Annals of Forest Science (2016) 73:77–103
DOI 10.1007/s13595-015-0488-3

REVIEW PAPER

Forest tree genomics: 10 achievements from the past 10 years and future prospects

Christophe Plomion[1,2] · Catherine Bastien[3] · Marie-Béatrice Bogeat-Triboulot[4,5] · Laurent Bouffier[1,2] · Annabelle Déjardin[3] · Sébastien Duplessis[6] · Bruno Fady[7] · Myriam Heuertz[1,2,8] · Anne-Laure Le Gac[9] · Grégoire Le Provost[1,2] · Valérie Legué[10,11] · Marie-Anne Lelu-Walter[3] · Jean-Charles Leplé[3] · Stéphane Maury[9] · Alexandre Morel[3] · Sylvie Oddou-Muratorio[7] · Gilles Pilate[3] · Léopoldo Sanchez[3] · Ivan Scotti[7] · Caroline Scotti-Saintagne[7] · Vincent Segura[3] · Jean-François Trontin[12] · Corinne Vacher[1,2]

Received: 30 October 2014 / Accepted: 18 May 2015 / Published online: 10 June 2015

Abstract

• ***Key message*** **This review highlights some of the discoveries and applications made possible by "omics" technologies over the last 10 years and provides perspectives for pioneering research to increase our understanding of tree biology.**

• ***Context*** A decade after the first forest tree genome sequence was released into the public domain, the rapidly evolving genomics and bioinformatics toolbox has advanced our understanding of the structure, functioning, and evolution of forest tree genomes.

• ***Aims and methods*** This review highlights some of the discoveries and applications that "omics" technologies have made possible for forest trees over the past 10 years.

• ***Results*** In this review, we start by our current understanding of genome evolution and intricacies of gene regulation for reproduction, development, and responses to biotic and abiotic stresses. We then skim over advances in interactome analysis and epigenomics, the knowledge of the extent of genetic variation within and between species, revealing micro- and macro-evolutionary processes and species history, together with the complex architecture of quantitative traits. We finally end with applications in genetic resource conservation and breeding.

• ***Conclusion*** The knowledge gained through the use of these technologies has a huge potential impact for adapting forests to the main challenges they will have to face: changing demand from ecosystem services with potentially conflicting

Handling Editor: Jean-Michel Leban

Contribution of the co-authors CP coordinated this review, defined its content, wrote the abstract, introduction, and conclusion sections, and edited the whole manuscript. The following co-authors contributed to the following sections: JCL (2.1), GP (2.2.1), GLP (2.2.2), VL (2.2.3), MALW, AM, and JFT (2.2.4), MBBT (2.3.1), CB and SD (2.3.2), CV (2.3.3), AD (2.4), ALLG and SM (2.5), SOM and IS (2.6), MH and CSS (2.7), VS (2.8), BF (2.9), and LB and LS (2.10). All authors have read and approved the final manuscript.

Electronic supplementary material The online version of this article (doi:10.1007/s13595-015-0488-3) contains supplementary material, which is available to authorized users.

✉ Christophe Plomion
plomion@pierroton.inra.fr

Catherine Bastien
Catherine.Bastien@orleans.inra.fr

Marie-Béatrice Bogeat-Triboulot
triboulo@nancy.inra.fr

Laurent Bouffier
bouffier@pierroton.inra.fr

Annabelle Déjardin
Annabelle.Dejardin@orleans.inra.fr

Sébastien Duplessis
duplessi@nancy.inra.fr

Bruno Fady
bruno.fady@avignon.inra.fr

Myriam Heuertz
heuertzm@gmail.com

Anne-Laure Le Gac
annelaure.legac@gmail.com

Grégoire Le Provost
gregoire@pierroton.inra.fr

strategies in terms of conservation and use, as well as climate changes and associated threats. Genomics will undoubtedly play a major role over the next decade and beyond, not only to further understand the mechanisms underlying adaptation and evolution but also to develop and implement innovative management and policy actions to preserve the adaptability of natural forests and intensively managed plantations.

Keywords Genomics · Forest trees · Developmental biology · Epigenetics · Interactome · Micro-evolution · Breeding and conservation

1 Introduction

Forest system currently faces a number of major challenges, including increases in the demand for wood, pressure to conserve forest areas, and global climate change and associated threats. The adaptation of these systems in response to these challenges will require a multifaceted approach, in which genomic sciences have an important role to play. The aim is to apply a series of technologies from the fields of genetics, molecular and cell biology fostered by bioinformatics, and robotics to analyses of the structure, function, and evolution of sets of genes, right up to complete genomes, with the use of high-throughput methods. These methods include the structural characterization of entire genomes, genes, mRNA, and proteins, the genome mapping and sequencing (genomics), the study of mRNA, proteins, and metabolite abundances across different environmental conditions and/or developmental stages (functional genomics, i.e., transcriptomics, proteomics, metabolomics), the analysis of epigenetic modifications (epigenomics), the inference of evolutionary mechanisms (comparative and population genomics related to macro- and micro-evolutionary processes, respectively), and the study of complex biological systems (community genomics, metagenomics). Since the publication of the first tree genome sequence (black cottonwood, Tuskan et al. 2006), technologies and genetic resources facilitating research into forest tree genomics have advanced our understanding of the tree growth and development (resulting in large and long-lived organisms), the responses of trees to intrinsic (ontogenic phase change) and extrinsic (biotic and abiotic) factors, the remarkable buffering capacity (plasticity) of trees, enabling them to cope with chronic stresses and extreme events, the molecular basis of genetic variation within and between species, and the way in which this variation has been shaped by evolutionary forces and its relationship to phenotypic variation and adaptation. The objective of this review, which is not intended to be

Valérie Legué
valerie.LEGUE@univ-bpclermont.fr

Marie-Anne Lelu-Walter
Marie-anne.Lelu-Walter@orleans.inra.fr

Jean-Charles Leplé
Jean-Charles.Leple@orleans.inra.fr

Stéphane Maury
stephane.maury@univ-orleans.fr

Alexandre Morel
morel.alx@gmail.com

Sylvie Oddou-Muratorio
sylvie.muratorio@avignon.inra.fr

Gilles Pilate
Gilles.Pilate@orleans.inra.fr

Léopoldo Sanchez
Leopoldo.Sanchez@orleans.inra.fr

Ivan Scotti
Ivan.Scotti@paca.inra.fr

Caroline Scotti-Saintagne
Caroline.Scotti@paca.inra.fr

Vincent Segura
vincent.segura@orleans.inra.fr

Jean-François Trontin
Jean-Francois.TRONTIN@fcba.fr

Corinne Vacher
corinne.vacher@pierroton.inra.fr

[1] INRA, UMR1202 BIOGECO, 33612 Cestas, France

[2] University of Bordeaux, BIOGECO, UMR 1202, 33615 Pessac, France

[3] INRA, UR0588 AGPF, Amélioration, Génétique et Physiologie Forestières, 2163 Avenue de la Pomme de Pin, CS 40001 Ardon, 45075 Cedex 2, Orléans, France

[4] INRA, UMR Ecologie et Ecophysiologie Forestière, 25420 Champenoux, France

[5] UMR Ecologie et Ecophysiologie Forestière, Université de Lorraine, BP 239, 54506 Vandoeuvre, France

[6] INRA, IAM, UMR 1136, 54280 Champenoux, France

[7] INRA, UR629, URFM, Ecologie des Forêts Méditerranéennes, Domaine St Paul, 84914 Avignon, France

[8] INIA Forest Research Centre, Carretera de A Coruña km 7.5, 28040 Madrid, Spain

[9] USC 1328 INRA, Laboratoire de Biologie des Ligneux et des Grandes Cultures, University of Orléans EA 1207, 45067 Orléans, France

[10] Clermont Université, Université Blaise-Pascal, UMR 547 PIAF, BP 10448, 63000 Clermont-Ferrand, France

[11] INRA, UMR 547 PIAF, 63100 Clermont-Ferrand, France

[12] FCBA, Pôle Biotechnologies et Sylviculture avancée, 71 route d'Arcachon, 33610 Cestas, France

exhaustive, is to outline the tremendous progress achieved in forest tree genomics over the last 10 years. This progress is illustrated by considering a series of 10 achievements. We will begin by skimming over the knowledge acquired from the sequencing of forest tree genomes. We will then move on to the key developmental traits underlying the biology of woody perennials, the molecular mechanisms driving the responses of trees to biotic and abiotic stresses, and we will tackle two emerging fields (molecular interactions, epigenetics) that promise to improve our understanding of the functioning of tree genomes considerably. Finally, we consider the knowledge gained from the description and interpretation of naturally occurring genetic variation within and between species. This knowledge has provided us with an understanding of the contemporary and historical evolutionary processes that have contributed to the observed patterns of geographic and phenotypic variation and the molecular basis of quantitative trait variation. We will conclude by considering the potential impact of genomic studies on the conservation and improvement of forest tree genetic resources. This review is accessible to readers from diverse backgrounds keen to acquire a basic understanding of the opportunities for tacking the complex issues facing the world's forests provided by this discipline. A glossary of terms for readers that are not necessarily familiar with genomics or other specialized jargon was added as electronic supplemental material (ESM_1.doc).

2 Accomplishments in forest tree genomics

2.1 What have we learnt from the sequencing of tree genomes?

Over the last 10 years, forest tree genomics has benefited considerably from advances in next-generation sequencing technologies, making it possible to investigate the role of hundreds of genes, to access sequence-based markers for breeding at the genome scale, and to study the evolutionary history of tree species (Neale and Kremer 2011). Even if the sequencing of forest tree genomes lags far behind that of fruit trees and annual crops (http://en.wikipedia.org/wiki/List_of_sequenced_plant_genomes), more than 20 tree genome sequencing projects are currently underway (Neale et al. 2013), and six completed forest tree genomes have yet been released and published. In 2002, the US Department of Energy (DOE, USA) set up an international initiative to sequence the *Populus trichocarpa* genome, the first tree genome (Tuskan et al. 2006) and only the third plant genome, after *Arabidopsis thaliana*, the model dicot for plant biology (Arabidopsis Genome Initiative 2000), and rice, the major economic crop that became a model monocot (Goff et al. 2002; Yu et al. 2002; IRGSP 2005). An additional poplar species has since been sequenced, *Populus euphratica* (Ma et al. 2013), together with a second angiosperm tree species, *Eucalyptus* (Myburg et al. 2014), and three conifers, Norway spruce (Nystedt et al. 2013), white spruce (Birol et al. 2013), and loblolly pine (Neale et al. 2014; Zimin et al. 2014). In the first section of this review, we will briefly summarize the new findings obtained by exploring these first forest tree genomes, in terms of specific aspects of tree biology, such as angiosperm/gymnosperm wood formation and life history traits.

The *Populus* genome sequence provided the first insight into the genome structure and functional biology of a specific tree. The poplar genome is about four times larger than that of *Arabidopsis*. This larger size results mostly from a complex history of whole-genome duplications, chromosomal rearrangements, and tandem duplications, as shown by Tuskan et al. (2006). The poplar genome contains 1.6 times as many genes as the *Arabidopsis* genome, and the expansion of several gene families associated with tree-specific traits may also have contributed to the evolution of tree biology. The annual formation of wood is just one of a number of processes exclusive to trees (Plomion et al. 2001). Several genes associated with cellulose or lignin biosynthesis were found to occur in duplicated pairs in *Populus*, whereas only single copies were present in *Arabidopsis*. For example, the oxidative polymerization of monolignols, the precursors of lignins, involves two different, non-redundant types of oxidative enzymes—laccases and peroxidases—as recently demonstrated for *Arabidopsis* vascular development (Zhao et al. 2013a). The laccase gene family of *Populus* is much larger than that of *Arabidopsis* (51 versus 17 genes, respectively; Berthet et al. 2012). Moreover, phylogenetic and expression studies of these two multigene families showed that the 16 orthologs of the three *Arabidopsis* lignin-related laccases (AtLAC4, AtLAC11, and AtLAC17) present in poplar were expressed mostly in the stem xylem, but with some exceptions (expression in the root or other organs), suggesting a conservation of function, but with subfunctionalization for some duplicated poplar genes.

If we take a closer look at the process of wood formation, we find that secondary xylem development evolved between gymnosperms and angiosperms. One of the two main differences between these two groups is the lack of S units in conifer lignins. All the genes involved in monolignol biosynthesis in angiosperms were identified in the pine genome, except for the F5H homolog, which is crucial for the biosynthesis of S units in plants (Bonawitz and Chapple 2010, for review; Neale et al. 2014). The second difference concerns water transport and mechanical support functions, both of which are mediated principally by tracheids in conifers, whereas angiosperms have fibers for mechanical support and vessels for water transport. The number of genes encoding vascular-related NAC domain transcription factors, key regulators of xylem vessel differentiation in angiosperms, is much smaller (only two

genes identified) in Norway spruce (Nystedt et al. 2013) than in *Arabidopsis* and poplar, in which 7 and 16 such genes, respectively, have been identified (Ohtani et al. 2011) suggesting a possible expansion and subfunctionalization of the NAC domain transcription factor family in angiosperms, particularly in woody angiosperms.

Unlike annual plants, which have a short lifespan, trees have had to develop long-term defense strategies against insects and pathogens. Interestingly, although the total number of NBS-LRR pathogen resistance genes is similar in *Arabidopsis* and poplar, some subfamilies are considerably larger in *Populus* (Tuskan et al. 2006; Kohler et al. 2008; Bresson et al. 2011). Interestingly, the TIR class of NB-LRR proteins (TNLs) is also much larger in conifers (780 genes in loblolly pine and 180 genes in Norway spruce) than in *Arabidopsis* (3 genes).

In addition to revealing functional differences between genomes, these recent works have increased our understanding of the life histories of angiosperm/gymnosperm trees. *Populus*, like *Eucalyptus* and grapevine, has undergone at least one whole-genome duplication (WGD) event. Analysis of the *Populus* genome led to the identification of three separate WGD events. The most recent burst of gene creation happened 60 to 65 million years (Myr) ago, due to a single whole-genome event. A second duplication, in a common ancestor of *Populus* (Fabid, also known as eurosid I) and *Arabidopsis* (Malvid, also known as eurosid II), was found to have occurred about 100 to 120 Myr ago. A weaker signal was also found for a third, more ancient duplication event, but this event was not dated by the authors (it probably corresponds to the palaeohexaploidy event, see below). A comparison of the estimated molecular clocks of *Populus* and *Arabidopsis* revealed a markedly slower rate of sequence evolution in *Populus*, possibly due to its perennial status, leading to recurrent contributions of "ancient gametes" from old individuals (Tuskan et al. 2006). An analysis of the recently sequenced *Eucalyptus* genome and phylogenetic studies based on 17 species have suggested that *Populus* should be reclassified as a Malvid and that *Eucalyptus* should be placed in a sister taxon to the eurosids (Myburg et al. 2014). A study of the evolutionary history of the *Eucalyptus* genome also revealed the ancient palaeohexaploidy event (~130 to 150 Myr), which has also been discovered in grapevine and is common to all eudicots (Jaillon et al. 2007). This event has also been shown to be superimposed over a more recent lineage-specific palaeotetraploidy event (~110 Myr). WGD is an important mechanism of genome size expansion. As conifers have genomes about 20 to 30 times larger than those of angiosperms, it was intriguing that no evidence for such a mechanism was found in Norway spruce (Nystedt et al. 2013); only the trace of a very ancient WGD predating the divergence of angiosperms and gymnosperms was detected by the authors (~350 Myr), consistent with previous phylogenetic analysis (Jiao et al. 2011). Instead, it appears that conifer genomes grew mostly through the insertion of repeated elements, principally long terminal repeat-retrotransposons (LTR-RT; essentially the Ty3/Gypsy and Ty1/Copia superfamilies) (Nystedt et al. 2013). The authors proposed a model for genome evolution in conifers (six conifer species were investigated), in which retrotransposon activity began early in evolution and was not countered as efficiently as in angiosperms, resulting in larger gene/pseudogene numbers, and numerous long introns, with genes separated by large regions of transposable element-rich, highly polymorphic DNA in conifers but with the maintenance of synteny over large phylogenetic distances.

2.2 Developmental genomics

As reported in Section 2.1, the development of genomics-enabled research has been instrumental in the identification of genes underlying the traits characteristic of the biology of woody perennials. In this section, we review the progress made by the forest tree genomics community towards elucidating wood formation, bud dormancy, root development in interaction with microbial symbionts, and embryo development in conifers.

2.2.1 Wood formation

Wood or secondary xylem formation results from cambium activity. The secondary xylem fulfills highly specialized functions critical for tree growth and development, such as water conduction from the roots to the crown and mechanical support. Xylem cells—vessels, fibers, and tracheids—specialize in one or both of these functions. These cells differentiate, developing thick and lignified secondary cell walls, before undergoing programmed cell death. In addition, ray cells generated by the cambium connect the phloem to the inner wood. They have storage functions and are also involved in heartwood formation. Trees can give rise to different types of wood—juvenile and mature wood, early and late wood, and reaction wood—with very different properties, depending on their stage of development or in response to environmental cues (Déjardin et al. 2010; Plomion et al. 2001). Wood formation is, therefore, a highly regulated process, and our understanding of which has greatly improved, thanks to advances in genomics.

Genomic studies of wood formation in angiosperms and gymnosperms began more than 15 years ago (Allona et al. 1998; Sterky et al. 1998), with comparisons of transcriptomes and proteomes between different types of wood, leading to the establishment of correlations between the expression of a number of specific genes and specific wood characteristics, such as cell wall composition and properties (Plomion et al. 2000; Déjardin et al. 2004; Gion et al. 2005; Paux et al. 2005;

Andersson-Gunneras et al. 2006; Qiu et al. 2008; Paiva et al. 2008a, b; Villalobos et al. 2012).

Furthermore, causal links were established through studies of natural mutants or, more frequently, genetically modified (GM) plants with modified expression of genes involved in cell wall component metabolism: mostly for lignin (van Holme et al. 2013; van Acker et al. 2014) but also for cellulose (Coleman et al. 2007; Coleman et al. 2009), hemicelluloses (Park et al. 2004; Baba et al. 2009; Nishikubo et al. 2011; Zhao et al. 2013b), pectins (Biswal et al. 2014), and other cell wall proteins, such as expansin (Gray-Mitsumune et al. 2008). Nowadays, such modified plants would ideally be characterized by global omics analyses, leading to the establishment of relationships between the phenotype and networks of molecular components, and this approach will undoubtedly lead to the construction of increasingly realistic models explaining wood formation.

The regulation of wood formation has been investigated in a number of studies, leading to the identification of transcription factors involved in secondary cell wall differentiation: this is the case, for example, for ARBORKNOX2 and KNAT7, two *Populus* homeobox genes (Du et al. 2009; Li et al. 2012), and for *PtrMYB152* (Wang et al. 2014a), *PtrHB7*, a class III HD-Zip gene (Zhu et al. 2013b), and *FPF1*, a gene also involved in the regulation of flowering (Hoenicka et al. 2012). Likewise, Zhong et al. (2011) demonstrated a key role of PtrWNDs, wood-associated NAC domain transcription factors, as master switches regulating a battery of downstream transcription factors forming a transcriptional network controlling secondary cell wall biosynthesis during wood formation. As described in Section 2.4, the conjunction of several complementary techniques, such as xylem protoplast transfection, RNA-Seq, and ChIP-Seq, has paved the way for the identification of hierarchical gene regulatory networks directed by master transcription factors in wood formation (Lin et al. 2013).

MicroRNAs (miRNAs) have also been implicated in the regulation of wood formation. These miRNAs are short noncoding RNAs with critical regulatory functions. In a pioneering study, Lu et al. (2005) described a number of miRNAs regulated in wood in response to tension and compression stresses. The use of next-generation sRNA sequencing has led to the identification of several miRNAs that are particularly abundant in the xylem: these miRNAs are predicted to target genes known to be important in secondary growth, including the critical reaction wood enzyme xyloglucan endotransglycosylase/hydrolase and vascular system-related transcription factors (Puzey et al. 2012). Furthermore, transgenic poplar trees expressing either a miRNA-resistant POPCORONA or a synthetic miRNA targeting POPCORONA have been used to infer the function of POPCORONA during secondary growth: the synthetic miRNA-mediated knockdown of POPCORONA expression resulted in abnormal lignification in pith cells, whereas the overexpression of a miRNA-resistant POPCORONA delayed the lignification of xylem and phloem fibers during secondary growth. The misexpression of POPCORONA was also shown to result in a coordinated change in gene expression within a previously described transcriptional network regulating cell differentiation and cell wall biosynthesis and in the expression of hormone-related genes associated with fiber differentiation (Du et al. 2011).

