high degree of genetic complexity in *P. trichocarpa* involving over 70 genes,

each one explaining 1–13% of the phenotypic variance of each trait (McKown, Klápště, et al., 2014). Other studies identified genes underlying fungal resistance, leaf drop and cellulose content (*REVOLUTA*; Porth, Klápště, Skyba, Lay et al., 2013), or stomatal variation (*PHABULOSA*, *BRASSINOSTEROID-INSENSITIVE 2*) and wound/disease response (*GLUTAMATE–CYSTEINE LIGASE*; McKown, Guy, et al., 2014), while a GWAS involving 29,333 SNPs across 3543 candidate genes reported significant associations with cell wall sugars, lignin and ultrastructure traits for 141 SNPs distributed across all 19 *Populus* chromosomes (Porth, Klápště, Skyba, Hannemann et al., 2013). This large number of associations, each explaining a small portion of the phenotypic variance (4–6%), is consistent with the highly polygenic architecture for wood characteristics.

The strong genetic basis of the traits above suggests that prospects are excellent for ongoing GWAS studies on whole genome re-sequencing in species that are candidates for woody biomass production and that have newly developed reference genome sequences, such as willow (*Salix purpurea* v1.0, DOE-JGI) or *Eucalyptus* (Myburg et al., 2014). However, GWAS can also be used to disentangle the complex genetic architecture underlying phenotypic traits when a reference genome is not available. For example, Pardchan et al. (2012) used 95,000 SNPs identified by genotyping by sequencing and Bayesian GLM to identify 11 loci associated with serotiny in lodgepole pine (*Pinus contorta*). Serotiny is considered a key adaptation in conifers inhabiting environments with frequent fire regimes, and its variability and inheritance seems to be far more complex than previously thought (Pardchan et al., 2012).

Association tests using fewer markers (i.e. a few 100–~1000 SNPs) have also retrieved interesting associations with phenotypic traits. For example, six SNPs were related to wood quality and pollen cone production in *Cryptomeria japonica* (Uchiyama et al., 2013), while 24 markers exhibited associations with the phenotypic variation of 20 secondary metabolites in *P. taeda* (Eckert et al., 2012). When considered jointly in a multi-locus approach, these 24 SNPs explained more than 50% of the phenotypic variation of some compounds' concentration, thus indicating that the interaction of genes may be more important to explain metabolome phenotypes than individual genes. In *Eucalyptus*, Cappa et al. (2103) identified 18 loci related to growth and wood properties, two of which physically mapped closely to a gene coding for the enzyme *F5H* (coniferaldehyde 5-hydroxylase) that has a key role in the lignin synthesis pathway.

Studies using a more modest quantity of markers (i.e. tens to hundreds of SNPs) can also provide promising results when these are located in carefully selected candidate genes or in regions encompassing QTL. For example, SNPs located in three candidates from the photoperiod pathway accounted for up to 15% of the phenotypic variation in growth cessation in *Populus tremula* (Ma, Hall, Onge, Jansson, & Ingvarsson, 2010), while SNPs in the *CAD* and *SAM*-2 genes accounted for 7–20% of the total genetic variance for earlywood specific gravity in *P. taeda* (González-Martínez, Wheeler, Ersoz, Nelson, & Neale, 2007). Similarly, haplotype variation in a single gene (*CESA*4) explained 2–5% of the variance of growth and some wood properties in *Populus tomentosa* (Du et al., 2013), while another gene (*CCR*) explained over 3% of the lignin composition in *Eucalyptus urophylla* (Mandrou et al., 2012). This last gene also exhibited strong associations with microfibril angle in two other *Eucalyptus* taxa (Thumma, Nolan, Evans, & Moran, 2005). The candidate gene approach can also deliver interesting results for more complex characters when a larger number of genes are considered. For instance, in *P. taeda*, associations of 1–2% per SNP were observed when the variation at 41 stress-inducible genes was correlated with carbon isotope discrimination (González-Martínez, Huber, Ersoz, Davis, & Neale, 2008). However, for such traits involving countless metabolic pathways, a whole-genome re-sequencing in a much larger sample of individuals might be required to account for larger portions of the phenotypic variance.

5. GENOME SCANS

Association methods can also be used at the range-wide level to identify clinal variation between allele frequencies and environmental variables. This approach is often complemented by methods that test markers for extreme population genetic structure (i.e. outlier methods). Both kinds of strategies rely on the premise that differential selective pressures in populations subjected to contrasting environmental conditions result in a significant differentiation of allele frequencies at the loci targeted by selection (and their linked loci). The outlier methods address this by estimating genetic differentiation (e.g. F_{ST}) across loci and looking for those exhibiting significantly higher or lower values than those expected under various stochastic neutral models; these outlier loci are then assumed to be potentially affected by directional or balancing selection, respectively (Luikart, England, Tallmon, Jordan, & Taberlet, 2003). Genotype–environment association (GEA)

methods, on the other hand, allow the inference of local adaptation along environmental gradients, perhaps even when selection is not strong enough to produce F_{ST} outliers (Joost et al., 2007). Both, the outlier and GEA methods have the drawback of being prone to significant bias when populations are not at demographic equilibrium, for example when they are hierarchically structured, show isolation by distance and/or when mutation rates are highly heterogeneous across loci (e.g. De Mita et al., 2013). Some approaches are available to correct some of this potential bias (e.g. Günther & Coop, 2013), but currently, most authors prefer to combine different strategies when scanning the genome for putatively adaptive genes and give greater confidence to those loci that are identified by multiple methods (e.g. Alberto et al., 2013; Prunier, Gerardi, Beaulieu, & Bousquet, 2012; Tsumura et al., 2012).

Another caveat that both the outlier and GEA approaches have in common is that they are most sensitive to detecting locally adapted alleles that are evolving through antagonistic pleiotropy; that is, alleles that are beneficial in one environment but deleterious in another one (Tiffin & Ross-Ibarra, 2014). However, studies in model species have shown that locally favoured alleles actually tend to be neutral elsewhere in the species range (e.g. Fournier-Level et al., 2011; Lowry, Hall, Salt, & Willis, 2009). In such conditional neutrality circumstances, the power for detecting local adaptation will be greatly diminished for both kinds of methods (see Tiffin and Ross-Ibarra (2014)). A potential way to circumvent this problem could be to test for environmental association and deviations from neutral expectations on individual genes, populations or gene clusters (e.g. Grivet, Sebastiani, González-Martínez, & Vendramin, 2009; Jaramillo-Correa et al., 2015; Prunier et al., 2012).

The first genome scan that employed an outlier approach in forest trees was performed to identify regions potentially involved in the divergence of two closely related oak species in Europe (Scotti-Saintagne et al., 2004). In this study, the authors pinpointed over 47 loci (a combination of AFLPs, isozymes and SSRs) exhibiting significantly high F_{ST} values between taxa, and which were distributed in most of the linkage groups of the species pair. Other early surveys in forest trees include one in *P. abies*, where 37 outliers (AFLPs, ESTPs and SSRs) distributed across its genome were detected (Acheré, Favre, Besnard, & Jeandroz, 2005), another one in *C. japonica*, that found five CAPs (cleaved amplified polymorphisms developed from ESTs) exhibiting significantly high F_{ST} values (Tsumura et al., 2007), and a survey in populations of *P. glauca* from eastern Canada, that

reported a list of 49 outlier SNPs from coding genes (Namroud, Beaulieu, Jaquish, Laroche, & Bousquet, 2008), among others. However, as these early studies did not take population structure or differences in the mutation rate across loci into account, the candidates they report should be taken with caution.

More recent studies (summarized in Table 1) that do correct for such potential bias tend to combine outlier and environmental association approaches and report lists of candidates that overlap across methods. In conifers, these surveys have been performed for different species in genus *Pinus* and *Picea*, and for *C. japonica* (Table 1); although an interesting comparative study in coniferous taxa from the Alps also included *Abies alba* and *Larix decidua* (Mosca et al., 2012). In angiosperms other than oaks, genome scans have been carried out in *Fagus*, *Eucalyptus*, *Populus*, *Alnus* and in the tropical tree *Eperua falcata*, among others (Table 1). The identity of candidates generally differs across studies and species, even when they are co-distributed in the same region. For instance, in the above-mentioned survey by Mosca et al. (2012) no candidate outlier SNPs were shared by the four Alpine taxa analyzed, while there were only seven loci in common between species pairs, only one of which (located in a heat-shock protein 101) was associated with the same environmental variable in two pines (Mosca et al., 2012). Similarly, two spruce species co-distributed in eastern Canada only shared three of the 25 and 49 outlier genes respectively detected for each taxa (Prunier et al., 2011), while two environmental gradients surveyed for associations in *Quercus petrae* only had a few SNPs in common (Alberto et al., 2013).

When candidate comparisons across taxa and studies are expanded to the gene family or function levels, several interesting patterns begin to arise. For instance, transcription factors, especially those of the *MYB*, *C3HL*, *C2H2* zinc finger and *HD-ZIP* gene families recurrently appear as outliers in species as varied as *P. mariana*, *Q. petrae* or *C. japonica*, among others, where they further show significant correlations with temperature and precipitation (Alberto et al., 2013; Prunier et al., 2011; Tsumura et al., 2012). Other gene families such as the dehydrins, the 60*S* ribosomal and heat-shock proteins or *4CL* also have shown associations with environment across taxa (e.g. González-Martínez et al., 2007; Grivet et al., 2009), while other genes coding for proteins involved in membrane transport, abiotic stress response, lignin biosynthesis or plant development are also recurrent outliers in many studies (Table 2).

