

The Genomics of Wood Formation in Angiosperm Trees

Xinqiang He and Andrew T. Groover

Abstract Advances in genomic science have enabled comparative approaches that can evaluate the evolution of genes and mechanisms underlying phenotypic traits relevant to angiosperm forest trees. Wood formation is an excellent subject for comparative genomics, as it is an ancestral trait for angiosperms and has undergone significant modification in different angiosperm lineages. This chapter discusses some of the traits associated with wood formation, what is currently known about the genes and mechanisms regulating these traits, and how comparative evolutionary genomic studies can be undertaken to provide more comprehensive views of the evolution and development of wood formation in angiosperms.

Keywords Wood formation • Wood development • Wood evolution • Transcriptional regulation • Epigenetics • Comparative genomics

Genomic Perspectives on the Evolutionary Origins and Variation in Angiosperm Wood

Introduction

The evolutionary and developmental biology of wood are fundamental to understanding the amazing diversification of angiosperms. While flower morphology and reproductive characters associated with angiosperms have been the focus of numerous genetic and genomic evo-devo studies, wood development is a relatively neglected trait of angiosperm evolution and development that is now tractable for comparative and evolutionary genomic studies. The ability to produce wood from a

X. He

School of Life Sciences, Peking University, Beijing 100871, China

A.T. Groover (✉)

Pacific Southwest Research Station, US Forest Service, Davis, 95618 CA, USA

Department of Plant Biology, University of California Davis, Davis, 95618 CA, USA

e-mail: agroover@fs.fed.us

vascular cambium is an ancestral trait of the angiosperms, but has been extensively modified in both basal lineages as well as in more recent, derived taxa. Interestingly, some wood development traits have arisen independently in unrelated taxa, presenting the question of whether common or multiple mechanisms can produce similar developmental changes.

Wood is the product of the vascular cambium. The vascular cambium is a lateral meristem whose initials divide to provide daughter cells that differentiate into secondary xylem (wood) towards the center of the stem, or secondary phloem (inner bark) towards the outside of the stem (Larson 1994). This radial growth from a vascular cambium in woody stems is collectively referred to as secondary growth. For most angiosperms, there are two types of cambial initials. Fusiform initials produce the axial tissues and cell types that can include water conducting tracheary elements, fibers, and xylem parenchyma. Ray initials produce the radial tissues, the rays. Rays are believed to serve various functions including radial transport across the stem, storage of water and nutrients, and biochemical functions. The axial tissues function primarily in water and nutrient conduction, and mechanical support. Significant quantitative, biochemical, and morphological variation exists among angiosperms for many aspects of secondary growth. For example, striking examples of morphological variation include gain and loss of the cambium, successive cambia, xylem furrowed by phloem, and interxylary phloem (Spicer and Groover 2010). The mechanisms underlying this variation are only recently being revealed.

Wood serves multiple functions, including water conduction, water and nutrient storage, and mechanical support. Wood development thus reflects developmental mechanisms that integrate environmental cues to produce highly complex tissues (Arnold and Mauseth 1999). A familiar example of this integration is the annual rings of many temperate tree species, which reflect modification of wood development in response to the changing environmental conditions within and between growing seasons. How this integration occurs is still opaque, in part because of traditional divisions among the different disciplines associated with wood development. Indeed, wood development has been studied since the earliest days of botany and has provided important insights into the evolution and development of plants in general, and of trees in particular (Baas 1982; Groover and Cronk 2013). More recently, advanced genomic approaches have been applied to wood development, including studies in a limited number of angiosperm model tree species with fully sequenced genomes. Integrating genomic studies with classical literature and traditional disciplines including anatomy and physiology could be extremely effective for comparative studies. An ultimate goal is to describe the molecular mechanisms regulating wood formation, structure and physiological functions, including how these biological processes have been modified in different angiosperm lineages and how environmental cues are integrated into developmental processes.

Arguably, wood formation is an excellent subject for comparative and evolutionary genomic studies because (1) wood development is experimentally tractable; (2) wood is of extreme economic and ecologic importance; (3) wood shows incredible phenotypic variation across angiosperm taxa; (4) there is an extensive literature detailing angiosperm wood anatomy; and (5) insights are being made into at least some of the regulatory mechanisms regulating wood development.