In parallel to these important advances in molecular biology, progress has also been made towards the high-throughput phenotyping of cell wall properties. Optimization of the MS-based sequencing of lignin oligomers (Morreel et al. 2010), together with the phenol profiling of plants with lignin modifications, has provided a clear overview of the lignin biosynthesis gene network (van Holme et al. 2012a, b). Chemometric analyses of 2D NMR spectra for cell walls have made it possible to identify changes in cell wall components (Hedenström et al. 2009). Likewise, in situ images of the different chemotypes present in poplar cell walls can be obtained by MicroFTIR analysis (Gorzsás et al. 2011). Finally, a wealth of information has been generated by the development of cell wall polymer mapping with microarrays, using specific monoclonal antibodies and cellulose-binding modules (Moller et al. 2007).

The last few years have seen major technical improvements, leading to the generation of huge amounts of data and opening up new possibilities for network analysis. For example, text mining, co-expression network analysis, and comparative genomics are providing an ever-increasing number of opportunities to identify candidate genes for cell wall biosynthesis (Yang et al. 2011). A genome-wide metabolic pathway database was recently created for *P. trichocarpa* with pathway analysis tools (Zhang et al. 2010). In the future, a combination of microdissection and high-throughput analysis should make it possible to acquire large amounts of data for increasingly precisely defined samples corresponding to single cell types or to very specific developmental steps. Finally, studies of interactions between molecular components should greatly advance our understanding of wood formation.

2.2.2 Vegetative bud phenology

Perennial plants are immobile organisms that cannot migrate to cope with unfavorable winter conditions. They have developed a strategy for synchronizing their growth and reproductive phases with the favorable environmental conditions. This strategy, dormancy, enables trees both to protect themselves from cold injuries and to ensure an appropriate architecture (reviewed by Preston and Sandve 2013). The term "dormancy" was defined in 1987 by Lang et al. (1987) as a phase during which growth is temporarily suspended, in any plant structure containing a meristem (i.e., meristematic cells). In trees,

dormancy occurs mostly in the buds or vascular cambium and can be divided into three main phases according to the factors controlling growth cessation (Chao et al. 2007). The first kind of dormancy is called paradormancy. During paradormancy, growth is stopped by physiological factors external to the affected structure (i.e., apical dominance or correlative inhibition). Paradormancy is followed by "winter dormancy," which is divided into two main stages: endodormancy, the deepest phase of dormancy (reviewed by Horvath et al. 2003) induced by a decrease in day length, and ecodormancy, which is imposed by unfavorable environmental conditions. Endodormancy is maintained by internal factors specific to the bud itself. Endodormancy generally begins at the end of the summer and steadily increases in intensity during the fall, peaking at some time in November, the precise timing of the peak depending on the species considered (Naor et al. 2003). Endodormancy is broken when chilling requirements are fulfilled. Ecodormancy is established at the release of endodormancy. Ecodormancy occurs during late winter and early spring and is imposed exclusively by environmental factors unfavorable for growth (essentially cold temperatures).

Dormancy (i.e., endo- and ecodormancy) is tightly controlled by both photoperiod and temperature. Photoperiod strongly influences the induction of dormancy, whereas temperature is involved in its release (for both endodormancy and ecodormancy). Temperature thus plays a major role in the phenological cycle of these species. Global warming may strongly affect phenology, because increases in temperatures may extend the growing seasons of trees (reviewed by Menzel and Fabian 1999) and prevent endodormancy release if chilling requirements are not fulfilled during early winter. They may also favor ecodormancy release by accelerating bud cell growth during late winter and/or early spring. There is therefore an urgent need to decipher the mechanisms underpinning dormancy and to identify the genes/polymorphisms that matter for adaptation, because these changes may have severe effects in forest ecosystems, increasing the risk of early frost damage or of exposure to new pathogens.

Molecular mechanisms involved in dormancy induction and release remain poorly characterized in forest trees. Gene expression profiling has been performed in poplar (Rhode et al. 2007), spruce (Yakovlev et al. 2006), oak (Ueno et al. 2013), peach (Bassett et al. 2006), and apricot (Yamane et al. 2008) and has led to the identification of a set of candidate genes for dormancy regulation. In peach, Bielenberg et al. (2008), using the ever-growing mutant, have shown that MADS box genes were relevant candidate genes for dormancy regulation. Indeed, the ever-growing mutant is characterized by continuous growth of its apical meristem and does not respond to short-day signaling or low temperature. Ruttinkt et al. (2007) carried out a combination of transcriptome and metabolome profiling in poplar, to obtain the first molecular time table of apical bud formation and dormancy induction in a forest tree species. More recently, Ueno et al. (2013) provided the first insight into the gene networks involved in endo- and ecodormancy in European white oaks. They reported that genes overexpressed during endodormancy were related to dehydration, high light intensity, and abscisic acid, whereas those most strongly overexpressed during ecodormancy were related to metal ion binding, cellular transition, and fatty acid binding.

2.2.3 Root development

In woody perennial species, the optimal adaptation of root architecture to the soil is crucial to ensure solid anchorage of the plant in the soil and the efficient acquisition of water and nutrients. The mature root system of trees has a typical root architecture, including primary and lateral roots, resulting from the integration of multiple environmental signals such as symbiosis associations and water availability. The need to integrate signals from multiple pathways, therefore, complicates the dissection of the transduction pathways involved in root development. Only a few transcript profiling analyses have been conducted in woody plants, to investigate the events regulated during root development, and there is still no transcriptional roadmap. The development of genomic and transcriptomic tools, such as EST sequencing, microarrays, and next-generation sequencing technologies, and their application to different plant–fungal ectomycorrhizal associations have resulted in highly valuable information being obtained, including the identification of key genes involved in root architecture. For example, the research performed on the poplar/*Laccaria bicolor* association (Felten et al. 2009) and, more recently, on the oak/*Pisolithus tinctorius* association (Sebastiana et al. 2014) has improved our understanding of auxin-dependent pathways in roots.

Transcriptome analysis requires the detailed and precise annotation of sequenced plant genomes and thus provides an opportunity to identify genes regulated differently in the plant studied. The genomic sequences of *P. trichocarpa* (Tuskan et al. 2006) and, more recently, of *E. grandis* (Myburg et al. 2014) have been used to identify the members of gene families and to compare entire gene families between species. For example, some studies have reported a lack of expansion in poplar or in eucalyptus of certain gene families involved in hormone homeostasis and signal transduction. This is the case for families encoding cytokinin-related enzymes (Ramírez-Carvajal et al. 2008) and *auxin response factor* (*ARF*) gene family (Yu et al. 2014). Conversely, genome-wide analysis of the *Populus* PIN (auxin efflux transporters) family highlighted a more diversified expansion of this family in *Populus* than in *Arabidopsis*, indicating a potential role of these transporters in tree growth and development and, more importantly, in the development of roots and leaves (Liu et al. 2014).

Transcriptomic analyses performed at different stages of adventitious root formation in woody plants have revealed significant transcriptome remodeling during the formation of adventitious roots in *Pinus taeda* (Brinker et al. 2004), *Populus* sp. (Ramírez-Carvajal et al. 2009; Rigal et al. 2012), and *Eucalyptus grandis* (Abu-Abied et al. 2012) and allowed the identification of key genes involved in these developmental events. For example, recent studies on several woody species have generated interesting data, strongly implicating transcription factors in the successive steps of adventitious root establishment (reviewed by Legué et al. 2014) including GRAS (Sanchez et al. 2007) and AP2/ERF families (Ramírez-Carvajal et al. 2009; Rigal et al. 2012).

We also need to determine the links between changes in gene expression and alterations of biochemical and physiological functions and, ultimately, root development and the way in which the expression patterns of different genes are interconnected. The determination of gene function will require the construction and production of transgenic tree lines. This will involve a considerable technical effort on the part of laboratories. Several ambitious projects managed by several teams have been successfully carried out, providing the scientific community with a number of lines in which gene expression has been enhanced or decreased by RNA interference and gene tagging (Busov et al. 2005, 2010). Studies aiming to identify genes on the basis of their expression patterns, using enhancer-trap and gene-trap insertion lines, are also very useful, but fewer studies of this type have been carried out.

The transcriptional findings and integrative databases already available constitute a fundamental resource for future studies of molecular events and for the identification of key proteins involved in developmental processes. In parallel, transgenic lines are crucial for functional genomics studies and should be made more readily available to the research community.

2.2.4 Conifer embryo development

Conifers are the primary source for wood production worldwide (Canales et al. 2014). By unraveling the complexity of the regulated gene network involved in conifer embryo development, it should be possible to develop genomic and epigenomic tools for the early selection of improved varieties. This is a critical issue to face rapid socioeconomic and environmental changes. Plant growth can be determined early during embryogenesis (Yakovlev et al. 2014). This knowledge is also required to develop efficient clonal propagation methods of selected trees, such as somatic embryogenesis (Klimaszewska et al. 2011). This process has great potential for the deployment of new varieties in plantation forestry (Lelu-Walter et al. 2013; Klimaszewska et al. 2015).

It is difficult to sample manageable quantities of embryogenic masses during early zygotic embryogenesis (de Vega-Bartol et al. 2013; Elhiti et al. 2013). As somatic embryos (SEs) closely mimic zygotic embryos (ZEs) during maturation from early to late embryogenesis, they are considered as a model in vitro system to study the molecular biology of embryo development in conifers (Vestman et al. 2011; Yakovlev et al. 2014). Classical genetic approaches for the identification of embryogenesis-related genes are impracticable in conifers, due to their long generation time and large genome size. Since the discovery of somatic embryogenesis 30 years ago (Hakman et al. 1985; Klimaszewska et al. 2015), our knowledge of regulated genes in this system has thus remained highly fragmented and mostly based on expression studies of a few candidate genes (ESM_2A.doc). Current somatic embryogenesis protocols also essentially resulted from tedious "trial and error" strategies with low inputs from molecular studies. Microarray and RNA sequencing methods have recently provided critical advances for the genome-wide profiling of gene expression. Transcriptomics is developing rapidly in conifers, with the recent advent of large genomic resources (Lorenz et al. 2012; Raherison et al. 2012; Canales et al. 2014), including draft genomes for *Picea* (Birol et al. 2013; Nystedt et al. 2013) and *Pinus* (Neale et al. 2014).

Transcriptomic profiling is generating a growing body of information about coordinated gene expression during conifer embryo development (ESM_2B.doc, reviewed in Trontin et al. 2015). The number of transcribed genes appears to be 30 to 40 % larger than in any other tissue (Cairney and Pullman 2007; Yakovlev et al. 2014) with high relevance to the gene network in *A. thaliana* (300–450 genes, Cairney and Pullman 2007; Zhang et al. 2012a). Transcript profiles during zygotic embryogenesis are highly correlated between *Pinus pinaster* and *A. thaliana*, with only 3 % of the transcripts estimated to be gymnosperm-specific (de Vega-Bartol et al. 2013). Differences between angiosperms and gymnosperms are thought to arise principally from spatiotemporal variations in gene expression resulting partly from epigenetic modifications, which may act as an adaptive mechanism in such long-lived species (Cairney and Pullman 2007; Vestman et al. 2011; de Vega-Bartol et al. 2013). Yakovlev et al. (2014) specifically reported temperature-dependent differential transcriptomes in *Picea abies* embryogenic masses potentially associated with the formation of an epigenetic memory, with a delayed, persistent impact on seedling growth. Zhang et al. (2012b) also demonstrated the widespread occurrence of microRNAs in *Larix kaempferi* embryogenic masses, with predicted target genes involved in SE development. Transcriptome profiling confirmed that important processes are conserved in higher plants, including the apical–basal embryo patterning driven by polar auxin transport and the activation of the auxin-mediated response machinery during radial embryo patterning (Vestman et al. 2011; de Vega-Bartol et al. 2013). Transcriptomics has also highlighted the complexity of the processes and genes involved in the spatiotemporal development of conifer embryos

from embryogenic induction (Elhiti et al. 2013; Rutledge et al. 2013) to the switch from embryonic phase to vegetative growth (Stasolla et al. 2003, 2004; Vestman et al. 2011; Morel et al. 2014a). An impressive picture of coordinated functions and genes has been obtained for SE maturation in *P. abies* (Stasolla et al. 2003, 2004; Vestman et al. 2011) and ZE formation in *P. pinaster* (de Vega-Bartol et al. 2013). It should be possible to model embryo development, by interpreting and integrating the large transcript, protein, and metabolite datasets (Kell et al. 2005; Vanderschuren et al. 2013; Wolfender et al. 2013). However, proteomic and metabolomic studies are currently scarce for conifers (ESM_2B.doc, reviewed in Trontin et al. 2015). There are still many limitations, at both experimental (Abril et al. 2011) and interpretational levels (Lippert et al. 2005; Saghatelian and Cravatt 2005; Teyssier et al. 2011, 2014). A recent study of embryo development in *P. pinaster* demonstrated the applicability of integrated approaches for the production of robust data (Morel et al. 2014a).

"Omics" approaches could offer practical offshoots, such as diagnostic tools for checking embryogenic potential or embryo quality. Various miRNAs with stage-specific expression have been described in *L. kaempferi*, suggesting possible modulation of embryogenic potential (Zhang et al. 2012b; Li et al. 2014a). The metabolic signature has been shown to accurately predict embryogenic potential in *P. taeda* (Robinson et al. 2009). Transcriptomics in *P. taeda* (Pullman et al. 2003) and proteomics in *P. pinaster* (Morel et al. 2014b) have provided strong evidence of differences between SEs and fully mature ZEs. Transcriptomics has also proved to be of practical value in *Picea glauca* to check SE quality in different maturation conditions (Stasolla et al. 2003). The transcriptomic and proteomic profiling of early maturing SEs in *P. pinaster* has also yielded robust diagnostic tools for detecting disturbances in pathways critical for normal embryo development (ESM_2C.doc, Morel et al. 2014a).

2.3 Molecular mechanisms involved in biotic and abiotic stress responses

We illustrate in this section how genomic technologies have not only improved our understanding of the structure and evolution of forest tree genomes but have also provided a suitable platform for obtaining knowledge about the molecular mechanisms involved in responses to biotic and abiotic cues. Considering that trees are not standing by themselves, we also considered in this section emerging researches aiming at drawing a holistic picture of the interactions between forest trees and their microbiome.

2.3.1 Abiotic stresses

Abiotic stresses decrease the growth and productivity of crops and forests. The physiological mechanisms of acclimation to environmental stresses have been extensively studied, but the analysis of their molecular bases started more recently (a decade ago in forest trees). Understanding these molecular mechanisms is of particular relevance in the frame of climate change, to achieve more rapid genetic gains in abiotic stress resistance by molecular breeding. As pointed out by Dubos et al. (2003), the molecular mechanisms involved in stress responses in trees, which have a long life cycle and specific tissues, should be considered separately and in addition to those of crops and model plants. Over the last 10 years, the number of studies analyzing the molecular basis of the response/acclimation/adaptation of trees to abiotic stresses has increased steadily. The release of the *P. trichocarpa* genome (Tuskan et al. 2006) led to poplar becoming the model tree species for functional genomics. About three quarters of "omics" publications concerning the response of trees to abiotic stresses concern a wide range of species from the genus *Populus* (ESM_3.xls). The remaining studies concern eucalyptus, pine, and, to a lesser extent, oak, beech, and Douglas fir. The context of global change may account for water deficit being the stress most frequently studied in functional genomics. Salinity, high atmospheric CO_2 concentration, hypoxia, heat, cold, ozone, nitrogen deprivation, and metal toxicity have also been studied, to a lesser extent.

Comparative approaches have been widely used through "omics" or quantitative trait locus (QTL) approaches. The genotype specificity of transcriptomic, proteomic, or metabolomic responses has been highlighted by many studies. Comparing six genotypes of *Populus balsamifera*, Hamanishi et al. (2010) showed that the growth response to drought was correlated with genetic responsiveness. In *P. deltoides* × *nigra*, the phenotypic response to moderate drought was found to be very similar for a stress-tolerant and a stress-sensitive genotype, and the differences in transcriptional responses were therefore attributed to intrinsic divergences in genome functioning (Cohen et al. 2010). In a similar study on *Eucalyptus*, the drought-tolerant hybrid displayed changes in the expression of a broader set of genes in response to water deficit, and the stress signaling cascade differed between the two genotypes studied (Villar et al. 2011). Similar genotype × environment interactions have been reported for proteomic studies (Bonhomme et al. 2009; Xiao et al. 2009; Bedon et al. 2012). The results obtained in a field experiment suggested that the better maintenance of productivity during the dry season by a drought-tolerant *Eucalyptus* genotype involved cell wall modification, ROS detoxification, and osmoregulation (Bedon et al. 2012). A comparison of the metabolomes of two *Eucalyptus* species showed that many low-abundance compounds may help plants to cope with water stress through non-osmotic functions (Warren et al. 2012). Comparisons of the transcriptomes and metabolomes of mature unstressed leaves from *P. euphratica* (salt-tolerant) and *Populus canescens* (salt-sensitive) suggested that the evolutionary

adaptation of *P. euphratica* to saline environments involved the permanent activation of control mechanisms for osmotic adjustment, ion compartmentalization, and the detoxification of reactive oxygen species (Janz et al. 2010). Other molecular players and networks involved in acclimation to abiotic stresses are detailed in a recent review by Harfouche et al. (2014).

Few studies have investigated the effects of releasing an environmental stress. In *P. euphratica*, full re-irrigation after water deficit was found to lead to the recovery of most phenotypic traits and the reversal of transcriptional changes (Bogeat-Triboulot et al. 2007). In a similar experiment on oak, recovery following re-irrigation occurred in two steps, in which the observed transcriptional remodeling was consistent with physiology and growth (Spiess et al. 2012).

Other studies have provided interesting information about less studied abiotic stresses. Large differences were found between poplar and *Arabidopsis*, which is flood-sensitive, in terms of metabolite and transcript patterns in response to hypoxia, accounting for the ability of poplar to maintain its carbon and energy metabolism and, thus, its flood tolerance (Kreuzwieser et al. 2009). A QTL analysis associated with a study of transcriptional responses to ozone showed the involvement of key genes relating to ethylene production and response (Street et al. 2011). As for model plants, most molecular studies of the response of trees to abiotic stresses have been conducted on leaf material (ESM_3.xls). Those comparing transcriptomic or proteomic changes in different organs have highlighted the tissue specificity of the response (Bogeat-Triboulot et al. 2007; Cohen et al. 2013; Bedon et al. 2012). The proteins and genes identified are potential markers and targets for molecular breeding, but the diverse requirements for protecting and maintaining the function of different plant organs may render the engineering of stress tolerance in plants more difficult (Polle et al. 2006).

More attention should be paid to the molecular aspects of wood response, including xylem hydraulic adaptation to salt stress (Janz et al. 2012), drought and embolism (Berta et al. 2010; Secchi et al. 2011), and cadmium accumulation in bark and phytoremediation (He et al. 2013). Greenhouse and laboratory experiments are useful as they allow the control of environmental conditions, but field studies remain exception, despite their high value, and this approach should be developed further (Villar et al. 2011; Pandey et al. 2013). In addition to transcriptomic, proteomic, and metabolomic changes, epigenetic responses and non-coding microRNAs contribute to acclimation or adaptation to abiotic stresses. There is currently a lack of studies in this area (see Section 2.5), and further research is required (Harfouche et al. 2014). More integrated research will be necessary, if we are to unravel the complex molecular mechanisms and pathways underlying responses to abiotic stresses (Castell and Ernst 2012; Harfouche et al. 2014). The release of additional completed tree genome sequences (Myburg et al. 2014; Zimin et al. 2014) and next-generation sequencing should expand the range of research and help us to decipher the molecular mechanisms underlying the response of forest trees to abiotic stresses.

2.3.2 Biotic stresses

Several disturbances to natural forest ecosystems and forest plantations due to outbreaks of insect or pathogens have had major impacts, at the regional scale, on timber production or tree population sustainability (Sturrock et al. 2011). With current predictions of climate change, tree populations may become even more susceptible to outbreaks of existing and new pests and diseases. Structural and functional genomic approaches in forest trees have revealed a large diversity in genetic control of induced responses in the host and the complexity of molecular communications between the two partners (Duplessis et al. 2009) which both contrast with the Flor's gene-for-gene model.