Table 1 Number of Candidate Gene-Makers Detected with Outlier (i.e. F_{ST} Based) and Correlation (i.e. Genotype–Environment Associations; GEA) Methods in Various Studies in Forest Trees. The F_{ST} Range and the Identity of the Environmental Factors Associated with the Markers Are also Included

Species	No. Outliers (F_{ST})[a]/ No. Markers	F_{ST} Range	No. Candidates (GEA)[a]	Environmental Factor	References
Angiosperms					
Alnus glutinosa	57/1990	0.05–0.12	20 2	Temperature Precipitation	De Kort et al. (2014)
Castanopsis fargesii	7/32	0.15–0.20	5	Precipitation	Li, Sun, Huang, and Cannon (2014)
Eperua falcata	8/74	0.04–0.15	4	Habitat type	Audigeos, Brousseau, Traissac, Scotti-Saintagne, and Scotti (2013)
Eucalyptus camaldulensis	5/76	0.21–0.42	4	Climate	Dillon et al. (2014)
Fagus sylvatica	5/546	0.27–0.30	–	Elevation	Csilléry et al. (2014)
Populus balsamifera	26/747	0.22/0.94	45	Climate	Keller, Levsen, Olson, and Tiffin (2012)
Populus trichocarpa	1130/ $\approx 18 \times 10^6$	0.12–0.37	194 157	Temperature Precipitation	Evans et al. (2014)
Quercus petraea	28/384	0.01–0.11	5 3	Elevation Latitude	Alberto et al. (2013)
Quercus robur–Q. petraea	60/262	0.22–0.93	–	–	Gauchoux et al. (2013)
Conifers					
Abies alba	N.A./249	N.A.	4	Climate	Mosca et al. (2012)
Cryptomeria japonica	20/1026	0.11–0.22	9 2 4	Temperature Precipitation Sun radiation	Tsumura et al. (2012)
Larix decidua	N.A./267	N.A.	6	Climate	Mosca et al. (2012)

(*Continued*)

Table 1 Number of Candidate Gene-Makers Detected with Outlier (i.e. F_{ST} Based) and Correlation (i.e. Genotype—Environment Associations; GEA) Methods in Various Studies in Forest Trees. The F_{ST} Range and the Identity of the Environmental Factors Associated with the Markers Are also Included—cont'd

Species	No. Outliers (F_{ST})[a]/ No. Markers	F_{ST} Range	No. Candidates (GEA)[b]	Environmental Factor	References
Picea abies	29/445	0.10—0.17	18	Latitude	Chen et al. (2012)
Picea glauca	20/768	0.04—0.13	—	—	Namroud et al. (2008)
P. glauca—Picea engelmannii	7/311	0.08—0.18	17	Climate	De la Torre, Roberts, and Aitken (2014)
Picea mariana	10/99	0.04—0.16	14	Precipitation	Prunier et al. (2012)
			10	Temperature	
Picea obovata	13/356	0.03—0.05	22	Latitude	Chen et al. (2014)
Picea rubens	4/33	0.03—0.12	7	Climate	Bashalkhanov, Eckert, and Rajora (2013)
			3	Pollution	
Pinus banksiana	1/472	N.A.	3	Temperature	Cullingham et al. (2014)
Pinus cembra	N.A./459	N.A.	11	Climate	Mosca et al. (2012)
Pinus contorta	8/472	N.A.	14	Temperature	Cullingham et al. (2014)
Pinus mugo	N.A./693	N.A.	18	Climate	Mosca et al. (2012)
Pinus taeda	24/3059	0.14—0.27	5	Aridity	Eckert et al. (2010)

[a]Only markers detected with Bayesian (i.e. Bayescan) methods at $P < 0.01$ (i.e. BF = 2) or higher are reported.
[b]Only markers detected with regression methods that correct for population structure are included.

Table 2 Short List of Candidate Genes or Gene Families Associated with Adaptive Traits in Various (i.e. At least Two) Forest Tree Species through Some of the Methods Discussed in the Text

Gene/Gene Family	Biological Function	Species	Associated Trait	Evidence[a]	References
4CL	Plant defence/lignin biosynthesis	*Picea glauca*	Growth, pest resistance	Gene expression/ QTL mapping	Porth et al. (2011)
		Picea rubens	Air pollution	GWAS	Bashalkhanov et al. (2013)
		Picea sitchensis	Cold hardiness	Gene expression	Holliday et al. (2008)
		Pinus pinaster	Drought resistance	GEA	Jaramillo-Correa et al. (2015)
Aquaporin	Membrane transport	*Quercus petraea*	Bud phenology	GWAS	Alberto et al. (2013)
		Eucalyptus camaldulensis	Drought resistance	GEA	Dillon et al. (2014)
AUX-IAA	Growth regulation	*Picea mariana*	Bud set	GWAS	Prunier et al. (2013)
		Picea sitchensis	Bud set	GWAS	Holliday, Ritland, and Aitken (2010)
		Picea sitchensis	Cold hardiness	GWAS	Holliday et al. (2010)
α- and β-*TUBULIN*	Membrane transport	*Picea glauca*	Wood quality	GWAS, gene expression	Beaulieu et al. (2011)
		Pinus pinaster	—	GEA	Jaramillo-Correa et al. (2015)
		Pinus taeda	Wood quality	GWAS, NT	Gonzalez-Martinez et al. (2007)
		Populus nigra	Wood quality	GWAS	Guerra et al. (2013)
CAD	Plant defence/lignin biosynthesis	*Picea glauca*	Growth, pest resistance	Gene expression/ QTL mapping	Porth et al. (2011)
		Pins taeda	Wood quality	GLM, NT	Gonzalez-Martinez et al. (2007)

(*Continued*)

Table 2 Short List of Candidate Genes or Gene Families Associated with Adaptive Traits in Various (i.e. At least Two) Forest Tree Species through Some of the Methods Discussed in the Text—cont'd

Gene/Gene Family	Biological Function	Species	Associated Trait	Evidence*	References
CCoAOMT	Plant defence	*Picea glauca*	Growth, pest resistance	Gene expression/ QTL mapping	Porth et al. (2011)
		Pinus taeda	Drought resistance	GWAS	González-Martínez et al. (2008)
		Populus nigra	Wood quality	GWAS	Guerra et al. (2013)
		Populus trichocarpa	Bud set, growth	GWAS	McKown, Guy, et al. 2014
CCR	Plant defence/lignin biosynthesis	*Eucalyptus grandis*	Wood quality	LD-mapping, GWAS	Thumma et al. (2005)
		Eucalyptus nitens	Wood quality	LD-mapping, GWAS	Thumma et al. (2005)
		Eucalyptus urophylla	Wood quality	GWAS	Mandrou et al. (2012)
CESA	Cellulose synthesis/ plant defence	*Pinus pinaster*	—	NT	Pot et al. (2005)
		Pinus radiata	—	NT	Pot et al. (2005)
		Populus nigra	Wood quality	GWAS	Guerra et al. (2013)
		Populus tomentosa	Growth, wood quality	GWAS	Du et al. (2013)
		Populus trichocarpa	Bud set, growth, biomass	GWAS	McKown, Klápště, et al. (2014)
DHN	Macromolecule protection	*Eucalyptus gunnii*	Cold hardiness	Gene expression	Keller et al. (2009)
		Picea sitchensis	Cold hardiness	Gene expression	Holliday et al. (2008)
		Pinus pinaster	Drought resistance	GWAS	Eveno et al. (2008)
		Pinus taeda	Photosynthesis efficiency	GWAS	González-Martínez et al. (2008)
		Populus tremula × *Populus tremuloides*	Cold hardiness	Gene expression	Ruttink et al. (2007)
		Populus trichocarpa	Leaf drop	GWAS	McKown, Klápště, et al. (2014)
		Prunus persica	Cold hardiness	Gene expression	Arora et al. (1997)

		Pseudotsuga menziesii	—	NT	Eckert et al. (2009)
α- and β-*EXPANSIN*	Growth regulation	*Picea glauca*	Wood quality	GWAS, gene expression	Beaulieu et al. (2011)
		Picea sitchensis	Cold hardiness	GWAS	Holliday et al. (2010)
		Pinus radiata	Wood quality	GWAS	Dillon et al. (2009)
FT2/FTL2/PTFT	Photoperiod response, flowering time	*Picea abies* *Picea obovata*	Bud set	GEA, GWAS	Chen et al. (2012, 2014)
		Picea abies	Bud set	Gene expression	Gyllenstrand et al. (2007)
		Picea glauca	Bud set	Gene expression	El Kayal et al. (2011)
		Pinus sylvestris	Bud set	Gene expression/GWAS	Avia, Kärkkäinen, Lagercrantz, and Savolainen (2014)
		Populus tremula × *P. tremuloides*	Growth, bud phenology, flowering	Transgenic plants, gene expression	Böhlenius et al. (2006)
		Populus trichocarpa	Bud set, growth	GEA/GWAS/Gene expression	Evans et al. (2014)
		Prunus domestica	Growth, bud phenology, flowering	Transgenic plants	Srinivasan et al. (2012)
GIGANTEA	Circadian clock	*Picea abies* *Picea obovata*	Bud set	GEA, GWAS	Chen et al. (2012, 2014)
		Picea sitchensis	Bud set	GWAS	Holliday et al. (2010)
		Populus balsamifera		GWAS	

(*Continued*)

Table 2 Short List of Candidate Genes or Gene Families Associated with Adaptive Traits in Various (i.e. At least Two) Forest Tree Species through Some of the Methods Discussed in the Text—cont'd

Gene/Gene Family	Biological Function	Species	Associated Trait	Evidence[a]	References
			Bud set, cold hardiness		Keller et al. (2012); Olson et al. (2013)
Heat-shock protein	Drought resistance	*Picea rubens*	Air pollution	GWAS	Bashalkhanov et al. (2013)
		Picea mariana	Growth	GWAS	Prunier et al. (2013)
		Picea sitchensis	Cold hardiness	Gene expression	Holliday et al. (2008)
		Pinus pinaster	Drought resistance	GEA	Jaramillo-C. et al. (2015)
MYB	Gene regulation	*Cryptomeria japonica*	Growth	GEA	Tsumura et al. (2012)
		Eucalyptus gunnii	Cold hardiness	Gene expression	Keller et al. (2009)
		Picea glauca	Growth, pest resistance	Gene expression/ QTL mapping	Porth et al. (2011)
		Picea glauca	Growth, plant defense	Gene expression/ transgenic plants	Bedon et al. (2010)
		Picea mariana	Bud set, growth	GWAS	Prunier et al. (2013)
		Picea sitchensis	Cold hardiness	Gene expression	Holliday et al. (2008)
		Pinus pinaster	Drought resistance, photosynthesis efficiency	Gene expression/ QTL mapping	De Miguel et al. (2014)
		Populus tremula × *P. tremuloides*	Cold hardiness	Gene expression	Ruttink et al. (2007)
		Populus trichocarpa	Bud set, growth	GWAS	McKown, Guy, et al. (2014)
PAL	Plant defence/lignin biosynthesis	*Picea sitchensis*	Bud set	GWAS	Holliday et al. (2010)
		Picea glauca	Growth/Pest resistance	Gene expression/ QTL mapping	Porth et al. (2011)

		Pinus radiata	Wood density	GWAS	Dillon et al., (2009)
PHYA/PHYB	Circadian clock	*Picea sitchensis*	Bud set	GWAS	Holliday et al. (2010)
		Populus tremula	Bud set	GWAS, NT	Ingvarsson et al. (2006, 2008)
		Populus tremula × *P. tremuloides*	Gorwth, cold hardiness	Transgenic plants	Olsen et al. (1997)
PRR/PaRR	Circadian clock	*Picea abies/Picea obovata*	Bud set	GEA, GWAS	Chen et al. (2012, 2014)
		Picea abies	–	NT	Kallman et al. (2014)
		Populus trichocarpa	Temperature response	GEA	Geraldes et al. (2014)
		Popuus trichocarpa	Bud set, growth, biomass	GWAS	McKown, Guy, et al. (2014)
Xiloglucan glucosyl transferase	Plant defence	*Picea sitchensis*	Bud set	GWAS	Holliday et al. (2010)
		Picea sitchensis	Cold hardiness	GWAS	Holliday et al. (2010)
		Populus tremula × *P. tremuloides*	Wood formation	Gene expression/ transgenic plants	Derba-Maceluch et al. (2015)
		Popuus trichocarpa	Bud set, growth, biomass	GWAS	McKown, Guy, et al. (2014)
Ribosomal protein	Gene regulation	*Eucalyptus gunnii*	Cold hardiness	Gene expression	Keller et al. (2009)
		Quercus petraea	Bud phenology	GWAS	Alberto et al. (2013)
		Pinus pinaster	Drought resistance	GEA	Jaramillo-Correa et al. (2015)

[a]GEA = Genotype-environment associations (e.g. Bayenv, MatSAM, etc.); GWAS = Genome-wide association studies (e.g. general linear models, linear-mixed models, unfied-mixed models, etc.); NT = Neutrality tests; QTL = Quantitative trait loci.