Additionally, detailed angiosperm phylogenies combined with new sequencing technologies make possible the comparison of gene function across diverse angiosperm taxa, even for non-model species that display unique anatomical or developmental attributes. This chapter presents some of the fundamental information and concepts underlying current and future comparative and evolutionary genomic approaches for wood development in angiosperms.

Expectations for Evolution of Regulatory Mechanisms Regulating Wood Development

Wood development can be subdivided into interrelated processes including the regulation of cell division in the cambial zone, cell expansion and differentiation in the developing xylem, tissue patterning (e.g. the location of vessels within the wood), and numerous other traits. Wood development is also highly plastic, and varies based on stage of development, time of year, and in response to myriad environmental and physiological cues. Anatomical variation has been well-described for large numbers of angiosperms for many wood-related traits in the classical wood anatomy and wood paleobotanical literature (Carlquist 2001). At the same time, increasingly robust phylogenies for angiosperms are being produced using DNA sequence (The Angiosperm Phylogeny 2009). Designing comparative and evolutionary genomic studies of wood formation can take advantage of this information. As discussed later in this review, different phylogenetic scales can be considered for such studies, ranging from ancestral traits that have been modified in different angiosperm lineages, to traits that have arisen independently through convergent evolution in different taxa.

There are a number of wood-related traits that were present in ancestral angiosperms that could be excellent case studies for comparative genomic studies – for example, the tracheary element. Tracheary elements are water-conducting cells that undergo an elaborate differentiation process that supports biosynthesis of a lignified secondary cell wall, and ultimately ends in programmed cell death and lysis to produce a hollow cell corpse (Groover and Jones 1999; Escamez and Tuominen 2014). There are two major forms of water-conducting tracheary elements: tracheids, which are elongated cells with pits that facilitate water transport, and vessel elements, which tend to have larger diameters than tracheids and have openings at each end of the cell, termed perforation plates, that vary from ladder-like scalariform perforations to a simple opening (Esau 1977; Bliss 1921). Within the angiosperms, extensive variation can be seen for tracheary element morphology at different taxonomic levels, among organs within a species, among individuals within a species in response to different environmental conditions, and within individuals at different stages of development (Carlquist and Schneider 2002). This variation in tracheary element morphology and associated functional traits has been the basis of major hypotheses of early angiosperm evolution and diversification (Carlquist 2001, 2009; Bailey 1944). Early angiosperms are characterized by specific wood traits including tracheids, long cambial initials, and upright ray cells. The wood of more derived angiosperms possess vessel elements (or vessel

elements and tracheids, or tracheary elements with features of both vessel elements and tracheids), contains parenchyma that can serve to refill embolized vessels, and have procumbent ray cells (Carlquist 2009). Modification of wood anatomies and associated physiological conditions in different lineages likely played a primary role in angiosperm diversification and expansion into extremely diverse habitats. Indeed, xylem heterochrony is a primary factor in understanding the evolution and diversification of early angiosperms (Carlquist 2009).

There are striking innovations in wood anatomy among angiosperms. For example, for at least for some if not all species with successive cambia, a “master cambium” produces conjunctive tissue along with new cambia that then produce secondary xylem and phloem (Carlquist 2007). The result is a stem with repeating layers of secondary xylem and phloem embedded in a matrix of conjunctive tissue. Successive cambia have evolved within 14 orders scattered through the eudicots, and may enable flexibility in woody stems and lianas (Carlquist 2007; Spicer and Groover 2010).

Wood properties can also be examined from an ecological evolutionary perspectives. Since wood is the water conducting tissue of stems, wood properties ultimately determine the ability of a species or individual to grow in different habitats or respond to environmental stresses such as drought. For example, changes in wood properties (specifically, smaller tracheary elements less prone to cavitation) was a major factor in the successful movement of woody angiosperms into freezing environments (Zanne et al. 2014). Interestingly, woody perennial growth has also been found to be associated with reduced exploration of climate space (Smith and Beaulieu 2009) and slower rates of molecular evolution (Smith and Donoghue 2008). Parenchymatous woods illustrate how anatomy affects physiology and adaptive traits, for example in cacti (Mauseth and Plemons-Rodriguez 1998). The number and diameter of tracheary elements produced can vary dramatically in response to water availability, although the mechanisms controlling these responses are poorly understood. These and related topics regarding wood evolution and development have direct relevance to understanding how different species will respond to ongoing climate change.