Like many other plants, forest trees have evolved two strategies for recognizing microorganisms including pathogens, symbionts, and endophytes (Guttman et al. 2014). Hereafter, the focus will be laid on local and systemic responses induced by a pathogen attack. Conserved microbial elicitors (PAMPs) and more specific pathogen effectors are recognized by host receptor proteins (PRR) present in the plasma membrane of tree cells and by-products of disease resistance genes (R genes) present in the cytosol, respectively (Jones and Dangl 2006). Several published studies on forest trees have focused on the effector-triggered immunity (ETI) initiated by these R genes, which leads to the hypersensitive response and a disease-free phenotype at whole-plant level (Rinaldi et al. 2007; Liu et al. 2013). Transcriptome analyses for the poplar–*Melampsora* interaction have shown that the defense reaction is triggered by specific signaling systems and includes the accumulation of transcripts encoding pathogenesis-related proteins (PRs), gluthathione *S*-transferases (GSTs), and a rust-induced secreted protein (RISP) specific to *Populus* (Duplessis et al. 2009).

Several such qualitative resistances with oligogenic control and a major impact on damage levels were genetically mapped for different forest pathosystems soon after the development of the first genetic maps: *Pinus* sp.–*Cronartium* sp. (Wilcox et al. 1996), *Eucalyptus–Puccinia* (Junghans et al. 2004), and *Poplar* sp.–*Melampsora* sp. (Newcombe et al. 1996). R genes are frequently overcome by pathogen populations (Kinloch et al. 2004; Dowkiw et al. 2010). Forest tree pathologists and breeders have, therefore, focused on the genetic and molecular basis of quantitative resistance, which is thought to be under more complex genetic control and to be more durable. The results of many fine mapping, candidate region sequencing, genome annotation, and transcriptional studies have blurred the distinction between qualitative and quantitative resistances: (i) major QTLs for quantitative resistance can account for more than 40 % of the observed

phenotypic variation in field or laboratory experiments (Jorge et al. 2005; Freeman et al. 2008); (ii) in artificial inoculation tests controlling for pathogen diversity, some major QTLs display significant strain specificity, resembling that of R genes (Dowkiw et al. 2010); (iii) qualitative resistance factors from *Populus deltoides* and quantitative resistance factors against leaf rust inherited from *P. trichocarpa* and *Salix viminalis* have been fine-mapped to genomic regions rich in R genes (Bresson et al. 2011; Samils et al. 2011); and (iv) the difference in the timing of the activation of similar defense responses between susceptibility and partial or full resistance in the poplar–*Melampsora* interaction is consistent with the signal conversion model described for *Arabidopsis* (Nimchuk et al. 2003; Duplessis et al. 2009).

Mapping studies have revealed a major contribution of both additive and non-additive (epistasis) genetic variation to disease resistance, supporting the hypothesis of complex interaction and possible successful clonal selection in which all genetic effects can be readily captured (Jorge et al. 2005; Alves et al. 2012). Thanks to association genetics, La Mantia et al. (2013) identified five variants in orthologs of *Arabidopsis* genes with known functions in plant defense each accounting for smaller proportions of phenotypic variation for leaf rust severity in *P. trichocarpa.*,

For tree–insect interactions, the availability of reference genomes for *Populus* and *Eucalyptus* led to attention being focused on comparative analyses of the PR protein-encoding genes induced by different pathogens and insects in forest trees (reviewed by Veluthakkal and Dasgupta 2010). Analyses of the diversity of the nucleotide sequences encoding protease inhibitors (PIs) in *P. balsamifera* and *Populus tremula* revealed few signs of selection but differences in adaptive histories both within a single species and between closely related species (Neiman et al. 2009; Bernhardsson and Ingvarsson 2012).

In association with other omics studies, the completion of several new tree genomes and of new pathogen and insect genomes should contribute in the near future to the discovery of processes underlying severe disturbances such as mountain pine beetle attack involving different pine host species, a bark beetle species, and a tree-killing fungus (http://www.thetriaproject.ca/).

2.3.3 Trees are holobionts

Almost all plant tissues harbor microorganisms (Turner et al. 2013). Trees do not escape the rule (Hacquard and Schadt 2015) and can be considered as "superorganisms" or holobionts (Margulis 1991). Over the past decade, high-throughput sequencing technologies have unraveled the huge diversity of microbial communities associated to trees and forest ecosystems. Buée et al. (2009), Öpik et al. (2009), and Jumpponen and Jones (2009) pioneered in this field by using 454 pyrosequencing of barcode regions to study fungal diversity associated to forest soils, roots, and leaves, respectively. They described these communities in unprecedented detail because high-throughput sequencing technologies enable the detection of non-cultivable microorganisms (Hibbett et al. 2009). They opened the way to numerous studies highlighting variations in tree-associated microbial communities at various spatial scales. For instance, a fine-scale study showed that bacterial communities of oak rhizosphere differ from those of the surrounding soil (Uroz et al. 2010). Less contrasted results were obtained for fungal communities of beech rhizosphere (Coince et al. 2013). Fine-scale studies were also conducted for above-ground communities. For instance, Cordier et al. (2012a) showed within-canopy variations in foliar fungal communities of beech. Leff et al. (2015) obtained similar results in the case of bacterial communities associated to leaf and bark of *Gingko biloba* trees. By opening the microbial world to ecologists, the next-gen revolution also gave rise to larger scale studies, aimed at assessing the influence of climate and other global change components on tree-associated microbial communities. The analysis of beech-associated fungal communities along elevation gradients suggested that the air temperature is a major structuring factor of foliar communities (Cordier et al. 2012b) but not of their below-ground counterparts (Coince et al. 2014). A comparison between urban and non-urban stands revealed significant effects of anthropogenic activities on foliar fungal communities of *Quercus macrocarpa* (Jumpponen and Jones 2010). High-throughput sequencing technologies have also fostered the study of the relationship between the genetic variability of trees and the variability of their microbial communities. There is now a large body of evidence showing that microbial communities are influenced by the genetic variability of host trees, both at intra-specific (Redford et al. 2010; Cordier et al. 2012a; Bálint et al. 2013) and inter-specific (Redford et al. 2010; Kembel et al. 2014; Kembel and Mueller 2014) levels. A current challenge is to assess relative effects of tree genotype and environment by using common-garden experiments (Bálint et al. 2015) and to decipher the genetic architecture of microbial communities. Significant advances have been made recently on this latter topic on the model plant *A. thaliana*. The use of mutant lines and genome-wide association mapping revealed that plant loci responsible for defense, cell wall integrity, and cuticular wax composition influence the foliar microbiota (Riesberg et al. 2013; Horton et al. 2014). Another challenge is to go beyond the identification of the genes and environmental factors structuring microbial communities, by elucidating the effects of these hyperdiverse communities on tree growth, health, and reproduction. We are indeed far from a proper understanding of the outcomes of the interactions between microorganisms and their host (Borer et al. 2013), even for model tree species such as poplars (Hacquard and Schadt 2015). It is thus time to move from meta-barcoding

approaches to meta-"omics" approaches enabling the identification of gene transcripts, proteins, and metabolites expressed by tree-associated microbial communities. Such functional approaches are still rare for microbial communities associated to plants and soils (Knief et al. 2012; Damon et al. 2012), and deserve further investigation.

2.4 Molecular interactions

The completion and annotation of plant genome sequences have revealed that the functions of most genes are still unknown. As most proteins function in macromolecular complexes, the identification of protein partners or other molecular interactors (nucleic acids, carbohydrates, lipids, etc.) is particularly relevant when assigning a function to a given protein. Post-genomic high-throughput studies have been carried out, to accelerate functional studies through the identification of plant interactomes, i.e., the whole set of molecular interactions in a given cell or tissue, including both physical interactions between molecules and indirect interactions between genes (genetic interactions). For example, the first experimental *Arabidopsis* interactome (AI-1) led to the identification of 2,700 proteins and 6,200 interactions, generating hypotheses about the molecular functions of several thousand unknown proteins (Braun et al. 2011). The molecular interactions identified can then be displayed as networks. Depending on the underlying biological question and the data available, these networks may illustrate transcriptional and post-transcriptional regulations (transcription factor/DNA or miRNA/target gene interactions), signal transduction pathways, or fluxes in metabolic pathways. Most traits of interest are quantitative and, therefore, particularly suitable for modeling by network approaches.

Both high-throughput experimental techniques and computational predictions are required to decipher the interactome. Experimental data remain very scarce in the plant kingdom, with the exception of the model plant *Arabidopsis*. Indeed, several large-scale protein interaction studies have been carried out with different experimental approaches, mostly based on yeast two-hybrid (Y2H) screening with a cloned ORFeome and tandem affinity purification/mass spectrometry (for a review, see Braun et al. 2013). The *Arabidopsis* interactome was further expanded by prediction methods based on statistical learning methodology and/or the transfer of interaction annotation based on homology with other species (human, yeast, nematode, fruitfly) (De Bodt et al. 2009; Geisler-Lee et al. 2007; Lin et al. 2011). In tree species, no large-scale systematic study has yet been undertaken. However, for *P. trichocarpa*, a dedicated biomass ORFeome was cloned from 374 selected ORFs found to be more strongly expressed in xylem tissues than in phloem and was used in a binary Y2H or a Y2H cDNA library screening: interacting proteins were found for 74 baits (http://xylome.vbi.vt.edu/index.html), but further validation, with alternative methods like co-immunoprecipitation or bimolecular fluorescence complementation, is required as Y2H is prone to artifacts. For monolignol biosynthesis, a predictive kinetic metabolite-flux model has been recently established for the 21 enzymes and 24 metabolites of the pathway in *P. trichocarpa* differentiating secondary xylem, based on both in vivo mass spectrometry quantification of all the isoforms in the pathway and kinetic parameters measured in vitro for functional recombinant proteins (Wang et al. 2014b). Predictions derived from the model were validated in transgenic poplars with altered monolignol biosynthesis. For transcription factors, it is now possible to identify DNA targets at the genome level, by combining the immunoprecipitation of chromatin with NGS sequencing (ChIP-SEQ), thus facilitating the construction of regulation networks at the genome scale. The ChIP-SEQ technique has been shown to be applicable on poplar cambial or xylem tissues and will certainly provide informative data about gene regulation networks in years to come (Li et al. 2014b). ChIP-PCR was successfully used on poplar secondary xylem, to validate potential DNA targets of secondary wall-associated NAC domain 1 (SND1), a transcription factor controlling wood formation (Lin et al. 2013). Predictive approaches have proved highly successful for tree species. Rodgers-Melnick et al. (2013) presented the first predicted interactome for *P. trichocarpa*, generated by the computational prediction of protein–protein interactions from primary sequence data only (conserved protein domains and predicted subcellular localization as input features).

The deciphering of molecular interactions is still in its infancy for tree species but will certainly increase in importance in the near future. We must first establish the critical datasets required, together with experimental evidence, to accelerate the inference of gene function and the construction of gene networks. Interactome data are not particularly straightforward to obtain, particularly for non-model species, but they can also be coupled to transcriptomic data, for the generation of informative biological networks. Biological network approaches have highlighted the existence of highly connected proteins, called hubs: these proteins are of particular interest in functional studies as they must play a key role in plant biology. In addition, biological network modeling could be used to predict the phenotypic changes resulting from changes to gene expression, thereby accelerating hypothesis-driven research for the development of new breeding applications, as already described in the field of medical research (Hood et al. 2004).

2.5 Epigenomics in trees: a new dimension to phenotype prediction in a changing environment

In addition to the genetic component, epigenetic variation is now proposed to contribute to phenotypic plasticity, adaptive

capacity, and evolutionary trajectories in both natural and cultivated plant populations (Bossdorf et al. 2008; Nicotra et al. 2010). Indeed, genome sequencing and genotyping were expected to identify the genetic components of common traits, but these approaches were not entirely successful, suggesting that there must be other sources of the missing heritability (Maher 2008). Epigenetics is the study of meiotically or mitotically heritable changes in gene function that do not result from changes in DNA sequence (Allis et al. 2007). At the molecular level, epigenetic phenomena are mediated by reversible marks, such as DNA methylation and histone modifications, including methylation, acetylation, phosphorylation, and ubiquitination, and by small RNAs that can alter regulatory states of genes or genomic regions. The study of epigenetic patterns at the genome-wide level is referred to as "epigenomics." As for spontaneous mutations in DNA, errors in the maintenance of methylation state result in the accumulation of single methylation polymorphisms (SMPs) over an evolutionary timescale. If the rates of SMP formation are orders of magnitude greater than those of spontaneous mutations, changes in regional methylation levels occur at similar frequencies. The regions concerned, which are known as differentially methylated regions (DMRs), correspond to genomic regions with different methylation profiles in different samples (tissues, cells, individuals, or others). DMRs are regarded as possible functional regions involved in the regulation of gene transcription and could act as QTLepi in natural populations, thereby constituting a measureable component of the so-called missing heritability (Cortijo et al. 2014). Epigenomic data have recently been reported in forest trees (Bräutigam et al. 2013).

DNA methylation in trees was first evaluated by determining global DNA methylation percentages by HPLC or HPCE, after the hydrolysis of DNA to generate nucleosides or nucleotides (Gentil and Maury 2007). Variations of global DNA methylation have been reported in several tree species, in different populations with different origins, in various organs, at different developmental stages, in different culture conditions, and in response to several environmental constraints, such as water availability and temperature (Hasbún et al. 2008; Monteuuis et al. 2009; Gourcilleau et al. 2010; Mankessi et al. 2011; Teyssier et al. 2014). However, it was not possible to identify the genomic context of these variations in these analyses. In recent years, epigenomics has emerged in parallel to the development of next-generation sequencing (NGS) techniques, genomic resources, and associated bioinformatics and biostatistics packages, together with the use of specific methods to identify DNA methylation marks. Three main methods for detecting DNA methylation at the genome-wide level have been applied to trees (Mensaert et al. 2014; Rodriguez et al. 2012; ESM_4.doc, ESM_5.doc): (1) methylation-sensitive amplification polymorphism (MSAP), (2) methylated DNA immunoprecipitation (MeDIP), and (3) whole-genome bisulfite sequencing (WGBS).

MSAP revealed DNA methylation polymorphism in different families of trees, between individuals/populations, in response to environmental changes, and during in vitro culture (ESM_4.doc). MeDIP and WGBS have been used to identify thousands of DMRs at the genome-wide level but only in *Populus*, the first tree to be sequenced (Tuskan et al. 2006). These epigenomic studies have significantly advanced our knowledge of structural and functional aspects of genomics. Indeed, determination of the poplar methylome revealed features particular to this tree, such as a higher CHG methylation level than reported for other plants (Feng et al. 2010) and low levels of methylation at DNA recombination hotspots (Slavov et al. 2012). Furthermore, gene-body DNA methylation is extensive in the open chromatin state, linked to structural gene characteristics (gene size and copy number) and correlated with tissue-specific gene expression (Vining et al. 2012; Lafon-Placette et al. 2013; Vining et al. 2013). DNA methylation is also involved in regulating stress response genes (Liang et al. 2014; Maury and Lafon-Placette, unpublished data). In addition, the methylation patterns of the parents are partially and dynamically passed onto their hybrid offspring (Gao et al. 2014). Finally, the first hypomethylated poplar trees (ESM_4.doc; Zhu et al. 2013a) were obtained by using a RNAi strategy to silence the *DDM1* gene, the product of which mediates the methylation of transposable elements and genes. These mutants represent an interesting model for investigating the role of DNA methylation under conditions of environmental variation. Insights into epigenomics will improve our understanding of adaptive tree responses to fluctuations in the environment, particularly in a context of global climate change.

2.6 Ecological genomics: genomic answers to ecological questions

Ever since Darwin (1859) and Ford (1964), the question of the heritable basis of adaptation to environmental conditions has been a fundamental issue in population genetics and ecology. Environmental parameters vary continually over space and time, so "adaptation" is intrinsically "local." Forest trees have large, often continuous populations and long life cycles. The question of how they adapt to changing environmental conditions is thus of considerable fundamental importance in addition to having implications for forest management. Genomics constitutes a powerful approach for obtaining answers to long-standing ecological questions, by shedding light on the role of selection in shaping patterns of genetic diversity and identifying environmental drivers of selection, in particular.

The power of genomics in forestry has already been described by González-Martínez et al. (2006). The relative durability of tree stands makes it possible to associate ecological variables with genotype frequency patterns in a reliable manner (Sork et al. 2013). Genomic methods can then be used to

screen large numbers of loci for association with environmental parameters. This approach has led to the identification of polymorphisms associated with environmental gradients at the regional scale in *P. taeda* (Eckert et al. 2010) and at the local scale (Jump et al. 2006; Csilléry et al. 2014) in *Fagus sylvatica*. Additional studies are underway in several species (*Pinus halepensis*, *Eperua falcata*, *P. pinaster*, *Symphonia globulifera*, etc.); preliminary results for candidate loci and traits have already demonstrated signatures of microgeographic adaptation in *E. falcata* (Audigeos et al. 2013; Brousseau et al. 2013) and *P. halepensis* (Hernandez-Serrano et al. 2013). Generally speaking, two kinds of approach have been applied to data analysis, depending on the nature of the information about the structure of environmental variation. In cases in which variation is considered to be *continuous*, the methods applied are mostly based on regression between environmental variables and allele frequencies (e.g., SAM; Joost et al. 2007). By contrast, in cases in which environmental variation is considered to be discrete (or is discretized during data collection, e.g., by sampling at different positions along a continuous gradient), the preferred methods are those based on the detection of divergence outliers (e.g., Beaumont and Balding 2004). However, it should be noted that the consideration of environmental variation as continuous or discrete often depends on the experimental design and constraints. With the exception of particular cases, in which one or more environmental variables change abruptly over space (e.g., over a cliff), or in which the patches available for tree growth are themselves discontinuous (e.g., on islands), environmental variation is generally continuous. This raises the question of a further merger, between ecological genetics (or genomics) and landscape genetics (or genomics). Depending on the extent to which ecological contrasts give rise to continuous patterns, there are good reasons for treating dispersal, migration, and adaptation as a unified process (Sork et al. 2013; Schoville et al. 2012).

Most early population genomic studies on forest trees focused on a limited number of polymorphisms of a few candidate genes. However, more realistic mechanisms of evolutionary change, such as polygenic and epistatic selection, must be considered in selection tests. To this end, population genomics will benefit from the ongoing development of pan-genomic approaches.

2.7 Integrating genomics into phylogeography and phylogeny

The phylogenetic relationships between species, the delimitation of closely related species, and the genetic structure of populations within species provide key information for decision-making in the conservation and sustainable use of forest tree germplasm. Early genetic studies on isozymes or plastid DNA revealed major footprints of past range dynamics in temperate trees (e.g., Petit et al. 2003), which found applications in genetic provenance discrimination of forest reproductive material. Plastid DNA markers were also applied for species delimitation in trees (Kress and Erickson 2008); however, their power remained limited because of weak reproductive barriers and frequent interspecific gene flow in trees (Petit and Hampe 2006). In the last decade, the increased use of multiple nuclear genetic markers and high-throughput genomics has made it possible to provide more precise information on past population history, more efficient delimitation of species, and inference of phylogenetic relationships between them and to characterize adaptive evolution at the molecular level in forest trees. This revolution in genetic markers has been accompanied by major transitions in data analysis, notably a shift from descriptive to hypothesis testing approaches, yielding valuable information for a more informed management of forest genetic resources.

In the last decade, phylogeographic studies have increasingly used multiple unlinked nuclear loci in combination with population genetic models using coalescent theory (Nielsen and Beaumont 2009). This has allowed discriminating between alternative scenarios of population genetic history within species and in closely related species (Heuertz et al. 2006; Gao et al. 2012; Cornille et al. 2013) and has permitted a reliable estimation of population genetic parameters, such as divergence times between lineages (Budde et al. 2013; Couvreur et al. 2008; Morris et al. 2008; Scotti-Saintagne et al. 2013a; Scotti-Saintagne et al. 2013b). The inferred demographic history has been used as a robust baseline information to detect gene loci under adaptive evolution (Grivet et al. 2011; Källman et al. 2014). Comparative phylogeographic approaches have given insights into the congruence of demographic history across species through time, first in temperate tree species (e.g., Jaramillo-Correa et al. 2010; Petit et al. 2003) and more recently in tropical species (Dauby et al. 2014; Heuertz et al. 2014; Jones et al. 2013; van der Merwe et al. 2014). Important predictive power has been gained in phylogeography by integrating species distribution modeling based on spatially interpolated climatic data for different time periods (Carstens and Richards 2007; Cornille et al. 2013). Significant advances in phylogeography could further come from new disciplines, such as "geogenomics," which involves the use of large-scale genetic data to constrain geological hypotheses (Baker et al. 2014).