As briefly mentioned above, genome scans are powerful tools for studying the introgression and divergence patterns between closely related species pairs. For instance, other than the early survey by Scotti-Saintagne et al. (2004), introgression between the early successional *Q. robur* and the late-successional *Q. petrae* was analysed with outlier and non-outlier loci, with the former showing a pattern that is coincident with the predicted introgression from the late-successional species into the early successional one, and the later showing no particular pattern (Guichoux et al., 2013). An analogous study in the hybrid zone between the coastal *Picea sitchensis* and the continental *P. glauca* in western Canada revealed a far more complex story, with separate sets of outliers introgressing more than expected under a neutral pattern from the coast into the continent and vice versa, and other loci exhibiting less introgression than predicted (Hamilton, Lexer, et al., 2013). Similarly complex patterns were observed on the other side of the Canadian Rockies between the mountain species *Pinus contorta* var. *latifolia* and the interior taxa *P. baksiana* (Cullingham, Cooke, & Coltman, 2014), while in Europe, an in-depth genome scan in the sympatric but ecologically divergent *Populus alba* and *P. tremula* revealed long shared genomic blocks, particularly within a region containing an incipient sex chromosome, which suggests that introgression is being favoured by balancing selection in this species pair (Stölting et al., 2013).

6. SURVEYING SEQUENCE VARIATION AT CANDIDATE GENES

Much of the studies discussed so far focused on a genome-wide strategy, in which a large amount (i.e. from several hundreds to some thousands) of markers are surveyed to identify either genomic regions (i.e. QTL mapping) or specific polymorphisms covarying with phenotypic traits or environmental conditions (GWAS, GEA). However, most pinpointed regions/polymorphisms are usually assumed to be linked to instead of being the causal variant targeted by selection, which often remains unknown (Neale & Savolainen, 2004). The candidate gene approach is a powerful way to fill this gap, as it focuses on these specific regions, particularly those containing genes of known functions. In addition, this strategy can be directly applied in natural populations of undomesticated species with no other prior knowledge than the location (i.e. flanking DNA sequence) of the region/gene itself. The candidate gene approach has been favoured in

forest trees because of their mostly outcrossing nature and high genome-wide recombination rates, which makes LD decay rapidly, mostly within gene limits (Neale & Kremer, 2011). Therefore, if one assumes that genomes evolve in haplotype blocks, the blocks to explore should be smaller in forest trees than in other species with different biological traits, like selfers or restricted endemics (Neale & Savolainen, 2004).

Early candidate gene studies focused on model organisms and the knowledge of biochemical and regulatory basis of specific traits to target the genes involved (Zhu & Zhao, 2007). However, in non-model organisms these methods are not easily applied, and candidate gene discovery has taken several complementary approaches for these taxa, like QTL cloning, or comparative and functionally based genomics. QTL cloning focusses on the extensive sequencing of regions previously associated with quantitative trait variation in linkage mapping surveys. This approach has been useful in species with small and completely sequenced genomes, but for taxa with larger genomes, this method can be applied only when LD is strong (Zhu & Zhao, 2007) and QTL intervals are narrow enough to contain only a few genes. Unfortunately, the size of such intervals in trees make positional cloning impractical, even for species with completely sequenced genomes like *Eucalyptus*, *Populus*, *Picea* or *Pinus*. So far, only the annotation of genes contained in QTLs can provide short lists of candidates to survey by direct sequencing (e.g. Ingvarsson et al., 2006; Thumma et al., 2005).

Another strategy is to survey the genomic resources already available for forest trees (e.g. annotated genes and transcriptomes) and identify potential candidate genes by comparison with model species. For instance, if one assumes that the orthologous gene coding for an enzyme involved in a specific biochemical pathway in a model species, like the biosynthesis of lignin or other secondary metabolites, could be affecting the variation of wood quality or pest resistance in a forest tree, then the genomic resources of the model taxa can be used to identify and design primers to test this ortholog for selection signals in the forest tree species (e.g. González-Martínez et al., 2007; Ingvarsson et al., 2006; Kulheim, Yeoh, Mainz, Foley, & Moran, 2009). This comparative genomics approach is, however, limited by the availability of genomic resources in the target species, the gene transferability (i.e. ortholog identification) across unrelated taxa and the prior knowledge of the biochemical pathways in the model species. Consequently, if a gene that plays a major adaptive role in a target species belongs to a gene family or a biochemical pathway that does not exist in a model species, it cannot be surveyed until it has been identified de novo through functional surveys.

For instance, the diversity and adaptive value of taxol-related genes in *Taxus* could only be evaluated once the genes involved in its biosynthetic pathways were completely outlined (e.g. Burgarella et al., 2012; Jennewein, Wildung, Chau, Walker, & Croteau, 2004).

Once a potential candidate gene has been identified, there is a number of methods or neutrality tests (NT) to infer if its variation deviates from the expectations of the neutral theory of molecular evolution (see Wright and Gaut (2005), González-Martínez et al. (2011) for reviews). Some involve an examination of its site-frequency spectrum (SFS) and compare different estimators of diversity (θ) with those produced by coalescent simulations that the take the underlying demographic history of the species/population of interest into account (e.g. Grivet et al., 2009), while others rely on the ratio of non-synonymous to synonymous substitution rates ($\omega = d_N/d_S$), and allow detecting selection at different evolutionary timescales and acting within or among species (e.g. Palmé, Pyhäjärvi, Wachowiak, & Savolainen, 2009).

The earliest candidate gene studies in forest trees date from about a decade ago and, as with previous approaches, they mostly concentrated in economically important characters, like growth, wood formation, stress tolerance and phenology, and focused on commercial taxa such as pines, spruces, *Eucalyptus* and poplars. For instance, Brown, Gill, Kuntz, Langley, and Neale (2004) tested for selection on the SFS (using Tajima's *D*) in 19 candidate genes in *P. taeda* and found no deviations from neutral expectations. Later on, some of these genes and other candidates for drought tolerance (selected after verifying homology with drought-response genes in model taxa) were surveyed in this same species through additional NT, and signals of selective sweeps and balancing selection were inferred for the *ERD*3 and *CCOAMT*-1 genes, respectively (González-Martínez et al., 2006). A similar study performed on wood formation candidates in two other pines (*Pinus pinaster* and *Pinus radiata*) also revealed the putative action of balancing selection in a cellulose-synthase gene (*CESA*3) and of positive selection in a member of the *KORRIGAN* gene family (Pot et al., 2005), while an extensive study based on compound tests and ABC simulations on photoperiod-related genes in *P. abies* showed that the circadian clock gene *PaPRR3* is a strong candidate for divergent selection and local adaptation (Källman et al., 2014). This same approach was used to infer positive selection in six cold hardiness candidates in *Pseudotsuga menziesii*, most of them encoding structural components or proteins associated with the cellular membrane (Eckert et al., 2009).

In *P. tremula*, the candidate gene approach uncovered strong associations between bud set and the *PHYB2* and the *LATE-ELONGATED HYPOCOTYL* genes involved in the circadian clock (paralogs *LHY1* and *LHY2*) (Ingvarsson et al., 2006, Ingvarsson, Garcia, Luquez, Hall, & Jansson, 2008; Ma et al., 2010). The *PHYB2* gene, which was selected from previous QTL surveys, additionally exhibited a strong latitudinal allele frequency cline that mirrored the phenotypic clinal variation in bud set measured in a common garden (Ingvarsson et al., 2006). Further comparisons across species, either on phylogenetic or coalescent frameworks, also allowed inferring the action of positive selection on particular candidate genes (see Table 2), such as *CCOAMT*-1 (caffeoylCoA O-methyltransferase) in *P. trichocarpa* and *P. nigra*, *CESA*3 (cellulose synthase A3) in *P. tomentosa* and several members of the dehydrin and *KNOX* families in pines and spruces, respectively (e.g. Du et al., 2013; Grivet et al., 2011; Guerra et al., 2013; Namroud, Guillet-Claude, Mackay, Isabel, Bousquet, 2010; Wegrzyn et al., 2010). Similarly, the comparison of 23 genes from four biosynthetic pathways of secondary metabolites across four *Eucalyptus* taxa revealed that two of them (both from the 1-deoxyxylulose-5-phosphate synthase family) were likely affected by purifying selection in all species (Külheim et al., 2009). Altogether, these results highlighted how subfunctionalization can foster the action of selection in recently duplicated genes by allowing the quick evolution of adaptation across taxa.

There is evidence suggesting that convergent evolution often involves separate genes and various evolutionary paths. For instance, closely related *Populus* species showed different genetic architecture for bud set, despite exhibiting similar phenotypic clines with latitude. Unlike the above-mentioned example in *P. tremula*, *P. balsamifera* showed little segregating covariation with bud set in phytochrome genes, including *PHYB2* (Keller, Levsen, Ingvarsson, Olson, & Tiffin, 2011); instead, significant associations were most prominent for SNPs in the *GIGANTEA* (*GI*), *EARLY FLOWERING 3* (*ELF3*) and *LEAFY* (*LFY*) genes (Olson et al., 2013). This curious case for parallel evolution became even more interesting when *P. trichocarpa* was surveyed (Evans et al., 2014; McKown, Guy, et al., 2014; McKown, Klápště, et al., 2014). This species has recently diverged from *P. balsamifera* (c. 75,000 years), with whom it still shares high amounts of ancestral polymorphism (Levsen, Tiffin, & Olson, 2012). However, in *P. trichocarpa* there was a notable absence of associations between bud set and the previously implicated genes in either *P. balsamifera* or the more distantly related *P. tremula* (*PHYB2*, *LHY1*, *LHY2*, *GI*, *ELF3*, *LFY*). A

similar case was observed for associations with climate in co-distributed boreal spruces, which only shared a few number of genes showing homologous adaptive responses in Canada and Eurasia (Chen et al., 2014; Prunier et al., 2011).