As discussed below, genes and mechanisms regulating wood formation have been identified that can now be examined in an evolutionary context. While this review cannot cover all of these subjects comprehensively, some examples of genes and mechanisms regulating wood formation are presented in the next section.

Transcriptional Regulation of Wood Formation

Transcription appears to be one of the primary points of regulation for wood formation. This can be inferred from studies in which gene expression was measured using microarrays for successive tissue samples were taken across a developing *Populus* stems, from the phloem across the cambium and into different stages of secondary xylem maturation (Schrader et al. 2004). In general, changes in gene expression across the tissue profiles are highly correlated to developmental events,

with obvious correlations seen between the function of genes associated with xylem development and processes occurring during progressive stages of tissue development (e.g. cell division, cell elongation, and cell wall biosynthesis). Additionally, as discussed below, key transcription factors have been identified that regulate specific aspects of secondary growth and wood formation.

A comprehensive view of the developmental mechanisms underlying wood formation and secondary growth is still emerging. As in other plants and animals, the transcriptional networks regulating gene expression are highly complex, but are beginning to be modeled through integration of different genomic data types using computational approaches. The complexity of transcriptional regulation is shown by ChIP-seq results indicating that individual transcription factors can bind many hundreds or thousands of loci throughout the *Populus* genome (Liu et al. 2015a). Additionally, there is relatively low correlation between the binding of the individual transcription factors surveyed to date and the expression of bound genes (Liu et al. 2015b), precluding simple models relying on strong correlation of transcription factor binding and transcriptional outcomes for target genes. The lack of strong correlation between transcription factor binding and transcription of target genes may reflect that most genes require the cooperative binding of multiple factors and permissive chromatin states to result in expression change. It may also reflect the type of transcription factors surveyed to date, and “master regulator” transcription factors such as NACs may show more direct correlations between binding and gene expression. Nonetheless, it is obvious that, as has been found by the humane ENCODE project (Consortium 2012), transcription factors tend to bind to large numbers of targets and predicting gene expression requires data describing the binding of multiple transcription factors and chromatin states to be robust.

Several genomic-based approaches have been taken to uncover genes and mechanisms regulating wood formation in angiosperms. Each approach has different strengths and caveats, and ultimately surveys genes affecting wood formation at different levels. For example, association mapping identified single nucleotide polymorphisms in *Populus* that were correlated with a number of wood chemistry and ultrastructure phenotypes for a large population-based survey (Porth et al. 2013b). Candidate genes identified with strongest effects on phenotypes were primarily structural genes (23 genes), but at least three transcription factors were identified. Limitations to association mapping include that only genes with allelic variation having significant effects on phenotypes can be identified, and the large majority of genes involved in wood development are transparent to this approach. However, the approach does provide insight into the mechanisms and specific genes that are being acted upon by selection. More comprehensively, gene expression during wood formation has been profiled using microarrays and sequencing-based approaches to provide complete catalogues of wood-related genes in species including poplar (Schrader et al. 2004), Eucalyptus (Etienne Paux 2005) and Arabidopsis (Zhao et al. 2005). These approaches do not directly provide functional insights into gene networks or natural genetic variation, however.

Characterization of individual transcription factors has provided important insights into the regulation of wood formation. Transcription factors regulating wood development have been most extensively characterized in Arabidopsis and

Populus. While the rosette-form of *Arabidopsis* lacks fundamental features of arborescent angiosperms such as rays, it does possess a vascular cambium and makes limited secondary xylem in the vegetative rosette. *Arabidopsis* can also produce limited secondary vascular tissue and interfascicular fibers in the inflorescence stem. Nonetheless, it should be noted that some aspects of wood formation may not be well represented in *Arabidopsis*.