Applications of massively parallel high-throughput sequencing (HTS) have emerged in recent years, facilitating cost-effective marker development (McPherson et al. 2013; Micheneau et al. 2011; Slavov et al. 2012). HTS technologies hold great promise for disentangling evolutionary relationships in complex groups, especially in tropical tree taxa in which botanical knowledge remains limited and cryptic species are common (Heuertz et al. 2014; Turchetto-Zolet et al. 2013). The first HTS studies reconstructed the complete

organellar genomes of tree taxa, which resulted in a higher resolution in phylogeography (McPherson et al. 2013; van der Merwe et al. 2014), taxonomy, and phylogenetics, particularly for tropical taxa (in Chrysobalanacae, Malé et al. 2014; in Malpighiales, Xi et al. 2012). HTS studies of the nuclear genome are still rare in phylogenetics and phylogeography of trees (but see Carstens et al. 2013; Stölting et al. 2013). They require the cost-effective parallel sequencing of hundreds or thousands of homologous DNA regions across hundreds of individuals, using appropriate technology for the targeted phylogenetic depth. Genome reduction methods involving enzymatic digestion are suitable at the within-species level (e.g., Stölting et al. 2013), whereas sequence capture for conserved nuclear regions is recommended for phylogenomics (reviewed in McCormack et al. 2013).

In phylogenomics, new Bayesian methods make it possible to infer phylogenetic relationships from multilocus nuclear data while accounting for intraspecies polymorphism and incomplete lineage sorting (Heled and Drummond 2010). In addition, Bayesian dating methods that incorporate uncertainty in evolutionary rate variation alongside time constraints based on fossils hold promise to improve dating (Dos Reis et al. 2012).

The characterization of genetic boundaries between closely related species remains challenging when reproductive barriers are incomplete. Interspecific gene flow can have important evolutionary consequences, such as the rapid introgression of beneficial variants (Morjan and Rieseberg 2004). This facilitates adaptation from standing genetic variation, which is particularly relevant during rapid range expansion (Keller et al. 2010; Lascoux and Petit 2010). Bayesian clustering algorithms (reviewed in François and Durand 2010) have proved useful for tree species delimitation (Duminil et al. 2012; Guichoux et al. 2013). However, population sampling must account for the presence of gene flow by including evolutionarily important areas, such as secondary contact zones or hybrid zones (Eckert et al. 2008; Scotti-Saintagne et al. 2013a).

Finally, phylogeographic and ecological genetics studies will undoubtedly become more integrated in the future. Historical inference is important not only for an understanding of the evolutionary past of a particular species or ecological community but also as a prerequisite to test for selection on sequence data (Carstens et al. 2013; Källman et al. 2014) and analyses of phenotypic evolution in different populations (Keir et al. 2011; Stone et al. 2011).

2.8 Genotype–phenotype association

Unraveling the genetic architecture of traits of economic or ecological importance, which are usually quantitative, is crucial for forest tree improvement and management. This can be achieved by identifying and localizing the genomic regions controlling the variation of quantitative traits (QTLs). Large-effect QTLs could theoretically be used in marker-assisted breeding schemes, but this approach has not yet been used in forest trees (Muranty et al. 2014), mostly because the most relevant traits for forest tree breeding are highly complex and probably controlled by many small-effect QTLs. Furthermore, most of the QTLs detected have large confidence intervals and have not been validated. Consequently, the identification of causal molecular polymorphisms controlling quantitative traits remains a great challenge that must be met before such information can be effectively transferred to breeding programs (ESM_6.pdf).

Since the first QTL mapping experiment in forest trees in the 1990s (Bradshaw and Stettler 1995), the techniques used to identify QTLs have changed radically. The advent of genomics and the democratization of sequencing due to NGS have generated thousands of single nucleotide polymorphism (SNPs) for genetic mapping, a prerequisite for QTL identification in controlled crosses. As a result, the number of markers on genetic maps has greatly increased, from a few hundred to several thousand, thanks to the extensive use of SNP arrays in many species (e.g., Chancerel et al. 2013) and, more recently, the direct use of sequencing for genotyping (Neves et al. 2014). Nevertheless, in order to improve the resolution of the QTLs mapped, it is necessary that the increase in the number of molecular markers is accompanied by an increase in the number of recombination events within the population studied, which can be achieved by increasing its size. This approach has been successfully used for the fine mapping of a major QTL for rust resistance in poplar (Bresson et al. 2011), but it remains time-consuming and expensive.

Association mapping (also known as linkage disequilibrium or LD mapping), which involves the detection of QTLs in more complex populations, offers an attractive alternative to the fine-mapping of QTLs in forest trees (Neale and Kremer 2011). Indeed, most forest trees are outcrossing species and this, together with their almost undomesticated status, implies a rapid decay of LD, making it possible to detect polymorphisms in the close physical vicinity of the causal variants or even the functional variants themselves. As a result, such approaches have become very popular in forest trees over the last decade, and many associations have been reported. However, despite the great promise of association mapping, most of the associations reported to date have accounted for only a very small proportion of the genetic variation, and this has greatly hindered their use in breeding programs. An obvious explanation for such disappointing results is that association studies in forest trees have not yet exhaustively screened the entire genome, as they have focused only on candidate genes and/or regions. Here again, the advent of NGS opens up new possibilities of screening for almost all SNPs within the gene space or even within the entire genome, resulting in an exhaustive genome-wide scan (Evans et al. 2014). Moreover, sequencing

approaches will enable a genome-wide discovery and typing of structural variation (presence/absence and copy number variants), with effects on quantitative traits that remain largely unexplored (Muchero et al. 2014).

Another way of gaining insight into the genetic determinism of complex traits is the genetic analysis of intermediate non-organismal phenotypes, such as transcriptomic, proteomic, or metabolomic data. Such systems biology approaches are appealing for highly complex traits, as they provide an intermediate step between the genotype and the phenotype, thereby facilitating the deciphering of genetic architecture for such traits. Only a few such studies have been carried out for forest trees, and these studies mostly involved QTL mapping for mRNA data obtained with microarrays (Kirst et al. 2005; Drost et al. 2010). With the advent of "omics" approaches, systems biology studies should and will undoubtedly become more widespread for forest trees in the near future. They will allow the detection of expression, protein, and metabolite QTLs and the construction of gene networks, which, together with genome-wide variation data, may make it possible to move from associations to causal links, through dedicated statistical modeling (Marjoram et al. 2014).

2.9 Conservation genomics

The science of conservation genomics is directly derived from conservation genetics, i.e., the use of genetic methods for understanding the impact of habitat modification on genetic structures and fitness and designing conservation strategies in practice, particularly in rare and endangered populations and species. Because quantitative genetic methods were complex and costly to implement, conservation genetics mostly focused on the use of molecular markers to decipher the demographic history of taxa and their phylogeography, thereby identifying groups of populations to be given priority for conservation efforts (see the concepts of ESU, evolutionary significant unit, as defined by Moritz 1994, and MU, management unit, as defined by Palsbøll et al. 2007).

Conservation schemes based on genetic methods have always recognized that neutral markers told only part of the evolutionary story (albeit an important one) and that access to parts of the genome undergoing selection was crucial. Before genomic tools became available, this was best achieved by comparing phenotypes of at least partly known ancestry in controlled environments (common gardens and reciprocal transplants). However, more frequently, environmental surrogates were used, to indicate the potential existence of natural selection. If both phenotypic or environmental divergence and neutral genetic or phylogeographic structure were found, such distinct populations were considered to be of high conservation priority (Lesica and Allendorf 1995; Allendorf et al. 2013). The European networks for the conservation of forest genetic resources are based on these strategies (Koskela et al. 2013; Lefevre et al. 2013).

With advances in genomics, and increasing access to many genes and potentially complete genome sequences for trees, it is becoming increasingly possible to compare surrogate phenotypic and ecological information with genetic information. Genomics is making it possible to revise our approach to conservation science and conservation strategies. In some cases, the change is just a question of degree, such as the much larger number of markers available now than in the past, making effective size and demographic estimates more precise. In other cases, the change is revolutionary, as, for example, for the comparison of DNA sequences responding to demography with those responding to selection (potentially in different ways, depending on the location of the populations within the ecological niche of the species) and for predicting the ability of populations to adapt to environmental changes and new patterns and thresholds of ecological disturbances (Allendorf et al. 2010). Even over short spatial scales (such as a single mountain), genomic studies have demonstrated the existence of local adaptation and significant differentiation for genes involved in key adaptive traits (phenology and resistance to drought and cold), of clear utility for the design of conservation strategies (Lalagüe et al. 2014).

There have been many calls for the use of genomic data for conservation purposes in forest trees (e.g., González-Martínez et al. 2006). However, few examples of the practical use of genomic data for strengthening conservation networks are available for forest trees as of yet, although many studies have demonstrated or confirmed that some populations or particular regions deserve protection. In *Pinus sylvestris* for example, genomic data from a relatively small number of adaptive genes have confirmed the high level of differentiation and unusual evolutionary history of populations from Scotland (Wachowiak et al. 2011). In *P. trichocarpa*, an extensive genome scan revealed previously unnoticed small- and large-scale geographic differentiation patterns in western North America (Slavov et al. 2012). As genome scan techniques are becoming cheaper and are now technically affordable for many different laboratories, they will undoubtedly soon be widely used for characterizing the evolutionary ecology and history of forest tree species, finally making it possible to include both demographic and adaptive processes in the design of conservation networks and to prioritize conservation actions. However, this era has yet to arrive.

2.10 Genomics and breeding

Most forest tree breeding programs were launched in the 1950s and have focused on species with relatively short rotation periods, either conifers (pine, spruce, larch, Douglas fir, etc.) or broadleaf trees (eucalyptus, poplar, wild cherry, sycamore, etc.). Considerable genetic gains have been achieved for

the economic and adaptive traits of most of these species (reviewed by Pâques 2013), despite the inherent difficulties involved in the breeding of these large, long-lived organisms. Given their key ecological and economic roles, forest trees have also been the object of rapid developments in genetic and association mapping for QTL discovery in the last 20 years (see Section 2.8). Despite these parallel developments in breeding and QTL mapping, few breakthrough applications for forest trees using a synergistic combination of these two approaches have been reported. This situation was quantitatively assessed and discussed in a review work by Muranty et al. (2014), which focused particularly on marker-assisted selection. Indeed, forest trees are among the genetically improved species most likely to benefit from the use of gene- or marker-based information in breeding. Indeed, reducing the long generation intervals of forest tree breeding programs by early marker-based evaluation is the most evident benefit, with other benefits including the limitation of phenotyping costs, an increase in the precision of evaluation of difficult traits (wood properties, phenology, biotic and abiotic responses to stresses, see previous sections), and the explicit management of genetic diversity. The large levels of polymorphism often harbored by forest tree species and their complex genomes have undoubtedly been limiting factors preventing the accumulation of genomic resources to the point required for the revolutionization of breeding programs (see Section 2.1). However, a number of initiatives worldwide are dealing with this issue, including dedicated European (reviewed in http://www.forestry.gov.uk/fr/euframeworkprojects) and North American (*American Conifer Translational Genomics Network*: https://dendrome.ucdavis.edu/ctgn/, *SMarTForests*: http://www.smartforests.ca/) projects.

A number of marker-based applications are paving the way for genomics-assisted breeding. Some of the simplest, in terms of marker requirements, do not require modifications to existing breeding programs and concern the management of breeding populations through fingerprinting (identity and pedigree checking, genetic diversity estimation) and the optimization of genetic gain deployment (seed quality control). Other applications have low requirements for genomic resources but potentially large impacts on existing breeding programs. These applications aim to minimize the uncertainties in the assessment of relatedness in pedigree-based genetic evaluations, through the use of marker-based estimates, and to improve precision. Markers can be used to recover full parentage in open or polymix mating regimes (El Kassaby and Lstiburek 2009), thereby increasing the precision of evaluations without the need for a costly control-cross regime. This strategy is currently being evaluated in several breeding programs, for Scots pine in Sweden (Rosvall 2011) and maritime pine in France (Bouffier et al. in preparation). In a more general approach known as G-BLUP, a marker-based relationship matrix is used alone or together with one from a pedigree, in the statistical mixed model for genetic evaluation. This approach can finely capture relatedness at within-family levels, thereby increasing the precision of evaluations.

Ultimately, the achievement of sufficiently high levels of genome coverage by dense genotyping or sequencing provides not only highly precise relatedness estimates but also information about any relevant genetic variation from underlying causal mutations. This is the principle behind genomic selection (GS, Meuwissen et al. 2001), which paves the way for the early evaluation of genotyped candidates without phenotype data. Simulation studies (Grattapaglia and Resende 2011, Iwata et al. 2011; Denis and Bouvet 2013) have already demonstrated the potential of GS for forest trees in diverse breeding and genetic scenarios. These theoretical studies have been complemented by the first empirical studies in forest trees (Zapata-Valenzuela et al. 2012; Zapata-Valenzuela et al. 2013; Resende et al. 2012a; Resende et al. 2012b), which have yielded medium- to high-level accuracies, even with limited numbers of markers (less than 5,000 SNPs). However, these preliminary results must be interpreted with caution, as they relate to populations with small effective sizes and thus favorably high linkage disequilibrium levels. In large populations, such as those in most forest tree breeding populations, GS may be less accurate, given the marker density currently attainable in these species (Beaulieu et al. 2014). Nevertheless, GS is clearly an alternative for the forest tree breeding programs of tomorrow.

3 Conclusion and perspectives

The sections above clearly illustrate how, in the 10 years since the DNA sequence of the poplar genome was made publicly available, the forest tree genomics community has fully embraced the rapidly evolving tools of genomics and bioinformatics (i) to study the genetic and molecular changes underlying the complex developmental traits characteristic of the biology of woody perennials, (ii) to determine how the information encoded in the genome of individual trees responds to external cues and to identify the evolutionary forces responsible for shaping the phenotypic variation we see at both the population and species levels in natural conditions, and (iii) to identify the causal genetic polymorphisms underlying phenotypic differences. By harnessing the power of genomics, this community has shown that it is committed to applying this knowledge in management practices, conservation, and breeding programs, to help natural and planted forests adapt to the rapid pace of current and projected climate changes. It is difficult to see where events will take us in the next 10 years, given the currently exponential increase in the amount of DNA sequence data, likely decreases in the cost of sequencing, and the

emergence of new technologies. This research community has matured, but its sustainable development will require the mobilization of adequate funding and human resources, improvements to international collaboration to address global challenges, and integration with other disciplines.

3.1 What is to come in the next 10 years?

Over the next decade, we will undoubtedly continue to see an accumulation of genomic resources (including reference genome and epigenome sequences) and basic understanding about genome structure and evolution, the distribution of gene, non-protein-coding transcribed fragments, and transposable elements across the genome and their interactions. The sequencing of large heterozygous tree genomes remains challenging, and work is still in the exploratory phase for most such genomes (e.g., chestnut and oak) or requires refinement (e.g., conifers) with new sequencing technologies and bioinformatic approaches (Faino and Thomma 2014). However, the recent results obtained in this area are more than encouraging and open the way for the resequencing of thousands of genotypes, a prerequisite for the description of sequence variation within and between species, the identification of causal variants underlying phenotypes of interest, and the provision of knowledge about the evolutionary history of tree populations. Beyond the apparent completeness of genome sequence information, much remains to be done concerning our understanding of gene regulation. Multidisciplinary groups are required to make use of this wealth of resources to derive fundamental insights into the biology of woody perennials through the integration of "omics" technologies into research activities. MicroRNAs constitute a specific class of noncoding molecules worthy of attention, given their fundamental biological role. Clearly, the next 10 years will see improvements in descriptions of the number and biological role of these RNAs, particularly as concerns tree-specific features. Moreover, systems biology, by providing a holistic approach, may help to uncover the molecular players, their complex networks of interaction and key hubs underpinning developmental processes, and responses to external disturbances. Advances in genomic technologies, statistical, mathematical, and computational methods are very promising and should enable us to meet this challenge. RNA interference-based screening and gene tagging approaches are still underdeveloped for forest trees but will probably be instrumental in achieving these ends. However, the logistic complexities associated with high-throughput screening in such large organisms may hamper the development of this area of research. Another field that will undoubtedly change our view on how trees develop, grow, and adapt throughout their extended life span is that of epigenetics. Reversible epigenetic marks contribute to phenotypic plasticity and, therefore, constitute an essential factor in the adaptive capacity of these long-lived organisms. This is clearly a hot topic in the framework of rapid climate change. Moreover, as epigenetic marks may differ between genotypes and may be heritable, a fraction of phenotypic variation shaped by epigenetic mechanisms may potentially be targeted by natural selection and, therefore, contribute to the evolutionary trajectory of populations. Clearly, assessing the relative contributions of epigenetic modifications and changes in allele frequency (the "classical" mechanism of adaptation operating at the population level) will be a major challenge in the coming years. Finally, two other research areas should benefit from the discovery of gene regulatory networks underlying trait variation. The hunt for the so-called missing heritability (proportion of phenotypic variance unaccounted for by single nucleotide polymorphisms) is one of these areas (Maher 2008). Understanding the nature of genetic adaptation to environmental heterogeneity over space and time is the other.

3.2 Will genomics change the landscape of forest tree breeding?

Proofs-of-concept of genomic prediction in forest trees have been obtained and show that this technology should work (i.e., it should improve estimated breeding value accuracy and, therefore, result in a genetic gain per unit time), and the genomic revolution holds great promise for economic benefit in the forest industry. However, further research is required to confirm that this technology could be readily implemented. We can see at least three complementary actions that could potentially favor a rapid and efficient implementation of genomic selection in the next decade. One is the construction of consensus reference populations for each species across countries, with shared efforts for the required phenotypic evaluation. Another action is the rational implementation of genotyping, notably through the use of different marker coverages at different population levels and of imputation for equaling density coverages. Useful lessons can be learned from existing success stories (as in dairy cattle and pig breeding), but the economic viability of incorporating genome-enabled selection into forest tree breeding programs remains to be demonstrated. Thus, a third action will be the demonstration that the magnitude of the estimated breeding value accuracy improvement will be large enough to counterbalance the genotyping/sequencing costs. The next decade will certainly see an accumulation of proof-of-concept studies but, more importantly, researchers and breeders will have to work together to demonstrate the economic viability of this methodology before its implementation.

Acknowledgments We would like to thank an anonymous reviewer for providing insightful and critical comments on the manuscript.

Funding The preparation of this article was supported by several projects funded by (i) the French National Research Agency (ARBRE: ANR-12-LABXARBRE-01, FLAG: ANR-12-ADAP-0007-01, GENOAK: 2011-BSV6-009-01, XYLOFOREST: ANR-10-EQPX-16, GENOQB: ANR-05-GPLA-027, SUSTAINPINE: ANR-09-KBBE-007, SYLVABIOM: ANR-08-BIOE-0006, SYBIOPOP: ANR-13-JSV6-0001), (ii) the European Community Seventh Framework Programme (ProCoGen: FP7-289841, WatBio: FP7-311929) and Eranet Biodiversa (LinkTree and TipTree), (iii) *Conseil Régional de la Région Centre* (EMBRYOME 33639), *Conseil Régional de la Région Aquitaine* (EMBRYO2011 09012579-045), (iv) the EFPA division of INRA (*AAP projets innovants*). ALLG received a PhD grant from the French *Ministère de la Recherche et Enseignement Supérieu*r, and (v) INRA, the SELGEN metaprogram (R2D2). AM received a PhD grant from *Région Centre*, France. MH acknowledges funding from the Spanish Ministry for Science and Innovation (Ramón y Cajal program, RYC2009-04537) and the Spanish Ministry for Economics and Competitiveness (AFFLORA project, CGL2012-40129-C02-02).