7. FUNCTIONAL GENETICS OF TREE ADAPTATION

Once a gene has been identified as candidate and has shown evidence to be affected by selective forces, it must be validated through different strategies. A powerful approach to validation is provided by functional genetics, which focuses on the expression levels or the phenotypic and physiological effects of specific genes to decipher the genetic basis of adaptation. For instance, functional genetics can be used to infer if a candidate gene changes its expression levels when an organism is submitted to particular conditions and/or to measure the extent of phenotypic or physiological variations in an individual that has the expression of this gene modified through transgenic manipulations (e.g. Flachowsky, Hnake, Peil, Strauss, & Hadung, 2009). The rationale behind functional genomics studies is that the genes involved in an adaptive response to a particular factor should be over-expressed during the exposition to this factor, while other genes should keep their expression levels constant or even see them diminished. On the other hand, individuals that have been modified to over-express these adaptive genes should perform better when exposed to such a factor than those with the same genes suppressed or expressing at normal levels (e.g. Flachowsky et al., 2009; Peña & Séguin, 2001; Séguin et al., 2014).

This last approach is usually employed with cloned plants, where the physiological responses can be evaluated using a constant genomic background except for the modified gene. Such a procedure always relies on previous knowledge regarding the candidate gene to evaluate (see above), which is usually gathered from model species like *Arabidopsis*. For instance, knowing that bud set timing is an important adaptive trait in boreal and temperate trees, Olsen et al. (1997) produced transgenic clones of *P. tremula* × *P. tremuloides* that over-expressed a gene (*PHYA*) known for its role in the photoperiod control of flowering in *Arabidopsis thaliana*. Similarly, the functional roles of genes involved in dormancy have been evaluated in *Prunus domestica* (Srinivasan, Dardick, Callahan, & Scorza, 2012), and those implicated in growth cessation and bacterial resistance have been surveyed in *Populus* spp. (e.g. Böhlenius et al., 2006; Peña & Séguin, 2001), while those

conferring resistance to insects have been studied in *Pinus radiata* (Grace, Charity, Gresham, Kay, & Walter, 2005), among many other examples (reviewed by Séguin et al. (2014)).

Changes in transcript accumulation through time is usually investigated among phenotypically different individuals submitted to specific controlled conditions, and by surveying various tissues to verify if the response occurs at the organ or organismal levels. Transcript accumulation of candidate genes has been studied through classic methods, like northern blot or in situ hybridization, or by using more recent ones, like reverse-transcription quantitative PCR (RT-qPCR). Examples (summarized in Table 2) include the expression of a dehydrin gene that had differential transcript accumulations in evergreen and deciduous genotypes of *Prunus persica* (Arora, Callahan, Bassett, & Wisniewski, 1997), some *FT*-like genes that are involved in bud phenology and temperature adaptation in conifers, which changed their transcript profiles under different photoperiod and thermal conditions in *Picea abies* (Gyllenstrand, Clapham, Källman, & Lagercrantz, 2007), or members of the subgroup 4 of the R2R3-MYBs that were over-expressed after wounding, jasmonic acid or cold treatments in *P. glauca* and *Pinus taeda* (Bedon et al., 2010), among many others.

Functional genomic approaches can also be used to discover new adaptive genes or those specific to non-model taxa, through microarray or whole transcriptome sequencing (EST or RNA-seq) studies. Experiments based on EST measurements can be implemented by characterizing differentially expressed EST libraries (e.g. Forth, Koch, Berenyi, Burg, & Burg, 2005), although the most currently used strategy is to first construct a gene catalog from EST libraries of a variety of tissues and development stages that cover as much genes as possible (e.g. Keller et al., 2009; Lesur et al., 2015; Pavy et al., 2005). Then, microarrays are developed by embedding probes that target all or most of these genes, and where RNA samples from individuals submitted to various conditions can be tested (Bouck & Vision, 2007). Such kind of resources are now available and have been used for a great variety of tree species of the genus *Pinus*, *Picea*, *Populus*, *Eucalyptus*, *Quercus* and *Castanea* (Neale & Kremer, 2011).

Among many interesting results, the microarray approach has allowed the clustering of genes in various expression profiles during the development of freezing and cold tolerance of *Populus tremula* × *P. alba* and *Eucalyptus gunnii*, the development of bud set in *Picea glauca* or the physiological responses that occurred between late summer and early winter in *Picea sitchensis* (El Kayal et al., 2011; Holliday et al., 2008; Keller et al., 2009; Rautiok et al.,

2007), among many others. These studies have pinpointed hundreds to thousands of genes involved in cold response, including dehydrins, metabolic proteins, signal transduction related genes, photoreceptors, transcription factors and other metabolism regulators, which help corroborate the results of previous GWAS or candidate gene-based studies (see Table 2). These surveys have also identified additional de novo genes with strong adaptive or commercial value. This is the case in a recent study in *P. glauca* that discovered a β-glucosidase gene, *Pgβglu-1*, that confers resistance against spruce budworm (*Choristoneura fumiferana*; Mageroy et al., 2015).

It must be noted though, that the study of transcriptome profiles quickly brings us to the concept of internal phenotype, which involves all internal states of an organism throughout its life, from the molecular and cellular levels to its tissue and physiological properties (see Houle, Govindaraju, and Omholt (2010)). As such, the rates of transcript variation can be seen as an internal phenotype, with its own heritability, and that is submitted to the same genetic and environmental controls as any quantitative trait. Nevertheless, microarray or RNA-seq studies rarely track the fraction of transcript expression that is transmitted across generations, which is where natural selection acts, thus leaving unexplored the genetic basis of individual differences in gene expression. However, as with any phenotype, the loci that co-segregate with expression variation can be pinpointed in a linkage map by surveying the expression levels of particular genes in a mapping population. In an adaptive context, these regulating hotspots can be seen as keystone genes having pleiotropic effects on a more complex external phenotype, for example, bud phenology. QTL mapping of these loci (eQTLs) is currently being conducted in trees. To date, a few hotspots related to defence response against white pine weevil (*Pissodes strobi*) have been found by eQTL mapping in *Picea* (Porth, Hamberger, White, & Ritland, 2011), but examples are still scarce in the forest trees literature. Hopefully, eQTL mapping will become more common in the near future since next-gen sequencing technologies applied to cDNA, such as RNA-seq, simultaneously yield transcript abundance and sequence polymorphisms of the genes expressed in a sample.

8. APPLICATIONS IN BREEDING AND TREE SELECTION

Breeding is one of the first steps to maximize the income potential of any renewable resource. In forestry and agriculture (among others), it is an

iterative process that starts with the identification of individuals with outstanding phenotypes for economically important or adaptive characters (Muranty et al., 2014), and then uses selective crossing to increase the frequency of desirable traits. However, as previously discussed, the long generation times of forest trees have made selection and breeding tedious and slow, particularly for complex characters with medium to low heritability such as height, diameter or wood quality (Neale & Savolainen, 2004). Traditionally, the selection of individuals is made based on their phenotypes; however, with the development of high-throughput genotyping techniques and the strategies for detecting candidate genes discussed so far, it has become possible to predict the breeding value of individuals at an earlier age based on genetic marker data (Isik, 2014). This is particularly appealing for forest tree breeders, given that marker-assisted selection (MAS) can shorten the rotation age by more than 30%, diminish the required area to establish field trials, limit phenotyping costs and open the possibility for controlling both the inbreeding and gene diversity within populations (Muranty et al., 2014).

There are several strategies for MAS (reviewed by Hospital (2009)), which range from simply detecting the individuals to be kept in a stand based on their genotypes (i.e. population screening), to specific frameworks of directed and recurrent crossings (or back crossings) of trees with particular genotypes aiming to maximize the number of alleles of interest in their progeny (e.g. gene pyramiding). These strategies can further involve the inclusion of marker and environmental information into predictive models for breeding values and expected phenotypes (Hospital, 2009), among others.

The use of MAS has been particularly successful in fruit trees, for which there are various examples of improved lines developed from candidate gene variation, particularly for agronomic characters in the Rosaceae (e.g. Dirlewanger et al., 2004). However, the empirical examples of MAS in forest trees are still scarce compared to their fruit crop counterparts. The issues underlying this lack of MAS programs in forest trees are thoroughly discussed elsewhere (Muranty et al., 2014), but causes include their early stage of domestication and the absence of sound cost/benefit analyses supporting the advantage of MAS over other silvicultural techniques. In addition, there are only a few examples of studies that report subsets of markers explaining large parts of the phenotypic variance in forest trees (see above), and their variation is seldom included in breeding programs (at least to our knowledge). Nevertheless, a recent study (Jaramillo-Correa et al., 2015) has

highlighted that MAS in forest trees can be a good alternative to more expensive strategies. In this work, a set of 18 SNPs associated with climate in *P. pinaster* was successfully used to predict survival in a test trial under extreme (hot and dry) conditions, which implies that variation at candidate genes can be effectively included in breeding and reforestation programmes.

The absence of markers explaining large parts of the phenotypic variance in forest trees is closely linked to the low levels of LD previously observed in these taxa, and implies that the number of markers required to pinpoint causative SNPs for a given phenotype is extremely large (Neale & Kremer, 2011). A way to circumvent this problem derives from current high-throughput genotyping technologies applied to populations of limited effective size (i.e. 'training populations'), and focuses on predicting phenotypes based on information from genome-wide markers without the prior knowledge of associations (Meuwissen, Hayes, & Goddard, 2001). Thus, with a dense enough coverage of the genome, many QTLs underlying a trait of interest can be captured in the model through their LD with genome-wide SNPs, without the need to identify individual SNP—trait associations. Simulations performed by assuming parameters similar to those found in forest trees breeding programmes (e.g. LD extent, breeding population sizes and number of QTLs) have shown that a marker density as low as two per cM can be enough to have over 50% of phenotype prediction accuracy for populations with an effective size lower than 30 individuals, even for traits with low heritability (i.e $h^2 = 0.2$) and controlled by as many as 100 QTLs (Grattapaglia & Resende, 2011). These first steps towards the application of genomic selection (GS) in forest trees enlighten a very promising path for tree breeding, as these accuracies are similar to those in programmes currently implemented in livestock (Resende, Muñoz, et al., 2012). For instance, predictions in *Pinus taeda* were highly variable, ranging from 23 to 75% for fusiform rust disease resistance, and growth (e.g. Resende, Muñoz, et al., 2012), while predictions between 50 and 70% were obtained for growth and wood traits in *Picea glauca* (Beaulieu, Doerksen, MacKay, Rainville, & Bousquet, 2014), and for wood quality in *Eucalyptus* (Resende, Resende, et al., 2012). However, given that some of these models were inconsistent across breeding populations, it is likely that, in contrast to MAS, GS might require population-specific predictive models (Beaulieu et al., 2014; Resende, Resende, et al., 2012).

9. PERSPECTIVES AND CONCLUSIONS

Population genomics is being implemented in an increasing range of forest tree species and is now helping relieve the genotyping bottleneck; large genome-wide SNP datasets are being collected that promise to reveal patterns of adaptive diversity with unprecedented detail. The challenge moving forward will be assembling the many fragmented insights provided by GWAS studies on the adaptive architecture of traits into something coherent, not unlike the assembly of many short sequence reads into a finished genome. This will require the integration of increasingly sophisticated approaches in marker-assisted and GS (Grattapaglia & Resende, 2011), including 'multi-omics' data sets (e.g. transcriptomics, proteomics, metabolomics), to generate more holistic predictions of plant phenotypes.