Developmental genetic, yeast one hybrid and other approaches have defined a hierarchy of NAC-domain and MYB-related transcription factors that are master regulators of tracheary element and fiber differentiation (Brady et al. 2011; Taylor-Teeples et al. 2015; Zhong et al. 2008; Zhong and Ye 2015). Generally, NAC-domain transcription factors occupy the top levels of this hierarchy, and regulate MYB-related transcription factors that in turn regulate genes encoding enzymes participating in secondary cell wall synthesis. Interestingly, misexpression of *VASCULAR-RELATED NAC-DOMAIN6* (*VND6*) results in the differentiation of parenchyma into metaxylem-like vessel elements, while misexpression of *VND7* results in differentiation of parenchyma into protoxylem-like vessel elements (Kubo et al. 2005). Furthermore, dominant repression of *VND6* and *VND7* specifically inhibits metaxylem and protoxylem vessel formation in roots, respectively. Other NAC family members including *NSD1*, *NST1* and *NST2* regulate the differentiation of fibers, and *nsd1/nst1/nst2* mutants fail to produce secondary cell walls in fibers (Zhong and Ye 2015; Zhong et al. 2006). Thus, NAC-domain transcription factors define mechanisms fundamental to the diversification of the ancestral tracheid cell type into vessels with different morphologies as well as fibers.

New insights into the regulation of the cambium and secondary vascular tissue differentiation was initially identified using the *in vitro* system for tracheary element differentiation in *Zinnia* (Ito et al. 2006). In this system, isolated mesophyll cells are induced to differentiate into tracheary elements. Differentiation is dependent on exogenous nutritive factors, hormones and a secreted peptide factor, TDIF (Ito et al. 2006). In the plant, TDIF is expressed in the secondary phloem and is perceived by the LRR-like receptor, TDIFRECEPTOR/PHLOEMINTERCALATED WITH XYLEM (TDR/PXY). TDR/PXY in turn modulates expression of a WOX-like transcription factor, WOX4, that regulates the rate of cambial cell division (Hirakawa et al. 2008 2010; Etchells and Turner 2010). Interestingly, this signaling circuit appears to be evolutionarily ancient, with TDIF activity conserved among extant euphyllophytes (Hirakawa and Bowman 2015). Variation in TDIF signaling is thus an attractive mechanism to survey in angiosperm tree species.

More obscure are the mechanisms that regulate the patterning of secondary vascular tissues and the rate of differentiation of cambial derivatives. Class I KNOX transcription factors have been identified that are expressed both in the shoot meristem and cambial zone of *Populus* (Du et al. 2009; Groover et al. 2006). Both *ARBORKNOX1* (*ARK1*) and *ARK2* alter the differentiation of lignified cells within secondary xylem when misregulated. This function is similar to better characterized orthologs in *Arabidopsis* (*SHOOTMERISTEMLESS* and *BREVIPEDACELLUS*) which generally act to repress differentiation within the shoot apical meristem (Long et al. 1996). The expression of these transcriptional regulators in both the

apical and cambial meristems presents an interesting opportunity to better understand the cooption of genes and mechanisms from the shoot apical meristem during the evolution of the cambium.

Another potential example of cooption of mechanisms to the cambium from the shoot apical meristem is given by Class III HD ZIP transcription factors. This small family of transcription factors includes members that function in the patterning and polarity of lateral organs and vascular tissues. For example, mutations in the *Arabidopsis* Class III HD ZIP, *REVOLUTA*, conditions phenotypes that include adaxialization of both lateral organs and vascular tissues. Mechanistically, *REVOLUTA* acts antagonistically with another small family of transcriptional regulators, *YABBYs*. In *Populus*, misexpression of the *REVOLUTA* ortholog results in mispatterning of secondary vascular tissues, including reversal of polarity of xylem and phloem and formation of ectopic cambium in the cortex (Robischon et al. 2011). The Class III HD ZIPs are evolutionarily ancient, but functions such as polarity regulation in secondary vascular tissues are derived (Floyd et al. 2006).

Promising advances are being made towards comprehensively modeling transcriptional networks, and identifying correlations between network features and phenotypes. For example, integration of genetic, genomic and phenotypic data was used to identify small, directed networks of genes affecting wood biochemistry (Porth et al. 2013a). Gene co-expression studies can be used to cluster genes into modules that show similar expression pattern across different genotypes, mutants, or in response to experimental treatments. Correlations can then be tested between phenotypic traits and the module eigengene (conceptually, the average) expression values. This approach has the advantage of reducing the dimensionality of the data and minimizes the problems associated with multiple testing commonly encountered in single-gene approaches. In one example, microarray data were used to identify gene modules associated with leaf development in *Populus*, including putative mechanisms conserved between *Populus* and *Arabidopsis* (Street et al. 2011).