References

Abril N, Gion JM, Kerner R, Müller-Starck G, Cerrillo RMN, Plomion C, Renaut J, Valledor L, Jorrin-Novo JV (2011) Proteomics research on forest trees, the most recalcitrant and orphan plant species. Phytochemistry 72:1219–1242

Abu-Abied M, Szwerdszarf D, Morde-Haev I, Levy A, Rogovoy O, Belausov E, Yaniv Y, Uliel S, Katzenellenbogen M, Riov J, Ophir R, Sadot E (2012) Microarray analysis revealed upregulation of nitrate reductase in juvenile cuttings of *Eucalyptus grandis*, which correlated with increased nitric oxide production and adventitious root formation. Plant J 71:787–799

Allendorf FW, Hohenlohe PA, Luikart GH (2010) Genomics and the future of conservation genetics. Nat Rev Genet 11:697–709

Allendorf FW, Luikart GH, Aitken SN (2013) Conservation and the genetics of populations, 2nd edn. Wiley-Blackwell, Oxford

Allis CD, Jenuwein T, Reinberg D, Caparros M-L (2007) Epigenetics, 1st edn. Cold Spring Harbor Laboratory Press, New York

Allona I, Quinn M, Shoop E, Swope K, Cyr SS, Carlis J, Riedl J, Retzel E, Campbell MM, Sederoff R, Whetten RW (1998) Analysis of xylem formation in pine by cDNA sequencing. Proc Natl Acad Sci U S A 95:9693–9698

Alves AA, Conçalves Rosado CC, Faria DA, da Silva Guimaraes LM, Lau D, Brommonschenkel SH, Grattapaglia D, Alfenas AC (2012) Genetic mapping provides evidence for the role of additive and non-additive QTLs in the response of inter-specific hybrids of *Eucalyptus* to *Puccinia psidii* rust infection. Euphytica 183:27–38

Andersson-Gunneras S, Mellerowicz EJ, Love J, Segerman B, Ohmiya Y, Coutinho PM, Nilsson P, Henrissat B, Moritz T, Sundberg B (2006) Biosynthesis of cellulose-enriched tension wood in Populus: global analysis of transcripts and metabolites identifies biochemical and developmental regulators in secondary wall biosynthesis. Plant J 45:144–165

Audigeos D, Brousseau L, Traissac S, Scotti-Saintagne C, Scotti I (2013) Molecular divergence in tropical tree populations occupying environmental mosaics. J Evol Biol 26:529–544

Baba K, Park YW, Kaku T, Kaida R, Takeuchi M, Yoshida M, Hosoo Y, Ojio Y, Okuyama T, Taniguchi T, Ohmiya Y, Kondo T, Shani Z, Shoseyov O, Awano T, Serada S, Norioka N, Norioka S, Hayashi T (2009) Xyloglucan for generating tensile stress to bend tree stem. Mol Plant 2:893–903

Baker PA, Fritz SC, Dick CW, Eckert AJ, Horton BK, Manzoni S, Ribas CC, Garzione CN, Battisti DS (2014) The emerging field of geogenomics: constraining geological problems with genetic data. Earth-Sci Rev 135:38–47

Bálint M, Tiffin P, Hallström B, O'Hara RB, Olson MS, Fankhauser JD, Piepenbring M, Schmitt I (2013) Host genotype shapes the foliar fungal microbiome of balsam poplar (*Populus balsamifera*). PLoS ONE 8:e53987

Bálint M, Bartha L, O'Hara RB, Olson MS, Otte J, Pfenninger M, Robertson AL, Tiffin P, Schmitt I (2015) Relocation, high-latitude warming and host genetic identity shape the foliar fungal microbiome of poplars. Mol Ecol 24:235–248

Bassett CL, Wisniewski ME, Artlip TS, Norelli JL, Renaut J, Farrell RE (2006) Global analysis of genes regulated by low temperature and photoperiod in peach bark. J Am Soc Hortic Sci 131:551–563

Beaulieu J, Doerksen T, Clément S, MacKay J, Bousquet J (2014) Accuracy of genomic selection models in a large population of open-pollinated families in white spruce. Heredity 113:343–352

Beaumont MA, Balding DJ (2004) Identifying adaptive genetic divergence among populations from genome scans. Mol Ecol 13:969–980

Bedon F, Villar E, Vincent D, Dupuy JW, Lomenech AM, Mabialangoma A, Chaumeil P, Barre A, Plomion C, Gion JM (2012) Proteomic plasticity of two *Eucalyptus* genotypes under contrasted water regimes in the field. Plant Cell Environ 35:790–805

Bernhardsson C, Ingvarsson PK (2012) Geographical structure and adaptive population differentiation in herbivore defence genes in European aspen (*Populus tremula* L., Salicaceae). Mol Ecol 21: 2197–2207

Berta M, Giovannelli A, Sebastiani F, Camussi A, Racchi ML (2010) Transcriptome changes in the cambial region of poplar (*Populus alba* L.) in response to water deficit. Plant Biol 12:341–354

Berthet S, Thevenin J, Baratiny D, Demont-Caulet N, Debeaujon I, Bidzinski P, Leple JC, Huis R, Hawkins S, Gomez LD, Lapierre C, Jouanin L (2012) Role of plant laccases in lignin polymerization. In: Jouann L, Lapierre C eds. Lignins: biosynthesis, biodegradation and bioengineering, 145–172

Bielenberg DG, Wang Y, Li Z, Zhebentyayeva T, Fan S, Reighard GL, Scorza R, Abbott AG (2008) Sequencing and annotation of the evergrowing locus in peach [*Prunus persica* (L.) Batsch] reveals a cluster of six MADS-box transcription factors as candidate genes for regulation of terminal bud formation. Tree Genet Genomes 4:495–507

Birol I, Raymond A, Jackman SD, Pleasance S, Coope R, Taylor GA, Yuen MM, Keeling CI, Brand D, Vandervalk BP, Kirk H, Pandoh P, Moore RA, Zhao Y, Mungall AJ, Jaquish B, Yanchuk A, Ritland C, Boyle B, Bousquet J, Ritland K, Mackay J, Bohlmann J, Jones SJ (2013) Assembling the 20 Gb white spruce (*Picea glauca*) genome from whole-genome shotgun sequencing data. Bioinformatics 29: 1492–1497

Biswal AK, Soeno K, Gandla ML, Immerzeel P, Pattathil S, Lucenius J, Serimaa R, Hahn MG, Moritz T, Jonsson LJ, Israelsson-Nordstrom M, Mellerowicz EJ (2014) Aspen pectate lyase PtxtPL1-27 mobilizes matrix polysaccharides from woody tissues and improves saccharification yield. Biotechnol Biofuels 7:11

Bogeat-Triboulot MB, Brosche M, Renaut J, Jouve L, Le Thiec D, Fayyaz P, Vinocur B, Witters E, Laukens K, Teichmann T, Altman A, Hausman JF, Polle A, Kangasjarvi J, Dreyer E (2007) Gradual soil water

depletion results in reversible changes of gene expression, protein profiles, ecophysiology, and growth performance in Populus euphratica, a poplar growing in arid regions. Plant Physiol 143:876–892

Bonawitz ND, Chapple C (2010) The genetics of lignin biosynthesis: connecting genotype to phenotype. Annu Rev Genet 44:337–363

Bonhomme L, Monclus R, Vincent D, Carpin S, Claverol S, Lomenech AM, Labas V, Plomion C, Brignolas F, Morabito D (2009) Genetic variation and drought response in two *Populus* x *euramericana* genotypes through 2-DE proteomic analysis of leaves from field and glasshouse cultivated plants. Phytochemistry 70:988–1002

Borer ET, Kinkel LL, May G, Seabloom EW (2013) The world within: quantifying the determinants and outcomes of a host's microbiome. Basic Appl Ecol 14:533–539

Bossdorf O, Richards CL, Pigliucci M (2008) Epigenetics for ecologists. Ecol Lett 11:106–115

Bradshaw HD, Stettler RF (1995) Molecular genetics of growth and development in Populus. IV. Mapping QTLs with large effects on growth, form, and phenology traits in a forest tree. Genetics 139: 963-973

Braun P, Carvunis AR, Charloteaux B, Dreze M, Ecker JR, Hill DE, Roth FP, Vidal M, Galli M, Balumuri P, Bautista V, Chesnut JD, Kim RC, Reyes C d l, Gilles P, Kim CJ, Matrubutham U, Mirchandani J, Olivares E, Patnaik S, Quan R, Ramaswamy G, Shinn P, Swamilingiah GM, Wu S, Byrdsong D, Dricot A, Duarte M, Gebreab F, Gutierrez BJ, MacWilliams A, Monachello D, Mukhtar MS, Poulin MM, Reichert P, Romero V, Tam S, Waaijers S, Weiner EM, Cusick ME, Tasan M, Yazaki J, Ahn YY, Barabasi AL, Chen HM, Dangl JL, Fan CY, Gai LT, Ghoshal G, Hao T, Lurin C, Milenkovic T, Moore J, Pevzner SJ, Przulj N, Rabello S, Rietman EA, Rolland T, Santhanam B, Schmitz RJ, Spooner W, Stein J, Vandenhaute J, Ware D, Arabidopsis Interactome Mapping Consortium (2011) Evidence for network evolution in an Arabidopsis interactome map. Science 333:601–607

Braun P, Aubourg S, Van Leene J, De Jaeger G, Lurin C (2013) Plant protein interactomes. Annu Rev Plant Biol 64:161–187

Bräutigam K, Vining KJ, Lafon-Placette C, Fossdal CG, Mirouze M, Marcos JG, Fluch S, Fraga MF, Guevara MA, Abarca D, Johnsen Ø, Maury S, Strauss SH, Campbell MM, Rohde A, Díaz-Sala C, Cervera MT (2013) Epigenetic regulation of adaptive responses of forest tree species to the environment. Ecol Evol 3:399–415

Bresson A, Jorge V, Dowkiw A, Guerin V, Bourgait I, Tuskan GA, Schmutz J, Chalhoub B, Bastien C, Faivre-Rampant P (2011) Qualitative and quantitative resistances to leaf rust finely mapped within two nucleotide-binding site leucine-rich repeat (NBS-LRR)-rich genomic regions of chromosome 19 in poplar. New Phytol 192: 151–163

Brinker M, van Zyl L, Liu W, Craig D, Sederoff RR, Clapham DH, von Arnold S (2004) Microarray analyses of gene expression during adventitious root development in *Pinus contorta*. Plant Physiol 135:1526–1539

Brousseau L, Bonal D, Cigna J, Scotti I (2013) Highly local environmental variability promotes intra-population divergence of quantitative traits: an example from tropical rainforest trees. Ann Bot 112:1169–1179

Budde KB, Gonzalez-Martinez SC, Hardy OJ, Heuertz M (2013) The ancient tropical rainforest tree *Symphonia globulifera* L. f. (Clusiaceae) was not restricted to postulated Pleistocene refugia in Atlantic Equatorial Africa. Heredity 111:66–76

Buée M, Reich M, Murat C, Morin E, Nilsson RH, Uroz S, Martin F (2009) 454 Pyrosequencing analyses of forest soils reveal an unexpectedly high fungal diversity. New Phytol 184:449–456

Busov VB, Brunner AM, Meilan R, Filichkin S, Ganio L, Gandhi S, Strauss SH (2005) Genetic transformation: a powerful tool for dissection of adaptive traits in trees. New Phytol 167:9–18

Busov V, Yordanov YS, Gou J, Meilan R, Ma C, Regan S, Strauss S (2010) Activation tagging is an effective gene tagging system in Populus. Tree Genet Genomes 7:91–101

Cairney J, Pullman GS (2007) The cellular and molecular biology of conifer embryogenesis. New Phytol 176:511–536

Canales J, Bautista R, Label P, Gómez-Maldonado J, Lesur I, FernÁndez-Pozo N, Rueda-López M, Guerrero-Fernández D, Castro-Rodríguez V, Benzekri H, Cañas RA, Guevara MA, Rodrigues A, Seoane P, Teyssier C, Morel A, Ehrenmann F, Le Provost G, Lalanne C, Noirot C, Klopp C, Reymond I, García-Gutiérrez A, Trontin JF, Lelu-Walter M-A, Miguel C, Cervera MT, Cantón FR, Plomion C, Harvengt L, Avila C, Claros MG, Cánovas FM (2014) De novo assembly of maritime pine transcriptome: implications for forest breeding and biotechnology. Plant Biotech J 12:286–299

Carstens BC, Brennan RS, Chua V, Duffie CV, Harvey MG, Koch RA, McMahan CD, Nelson BJ, Newman CE, Satler JD, Seeholzer G, Posbic K, Tank DC, Sullivan J (2013) Model selection as a tool for phylogeographic inference: an example from the willow Salix melanopsis. Mol Ecol 22:4014–4028

Carstens BC, Richards CL (2007) Integrating coalescent and ecological niche modeling in comparative phylogeography. Evolution 61: 1439–1454

Castell WZ, Ernst D (2012) Experimental "omics" data in tree research: facing complexity. Trees Struct Funct 26:1723–1735

Chancerel E, Lamy JB, Lesur I, Noirot C, Klopp C, Ehrenmann F, Boury C, Le Provost G, Label P, Lalanne C, Léger V, Salin F, Gion JM, Plomion C (2013) High density linkage mapping in a pine tree reveals a genomic region associated with inbreeding depression and provides clues to the extent and distribution of meiotic recombination. BMC Biol 11:50

Chao WS, Foley ME, Horvath DP, Anderson JV (2007) Signals regulating dormancy in vegetative buds. Int J Plant Dev Biol 1:49–56

Cohen D, Bogeat-Triboulot MB, Tisserant E, Balzergue S, Martin-Magniette ML, Lelandais G, Ningre N, Renou JP, Tamby JP, Le Thiec D, Hummel I (2010) Comparative transcriptomics of drought responses in Populus: a meta-analysis of genome-wide expression profiling in mature leaves and root apices across two genotypes. BMC Genomics 11:630

Cohen D, Bogeat-Triboulot MB, Vialet-Chabrand S, Merret R, Courty PE, Moretti S, Bizet F, Guilliot A, Hummel I (2013) Developmental and environmental regulation of aquaporin gene expression across *Populus* species: divergence or redundancy? PLoS ONE 8:e55506

Coince A, Caël O, Bach C, Lengellé J, Cruaud C, Gavory F, Morin E, Murat C, Marçais B, Buée M (2013) Below-ground fine-scale distribution and soil versus fine root detection of fungal and soil oomycete communities in a French beech forest. Fungal Ecol 6: 223–235

Coince A, Cordier T, Lengellé J, Defossez E, Vacher C, Robin C, Buée M, Marçais B (2014) Leaf and root-associated fungal assemblages do not follow similar elevational diversity patterns. PLoS ONE 9: e100668

Coleman HD, Canam T, Kang KY, Ellis DD, Mansfield SD (2007) Over-expression of UDP-glucose pyrophosphorylase in hybrid poplar affects carbon allocation. J Exp Bot 58:4257–4268

Coleman HD, Yan J, Mansfield SD (2009) Sucrose synthase affects carbon partitioning to increase cellulose production and altered cell wall ultrastructure. Proc Natl Acad Sci U S A 106:13118–13123

Cordier T, Robin C, Capdevielle X, Desprez-Loustau M-L, Vacher C (2012a) Spatial variability of phyllosphere fungal assemblages: genetic distance predominates over geographic distance in a European beech stand (*Fagus sylvatica*). Fungal Ecol 5:509–520

Cordier T, Robin C, Capdevielle X, Fabreguettes O, Desprez-Loustau M-L, Vacher C (2012b) The composition of phyllosphere fungal assemblages of European beech (*Fagus sylvatica*) varies significantly along an elevation gradient. New Phytol 196:510–519

Cornille A, Giraud T, Bellard C, Tellier A, Le Cam B, Smulders MJM, Kleinschmit J, Roldan-Ruiz I, Gladieux P (2013) Postglacial recolonization history of the European crabapple (*Malus sylvestris* Mill.), a wild contributor to the domesticated apple. Mol Ecol 22:2249–2263

Cortijo S, Wardenaar R, Colome-Tatche M, Gilly A, Etcheverry M, Labadie K, Caillieux E, Hospital F, Aury JM, Wincker P, Roudier F, Jansen RC, Colot V, Johannes F (2014) Mapping the epigenetic basis of complex traits. Science 343:1145–1148

Couvreur T, Chatrou L, Sosef M, Richardson J (2008) Molecular phylogenetics reveal multiple tertiary vicariance origins of the African rain forest trees. BMC Biol 6:54

Csillery K, Lalagüe H, Vendramin GG, Gonzalez-Martinez SC, Fady B, Oddou-Muratorio S (2014) Detecting local adaptation and epistatic selection in climate related candidate genes at a short spatial scale in European beech (*Fagus sylvatica*) populations. Mol Ecol 23:4696–4708

Damon C, Lehembre F, Oger-Desfeux C, Luis P, Ranger J, Fraissinet-Tachet L, Marmeisse R (2012) Metatranscriptomics reveals the diversity of genes expressed by eukaryotes in forest soils. PLoS ONE 7:e28967

Darwin C (1859) On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life. Murray, London

Dauby G, Duminil J, Heuertz M, Koffi GK, Stévart T, Hardy OJ (2014) Congruent phylogeographical patterns of eight tree species in Atlantic Central Africa provide insights into the past dynamics of forest cover. Mol Ecol 23:2299–2312

De Bodt S, Proost S, Vandepoele K, Rouze P, van de Peer Y (2009) Predicting protein-protein interactions in *Arabidopsis thaliana* through integration of orthology, gene ontology and co-expression. BMC Genomics 10:288

de Vega-Bartol JJ, Simões M, Lorenz WW, Rodrigues AS, Alba R, Dean JFD, Miguel C (2013) Transcriptomic analysis highlights epigenetic and transcriptional regulation during zygotic embryo development of Pinus pinaster. BMC Plant Biol 13:123

Déjardin A, Leplé J-C, Lesage-Descauses M-C, Costa G, Pilate G (2004) Expressed sequence tags from poplar wood tissues—a comparative analysis from multiple libraries. Plant Biol 6:55–64

Déjardin A, Arnaud D, Laurans F, Breton C, Pilate G, Leplé JC (2010) Wood formation in angiosperms. CR Biol 333:325–334

Denis M, Bouvet J-M (2013) Efficiency of genomic selection with models including dominance effect in the context of *Eucalyptus* breeding. Trees Geneti Genomes 9:37–51

Dos Reis M, Inoue J, Hasegawa M, Asher RJ, Donoghue PCJ, Yang Z (2012) Phylogenomic datasets provide both precision and accuracy in estimating the timescale of placental mammal phylogeny. Proc R Soc B 279:3491–3500

Dowkiw A, Voisin E, Bastien C (2010) Potential of Eurasian poplar rust to overcome a major quantitative resistance factor. Plant Pathol 59: 523–534

Drost DR, Benedict CI, Berg A, Novaes E, Novaes CRDB, Yu Q, Dervinis C, Maia JM, Yap J, Miles B, Kirst M (2010) Diversification in the genetic architecture of gene expression and transcriptional networks in organ differentiation of Populus. Proc Natl Acad Sci U S A 107:8492–8497

Du J, Mansfield SD, Groover AT (2009) The *Populus* homeobox gene ARBORKNOX2 regulates cell differentiation during secondary growth. Plant J 60:1000–1014

Du J, Miura E, Robischon M, Martinez C, Groover A (2011) The *Populus* class III HD ZIP transcription factor POPCORONA affects cell differentiation during secondary growth of woody stems. PLoS ONE 6:e17458

Dubos C, Le Provost G, Pot D, Salin F, Lalane C, Madur D, Frigerio JM, Plomion C (2003) Identification and characterization of water-stress-responsive genes in hydroponically grown maritime pine (*Pinus pinaster*) seedlings. Tree Physiol 23:169–179

Duminil J, Kenfack D, Viscosi V, Grumiau L, Hardy OC (2012) Testing species delimitation in sympatric species complexes: the case of an African tropical tree, *Carapa* spp. (Meliaceae). Mol Phylogenet Evol 62:275–285

Duplessis S, Major I, Martin F, Seguin A (2009) Poplar and pathogen interactions: insights from *Populus* genome-wide analyses of resistance and defense gene families and gene expression profiling. Crit Rev Plant Sci 28:309–334

Eckert AJ, Tearse BR, Hall BD (2008) A phylogeographical analysis of the range disjunction for foxtail pine (*Pinus balfouriana*, Pinaceae): the role of Pleistocene glaciation. Mol Ecol 17:1983–1997

Eckert AJ, Bower AD, González-MartÍnez SC, Wegrzyn JL, Coop G, Neale DB (2010) Back to nature: ecological genomics of loblolly pine (*Pinus taeda*, Pinaceae). Mol Ecol 19:3789–3805