Another focus of the future should be on comparative population genomics of forest trees, looking at the genetic architecture of adaption in parallel across lineages inhabiting similar selection environments or at how changing environments drive genomic modifications in interacting species that develop new phenotypes (and genotypes). The need for this multispecies approach is already being felt, and frameworks such systems genomics or landscape community genomics are being proposed (Barah & Bones, 2015; Hand, Lowe, Kovach, Muhfeld, & Lutkart, 2015); in forest trees these approaches could be implemented to study complex systems such as plant—insect interactions or drought adaptation, particularly for subtropical and tropical taxa, whose study is (again) lagging behind their temperate counterparts.

Connecting variability across the genome to functionally important trait variation will remain challenging, as phenotyping continues to be costly, time and space intensive, and the slow developmental pace of most trees test the patience of researchers. Relieving these bottlenecks will take both technological innovation, and the courage to leave the common garden in favour of in situ phenotyping in the field. On the technology side, high-resolution imaging and remote sensing analysis are promising sources of innovation for high-throughput phenotyping of a diverse array of ecophysiological traits, especially those associated with forest canopies (e.g. Brown et al., 2014). For example, hyperspectral and airborne sensors are, at least in principle, capable of phenotyping phenology and forest structure at the stand scale (Gleason & Im, 2012), as well as other properties of canopy reflectance associated with water stress, leaf nitrogen content, insect herbivory and other traits (Homolová, Malenovský, Clevers, García-Santos, &

Schaepman, 2013). Finer scale data such as LiDAR can provide resolution at the scale of individual trees, yielding phenotypic data on a diversity of structural and growth traits. Field-based sensors are also being developed that can simultaneously measure bud phenology and key environmental cues (e.g. heat sums) directly on individual trees growing in the field (Kleinknecht, Pers. Comm.). Together, this combination of field phenomics and genomics can spur the emergence of models for adaptive trait variation that go beyond genotype–phenotype mapping in trees growing under idealized common garden conditions. With continued innovation in sensor-based phenotyping and computational models for genomic prediction, forest tree genomics should be poised to take genotype-phenotype mapping out of the garden and into the field.

From the analytical point of view, the detection of candidate genes and their integration into predictive models will benefit from incorporating epistatic interactions between alleles. Some approaches are already being developed in this sense and showing promising results in model species data sets (e.g. Berg & Coop, 2014). Furthermore, the study of epigenetic factors implicated in adaptive responses and phenotypic plasticity must also be considered. Previous studies have already shown that epigenetic variation can account for a significant proportion of the variance in adaptation to climate in Norway spruce (Yakovlev, Fossdal, & Johnsen, 2010), while studies on the role of epigenetic factors have significantly expanded during the last decade, particularly in model species (Lisch, 2008). Lastly, forward predictive models that integrate genotypes as fitness predictors after environmental changes at different geographical scales are also being derived (e.g. Ralph & Coop, 2014), showing that local adaptation can quickly develop in species with rather limited dispersal, even when selection pressures are widespread. Such models could be thus readily applied to foretell the fate of modern natural forest tree populations and to establish mitigation measures on the onset of impeding environmental changes, once the drivers of adaptation have been correctly identified at different 'omic levels.

ACKNOWLEDGEMENTS

The authors thank I. Gamache, C. Plomion, and P. Canard for critical comments on the manuscript. This work was supported by grants from the Consejo Nacional de Ciencia y Tecnología (CB-2010-01-00153305) and the Dirección General de Asuntos del Personal Académico (IC200411 and IA201013) from the Universidad Nacional Autónoma de México to JPJ-C and the National Science Foundation (IOS-1461868 to S.R.K.).

REFERENCES

Acheré, V., Favre, J. M., Besnard, G., & Jeandroz, S. (2005). Genomic organization of molecular differentiation in Norway spruce (*Picea abies*). *Molecular Ecology, 14*, 3191–3201.

Akey, J., Jin, L., & Xion, M. (2001). Haplotypes vs single marker linkage disequilibrium tests: what do we gain? *European Journal of Human Genetics, 9*, 291–300.

Alberto, F. J., Derory, J., Boury, C., Frigerio, J.-M., Zimmermann, N. E., & Kremer, A. (2013). Imprints of natural selection along environmental gradients in phenology-related genes of *Quercus petraea*. *Genetics, 195*, 495–512.

Artlip, T. S., Callahan, A. M., Bassett, C. L., & Wisniewski, M. E. (1997). Seasonal expression of a dehydrin gene in sibling deciduous and evergreen genotypes of peach (*Prunus persica* [L.] Batsch). *Plant Molecular Biology, 33*, 61–70.

Audigeos, D., Brousseau, L., Traissac, S., Scotti-Saintagne, C., & Scotti, I. (2013). Molecular divergence in tropical tree populations occupying environmental mosaics. *Journal of Evolutionary Biology, 26*, 529–544.

Avia, K., Kärkkäinen, K., Lagercrantz, U., & Savolainen, O. (2014). Association of *FLOWERING LOCUS T/TERMINAL FLOWER1*-like gene *FTL2* expression with growth rhythm in Scots pine (*Pinus sylvestris*). *New Phytologist, 204*, 159–170.

Barah, P., & Bones, A. M. (2015). Multidimensional approaches for studying plant defence against insects: from ecology to omics and synthetic biology. *Journal of Experimental Botany, 66*, 479–493.

Bashalkhanov, S., Eckert, A. J., & Rajora, O. P. (2013). Genetic signatures of natural selection in response to air pollution in red spruce (*Picea rubens*, Pinaceae). *Molecular Ecology, 22*, 5877–5889.

Beaulieu, J., Doerksen, T., Boyle, B., Clement, S., Deslauriers, M., Beauseigle, S., et al. (2011). Association genetics of wood physical traits in the conifer white spruce and relationships with gene expression. *Genetics, 188*, 197–214.

Beaulieu, J., Doerksen, T. K., MacKay, J., Rainville, A., & Bousquet, J. (2014). Genomic selection accuracies within and between environments and small breeding groups in white spruce. *BMC Genomics, 15*, 1048.

Bedon, F., Bomal, C., Caron, S., Levasseur, C., Boyle, B., Mansfield, S. D., et al. (2010). Subgroup 4 R2R3-MYBs in conifer trees: gene family expansion and contribution to the isoprenoid- and flavonoid-oriented responses. *Journal of Experimental Botany, 61*, 3847–3864.

Benowicz, A., & El-Kassaby, Y. A. (1999). Genetic variation in mountain hemlock (*Tsuga mertensiana* Bong.): quantitative and adaptive attributes. *Forest Ecology and Management, 123*, 205–215.

Berg, J. J., & Coop, G. (2014). A population genetic signal of polygenic adaptation. *PLoS Genetics, 10*, e1004412.

Birol, I., Raymond, A., Jackman, S. D., Pleasance, S., Coope, R., Taylor, G. A., et al. (2013). Assembling the 20 Gb white spruce (*Picea glauca*) genome from whole-genome shotgun sequencing data. *Bioinformatics, 29*, 1492–1497.

Böhlenius, H., Huang, T., Charbonnel-Campaa, L., Brunner, A. M., Jansson, S., Strauss, S. H., et al. (2006). *CO/FT* regulatory module controls timing of flowering and seasonal growth cessation in trees. *Science, 312*, 1040–1043.

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

Bradshaw, H. D., & Stettler, R. F. (1995). Molecular genetics of growth and development in *Populus*. 4. Mapping QTLs with large effects on growth, form, and phenology traits in a forest tree. *Genetics, 139*, 963–973.

Brondani, R. P. V., Williams, E. R., Brondani, C., & Grattapaglia, D. (2006). A microsatellite-based consensus linkage map for species of Eucalyptus and a novel set of 230 microsatellite markers for the genus. *BMC Plant Biology, 6*, 20.

Brown, T. B., Cheng, R., Sirault, X. R. R., Rungrat, T., Murray, K. D., Trtilek, M., et al. (2014). TraitCapture: genomic and environment modelling of plant phenomic data. *Current Opinion in Plant Biology, 18*, 73–79.

Brown, G. R., Gill, G. P., Kuntz, R. J., Langley, C. H., & Neale, D. B. (2004). Nucleotide diversity and linkage disequilibrium in loblolly pine. *Proceedings of the National Academy of Sciences of the United States of America, 101*, 15255–15260.

Burgarella, C., Navascués, M., Zabal-Aguirre, M., Berganzo, E., Riba, M., Mayol, M., et al. (2012). Recent population decline and selection shape diversity of taxol-related genes. *Molecular Ecology, 21*, 3006–3021.

Campbell, R., Pawuk, W. A., & Harris, A. S. (1989). Microgeographic genetic variation of Sitka spruce in southeastern Alaska. *Canadian Journal of Forest Research, 19*, 1004–1013.

Cappa, E. P., El-Kassaby, Y. A., Garcia, M. N., Acuña, C., Borralho, N. M. G., Grattapaglia, D., et al. (2013). Impacts of population structure and analytical models in genome-wide association studies of complex traits in forest trees: a case study in *Eucalyptus globulus*. *PloS One, 8*, e81267.

Casasoli, M., Pot, D., Plomion, C., Monteverdi, M. C., Barreneche, T., Lauteri, M., et al. (2004). Identification of QTLs affecting adaptive traits in *Castanea sativa* Mill. *Plant, Cell & Environment, 27*, 1088–1101.

Chen, J., Källman, T., Ma, X., Gyllestrand, N., Zaina, G., Morgante, M., et al. (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.

Chen, J., Tsuda, Y., Stocks, M., Källman, T., Xu, N., Kärkkäinen, K., et al. (2014). Clinal variation at phenology-related genes in spruce: parallel evolution in *FTL2* and *Gigantea*? *Genetics, 197*, 1025–1038.

Csilléry, K., Lalagüe, H., Vendramin, G. G., González-Martínez, S. C., Sady, B., & Oddou-Muratorio, S. (2014). Detecting short spatial scale local adaptation and epistatic selection in climate-related candidate genes in European beech (*Fagus sylvatica*) populations. *Molecular Ecology, 23*, 4696–4708.

Cullingham, C. I., Cooke, J. E., & Coltman, D. W. (2014). Cross-species outlier detection reveals different evolutionary pressures between sister species. *New Phytologist, 204*, 215–229.

De Kort, H., Vandepitte, K., Bruun, H. H., Closset-Kopp, D., Honnay, O., & Mergay, J. (2014). Landscape genomics and a common garden trial reveal adaptive differentiation to temperature across Europe in the tree species *Alnus glutinosa*. *Molecular Ecology, 19*, 4709–4721.

De Miguel, M., Cabezas, J. A., de María, N., Sánchez-Gómez, D., Guevara, M. A., Vélez, M. D., et al. (2014). Genetic control of functional traits related to photosynthesis and water use efficiency in *Pinus pinaster* Ait. drought response: integration of genome annotation, allele association and QTL detection for candidate gene identification. *BMC Genomics, 15*, 464.