Network-based approaches were used for wood development in a recent study of tension wood development. Different *Populus* *ARK2* genotypes were subjected to gibberellic acid (GA) or control treatments, and then either placed horizontally to induced tension wood formation or left upright (Gerttula et al. 2015). Differentiating tension wood, opposite wood and normal wood from the trees and used for mRNA-sequencing. The resulting transcript abundance data were subjected to a co-expression analysis to place genes into modules, which were then correlated to wood properties, wood types and treatments. Genes within modules were then further analyzed to identify candidate transcription factors responsible for regulating the expression of other genes within each module. This proof of concept study could be expanded by including additional genetic or experimental perturbation, which would allow more precise assignment of genes to smaller gene modules. Importantly, these same approaches and concepts can be used to calculate and compare co-expression networks across species, supporting analyses capable of identifying ancestral mechanisms such as regulation of cambium division as well as taxa-specific derived traits.

Post-transcriptional Regulation of Wood Formation

Wood formation is regulated and modified post-transcriptionally by microRNA, by the regulation of protein abundance, and by a host of post-transcriptional protein modifications. The relative abundance of specific gene transcripts can be negatively regulated by non-coding microRNAs (miRNAs). At the protein level, transcript stability, rate of translation, protein stability and other factors can result in significant differences between relative rates of transcription and actual protein levels for a given gene. Additionally, phosphorylation, glycosylation, lipidation, ubiquitination and other post-translational modifications can have major effects on the abundance, localization, and activity of proteins. As discussed in this section, 'omics' approaches are defining the roles of these various types of post-transcriptional modifications during wood formation.

miRNA Regulation of Transcript Abundance During Wood Formation

miRNAs are short, nuclear-encoded non-coding RNAs involved post-transcriptional regulation of gene expression. In plants, the post-transcriptional regulation by miRNAs is achieved by first processing a miRNA precursor by Dicer-Like1 (Dcl1) into a 21 nucleotide miRNA/miRNA* duplex. The processed miRNA is then incorporated into an Argonaute-associated miRNA-induced silencing complex (miRISC). Specific complementary base pairing between a given miRNA and transcripts from target genes results in cleavage of the target transcript by the miRISC (Meng et al. 2011). miRNAs have been recognized to play crucial roles in diverse biological processes in plants and animals, including cambium differentiation and wood formation in *Arabidopsis*, as well as *Populus* and other tree species (Sun et al. 2012).

In a pioneering study of miRNAs in wood formation (Lu et al. 2005), miRNA families were identified by DNA sequencing of developing secondary xylem tissues of *Populus trichocarpa*. In this study, comparisons were made of tension wood and opposite wood from leaning stems, and normal wood from upright trees. Among the miRNAs identified, 12 families were either identical or very similar to *Arabidopsis* miRNAs, suggesting that these represent evolutionarily-conserved mechanisms. Interestingly, 10 *Populus* miRNA families were not conserved with *Arabidopsis*, and the majority of these non-conserved miRNAs were associated with tension wood and/or opposite wood formation. These results indicate that there are species-specific miRNAs, and that miRNAs may regulate tree-related traits such as tension wood formation (Lu et al. 2005).

Evolutionarily conserved miRNAs have been implicated in regulating specific processes underlying wood formation. miR165/166 directly targets transcripts encoding Class III HD-ZIP transcription factors that have been shown to play central roles in organ and vascular tissue polarity, and regulate xylem differentiation in

Arabidopsis roots (Carlsbecker et al. 2010) and shoots (Emery et al. 2003). The function and conserved role of Class III HD-ZIP genes has also been studied in *Populus*. Overexpression of miRNA-resistant Class III HD-ZIP *POPCORONA* resulted in delayed lignification of xylem and phloem fibers (Du et al. 2011), while overexpression of a miRNA-resistant form of *popREVOLUTA* resulted in the formation of ectopic cambial layers with reversed polarity within cortical parenchyma (Robischon et al. 2011). miRNAs also regulate structural genes involved in wood formation including laccase-encoding genes, whose protein products mediate the polymerization of monolignols during lignification. The *Populus* miRNA, Ptr-miR397, was found to directly target 29 of 49 predicted *Populus* laccase gene transcripts (Lu et al. 2013).

miRNAs also appear to be involved in perennial regulation of secondary growth and wood formation. Deep sequencing of short RNAs in the cambium zone of *Populus* stems identified more than 100 miRNAs with significant expression changes between active growth and dormancy, including developmental-, phytohormone- and stress-related miRNAs. Most of the development-related miRNAs were enriched in the active growth stage, such as miR164, miR396, miR168, miR319, miR171 (Ding et al. 2014). In contrast, miR166, which targets Class III HD-ZIP transcripts, was more abundant in dormancy and had low expression level during active growth (Ding et al. 2014; Ko et al. 2006), indicating that miR166 plays a role in regulating the annual growth cycle.