El Kassaby Y, Lstiburek M (2009) Breeding without breeding. Genet Res 91:111–120

Elhiti M, Stasolla C, Wang A (2013) Molecular regulation of plant somatic embryogenesis. In vitro Cell Dev Pl 49:631–642

Evans LM, Slavov GT, Rodgers-Melnick E, Martin J, Ranjan P, Muchero W, Brunner AM, Schackwitz W, Gunter L, Chen JG, Tuskan GA, DiFazio SP (2014) Population genomics of *Populus trichocarpa* identifies signatures of selection and adaptive trait associations. Nat Genet 46:1089–1096

Faino L, Thomma BP (2014) Get your high-quality low-cost genome sequence. Trends Plant Sci 19:288–291

Felten J, Kohler A, Morin E, Bhalerao RP, Palme K, Martin F, Ditengou FA, Legué V (2009) The ectomycorrhizal fungus Laccaria bicolor stimulates lateral root formation in poplar and Arabidopsis through auxin transport and signaling. Plant Physiol 151:1991–2005

Feng S, Cokus SJ, Zhang X, Chen PY, Bostick M, Goll MG, Hetzel J, Jain J, Strauss SH, Halpern ME, Ukomadu C, Sadler KC, Pradhan S, Pellegrini M, Jacobsen SE (2010) Conservation and divergence of methylation patterning in plants and animals. Proc Natl Acad Sci U S A 107:8689–8694

Ford EB (1964) Ecological genetics. Chapman and Hall, London

François O, Durand E (2010) Spatially explicit Bayesian clustering models in population genetics. Mol Ecol Res 10:773–784

Freeman JS, Potts BM, Vaillancourt RE (2008) Few Mendelian genes underlie the quantitative response of a forest tree, *Eucalyptus globulus*, to a natural fungal epidemic. Genetics 178:563–571

Gao JIE, Wang B, Mao J-F et al (2012) Demography and speciation history of the homoploid hybrid pine *Pinus densata* on the Tibetan plateau. Mol Ecol 21:4811–4827

Gao M, Huang Q, Chu Y, Ding C, Zhang B, Su X (2014) Analysis of the leaf methylomes of parents and their hybrids provides new insight into hybrid vigor in *Populus deltoides*. BMC Genet 15:S8

Geisler-Lee J, O'Toole N, Ammar R, Provart NJ, Millar AH, Geisler M (2007) A predicted interactome for *Arabidopsis*. Plant Physiol 145: 317–329

Gentil MV, Maury S (2007) Characterization of epigenetic biomarkers using new molecular approaches. In: Varshney R, Tuberosa R (eds) Springer, pp 351–370

Gion JM, Lalanne C, Le Provost G, Ferry-Dumazet H, Paiva J, Frigerio JM, Chaumeil P, Barré A, de Daruvar A, Brach J, Claverol S, Bonneu M, Plomion C (2005) The proteome of maritime pine wood forming tissue. Proteomics 5:3731–3751

Goff SA, Ricke D, Lan TH, Presting G, Wang R, Dunn M, Glazebrook J, Sessions A, Oeller P, Varma H, Hadley D, Hutchison D, Martin C, Katagiri F, Lange BM, Moughamer T, Xia Y, Budworth P, Zhong J, Miguel T, Paszkowski U, Zhang S, Colbert M, Sun WL, Chen L, Cooper B, Park S, Wood TC, Mao L, Quail P, Wing R, Dean R, Yu Y, Zharkikh A, Shen R, Sahasrabudhe S, Thomas A, Cannings R, Gutin A, Pruss D, Reid J, Tavtigian S, Mitchell J, Eldredge G, Scholl T, Miller RM, Bhatnagar S, Adey N, Rubano T, Tusneem N,

Robinson R, Feldhaus J, Macalma T, Oliphant A, Briggs S (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. japonica). Science 296:92–100

González-Martínez SC, Krutovsky KV, Neale DB (2006) Forest-tree population genomics and adaptive evolution. New Phytol 170:227–238

Gorzsás A, Stenlund H, Persson P, Trygg J, Sundberg B (2011) Cell-specific chemotyping and multivariate imaging by combined FT-IR microspectroscopy and orthogonal projections to latent structures (OPLS) analysis reveals the chemical landscape of secondary xylem. Plant J 66:903–914

Gourcilleau D, Bogeat-Triboulot M-B, Thiec D, Lafon-Placette C, Delaunay A, El-Soud WA, Brignolas F, Maury S (2010) DNA methylation and histone acetylation: genotypic variations in hybrid poplars, impact of water deficit and relationships with productivity. Ann Forest Sci 67:208

Grattapaglia D, Resende MDV (2011) Genomic selection in forest tree breeding. Tree Genet Genomes 7:241–255

Gray-Mitsumune M, Blomquist K, McQueen-Mason S, Teeri TT, Sundberg B, Mellerowicz EJ (2008) Ectopic expression of a wood-abundant expansin PttEXPA1 promotes cell expansion in primary and secondary tissues in aspen. Plant Biotechnol J 6:62–72

Grivet D, Sebastiani F, Alía R, Bataillon T, Torre S, Zabal-Aguirre M, Vendramin GG, González-Martínez SC (2011) Molecular footprints of local adaptation in two Mediterranean conifers. Mol Bio Evol 28: 101–116

Guichoux E, Garnier-Géré P, Lagache L, Lang T, Boury C, Petit RJ (2013) Outlier loci highlight the direction of introgression in oaks. Mol Ecol 22:450–462

Guttman DS, Mc Hardy AC, Schulze-Lefert P (2014) Microbial genome-enabled insights into plant-microorganism interactions. Nat Rev Genet 15:797–813

Hacquard S, Schadt CW (2015) Towards a holistic understanding of the beneficial interactions across the Populus microbiome. New Phytol 205:1424–1430

Hakman I, Fowke LC, von Arnold S, Eriksson T (1985) The development of somatic embryogenesis in tissue cultures initiated from immature embryos of *Picea abies* (Norway spruce). Plant Sci 38:53–59

Hamanishi ET, Raj S, Wilkins O, Thomas BR, Mansfield SD, Plant AL, Campbell MM (2010) Intraspecific variation in the *Populus balsamifera* drought transcriptome. Plant Cell Environ 33:1742–1755

Harfouche A, Meilan R, Altman A (2014) Molecular and physiological responses to abiotic stress in forest trees and their relevance to tree improvement. Tree Physiol (in press)

Hasbún R, Valledor L, Rodríguez JL, Santamaria E, Rıos D, Sanchez M, Canal MJ, Rodrıguez R (2008) HPCE quantification of 5-methyl-2′-deoxycytidine in genomic DNA: methodological optimization for chestnut and other woody species. Plant Physiol Biochem 46:815–822

He J, Li H, Luo J, Ma C, Li S, Qu L, Gai Y, Jiang X, Janz D, Polle A, Tyree M, Luo ZB (2013) A transcriptomic network underlies microstructural and physiological responses to cadmium in *Populus* x *canescens*. Plant Physiol 162:424–439

Hedenström M, Wiklund-Lindström S, Öman T, Lu F, Gerber L, Schatz P, Sundberg B, Ralph J (2009) Identification of lignin and polysaccharide modifications in populus wood by chemometric analysis of 2D NMR spectra from dissolved cell walls. Mol Plant 2:933–942

Heled J, Drummond AJ (2010) Bayesian inference of species trees from multilocus data. Mol Bio Evol 27:570–580

Hernandez-Serrano A, Verdu M, Gonzalez-Martinez SC, Pausas JG (2013) Fire structures pine serotiny at different scales. Am J Bot 100:2349–2356

Heuertz M, De Paoli E, Källman T, Larsson H, Jurman I, Morgante M, Lascoux M, Gyllenstrand N (2006) Multilocus patterns of nucleotide diversity, linkage disequilibrium and demographic history of Norway spruce [*Picea abies* (L.) Karst]. Genetics 174:2095–2105

Heuertz M, Duminil J, Dauby G, Savolainen V, Hardy OJ (2014) Comparative phylogeography in rainforest trees from lower Guinea, Africa. PLoS ONE 9:e84307

Hibbett DS, Ohman A, Kirk PM (2009) Fungal ecology catches fire. New Phytol 184:279–282

Hoenicka H, Lautner S, Klingberg A, Koch G, El-Sherif F, Lehnhardt D, Zhang B, Burgert I, Odermatt J, Melzer S, Fromm J, Fladung M (2012) Influence of over-expression of the FLOWERING PROMOTING FACTOR 1 gene (FPF1) from *Arabidopsis* on wood formation in hybrid poplar (*Populus tremula* L. x *P. tremuloides* Michx.). Planta 235:359–373

Hood L, Heath JR, Phelps ME, Lin BY (2004) Systems biology and new technologies enable predictive and preventative medicine. Science 306:640–643

Horton MW, Bodenhausen N, Beilsmith K, Meng D, Muegge BD, Nordborg M, Subramanian S, Vetter MM, Vilhja BJ, Gordon JI, Bergelson J (2014) Genome-wide association study of *Arabidopsis thaliana* leaf microbial community. Nat Commun 5:5320

Horvath DP, Anderson JV, Chao WS, Foley ME (2003) Knowing when to grow: signals regulating bud dormancy. Trends Plant Sci 8:534–540

Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408:796–815

IRGSP (2005) The map-based sequence of the rice genome. Nature 436: 793–800

Iwata H, Hayashi T, Tsumura Y (2011) Prospects for genomic selection in conifer breeding: a simulation study of Cryptomeria japonica. Tree Genet Genomes 7:747-758

Jaillon O, Aury JM, Noel B, Policriti A, Clepet C, Casagrande A, Choisne N, Aubourg S, Vitulo N, Jubin C, Vezzi A, Legeai F, Hugueney P, Dasilva C, Horner D, Mica E, Jublot D, Poulain J, Bruyere C, Billault A, Segurens B, Gouyvenoux M, Ugarte E, Cattonaro F, Anthouard V, Vico V, Del Fabbro C, Alaux M, Di Gaspero G, Dumas V, Felice N, Paillard S, Juman I, Moroldo M, Scalabrin S, Canaguier A, Le Clainche I, Malacrida G, Durand E, Pesole G, Laucou V, Chatelet P, Merdinoglu D, Delledonne M, Pezzotti M, Lecharny A, Scarpelli C, Artiguenave F, Pe ME, Valle G, Morgante M, Caboche M, Adam-Blondon AF, Weissenbach J, Quetier F, Wincker P (2007) The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature 449:463–467

Janz D, Behnke K, Schnitzler JP, Kanawati B, Schmitt-Kopplin P, Polle A (2010) Pathway analysis of the transcriptome and metabolome of salt sensitive and tolerant poplar species reveals evolutionary adaption of stress tolerance mechanisms. BMC Plant Biol 10:150

Janz D, Lautner S, Wildhagen H, Behnke K, Schnitzler JP, Rennenberg H, Fromm J, Polle A (2012) Salt stress induces the formation of a novel type of 'pressure wood' in two *Populus* species. New Phytol 194:129–141

Jaramillo-Correa JP, Grivet D, Terrab A et al (2010) The Strait of Gibraltar as a major biogeographic barrier in Mediterranean conifers: a comparative phylogeographic survey. Mol Ecol 19:5452–5468

Jiao Y, Wickett NJ, Ayyampalayam S, Chanderbali AS, Landherr L, Ralph PE, Tomsho LP, Hu Y, Liang H, Soltis PS, Soltis DE, Clifton SW, Schlarbaum SE, Schuster SC, Ma H, Leebens-Mack J, dePamphilis CW (2011) Ancestral polyploidy in seed plants and angiosperms. Nature 473:97–100

Jones DG, Dangl JL (2006) The plant immune system. Nature 444:323–329

Jones AF, Cerón-Souza I, Hardesty BD, Dick CW (2013) Genetic evidence of Quaternary demographic changes in four rain forest tree species sampled across the Isthmus of Panama. J Biogeogr 40:720–731

Joost S, Bonin A, Bruford MW, Després L, Conord C, Erhardt G, Taberlet P (2007) A spatial analysis method (SAM) to detect candidate loci for selection: towards a landscape genomics approach to adaptation. Mol Ecol 16:3955–3969

Jorge V, Dowkiw A, Faivre-Rampant P, Bastien C (2005) Genetic architecture of qualitative and quantitative *Melampsora larici-populin*a leaf rust resistance in hybrid poplar: genetic mapping and QTL detection. New Phytol 167:113–127

Jump AS, Hunt JM, Martínez-Izquierdo JA, Peñuelas J (2006) Natural selection and climate change: temperature-linked spatial and temporal trends in gene frequency in *Fagus sylvatica*. Mol Ecol 15:3469–3480

Jumpponen A, Jones KL (2009) Massively parallel 454 sequencing indicates hyperdiverse fungal communities in temperate *Quercus macrocarpa* phyllosphere. New Phytol 184:438–448

Jumpponen A, Jones KL (2010) Seasonally dynamic fungal communities in the *Quercus macrocarpa* phyllosphere differ between urban and nonurban environments. New Phytol 186:496–513

Junghans DT, Alfenas AC, Brommonschenkel SH, Oda S, Mello EJ, Grattapaglia D (2004) Resistance to rust *Puccinia psidii* winter in *Eucalyptus*: mode of inheritance and mapping of a major gene with RAPD markers. Theor Appl Genet 108:175–18

Källman T, De Mita S, Larsson H, Gyllenstrand N, Heuertz M, Parducci L, Suyama Y, Lagercrantz U, Lascoux M (2014) Patterns of nucleotide diversity at photoperiod related genes in Norway spruce [*Picea abies* (L.) Karst.]. PloS ONE 9:e95306

Keir KR, Bemmels JB, Aitken SN (2011) Low genetic diversity, moderate local adaptation, and phylogeographic insights in *Cornus nuttallii* (Cornaceae). Am J Bot 98:1327–1336

Kell DB, Brown M, Davey HM, Dunn WB, Spasic I, Oliver SG (2005) Metabolic footprinting and systems biology: the medium is the message. Nat Rev Microbiol 3:557–565

Keller SR, Olson MS, Silim S, Schroeder W, Tiffin P (2010) Genomic diversity, population structure, and migration following rapid range expansion in the balsam poplar, *Populus balsamifera*. Mol Ecol 19:1212–1226

Kembel SW, Mueller RC (2014) Plant traits and taxonomy drive host associations in tropical phyllosphere fungal communities. Botany 92:303–311

Kembel SW, Connor TKO, Arnold HK, Hubbell SP, Wright SJ (2014) Relationships between phyllosphere bacterial communities and plant functional traits in a neotropical forest. Proc Natl Acad Sci 111:13715–13720

Kinloch BB Jr, Sniezko RA, Dupper GE (2004) Virulence gene distribution and dynamics of the white pine blister rust pathogen in Western North America. Phytopathology 94:751–758

Kirst M, Basten CJ, Myburg AA, Zeng Z-B, Sederoff RR (2005) Genetic architecture of transcript-level variation in differentiating xylem of a *Eucalyptus* hybrid. Genetics 169:2295–2303

Klimaszewska K, Overton C, Stewart D, Rutledge RG (2011) Initiation of somatic embryos and regeneration of plants from primordial shoots of 10-year-old somatic white spruce and expression profiles of 11 genes followed during the tissue culture process. Planta 233:635–647

Klimaszewska K, Hargreaves C, Lelu-Walter M-A, Trontin J-F (2015). Advances in conifer somatic embryogenesis since year 2000. In: In vitro plant embryogenesis in higher plants, Germanà MA, Lambardi M (eds), Methods in molecular biology, Springer-Humana Press (in press)

Knief C, Delmotte N, Chaffron S, Stark M, Innerebner G, Wassmann R, von Mering C, Vorholt JA (2012) Metaproteogenomic analysis of microbial communities in the phyllosphere and rhizosphere of rice. ISME J 6:1378–1390

Kohler A, Rinaldi C, Duplessis S, Baucher M, Geelen D, Duchaussoy F, Meyers BC, Boerjan W, Martin F (2008) Genome-wide identification of NBS resistance genes in *Populus trichocarpa*. Plant Mol Biol 66:619–636

Koskela J, Lefèvre F, Schüler S, Kraigher H, Olrik DC, Hubert J, Longauer R, Bozzano M, Yrjänä L, Alizoti P, Rotach P, Vietto L, Bordács S, Myking T, Eysteinsson T, Souvannavong O, Fady B, De Cuyper B, Heinze B, von Wühlisch G, Ducousso A, Ditlevsen B (2013) Translating conservation genetics into management: pan-European minimum requirements for dynamic conservation units of forest tree genetic diversity. Biol Conserv 157:39–49

Kress WJ, Erickson DL (2008) DNA barcodes: genes, genomics, and bioinformatics. Proc Natl Acad Sci U S A 105:2761–2762

Kreuzwieser J, Hauberg J, Howell KA, Carroll A, Rennenberg H, Millar AH, Whelan J (2009) Differential response of gray poplar leaves and roots underpins stress adaptation during hypoxia. Plant Physiol 149:461–473

La Mantia J, Klápště J, El-Kassaby YA, Azam S, Guy RD, Douglas CJ, Mansfield SD, Hamelin R (2013) Association analysis identifies *Melampsora* x *Columbiana* poplar leaf rust resistance SNPs. PLoS ONE 8:e78423

Lafon-Placette C, Faivre-Rampant P, Delaunay A, Street N, Brignolas F, Maury S (2013) Methylome of DNase I sensitive chromatin in *Populus trichocarpa* shoot apical meristematic cells: a simplified approach revealing characteristics of gene-body DNA methylation in open chromatin state. New Phytol 197:416–430

Lalagüe H, Csilléry K, Oddou-Muratorio S, Safrana J, de Quattro C, Fady B, González-Martínez SC, Vendramin GG (2014) Nucleotide diversity and linkage disequilibrium at 58 stress response and phenology candidate genes in a European beech (*Fagus sylvatica* L.) population from southeastern France. Tree Genet Genomes 10:15–26

Lang GA, EarlyJD MGC, Darnell RL (1987) Endo-, para- and eco-dormancy: physiological, terminology and classification for dormancy research. HortSci 22:371–377

Lascoux M, Petit RJ (2010) The 'New Wave' in plant demographic inference: more loci and more individuals. Mol Ecol 19:1075–1078

Lefevre F, Koskela J, Hubert J, Kraigher H, Longauer R, Olrik DC, Schueler S, Bozzano M, Alizoti P, Bakys R, Baldwin C, Ballian D, Black-Samuelsson S, Bednarova D, Bordacs S, Collin E, De Cuyper B, De Vries SMG, Eysteinsson T, Frydl J, Haverkamp M, Ivankovic M, Konrad H, Koziol C, Maaten T, Paino EN, Ozturk H, Pandeva ID, Parnuta G, Pilipovic A, Postolache D, Ryan C, Steffenrem A, Varela MC, Vessella F, Volosyanchuk RT, Westergren M, Wolter F, Yrjana L, Zarina I (2013) Dynamic conservation of forest genetic resources in 33 European countries. Conserv Biol 27:373–384

Leff JW, Del Tredici P, Friedman WE, Fierer N (2015) Spatial structuring of bacterial communities within individual *Ginkgo biloba* trees. Environ Microbiol. doi:10.1111/1462-2920.12695

Legué V, Rigal A, Bhalerao RP (2014) Adventitious root formation in tree species: involvement of transcription factors. Physiol Plant 151:192–198

Lelu-Walter MA, Thompson D, Harvengt L, Sanchez L, Toribio M, Pâques LE (2013) Somatic embryogenesis in forestry with a focus on Europe: state-of-the-art, benefits, challenges and future direction. Tree Genet Genomes 9:883–899

Lesica P, Allendorf FW (1995) When are peripheral populations valuable for conservation? Conserv Biol 9:753–760

Li E, Bhargava A, Qiang W, Friedmann MC, Forneris N, Savidge RA, Johnson LA, Mansfield SD, Ellis BE, Douglas CJ (2012) The class II KNOX gene KNAT7 negatively regulates secondary wall formation in *Arabidopsis* and is functionally conserved in *Populus*. New Phytol 194:102–115

Li W, Lin YC, Li Q, Shi R, Lin CY, Chen H, Chuang L, Qu GZ, Sederoff RR, Chiang VL (2014a) A robust chromatin immunoprecipitation protocol for studying transcription factor–DNA interactions and histone modifications in wood-forming tissue. Nat Methods 9:2180–2193