De la Torre, A. R., Roberts, D. R., & Aitken, S. N. (2014). Genome-wide admixture and ecological niche modelling reveal the maintenance of species boundaries despite long history of interspecific gene flow. *Molecular Ecology, 23*, 2046–2059.

De Mita, S., Thuillet, A. C., Gay, L., Ahmadi, N., Manel, S., Ronfort, J., et al. (2013). Detecting selection along environmental gradients: analysis of eight methods and their effectiveness for outbreeding and selfing populations. *Molecular Ecology, 22*, 1383–1399.

Derba-Maceluch, M., Awano, T., Takahashi, J., Lucenius, J., Ratke, C., Kontro, I., et al. (2015). Suppression of xylan endotransglycosylase PtxtXyn10A affects cellulose microfibril angle in secondary wall in aspen wood. *New Phytologist, 205*, 666–681.

Derory, J., Scotti-Saintagne, C., Bertocchi, E., Le Dantec, L., Graignic, N., Jauffres, A., et al. (2010). Contrasting relationships between diversity of candidate genes and variation

of bud burst in natural and segregating populations of European oaks. *Heredity, 104*, 438–448.

Dillon, S., McEvoy, R., Balwin, D. S., Rees, G. N., Parsons, Y., & Southerton, S. (2014). Characterisation of adaptive genetic diversity in environmentally contrasted populations of *Eucalyptus camaldulensis* Dehnh. (river red gum). *PLoS One, 8*, e103515.

Dillon, S. K., Nolan, M., Li, W., Bell, C., Wu, H.-X., & Southerton, S. G. (2009). Allelic variation in cell wall candidate genes affecting solid wood properties in natural populations and land races of *Pinus radiata*. *Genetics, 185*, 1477–1487.

Dirlewanger, E., Graziano, E., Joobeur, T., Garriga-Calderé, F., Cosson, P., Howad, W., et al. (2004). Comparative mapping and marker-assisted selection in Rosaceae fruit crops. *Proceedings of the National Academy of Sciences of the United States of America, 101*, 9891–9896.

Du, Q., Xu, B., Pan, W., Gong, C., Wang, Q., Tian, J., et al. (2013). Allelic variation in a cellulose synthase gene (*PtoCesA4*) associated with growth and wood properties in *Populus tomentosa*. *Genes, Genomes and Genetics, 3*, 2069–2084.

Eckert, A. J., Bower, A. D., Wegrzyn, J. L., Pande, B., Jermstad, K. D., Krutovsky, K. V., et al. (2009). Association genetics of coastal Douglas-fir (*Pseudotsuga menziesii* var. *menziesii*, Pinaceae). I. Cold-hardiness related traits. *Genetics, 182*, 1289–1302.

Eckert, A. J., van Heerwaarden, J., Wegrzyn, J. L., Nelson, C. D., Ross-Ibarra, J., González-Martínez, S. C., et al. (2010). Patterns of population structure and environmental associations to aridity across the range of loblolly pine (*Pinus taeda* L., Pinaceae). *Genetics, 185*, 969–982.

Eckert, A. J., Wegrzyn, J. L., Cumbie, W. P., Goldfarb, B., Huber, D. A., Tolstikov, V., et al. (2012). Association genetics of the loblolly pine (*Pinus taeda*, Pinaceae) metabolome. *New Phytologist, 193*, 890–902.

El Kayal, W., Allen, C. C. G., Ju, C. J. T., Adams, E., King-Jones, S., Zaharia, L. I., et al. (2011). Molecular events of apical bud formation in white spruce, *Picea glauca*. *Plant Cell and Environment, 34*, 480–500.

Evans, L. M., Slavov, G. T., Rodgers-Melnick, E., Martin, J., Ranjan, P., Muchero, W., et al. (2014). Population genomics of *Populus trichocarpa* identifies signatures of selection and adaptive trait associations. *Nature Genetics, 46*, 1089–1096.

Eveno, E., Collada, C., Guevara, M. A., Léger, V., Soto, A., Díaz, L., et al. (2008). Contrasting patterns of selection at *Pinus pinaster* Ait. drought stress candidate genes as revealed by genetic differentiation analyses. *Molecular Biology and Evolution, 25*, 417–437.

Falconer, D. S., & MacKay, T. F. C. (1996). *Introduction to quantitative genetics* (4th ed.). Essex, UK: Pearson Education Limited-Prentice Hall.

Flachowsky, H., Hnake, M.-V., Peil, A., Strauss, S. H., & Fladung, M. (2009). A review on transgenic approaches to accelerate breeding of woody plants. *Plant Breeding, 128*, 217–226.

Fournier-Level, A., Korte, A., Cooper, M. D., Nordborg, M., Schmitt, J., & Wilczek, A. M. (2011). A map of local adaptation in *Arabidopsis thaliana*. *Science, 334*, 86–89.

Frewen, B. E., Chen, T. H. H., Howe, G. T., Davis, J., Rohde, A., Boerjan, W., et al. (2000). Quantitative trait loci and candidate gene mapping of bud set and bud flush in *Populus*. *Genetics, 154*, 837–845.

Geraldes, A., Farzaneh, N., Grassa, C., McKown, A. D., Guy, R. D., Mansfield, S. D., et al. (2014). Landscape genomics of *Populus trichocarpa*: the role of hybridization, limited gene flow, and natural selection in shaping patterns of population structure. *Evolution, 68*, 3260–3280.

Gleason, C. J., & Im, J. (2012). Forest biomass estimation from airborne LiDAR data using machine learning approaches. *Remote Sensing of Environment, 125*, 80–91.

González-Martínez, S. C., Dillon, S., Garnier-Géré, P. H., Krutovsky, K. V., Alía, R., Burgarella, C., et al. (2011). Patterns of nucleotide diversity and association mapping.

In C. Plomion, J. Bousquet, & C. Kole (Eds.), *Genetics, genomics and breeding of conifers* (pp. 239–275). Enfield, USA: Science Publishers-CRC Press.

González-Martínez, S. C., Ersoz, E., Brown, G. R., Wheeler, N. C., & Neale, D. B. (2006). DNA sequence variation and selection of tag single nucleotide polymorphisms at candidate genes for drought-stress response in *Pinus taeda* L. *Genetics, 172*, 1915–1926.

González-Martínez, S. C., Huber, D., Ersoz, E., Davis, J. M., & Neale, D. B. (2008). Association genetics in *Pinus taeda* L. II. Water use efficiency. *Heredity, 101*, 19–26.

González-Martínez, S. C., Wheeler, N. C., Ersoz, E., Nelson, C. D., & Neale, D. B. (2007). Association genetics in *Pinus taeda* L. I. Wood property traits. *Genetics, 175*, 399–409.

Grace, L. J., Charity, J. A., Gresham, B., Kay, N., & Walter, C. (2005). Insect-resistant transgenic *Pinus radiata*. *Plant Cell Reports, 24*, 103–111.

Grattapaglia, D., & Resende, M. D. V. (2011). Genomic selection in forest tree breeding. *Tree Genetics & Genomes, 7*, 241–255.

Grivet, D., Sebastiani, F., Alía, R., Bataillon, T., Torre, S., Zabal-Aguirre, et al. (2011). Molecular footprints of local adaptation in two Mediterranean conifers. *Molecular Biology and Evolution, 28*, 101–116.

Grivet, D., Sebastiani, F., González-Martínez, S. C., & Vendramin, G. G. (2009). Patterns of polymorphism resulting from long-range colonization in the Mediterranean conifer Aleppo pine. *New Phytologist, 184*, 1016–1028.

Groover, A. T. (2005). What genes make a tree a tree? *Trends Plant Sciences, 10*, 210–214.

Guerra, F. P., Wegrzyn, J. L., Sykes, R., Davis, M. F., Stanton, B. J., & Neale, D. B. (2013). Association genetics of chemical wood properties in black poplar (*Populus nigra*). *New Phytologist, 197*, 162–176.

Guichoux, E., Garnier-Géré, P., Lagache, L., Lang, T., Boury, C., & Petit, R. J. (2013). Outlier loci highlight the direction of introgression in oaks. *Molecular Ecology, 22*, 450–462.

Günther, T., & Coop, G. (2013). Robust identification of local adaptation from allele frequencies. *Genetics, 195*, 205–220.

Gyllenstrand, N., Clapham, D., Källman, T., & Lagercrantz, U. (2007). A Norway spruce *FLOWERING LOCUS T* homolog is implicated in control of growth rhythm in conifers. *Plant Physiology, 144*, 248–257.

Hamilton, J. A., Lexer, C., & Aitken, S. N. (2013). Genomic and phenotypic architecture of a spruce hybrid zone (*Picea sitchensis* × *P. glauca*). *Molecular Ecology, 22*, 827–841.

Hamilton, M. G., Williams, D. R., Tilyard, P. A., Pinkard, E. A., Wardlaw, T. J., Glen, M., et al. (2013). A latitudinal cline in disease resistance of a host tree. *Heredity, 110*, 372–379.

Hand, B. K., Lowe, W. H., Kovach, R. P., Muhfeld, C. C., & Luikart, G. (2015). Landscape community genomics: understanding eco-evolutionary processes in complex environments. *Trends in Ecology & Evolution, 30*, 161–168.

Holliday, J. A., Ralph, S. G., White, R., Bohlmann, J., & Aitken, S. N. (2008). Global monitoring of autumn gene expression within and among phenotypically divergent populations of Sitka spruce (*Picea sitchensis*). *New Phytologist, 178*, 103–122.

Holliday, J. A., Ritland, K., & Aitken, S. (2010). Widespread, ecologically relevant genetic markers developed from association mapping of climate-related traits in Sitka spruce (*Picea sitchensis*). *New phytologist, 188*, 501–514.

Homolová, L., Malenovský, Z., Clevers, J. G. P. W., García-Santos, G., & Schaepman, M. E. (2013). Review of optical-based remote sensing for plant trait mapping. *Ecological Complexity, 15*, 1–16.

Hospital, F. (2009). Challenges for effective marker-assisted selection in plants. *Genetica, 136*, 303–310.

Houle, D., Govindaraju, D. R., & Omholt, S. (2010). Phenomics: the next challenge. *Nature Reviews Genetics, 11*, 855–866.

Howe, G. T., Aitken, S. N., Neale, D. B., Jermstad, K. D., Wheeler, N. C., & Chen, T. C. C. (2003). From genotype to phenotype: unraveling the complexities of cold adaptation in forest trees. *Canadian Journal of Botany, 81*, 1247–1266.

Ingvarsson, P. K., García, M. V., Hall, D., Luquez, V., & Jansson, S. (2006). Clinal variation in *phyB2*, a candidate gene for day-length-induced growth cessation and bud set, across a latitudinal gradient in European aspen (*Populus tremula*). *Genetics, 172*, 1845–1853.

Ingvarsson, P. K., Garcia, M. V., Luquez, V., Hall, D., & Jansson, S. (2008). Nucleotide polymorphism and phenotypic associations within and around the phytochrome B2 locus in European aspen (*Populus tremula*, Salicaceae). *Genetics, 178*, 2217–2226.

Ingvarsson, P. K., & Street, N. R. (2011). Association genetics of complex traits in plants. *New Phytologist, 189*, 909–922.

Isik, F. (2014). Genomic selection in forest tree breeding: the concept and an outlook to the future. *New Forests, 45*, 379–401.

Jaramillo-Correa, J. P., Beaulieu, J., & Bousquet, J. (2001). Contrasting evolutionary forces driving population structure at expressed sequence tag polymorphisms, allozymes and quantitative traits in white spruce. *Molecular Ecology, 10*, 2729–2740.

Jaramillo-Correa, J. P., Rodríguez-Quilón, I., Grivet, D., Lepoittevin, C., Sebastiani, F., Heuertz, M., et al. (2015). Molecular proxies for climate maladaptation in a long-lived tree (*Pinus pinaster* aiton, Pinaceae). *Genetics, 199*, 793–807.

Jennewein, S., Wildung, M. R., Chau, M., Walker, K., & Croteau, R. (2004). Random sequencing of an induced *Taxus* cell cDNA library for identification of clones involved in taxol biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America, 101*, 9149–9154.

Jermstad, K. D., Bassoni, D. L., Wheeler, N. C., Anekonda, T. S., Aitken, S. N., Adams, W. T., et al. (2001). Mapping of quantitative trait loci controlling adaptive traits in coastal Douglas-fir. II. Spring and fall cold-hardiness. *Theoretical and Applied Genetics, 102*, 1152–1158.

Joost, S., Bonin, A., Bruford, M. W., Després, L., Conord, C., Erhardt, G., et al. (2007). A spatial analysis method (SAM) to detect candidate loci for selection: towards a landscape genomics approach to adaptation. *Molecular Ecology, 16*, 3955–3969.

Källman, T., De Mita, S., Larsson, H., Gyllenstrand, N., Heuertz, M., Parducci, L., et al. (2014). Patterns of nucleotide diversity at photoperiod related genes in Norway Spruce [*Picea abies* (L.) Karst.]. *PloS One, 9*, e95306.

Kang, H. M., Zaitlen, N. A., Wade, C. M., Kirby, A., Heckerman, D., Dalyk, M. J., et al. (2008). Efficient control of population structure in model organism association mapping. *Genetics, 178*, 1709–1723.

Kremer, A., & Le Corre, V. (2012). Decoupling of differentiation between traits and their underlying genes in response to divergent selection. *Heredity, 108*, 375–385.

Keller, S. R., Levsen, N., Ingvarsson, P. K., Olson, M. S., & Tiffin, P. (2011). Local selection across a latitudinal gradient shapes nucleotide diversity in balsam poplar, *Populus balsamifera* L. *Genetics, 188*, 941–952.

Keller, S. R., Levsen, N., Olson, M. S., & Tiffin, P. (2012). Local adaptation in the flowering-time gene network of balsam poplar, *Populus balsamifera*. *Molecular Biology and Evolution, 29*, 3143–3152.

Keller, G., Marchal, T., SanClemente, H., Navarro, M., Ladouce, N., Wincker, P., et al. (2009). Development and functional annotation of an 11,303-EST collection from *Eucalyptus* for studies of cold tolerance. *Tree Genetics & Genomes, 5*, 317–327.

Külheim, C., Yeoh, S. H., Maintz, J., Foley, W. J., & Moran, G. F. (2009). Comparative SNP diversity among four Eucalyptus species for genes from secondary metabolite biosynthetic pathways. *BMC Genomics, 10*, 452.

Latta, R. G. (1998). Differentiation of allelic frequencies at quantitative trait loci affecting locally adaptive traits. *The American Naturalist, 153*, 283–292.

Le Corre, V., & Kremer, A. (2003). Genetic variability at neutral markers, quantitative trait loci and trait in a subdivided population under selection. *Genetics, 164*, 2005–2019.

Leinonen, T., McCairns, R. J. S., O'Hara, R. B., & Merilä, J. (2013). Q_{ST}–F_{ST} comparisons: evolutionary and ecological insights from genomic heterogeneity. *Nature Reviews Genetics, 14*, 179–190.

Leinonen, T., O'Hara, R. B., Cano, J. M., & Merilä, J. (2008). Comparative studies of quantitative trait and neutral marker divergence: a meta-analysis. *Journal of Evolutionary Biology, 21*, 1–17.

Lespinasse, D., Grivet, L., Troispoux, V., Rodier-Goud, M., Pinard, F., & Seguin, M. (2000). Identification of QTLs involved in the resistance to South American leaf blight (*Microcyclus ulei*) in the rubber tree. *Theoretical and Applied Genetics, 100*, 975–984.

Lesur, I., Le Provost, G., Bento, P., Da Silva, C., Leplé, J.-C-, Murat, F., et al. (2015). The oak gene expression atlas: insights into Fagaceae genome evolution and the discovery of genes regulated during bud dormancy release. *BMC Genomics, 16*, 112.

Levsen, N. D., Tiffin, P., & Olson, M. S. (2012). Pleistocene speciation in the genus *Populus* (salicaceae). *Systematic Biology, 61*, 401–412.

Li, P., Beaulieu, J., Corriveau, A., & Bousquet, J. (1993). Genetic variation in juvenile growth and phenology in a white spruce provenance-progeny test. *Silvae Genetica, 42*, 52–60.

Li, P., Beaulieu, J., Daoust, G., & Plourde, A. (1997). Patterns of adaptive genetic variation in eastern white pine (*Pinus strobus*) from Quebec. *Canadian Journal of Forest Research, 27*, 199–206.

Li, C., Sun, Y., Huang, H. W., & Cannon, C. H. (2014). Footprints of divergent selection in natural populations of *Castanopsis fargesii* (Fagaceae). *Heredity, 113*, 533–541.

Lisch, D. R. (2008). Epigenetic regulation of transposable elements in plants. *Annual Review of Plant Biology, 60*, 43–66.

Lowry, D. B., Hall, M. C., Salt, D. E., & Willis, J. H. (2009). Genetic and physiological basis of adaptive salt tolerance divergence between coastal and inland Mimulus guttatus. *New Phytologist, 183*, 776–788.

Luikart, G., England, P. R., 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.

Ma, X.-F., Hall, D., Onge, K. R. S., Jansson, S., & Ingvarsson, P. K. (2010). Genetic differentiation, clinal variation and phenotypic associations with growth cessation across the *Populus tremula* photoperiodic pathway. *Genetics, 186*, 1033–1044.

Mageroy, M. H., Parent, G., Germanos, G., Giguère, I., Delvas, N., Maaroufi, H., et al. (2015). Expression of the β-glucosidase gene Pgβglu-1 underpins natural resistance of white spruce against spruce budworm. *The Plant Journal, 81*, 68–80.

Mandrou, E., Gherardi Hein, P. R., Villar, E., Vigneron, P., Plomion, C., & Gion, J.-M. (2012). A candidate gene for lignin composition in Eucalyptus: *cinnamoyl-CoA reductase (CCR). Tree Genetics and Genomes, 8*, 353–364.

Marguerit, E., Bouffier, L., Chancerel, E., Costa, P., Lagane, F., Guehl, J.-M., et al. (2014). The genetics of water-use efficiency and its relation to growth in maritime pine. *Journal of Experimental Botany, 65*, 4757–4768.

McKay, J. K., & Latta, R. G. (2002). Adaptive population divergence: markers, QTL and traits. *Trends in Ecology & Evolution, 17*, 285–291.

McKown, A. D., Guy, R. D., Klápště, J., Geraldes, A., Friedman, M., Cronk, Q. C. B., et al. (2014). Geographical and environmental gradients shape phenotypic trait variation and genetic structure in *Populus trichocarpa. New Phytologist, 201*, 1263–1276.

McKown, A., Klápště, J., Guy, R., Geraldes, A., Porth, I., Hannemann, J., et al. (2014). Genome-wide association implicates numerous genes underlying ecological trait variation in natural populations of *Populus trichocarpa*. *New Phytologist, 203*, 535–553.

Meuwissen, T., Hayes, B., & Goddard, M. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics, 157*, 1819–1829.

Mimura, M., & Aitken, S. N. (2007). Adaptive gradients and isolation-by-distance with post-glacial migration in *Picea sitchensis*. *Heredity, 99*, 224–232.

Morgenstern, E. K. (1969). Genetic variation in seedlings of *Picea mariana* (Mill.) BSP: II. Variation patterns. *Silvae Genetica, 18*, 161–167.

Mosca, E., Eckert, A. J., Pierro, Di, Rocchini, D., La Porta, N., Bellletti, P., et al. (2012). The geographical and environmental determinants of genetic diversity for four alpine conifers of the European Alps. *Molecular Ecology, 21*, 5530–5545.

Muranty, H., Jorge, V., Bastien, C., Lepoittevin, C., Bouffier, L., & Sánchez, L. (2014). Potential for marker-assisted selection for forest tree breeding: lessons from 20 years of MAS in crops. *Tree Genetics and Genomes, 10*, 1491–1510.

Myburg, A. A., Grattapaglia, D., Tuskan, G. A., Hellsten, U., Hayes, R. D., Grimwood, J., et al. (2014). The genome of *Eucalyptus grandis*. *Nature, 509*, 356–362.

Namroud, M.-C., 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.

Namroud, M.-C., Guillet-Claude, C., Mackay, J., Isabel, N., & Bousquet, J. (2010). Molecular evolution of regulatory genes in spruces from different species and continents: heterogeneous patterns of linkage disequilibrium and selection but correlated recent demographic changes. *Journal of Molecular Evolution, 70*, 371–386.

Neale, D. B., & Kremer, A. (2011). Forest tree genomics: growing resources and applications. *Nature Reviews Genetics, 12*, 111–122.

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

Neale, D. B., Wegrzyn, J., Stevens, K. A., Zimin, A. V., Puiu, D., Crepeau, M. W., et al. (2014). Decoding the massive genome of loblolly pine using haploid DNA and novel assembly strategies. *Genome Biology, 15*, R59.

Nei, M. (1977). F-statistics and analysis of gene diversity in subdivided populations. *Annals of Human Genetics, 41*, 225–233.

Niklas, K. J. (1997). *The evolutionary biology of plants*. Chicago: Univ. Chicago Press.

Nystedt, B., Street, N. R., Wetterbom, A., Zuccolo, A., Lin, Y. C., Scofield, D. G., et al. (2013). The Norway spruce genome sequence and conifer genome evolution. *Nature, 497*, 579–584.

Olsen, J. E., Junttila, O., Nilsen, J., Eriksson, M. E., Martinussen, I., Olsson, O., et al. (1997). Ectopic expression of oat phytochrome A in hybrid aspen changes critical daylength for growth and prevents cold acclimatization. *The Plant Journal, 12*, 1339–1350.