Li WF, Zhang SG, Han SY, Wu T, Zhang JH, Qi LW (2014b) The post-transcriptional regulation of LaSCL6 by miR171 during maintenance of embryogenic potential in *Larix kaempferi* (Lamb.) Carr. Tree Genet Genomes 10:223–229

Liang D, Zhang Z, Wu H, Huang C, Shuai P, Ye C-Y, Tang S, Wang Y, Yang L, Wang J, Yin W, Xia W (2014) Single-base-resolution methylomes of *Populus trichocarpa* reveal the association between DNA methylation and drought stress. BMC Genet 15:S9

Lin MZ, Zhou X, Shen XL, Mao CZ, Chen X (2011) The predicted Arabidopsis interactome resource and network topology-based systems biology analyses. Plant Cell 23:911–922

Lin YC, Li W, Sun YH, Kumari S, Wei HR, Li QZ, Tunlaya-Anukit S, Sederoff RR, Chiang VL (2013) SND1 transcription factor-directed quantitative functional hierarchical genetic regulatory network in wood formation in *Populus trichocarpa*. Plant Cell 25:4324–4341

Lippert D, Zhuang J, Ralph S, Ellis DE, Gilbert M, Olafson R, Ritland K, Ellis B, Douglas CJ, Bohlmann J (2005) Proteome analysis of early somatic embryogenesis in *Picea glauca*. Proteomics 5:461–473

Liu JJ, Sturrock RN, Benton R (2013) Transcriptome analysis of *Pinus monticola* primary needles by RNA-seq provides novel insight into host resistance to *Cronartium ribicola*. BMC Genomics 14:884

Liu LJ, Filkov V, Groover A (2014) Modeling transcriptional networks regulating secondary growth and wood formation in forest trees. Physiol Plant 151:156–163

Lorenz WW, Ayyampalayam S, Boedeaux JM, Howe GT, Jermstad KD, Neale DB, Rogers DL, Dean JFD (2012) Conifer DBMagic: a database housing multiple de novo transcriptome assemblies for 12 diverse conifer species. Tree Genet Genomes 8:1477–1485

Lu S, Sun Y-H, Shi R, Clark C, Li CVL (2005) Novel and mechanical stress-responsive microRNAs in *Populus trichocarpa* that are absent from *Arabidopsis*. Plant Cell 17:2186–2203

Ma T, Wang J, Zhou G, Yue Z, Hu Q, Chen Y, Liu B, Qiu Q, Wang Z, Zhang J, Wang K, Jiang D, Gou C, Yu L, Zhan D, Zhou R, Luo W, Ma H, Yang Y, Pan S, Fang D, Luo Y, Wang X, Wang G, Wang J, Wang Q, Lu X, Chen Z, Liu J, Lu Y, Yin Y, Yang H, Abbott RJ, Wu Y, Wan D, Li J, Yin T, Lascoux M, DiFazio SP, Tuskan GA, Wang J, Liu J (2013) Genomic insights into salt adaptation in a desert poplar. Nat Commun 4:2797

Maher B (2008) The case of the missing heritability. Nature 456:18–21

Malé P-JG, Bardon L, Besnard G, Coissac E, Delsuc F, Engel J, Lhuillier E, Scotti-Saintagne C, Tinaut A, Chave J (2014) Genome skimming by shotgun sequencing helps resolve the phylogeny of a pantropical tree family. Mol Ecol Res 14:966–975

Mankessi F, Saya AR, Favreau B, Doulbeau S, Conejero G, Lartaud M, Verdeil J-L, Monteuuis O (2011) Variations of DNA methylation in Eucalyptus urophylla × Eucalyptus grandis shoot tips and apical meristems of different physiological ages. Physiol Plant 143:178–187

Margulis L (1991) Symbiogenesis and symbionticism. In: Margulis L, Fester R (eds) Symbiosis as a source of evolutionary innovation: speciation and morphogenesis. MIT Press, Cambridge, pp 1–14

Marjoram P, Zubair A, Nuzhdin SV (2014) Post-GWAS: where next? More samples, more SNPs or more biology? Heredity 112:79–88

McCormack JE, Hird SM, Zellmer AJ, Carstens BC, Brumfield RT (2013) Applications of next-generation sequencing to phylogeography and phylogenetics. Mol Phylogenet Evol 66:526–538

McPherson H, van der Merwe M, Delaney S, Edwards MA, Henry RJ, McIntosh E, Rymer PD, Milner ML, Siow J, Rossetto M (2013) Capturing chloroplast variation for molecular ecology studies: a simple next generation sequencing approach applied to a rainforest tree. BMC Ecol 13:8

Mensaert K, Denil S, Trooskens G, Van Criekinge W, Thas O, DeMeyer T (2014) Next-generation technologies and data analytical approaches for epigenomics: next-generation technologies for epigenomics. Environ Mol Mutagen 55:155–170

Menzel A, Fabian P (1999) Growing season extended in Europe. Nature 397:659

Meuwissen THE, Hayes BJ, Goddard ME (2001) Prediction of total genetic value using genome-wide dense marker maps. Genetics 157:1819–1829

Micheneau C, Dauby G, Bourland N, Doucet JL, Hardy OJ (2011) Development and characterization of microsatellite loci in *Pericopsis elata* (Fabaceae) using a cost-efficient approach. Am J Bot 98:e268–e270

Moller I, Sørensen I, Bernal AJ, Blaukopf C, Lee K, Øbro J, Pettolino F, Roberts A, Mikkelsen JD, Knox JP, Bacic A, Willats WGT (2007) High-throughput mapping of cell-wall polymers within and between plants using novel microarrays. Plant J 50:1118–1128

Monteuuis O, Baurens FC, Goh DKS, Uimado MQ, Oulbeau SD, Verdeil JL (2009) DNA methylation in *Acacia mangium* in vitro and ex-vitro buds, in relation to their within-shoot position, age and leaf morphology of the shoots. Silvae Genet 58:287–292

Morel A, Teyssier C, Trontin JF, Pešek B, Eliášová K, Beaufour M, Morabito D, Boizot N, Le Metté C, Belal-Bessai L, Reymond I, Harvengt L, Cadene M, Corbineau F, Vágner M, Label P, Lelu-Walter MA (2014a) Early molecular events involved in *Pinus pinaster* Ait. somatic embryo development under reduced water availability: transcriptomic and proteomic analysis. Physiol Plant 152:184–201

Morel A, Trontin JF, Corbineau F, Lomenech A-M, Beaufour M, Reymond I, Le Metté C, Ader K, Harvengt L, Cadene M, Label P, Teyssier C, Lelu-Walter M-A (2014b) Cotyledonary somatic embryos of *Pinus pinaster* Ait. most closely resemble fresh, maturing cotyledonary zygotic embryos: biological, carbohydrate and proteomic analyses. Planta (in press)

Moritz C (1994) Defining 'evolutionarily significant units' for conservation. Trends Ecol Evol 9:373–375

Morjan CL, Rieseberg LH (2004) How species evolve collectively: implications of gene flow and selection for the spread of advantageous alleles. Mol Ecol 13:1341–1356

Morreel K, Dima O, Kim H, Lu F, Niculaes C, Vanholme R, Dauwe R, Goeminne G, Inzé D, Messens E, Ralph J, Boerjan W (2010) Mass spectrometry-based sequencing of lignin oligomers. Plant Physiol 153:1464–1478

Morris AB, Ickert-Bond SM, Brunson DB, Soltis DE, Soltis PS (2008) Phylogeographical structure and temporal complexity in American sweetgum (Liquidambar styraciflua; Altingiaceae). Mol Ecol 17: 3889–3900

Muchero W, Labbé J, Ranjan P, DiFazio S, Tuskan GA (2014) Genome resequencing in *Populus*: revealing large-scale genome variation and implications on specialized-trait genomics. In: Challenges and Opportunities for the World's Forests in the 21st Century. Springer, pp 587–595

Muranty H, Jorge V, Bastien C, Lepoittevin C, Bouffier L, Sanchez L (2014) Potential for marker-assisted selection for forest tree breeding: lessons from 20 years of MAS in crops. Tree Genet Genomes (in press)

Myburg AA, Grattapaglia D, Tuskan GA, Hellsten U, Hayes RD, Grimwood J, Jenkins J, Lindquist E, Tice H, Bauer D, Goodstein DM, Dubchak I, Poliakov A, Mizrachi E, Kullan AR, Hussey SG, Pinard D, van der Merwe K, Singh P, van Jaarsveld I, Silva-Junior OB, Togawa RC, Pappas MR, Faria DA, Sansaloni CP, Petroli CD, Yang X, Ranjan P, Tschaplinski TJ, Ye CY, Li T, Sterck L, Vanneste K, Murat F, Soler M, Clemente HS, Saidi N, Cassan-Wang H, Dunand C, Hefer CA, Bornberg-Bauer E, Kersting AR, Vining K, Amarasinghe V, Ranik M, Naithani S, Elser J, Boyd AE, Liston A, Spatafora JW, Dharmwardhana P, Raja R, Sullivan C, Romanel E, Alves-Ferreira M, Külheim C, Foley W, Carocha V, Paiva J, Kudrna D, Brommonschenkel SH, Pasquali G, Byrne M, Rigault P, Tibbits J, Spokevicius A, Jones RC, Steane DA, Vaillancourt RE, Potts BM, Joubert F, Barry K, Pappas GJ, Strauss SH, Jaiswal P, Grima-Pettenati J, Salse J, Van de Peer Y, Rokhsar DS, Schmutz J (2014) The genome of *Eucalyptus grandis*. Nature 510:356–362

Naor A, Flaishman M, Stern R, Moshe A, Erez A (2003) Temperature effect on dormancy completion of vegetative buds in apple. J Am Soc Hortic Sci 128:636–641

Neale DB, Kremer A (2011) Forest tree genomics: growing resources and applications. Nature Rev Genet 12:111–121

Neale D, Langley C, Salzberg S, Wegrzyn J (2013) Open access to tree genomes: the path to a better forest. Genome Biol 14:120

Neale D, Wegrzyn J, Stevens K, Zimin A, Puiu D, Crepeau M, Cardeno C, Koriabine M, Holtz-Morris A, Liechty J, Martinez-Garcia P, Vasquez-Gross H, Lin B, Zieve J, Dougherty W, Fuentes-Soriano S, Wu L-S, Gilbert D, Marcais G, Roberts M, Holt C, Yandell M, Davis J, Smith K, Dean J, Lorenz W, Whetten R, Sederoff R, Wheeler N, McGuire P, Main D, Loopstra C, Mockaitis K, deJong P, Yorke J, Salzberg S, Langley C (2014) Decoding the massive genome of loblolly pine using haploid DNA and novel assembly strategies. Genome Biol 15:R59

Neiman M, Olson MS, Tiffin P (2009) Selective histories of poplar protease inhibitors: elevated polymorphism, purifying selection, and positive selection driving divergence of recent duplicates. New Phytol 183:740–750

Neves LG, Davis JM, Barbazuk WB, Kirst M (2014) A high-density gene map of Loblolly pine (*Pinus taeda* L.) based on exome sequence capture genotyping. G3 4:29–37

Newcombe G, Bradshaw HD Jr, Chastagner GA, Stettler RF (1996) A major gene for resistance to *Melampsora medusa* f.sp. *deltoidae* in a hybrid poplar pedigree. Phytopathology 86:87–94

Nicotra AB, Atkin OK, Bonser SP, Davidson AM, Finnegan EJ, Mathesius U, Poot P, Purugganan MD, Richards CL, Valladares F, Van Kleunen M (2010) Plant phenotypic plasticity in a changing climate. Trends Plant Sci 15:684–692

Nielsen R, Beaumont MA (2009) Statistical inferences in phylogeography. Mol Ecol 18:1034–1047

Nimchuk Z, Eulgem T, Holt BF, Dangl JL (2003) Recognition and response in the plant immune system. Annu Rev Genet 37:579–609

Nishikubo N, Takahashi J, Roos AA, Derba-Maceluch M, Piens K, Brumer H, Teeri TT, Stalbrand H, Mellerowicz EJ (2011) Xyloglucan endo-transglycosylase-mediated xyloglucan rearrangements in developing wood of hybrid aspen. Plant Physiol 155: 399–413

Nystedt B, Street NR, Wetterbom A, Zuccolo A, Lin YC, Scofield DG, Vezzi F, Delhomme N, Giacomello S, Alexeyenko A, Vicedomini R, Sahlin K, Sherwood E, Elfstrand M, Gramzow L, Holmberg K, Hällman J, Keech O, Klasson L, Koriabine M, Kucukoglu M, Käller M, Luthman J, Lysholm F, Niittylä T, Olson A, Rilakovic N, Ritland C, Rosselló JA, Sena J, Svensson T, Talavera-López C, Theißen G, Tuominen H, Vanneste K, Wu ZQ, Zhang B, Zerbe P, Arvestad L, Bhalerao R, Bohlmann J, Bousquet J, Gil RG, Hvidsten TR, de Jong P, MacKay J, Morgante M, Ritland K, Sundberg B, Thompson SL, Van de Peer Y, Andersson B, Nilsson O, Ingvarsson PK, Lundeberg J, Jansson S (2013) The Norway spruce genome sequence and conifer genome evolution. Nature 497:579–584

Ohtani M, Nishikubo N, Xu B, Yamaguchi M, Mitsuda N, Goue N, Shi F, Ohme-Takagi M, Demura T (2011) A NAC domain protein family contributing to the regulation of wood formation in poplar. Plant J 67:499–512

Öpik M, Metsis M, Daniell TJ, Zobel M, Moora M (2009) Large-scale parallel 454 sequencing reveals host ecological group specificity of arbuscular mycorrhizal fungi in a boreo-nemoral forest. New Phytol 184:424–437

Paiva J, Garcés M, Alves A, Garnier-Géré P, Rodrigues J-C, Lalanne C, Porcon S, Le Provost G, da Silva D, Perez BJ, Frigerio JM, Claverol S, Barré A, Fevereiro P, Plomion C (2008a) Molecular and phenotypic profiling from the base to the crown in maritime pine wood forming tissue. New Phytol 178:283–301

Paiva J, Garnier-Géré PH, Rodrigues JC, Alves A, Santos S, Graça J, Le Provost G, Chaumeil P, da Silva-Perez D, Bosc A, Fevereiro P, Plomion C (2008b) Plasticity of maritime pine (*Pinus pinaster* Ait) wood forming tissues during a growing season. New Phytol 179:1080–1094

Palsbøll PJ, Bérubé M, Allendorf FW (2007) Identification of management units using population genetic data. Trends Ecol Evol 22:11–16

Pandey V, Haikio E, Lehesranta S, Koistinen K, Karenlampi S, Oksanen E (2013) Proteomic analysis of two hybrid aspen clones subjected to long-term chronic ozone exposure in open field. Curr Proteomics 10:67–74

Pâques (2013) Forest tree breeding in Europe (2013) Edited by L. Pâques. Springer, 527 p

Park YW, Ki B, Furuta Y, Iida I, Sameshima K, Arai M, Hayashi T (2004) Enhancement of growth and cellulose accumulation by overexpression of xyloglucanase in poplar. FEBS Lett 564:183–187

Paux E, Carocha V, Marques C, Mendes de Sousa A, Borralho N, Sivadon P, Grima-Pettenati J (2005) Transcript profiling of *Eucalyptus* xylem genes during tension wood formation. New Phytol 167:89–100

Petit RJ, Hampe A (2006) Some evolutionary consequences of being a tree. Annu Rev Ecol Evol Syst 37:187–214

Petit RJ, Aguinagalde I, de Beaulieu J-L, Bittkau C, Brewer S, Cheddadi R, Ennos R, Fineschi S, Grivet D, Lascoux M, Mohanty A, Müller-Starck G, Demesure-Musch B, Palmé A, Martín JP, Rendell S, Vendramin GG (2003) Glacial refugia: hotspots but not melting pots of genetic diversity. Science 300:1563–1565

Plomion C, Pionneau C, Brach J, Costa P, Baillères H (2000) Compression wood-responsive proteins in developing xylem of maritime pine (*Pinus pinaster* Ait). Plant Physiol 123:959–969

Plomion C, Le Provost G, Stokes A (2001) Wood formation in trees. Plant Physiol 127:1513–1523

Polle A, Altman A, Jiang XN (2006) Towards genetic engineering for drought tolerance in trees. Tree transgenesis: recent developments. Springer-Verlag Berlin, Berlin, pp 275–297

Preston JC, Sandve SR (2013) Adaptation to seasonality and winter freeze. Front Plant Sci 4:1–18

Pullman GS, Johnson S, Peter G, Cairney J, Xu N (2003) Improving Loblolly pine somatic embryo maturation: comparison of somatic and zygotic embryo morphology, germination, and gene expression. Plant Cell Rep 21:747–758

Puzey JR, Karger A, Axtell M, Kramer EM (2012) Deep annotation of *Populus trichocarpa* microRNAs from diverse tissue sets. PLoS ONE 7:e33034

Qiu D, Wilson IW, Gan S, Washusen R, Moran GF, Southerton SG (2008) Gene expression in *Eucalyptus* branch wood with marked variation in cellulose microfibril orientation and lacking G-layers. New Phytol 179:94–103

Raherison E, Rigault P, Caron S, Poulin PL, Boyle B, Verta JP, Giguère I, Bomal C, Bohlmann J, MacKay J (2012) Transcriptome profiling in conifers and the PiceaGenExpress database show patterns of diversification within gene families and interspecific conservation in vascular gene expression. BMC Genomics 13:434

Ramírez-Carvajal GA, Morse AM, Dervinis C, Davis JM (2008) Transcript profiles of the cytokinin response regulator gene family in *Populus* imply diverse roles in plant development. New Phytol 177:77–89

Ramírez-Carvajal GA, Morse AM, Dervinis C, Davis JM (2009) The cytokinin type-B response regulator PtRR13 is a negative regulator of adventitious root development in *Populus*. Plant Physiol 150:759–771

Redford AJ, Bowers RM, Knight R, Linhart Y, Fierer N (2010) The ecology of the phyllosphere: geographic and phylogenetic variability in the distribution of bacteria on tree leaves. Environ Microbiol 12:2885–2893