Olson, M. S., Levsen, N., Soolanayakanahally, R. Y., Guy, R. D., Schroeder, W. R., Keller, S. R., et al. (2013). The adaptive potential of *Populus balsamifera* L. to phenology requirements in a warmer global climate. *Molecular Ecology, 22*, 1214–1230.

Palmé, A. E., Pyhäjärvi, T., Wachowiak, W., & Savolainen, O. (2009). Selection on nuclear genes in a *Pinus* phylogeny. *Molecular Biology and Evolution, 26*, 893–905.

Paolucci, I., Gaudet, M., Jorge, V., Beritognolo, I., Terzoli, S., Kuzminsky, E., et al. (2010). Genetic linkage maps of *Populus alba* L. and comparative mapping analysis of sex determination across *Populus* species. *Tree Genetics & Genomes, 6*, 863–875.

Parchman, T. L., Gompert, Z., Mudge, J., Schilkey, F. D., Benkman, C. W., & Buerkle, C. A. (2012). Genome-wide association genetics of an adaptive trait in lodgepole pine. *Molecular Ecology, 21*, 2991–3005.

Parelle, J., Zapater, M., Scotti-Saintagne, C., Kremer, A., Jolivet, Y., Dreyer, E., et al. (2007). Quantitative trait loci of tolerance to waterlogging in a European oak (*Quercus robur* L.): physiological relevance and temporal effect patterns. *Plant, Cell & Environment, 30*, 422–434.

Pavy, N., Paule, C., Parsons, L., Crow, J. A., Morency, M-. J., Cooke, J., et al. (2005). Generation, annotation, analysis and database integration of 16 500 white spruce EST clusters. *BMC Genomics, 6*, 144.

Pelgas, B., Beauseigle, S., Achere, V., Jeandroz, S., Bousquet, J., Isabel, N., et al. (2006). Comparative genome mapping among *Picea glauca, P. mariana* × *P. rubens* and *P. abies*, and correspondence with other Pinaceae. *Theoretical and Applied Genetics, 113*, 1371–1393.

Pelgas, B., Bousquet, J., Meirmans, P. G., 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.

Peña, L., & Séguin, A. (2001). Recent advances in the genetic transformation of trees. *Trends in Biotechnology, 19*, 500–506.

Petit, R. J., & Hampe, A. (2006). Some evolutionary consequences of being a tree. *Annual Review of Ecology, Evolution and Systematics, 37*, 187–214.

Porth, I., Hamberger, B., White, R., & Ritland, K. (2011). Defense mechanisms against herbivory in *Picea*: sequence evolution and expression regulation of gene family members in the phenylpropanoid pathway. *BMC Genomics, 12*, 608.

Porth, I., Klápště, J., Skyba, O., Lay, B. S. K., Geraldes, A., Muchero, W., et al. (2013). *Populus trichocarpa* cell wall chemistry and ultrastructure trait variation, genetic control and genetic correlations. *New Phytologist, 197*, 777–790.

Porth, I., Klápště, J., Skyba, O., Hannemann, J., McKown, A. D., Guy, R. D., et al. (2013). Genome-wide association mapping for wood characteristics in *Populus* identifies an array of candidate single nucleotide polymorphisms. *New Phytologist, 200*, 710–726.

Porth, I., Koch, M., Berenyi, M., Burg, A., & Burg, K. (2005). Identification of adaptation-specific differences in mRNA expression of sessile and pedunculate oak based on osmotic-stress-induced genes. *Tree Physiology, 25*, 1317–1329.

Pot, D., McMillan, L., Echt, C., Le Provost, G., Garnier-Géré, P., Cato, S., et al. (2005). Nucleotide variation in genes involved in wood formation in two pine species. *New Phytologist, 167*, 101–112.

Prunier, J., Gerardi, S., Beaulieu, J., & Bousquet, J. (2012). Parallel and lineage-specific molecular adaptation to climate in boreal black spruce. *Molecular Ecology, 21*, 4270–4286.

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. *Molecular Ecology, 20*, 1702–1716.

Prunier, J., Pelgas, B., Gagnon, F., Desponts, M., Isabel, N., Beaulieu, J., et al. (2013). The genomic architecture and association genetics of adaptive characters using a candidate SNP approach in boreal black spruce. *BMC Genomics, 14*, 368.

Ralph, P. L., & Coop, G. (2014). The role of standing variation in geographic convergent adaptation. *bioRxiv*, 009803. http://dx.doi.org/10.1101/009803.

Rehfeldt, G. E., Ying, C. C., Spittlehouse, D. L., & Hamilton, D. A. (1999). Genetic responses to climate in *Pinus contorta*: niche breadth, climate change, and reforestation. *Ecological Monographs, 69*, 375–407.

Resende, M. F. R., Jr., Muñoz, P., Acosta, J. J., Peter, G. F., Davis, J. M., Grattapaglia, D., et al. (2012). Accelerating the domestication of trees using genomic selection: accuracy of prediction models across ages and environments. *New Phytologist, 193*, 617–624.

Resende, M. D. V., Resende, M. F. R., Jr., Sansaloni, C. P., Petroli, C. D., Missiaggia, A. A., Aguiar, A. M., et al. (2012). Genomic selection for growth and wood quality in Eucalyptus: capturing the missing heritability and accelerating breeding for complex traits in forest trees. *New Phytologist, 194*, 116–128.

Risterucci, A., Paulin, D., Ducamp, M., N'goran, J., & Lanaud, C. (2003). Identification of QTLs related to cocoa resistance to three species of *Phytophthora*. *Theoretical and Applied Genetics, 108*, 168–174.

Ritland, K., Krutovsky, K. V., Tsumura, Y., Pelgas, B., Isabel, N., & Bousquet, J. (2011). Genetic mapping in conifers. In C. Plomion, J. Bousquet, & C. Kole (Eds.), *Genetics, genomics and breeding of conifers* (pp. 196–238). Enfield, USA: Science Publishers-CRC Press.

Ruttink, T., Arend, M., Morreel, K., Storme, V., Rombauts, S., Fromm, J., et al. (2007). A molecular timetable for apical bud formation and dormancy induction in poplar. *The Plant Cell Online, 19*, 2370–2390.

Savolainen, O., Lascoux, M., & Merilä, J. (2013). Ecological genomics of local adaptation. *Nature Reviews Genetics, 14*, 807–820.

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

Scalfi, M., Troggio, M., Piovani, P., et al. (2004). A RAPD, AFLP and SSR linkage map, and QTL analysis in European beech (*Fagus sylvatica* L.). *Theoretical and Applied Genetics, 108*, 433–441.

Scotti-Saintagne, C., Mariette, S., Porth, I., Goicoechea, P. G., Barreneche, T., Bodénès, C., et al. (2004). Genome scanning for interspecific differentiation between two closely related oak species [*Quercus robur* L. and *Q. petraea* (Matt.) Liebl.]. *Genetics, 168*, 1615–1626.

Séguin, A., Lachance, D., Désjardins, A., Leplé, J.-C., & Pilate, G. (2014). Scientific research related to genetically modified trees. In T. Fenning (Ed.), *Challenges and opportunities for the World's forests in the 21st century*. Dordrecht: Springer.

Soto-Cerda, B., & Cloutier, S. (2010). Association mapping in plant genomes. In M. Çalişkan (Ed.), *Genetic diversity in plants* (pp. 29–54). Rijeka: InTech.

Srinivasan, C., Dardick, C., Callahan, A., & Scorza, R. (2012). Plum (*Prunus domestica*) trees transformed with poplar *FT1* result in altered architecture, dormancy requirement, and continuous flowering. *PLoS One, 7*, e40715.

Stölting, K. N., Nipper, R., Lindtke, D., Caseys, C., Waeber, S., Castiglioni, S., et al. (2013). Genomic scan for single nucleotide polymorphisms reveals patterns of divergence and gene flow between ecologically divergent species. *Molecular Ecology, 22*, 842–855.

Thumma, B. R., Nolan, M. F., Evans, R., & Moran, G. F. (2005). Polymorphisms in *Cinnamoyl CoA Reductase (CCR)* are associated with variation in microfibril angle in *Eucalyptus* spp. *Genetics, 171*, 1257–1265.

Tiffin, P., & Ross-Ibarra, J. (2014). Advances and limits of using population genetics to understand local adaptation. *Trends Ecology and Evolution, 29*, 673–680.

Tsumura, Y., Kado, T., Takahashi, T., Tani, N., Ujino-Ihara, T., & Iwata, I. (2007). Genome scan to detect genetic structure and adaptive genes of natural populations of *Cryptomeria japonica*. *Genetics, 176*, 2393–2403.

Tsumura, Y., Uchiyama, K., Moriguchi, Y., Ueno, S., & Ihara-Ujino, T. (2012). Genome scanning for detecting adaptive genes along environmental gradients in the Japanese conifer, *Cryptomeria japonica*. *Heredity, 109*, 349–360.

Tuskan, G. A., DiFazio, S., Jansson, S., Bohlman, J., Grigoriev, I., Hellsten, U., et al. (2006). The genome of black cottonwood, *Populus trichocharpa* (Torr. & Gray). *Science, 313*, 1596–1604.

Uchiyama, K., Iwata, H., Moriguchi, Y., Ujino-Ihara, T., Ueno, S., Tagushi, Y., et al. (2013). Demonstration of genome-wide association studies for identifying markers for wood property and male strobili traits in *Cryptomeria japonica*. *PloS One, 8*, e79866.

Wegrzyn, J. L., Eckert, A. J., Choi, M., Lee, J. M., Stanton, B. J., Sykes, R., et al. (2010). Association genetics of traits controlling lignin and cellulose biosynthesis in black cottonwood (*Populus trichocarpa*, Salicaceae) secondary xylem. *New Phytologist, 188*, 515–532.

Whiteman, A. (2014). Global trends and outlook for forest resources. In T. Fenning (Ed.), *Challenges and opportunities for the World's forests in the 21st century*. Dordrecht: Springer.
Wright, S. (1951). The genetical structure of populations. *Annals of Eugenics, 15*, 323–354.
Wright, S. I., & Gaut, B. S. (2005). Molecular population genetics and the search for adaptive evolution in plants. *Molecular Biology and Evolution, 22*, 506–519.
Xu, S. (2003). Theoretical basis of the Beavis effect. *Genetics, 165*, 2259–2268.
Yakovlev, I. A., Fossdal, C. G., & Johnsen, Ø. (2010). MicroRNAs, the epigenetic memory and climate adaptation in Norway spruce. *New Phytologist, 187*, 1154–1169.
Yu, J., Pressoir, G., Briggs, W. H., Bi, I. V., Yamasaki, M., Doebley, J. F., et al. (2006). A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nature Genetics, 38*, 203–208.
Zhu, M., & Zhao, S. (2007). Candidate gene identification approach: progress and challenges. *International Journal of Biological Sciences, 3*, 420–427.