Resende MDV, Resende MFR, Sansaloni CP, Petroli CD, Missiaggia AA, Aguiar AM, Abad JM, Takahashi EK, Rosado AM, Faria DA, Pappas GJ, Kilian A, Grattapaglia D (2012a) Genomic selection for growth and wood quality in *Eucalyptus*: capturing the missing

heritability and accelerating breeding for complex traits in forest trees. New Phytol 194:116–128
Resende MFR, Munoz P, Acosta JJ, Peter GF, Davis JM, Grattapaglia D, Resende MDV, Kirst M (2012b) Accelerating the domestication of trees using genomic selection: accuracy of prediction models across ages and environments. New Phytol 193:617–624
Rhode A, Ruttink T, Hostyn V, Sterck L, Van Driessche K, Boerjan W (2007) Gene expression during the induction, maintenance, and release of dormancy in apical buds of poplar. J Exp Bot 58:4047–4060
Riesberg EE, Hildebrandt U, Riederer M, Hentschel U (2013) Distinct phyllosphere bacterial communities on Arabidopsis wax mutant leaves. PLoS ONE 8:e78613
Rigal A, Yordanov YS, Perrone I, Karlberg A, Tisserant E, Bellini C, Busov VB, Martin F, Kohler A, Bhalerao R, Legué V (2012) The AINTEGUMENTA LIKE1 homeotic transcription factor PtAIL1 controls the formation of adventitious root primordia in poplar. Plant Physiol 160:1996–2006
Rinaldi C, Kohler A, Frey P, Duchaussoy F, Ningre N, Couloux A, Wincker P, Le Thiec D, Fluch S, Martin F, Duplessis S (2007) Transcript profiling of poplar leaves upon infection with compatible and incompatible strains of the foliar rust *Melampsora larici-populina*. Plant Physiol 144:347–366
Robinson AR, Dauwe R, Ukrainetz NK, Cullis IF, White R, Mansfield SD (2009) Predicting the regenerative capacity of conifer somatic embryogenic cultures by metabolomics. Plant Biotech J 7:952–963
Rodgers-Melnick E, Culp M, DiFazio SP (2013) Predicting whole genome protein interaction networks from primary sequence data in model and non-model organisms using ENTS. BMC Genomics 14: 608
Rodriguez JL, Pascual J, Viejo M, Valledor, L, Meijón M, Hasbún R, Yrei NY, Santamaría ME, Pérez M, Fraga MF, Berdasco M, Fernández RR, Cañal MJ (2012) Basic procedures for epigenetic analysis in plant cell and tissue culture. Methods Mol Biol 877:325–341
Rosvall O (2011) Review of the Swedish tree breeding programme: http://www.skogforsk.se/PageFiles/57629/Tree%20breeding.pdf
Rutledge RG, Stewart D, Caron S, Overton C, Boyle B, MacKay J, Klimaszewska K (2013) Potential link between biotic defense activation and recalcitrance to induction of somatic embryogenesis in shoot primordia from adult trees of white spruce (*Picea glauca*). BMC Plant Biol 13:116
Ruttinkt AM, Morreel K, Storme V, Rombauts S, Fromm J, Bhalerao RP, Boerjan W, Rhode A (2007) A molecular timetable for apical bud formation and dormancy induction in poplar. Plant Cell 19:2370–2390
Saghatelian A, Cravatt BF (2005) Global strategies to integrate the proteome and metabolome. Curr Opin Chem Biol 9:62–68
Samils B, Ronnberg-Wastljung AC, Stenlid J (2011) QTL mapping of resistance to leaf rust in *Salix*. Tree Genet Genomes 7:1219–1235
Sanchez C, Vielba JM, Ferro E, Covelo G, Sole A, Abarca D, de Mier BS, Diaz-Sala C (2007) Two SCARECROW-LIKE genes are induced in response to exogenous auxin in rooting-competent cuttings of distantly related forest species. Tree Physiol 27:1459–1470
Schoville SD, Bonin A, François O, Lobreaux S, Melodelima C, Manel S (2012) Adaptive genetic variation on the landscape: methods and cases. Annu Rev Ecol Evol S 43:23–43
Scotti-Saintagne C, Dick CW, Caron H, Vendramin GG, Guichoux E, Buonamici A, Duret C, Sire P, Valencia R, Lemes MR, Gribel R, Scotti I (2013a) Phylogeography of a species complex of lowland Neotropical rain forest trees (*Carapa*, Meliaceae). J Biogeogr 40: 676–692
Scotti-Saintagne C, Dick CW, Caron H, Vendramin GG, Troispoux V, Sire P, Casalis M, Buonamici A, Valencia R, Lemes MR, Gribel R, Scotti I (2013b) Amazon diversification and cross-Andean dispersal of the widespread Neotropical tree species *Jacaranda copaia* (Bignoniaceae). J Biogeogr 40:707–719
Sebastiana M, Vieira B, Lino-Neto T, Monteiro F, Figueiredo A, Sousa L, Pais MS, Tavares R, Paulo OS (2014) Oak root response to ectomycorrhizal symbiosis establishment: RNA-Seq derived transcript identification and expression profiling. PLoS ONE 9:e98376
Secchi F, Gilbert ME, Zwieniecki MA (2011) Transcriptome response to embolism formation in stems of *Populus trichocarpa* provides insight into signaling and the biology of refilling. Plant Physiol 157: 1419–1429
Slavov GT, DiFazio SP, Martin J, Schackwitz W, Muchero W, Rodgers-Melnick E, Lipphardt MF, Pennacchio CP, Hellsten U, Pennacchio LA, Gunter LE, Ranjan P, Vining K, Pomraning KR, Wilhelm LJ, Pellegrini M, Mockler TC, Freitag M, Geraldes A, El-Kassaby YA, Mansfield SD, Cronk QCB, Douglas CJ, Strauss SH, Rokhsar D, Tuskan GA (2012) Genome resequencing reveals multiscale geographic structure and extensive linkage disequilibrium in the forest tree *Populus trichocarpa*. New Phytol 196:713–725
Sork VL, Aitken SN, Dyer RJ, Eckert AJ, Legendre P, Neale DB (2013) Putting the landscape into the genomics of trees: approaches for understanding local adaptation and population responses to changing climate. Tree Genet Genomes 9:901–911
Spiess N, Oufir M, Matusikova I, Stierschneider M, Kopecky D, Homolka A, Burg K, Fluch S, Hausman JF, Wilhelm E (2012) Ecophysiological and transcriptomic responses of oak (*Quercus robur*) to long-term drought exposure and rewatering. Environ Exp Bot 77:117–126
Stasolla C, van Zyl L, Egertsdotter U, Craig D, Liu W, Sederoff RR (2003) The effects of polyethylene glycol on gene expression of developing white spruce somatic embryos. Plant Physiol 131:49–60
Stasolla C, Bozhkov PV, Chu TM, Van Zyl L, Egertsdotter U, Suarez MF, Craig D, Wolfinger RD, Von Arnold S, Sederoff RR (2004) Variation in transcript abundance during somatic embryogenesis in gymnosperms. Tree Physiol 24:1073–1085
Sterky F, Regan S, Karlsson J, Hertzberg M, Rohde A, Holmberg A, Amini B, Bhalerao R, Larsson M, Villaroel R, van Montagu M, Sandberg G, Olsson O, Teeri T, Boerjan W, Gustafsson P, Uhlen M, Sundberg B, Lundeberg J (1998) Gene discovery in the wood-forming tissues of poplar: analysis of 5,692 expressed sequence tags. Proc Natl Acad Sci USA 95:13330–13335
Stölting KN, Nipper R, Lindtke D, Caseys C, Waeber S, Castiglione S, Lexer C (2013) Genomic scan for single nucleotide polymorphisms reveals patterns of divergence and gene flow between ecologically divergent species. Mol Ecol 22:842–855
Stone GN, Nee S, Felsenstein J (2011) Controlling for non-independence in comparative analysis of patterns across populations within species. Phil T R Soc B 366:1410–1424
Street NR, Tallis MJ, Tucker J, Brosche M, Kangasjarvi J, Broadmeadow M, Taylor G (2011) The physiological, transcriptional and genetic responses of an ozone- sensitive and an ozone tolerant poplar and selected extremes of their F-2 progeny. Environ Pollut 159:45–54
Sturrock RN, Frankel SJ, Brown AV, Hennon PE, Kliejunas JT, Lewis KJ, Worrall JJ, Woods AJ (2011) Climate change and forest diseases. Plant Pathol 60:133–149
Teyssier C, Grondin C, Bonhomme L, Lomenech AM, Vallance M, Morabito D, Label P, Lelu-Walter MA (2011) Increased gelling agent concentration promotes somatic embryo maturation in hybrid larch (*Larix* x *eurolepsis*): a 2-DE proteomic analysis. Physiol Plant 141:152–165
Teyssier C, Maury S, Beaufour M, Grondin C, Delaunay A, Le Metté C, Ader K, Cadene M, Label P, Lelu-Walter M-A (2014) In search of markers for somatic embryo maturation in hybrid larch (*Larix* x *eurolepis*): global DNA methylation and proteomic analyses. Physiol Plant 150:271–291
Trontin J-F, Klimaszewska K, Morel A, Hargreaves C, Lelu-Walter M-A (2015). Molecular aspects of conifer zygotic and somatic embryo development: a review of genome-wide approaches and recent insights. In: In vitro plant embryogenesis in higher plants, Germanà

MA, Lambardi M (eds), Methods in molecular biology, Springer-Humana Press (in press)

Turchetto-Zolet AC, Pinheiro F, Salgueiro F, Palma-Silva C (2013) Phylogeographical patterns shed light on evolutionary process in South America. Mol Ecol 22:1193–1213

Turner TR, James EK, Poole PS (2013) The plant microbiome. Genome Biol 14:209

Tuskan GA, DiFazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, Ralph S, Rombauts S, Salamov A, Schein J, Sterck L, Aerts A, Bhalerao RR, Bhalerao RP, Blaudez D, Boerjan W, Brun A, Brunner A, Busov V, Campbell M, Carlson J, Chalot M, Chapman J, Chen GL, Cooper D, Coutinho PM, Couturier J, Covert S, Cronk Q, Cunningham R, Davis J, Degroeve S, Dejardin A, Depamphilis C, Detter J, Dirks B, Dubchak I, Duplessis S, Ehlting J, Ellis B, Gendler K, Goodstein D, Gribskov M, Grimwood J, Groover A, Gunter L, Hamberger B, Heinze B, Helariutta Y, Henrissat B, Holligan D, Holt R, Huang W, Islam-Faridi N, Jones S, Jones-Rhoades M, Jorgensen R, Joshi C, Kangasjarvi J, Karlsson J, Kelleher C, Kirkpatrick R, Kirst M, Kohler A, Kalluri U, Larimer F, Leebens-Mack J, Leple JC, Locascio P, Lou Y, Lucas S, Martin F, Montanini B, Napoli C, Nelson DR, Nelson C, Nieminen K, Nilsson O, Pereda V, Peter G, Philippe R, Pilate G, Poliakov A, Razumovskaya J, Richardson P, Rinaldi C, Ritland K, Rouze P, Ryaboy D, Schmutz J, Schrader J, Segerman B, Shin H, Siddiqui A, Sterky F, Terry A, Tsai CJ, Uberbacher E, Unneberg P, Vahala J, Wall K, Wessler S, Yang G, Yin T, Douglas C, Marra M, Sandberg G, Van de Peer Y, Rokhsar D (2006) The genome of black cottonwood, *Populus trichocarpa* (Torr. and Gray). Science 313:1596–1604

Ueno S, KloppC LJC, Derory J, Noirot C, Leger V, Prince E, Kremer A, Plomion C, Le Provost G (2013) Transcriptional profiling of bud dormancy induction and release in oak by next-generation sequencing. BMC Genomics 14:236–250

Uroz S, Buée M, Murat C, Frey-Klett P, Martin F (2010) Pyrosequencing reveals a contrasted bacterial diversity between oak rhizosphere and surrounding soil. Environ Microbiol Rep 2:281–288

van Acker R, Leplé JC, Aerts D, Storme V, Goeminne G, Ivens B, Piens K, Van Montagu M, Santoro N, Foster C, Ralph J, Soetaert W, Pilate G, Boerjan W (2014) Improved saccharification and ethanol yield from field-grown transgenic poplar deficient in cinnamoyl-CoA reductase. Proc Natl Acad Sci U S A 111:845–850

van der Merwe M, McPherson H, Siow J, Rossetto M (2014) Next-Gen phylogeography of rainforest trees: exploring landscape-level cpDNA variation from whole-genome sequencing. Mol Ecol Res 14:199–208

van Holme R, Morreel K, Darrah C, Oyarce P, Grabber JH, Ralph J, Boerjan W (2012a) Metabolic engineering of novel lignin in biomass crops. New Phytol 196:978–1000

van Holme R, Storme V, van Holme B, Sundin L, Christensen JH, Goeminne G, Halpin C, Rohde A, Morreel K, Boerjan W (2012b) A systems biology view of responses to lignin biosynthesis perturbations in *Arabidopsis*. Plant Cell 24:3506–3529

van Holme B, Cesarino I, Goeminne G, Kim H, Marroni F, van Acker R, van Holme R, Morreel K, Ivens B, Pinosio S, Morgante M, Ralph J, Bastien C, Boerjan W (2013) Breeding with rare defective alleles (BRDA): a natural *Populus nigra* HCT mutant with modified lignin as a case study. New Phytol 198:765–776

Vanderschuren H, Lentz E, Zainuddin I, Gruissem W (2013) Proteomics of model and crop plant species: status, current limitations and strategic advances for crop improvement. J Proteomics 93:5–19

Veluthakkal R, Dasgupta MG (2010) Pathogenesis-related genes and proteins in forest tree species. Trees 24:993–1006

Vestman D, Larsson E, Uddenberg D, Cairney J, Clapham D, Sundberg E, von Arnold S (2011) Important processes during differentiation and early development of somatic embryos of Norway spruce as revealed by changes in global gene expression. Tree Genet Genomes 7:347–362

Villalobos DP, Díaz-Moreno SM, Said ES, Cañas RA, Osuna D, van Kerckhoven SHE, Bautista R, Claros MG, Cánovas FM, Cantón MR (2012) Reprogramming of gene expression during compression wood formation in pine: coordinated modulation of S-adenosylmethionine, lignin and lignan related genes. BMC Plant Biol 12:100

Villar E, Klopp C, Noirot C, Novaes E, Kirst M, Plomion C, Gion JM (2011) RNA-Seq reveals genotype-specific molecular responses to water deficit in eucalyptus. BMC Genomics 12:538

Vining KJ, Pomraning KR, Wilhelm LJ, Priest HD, Pellegrini M, Di Y, Mockler TC, Freitag M, Strauss SH (2012) Dynamic DNA cytosine methylation in the *Populus trichocarpa* genome: tissue-level variation and relationship to gene expression. BMC Genomics 13:27

Vining K, Pomraning KR, Wilhelm LJ, Ma C, Pellegrini M, Di Y, Mockler TC, Freitag M, Strauss SH (2013) Methylome reorganization during in vitro dedifferentiation and regeneration of Populus trichocarpa. BMC Plant Biol 13:92

Wachowiak W, Salmela MJ, Ennos RA, Iason G, Cavers S (2011) High genetic diversity at the extreme range edge: nucleotide variation at nuclear loci in Scots pine (*Pinus sylvestris* L.) in Scotland. Heredity 106:775–787

Wang JP, Naik PP, Chen HC, Shi R, Lin CY, Liu J, Shuford CM, Li QZ, Sun YH, Tunlaya-Anukit S, Williams CM, Muddiman DC, Ducoste JJ, Sederoff RR, Chiang VL (2014a) Complete proteomic-based enzyme reaction and inhibition kinetics reveal how monolignol biosynthetic enzyme families affect metabolic flux and lignin in *Populus trichocarpa*. Plant Cell 26:894–914

Wang S, Li E, Porth I, Chen J-G, Mansfield SD, Douglas CJ (2014b) Regulation of secondary cell wall biosynthesis by poplar R2R3 MYB transcription factor PtrMYB152 in *Arabidopsis*. Sci Rep 4: 05054

Warren C, Aranda I, Cano FJ (2012) Metabolomics demonstrates divergent responses of two *Eucalyptus* species to water stress. Metabolomics 8:186–200

Wilcox PL, Amerson HV, Kuhlman EG, Liu BH, O'Malley DM, Sederoff RR (1996) Detection of a major gene for resistance to fusiform rust disease in loblolly pine by genomic mapping. Proc Natl Acad Sci U S A 93:3859–3864

Wolfender JL, Rudaz S, Choi YH, Kim HK (2013) Plant metabolomics: from holistic data to relevant biomarkers. Curr Med Chem 20:1056–1090

Xi Z, Ruhfel BR, Schaefer H, Amorim AM, Sugumaran M, Wurdack KJ, Endress PK, Matthews ML, Stevens PF, Mathews S, Davis CC (2012) Phylogenomics and a posteriori data partitioning resolve the Cretaceous angiosperm radiation Malpighiales. Proc Natl Acad Sci U S A 109:17519–17524

Xiao X, Yang F, Zhang S, Korpelainen H, Li C (2009) Physiological and proteomic responses of two contrasting Populus cathayana populations to drought stress. Physiol Plant 136:150–168

Yakovlev IA, Fossdal CG, Johnsen O, Junttila O, Skroppa T (2006) Analysis of gene expression during bud burst initiation in Norway spruce via ESTs from subtracted cDNA libraries. Tree Genet Genomes 2:39–52

Yakovlev IA, Lee Y, Rotter B, Olsen JE, Skrøppa T, Johnsen Ø, Fossdal CG (2014) Temperature-dependent differential transcriptomes during formation of an epigenetic memory in Norway spruce embryogenesis. Tree Genet Genomes 10:355–366

Yamane H, Kashiwa Y, Ooka T, Tao R, Yonemori K (2008) Suppression subtractive hybridization and differential screening reveals endodormancy-associated expression of an SVP/AGL24-type MADS-box gene in lateral vegetative buds of Japanese apricot. J Am Soc Hortic Sci 133:708–716

Yang X, Ye C-Y, Bisaria A, Tuskan GA, Kalluri UC (2011) Identification of candidate genes in Arabidopsis and Populus cell wall biosynthesis using text-mining, co-expression network analysis and comparative genomics. Plant Sci 181:675–687

Yu J, Hu S, Wang J, Wong GK, Li S, Liu B, Deng Y, Dai L, Zhou Y, Zhang X, Cao M, Liu J, Sun J, Tang J, Chen Y, Huang X, Lin W, Ye C, Tong W, Cong L, Geng J, Han Y, Li L, Li W, Hu G, Huang X, Li W, Li J, Liu Z, Li L, Liu J, Qi Q, Liu J, Li L, Li T, Wang X, Lu H, Wu T, Zhu M, Ni P, Han H, Dong W, Ren X, Feng X, Cui P, Li X, Wang H, Xu X, Zhai W, Xu Z, Zhang J, He S, Zhang J, Xu J, Zhang K, Zheng X, Dong J, Zeng W, Tao L, Ye J, Tan J, Ren X, Chen X, He J, Liu D, Tian W, Tian C, Xia H, Bao Q, Li G, Gao H, Cao T, Wang J, Zhao W, Li P, Chen W, Wang X, Zhang Y, Hu J, Wang J, Liu S, Yang J, Zhang G, Xiong Y, Li Z, Mao L, Zhou C, Zhu Z, Chen R, Hao B, Zheng W, Chen S, Guo W, Li G, Liu S, Tao M, Wang J, Zhu L, Yuan L, Yang H (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). Science 296:79–92

Yu H, Soler M, Mila I, San Clemente H, Savelli B, Dunand C, Jorge APP, Myburg A, Bouzayen M, Grima-Pettenati J, Cassan-Wang H (2014) Genome-wide characterization and expression profiling of the *AUXIN RESPONSE FACTOR* (*ARF*) gene family in *Eucalyptus grandis*. PLoS ONE 9:e108906

Zapata-Valenzuela J, Isik F, Maltecca C, Wegrzyn J, Neale D, McKeand S, Whetten R (2012) SNP markers trace familial linkages in a cloned population of *Pinus taeda*—prospects for genomic selection. Tree Genet Genomes 8:1307–1318

Zapata-Valenzuela J, Whetten RW, Neale DB, McKeand SE, Isik F (2013) Genomic estimated breeding values using genomic relationship matrices in a cloned population of loblolly pine. Genes Genomes Genet 3:909–916

Zhang P, Dreher K, Karthikeyan A, Chi A, Pujar A, Caspi R, Karp P, Kirkup V, Latendresse M, Lee C, Mueller LA, Muller R, Rhee SY (2010) Creation of a genome-wide metabolic pathway database for *Populus trichocarpa* using a new approach for reconstruction and curation of metabolic pathways for plants. Plant Physiol 153:1479–1491

Zhang J, Zhang S, Han S, Wu T, Li X, Li W, Qi L (2012a) Genome-wide identification of microRNAs in larch and stage-specific modulation of 11 conserved microRNAs and their targets during somatic embryogenesis. Planta 236:647–657

Zhang Y, Zhang S, Han S, Li X, Qi L (2012b) Transcriptome profiling and in silico analysis of somatic embryos in Japanese larch (*Larix leptolepis*). Plant Cell Rep 31:1637–1657

Zhao Q, Nakashima J, Chen F, Yin Y, Fu C, Yun J, Shao H, Wang X, Wang ZY, Dixon RA (2013a) Laccase is necessary and nonredundant with peroxidase for lignin polymerization during vascular development in *Arabidopsis*. Plant Cell 25: 3976–3987

Zhao YJ, Song DL, Sun JY, Li LG (2013b) *Populus* endo-beta-mannanase PtrMAN6 plays a role in coordinating cell wall remodeling with suppression of secondary wall thickening through generation of oligosaccharide signals. Plant J 74:473–485

Zhong R, McCarthy RL, Lee C, Ye ZH (2011) Dissection of the transcriptional program regulating secondary wall biosynthesis during wood formation in poplar. Plant Physiol 157:1452–1468

Zhu Y, Song D, Sun J, Wang X, Li L (2013a) PtrHB7, a class III HD-Zip gene, plays a critical role in regulation of vascular cambium differentiation in *Populus*. Mol Plant 6:1331–1343

Zhu R, Shevchenko O, Ma C, Maury S, Freitag M, Strauss SH (2013b) Poplars with a PtDDM1-RNAi transgene have reduced DNA methylation and show aberrant post-dormancy morphology. Planta 237: 1483–1493

Zimin A, Stevens KA, Crepeau MW, Holtz-Morris A, Koriabine M, Marcais G, Puiu D, Roberts M, Wegrzyn JL, de Jong PJ, Neale DB, Salzberg SL, Yorke JA, Langley CH (2014) Sequencing and assembly of the 22-gb loblolly pine genome. Genetics 196: 875–890