Profiling miRNAs expressed during wood formation has also been undertaken in a handful of other tree species. In *Eucalyptus*, the investigation of miRNAs from xylem, phloem and leaves identified 20 miRNAs of 5 families known in other plants and 28 novel miRNAs of 8 additional families (Victor 2006). In *Acacia*, 6 highly conserved miRNA families were identified as the potential regulators of secondary wall formation in xylem development, including miR166, miR172, miR168, miR159, miR394 and miR156.

Regulation of Protein Abundance and Post-translational Modification During Wood Formation

The term proteome refers to the total set of proteins and post-translational modifications found in a biological sample. A primary goal of proteomics for wood formation is to determine where and when thousands of individual proteins are produced in secondary vascular tissues, and how they are modified and interact.

Different approaches have been used to catalogue the proteins involved in secondary growth and wood formation. In an early wood-related proteomics study, 15 proteins were identified from differentiating xylem, mature xylem and bark of *Populus* by 2-Dimensional Electrophoreses (2-DE) coupled with mass spectrometry (MS), including key participants of the phenylpropanoid pathway and lignification (Mijnsbrugge et al. 2000). In addition, highly abundant proteins with unknown function were identified, underscoring that many genes and proteins vital to wood

formation remain anonymous. In *E. grandis*, the cambial region proteome was compared for three ages of growth, and 240 proteins of various putative functions were identified using a 2-DE-LC-MS/MS strategy (Fiorani et al. 2007). From xylem sap of *P. trichocarpa* × *P. deltoides*, 97 proteins were identified, including metabolic and glycolytic enzymes and defense proteins (Dafoe and Constabel 2009). The regeneration of secondary vascular tissues after removal of bark provides another useful experimental system for the study of secondary growth. Using this system, 2-DE was used in combination with MALDITOF-MS to identify 244 differentially expressed proteins during the secondary vascular tissue regeneration in *Populus*, 199 of which were assigned and classified under different functional classes including metabolism, signaling, cytoskeleton functions, cell cycle and secondary cell wall formation (Du et al. 2006). In another study performed by shotgun proteomics on xylem and phloem tissues from two *Populus* species: *P. deltoides* and *P. tremula* × *alba*, 7505 proteins were identified, 2627 were confidently identified in both xylem and phloem, 606 unique in xylem and 461 in phloem (Abraham et al. 2012). As can be seen from these studies, technical advances in proteomics are quickly providing a more accurate and informative characterizations of the proteins involved in wood formation.

Proteomic approaches have also been applied to more specific aspects of wood formation in angiosperm trees, including reaction wood formation. Using 2-DE, 140 protein species were identified in the upper side of a leaning stems of *E. gunnii*, with 12 proteins significantly associated with tension wood formation (Plomion et al. 2003). A proteomic approach of tension wood formation in *Populus* revealed 39 proteins, primarily cell wall-related, in the G-layer in mature xylem by 2-DE and gel-free MS methods (Kaku et al. 2009). A quantitative proteomic and phosphoproteomic analysis of tension wood formation in *Populus* revealed remarkable developmental plasticity, identifying 1155 proteins and phosphorylation events in comparison of normal wood and tension wood (Mauriat et al. 2015). These results underscores the importance of rapid and reversible post-translational modifications through phosphorylation during wood formation.

Other studies have sought to identify proteins that are uniquely expressed during secondary vascular development. An investigation of plasma membrane proteins from leaves, xylem, and cambium/phloem in *Populus* found that proteins involved in cell wall and carbohydrate metabolism, membrane trafficking were most abundant in the xylem plasma membranes, in agreement with the large role of cell wall biosynthesis in wood formation. Interestingly, the proteins uniquely found in xylem plasma membranes included enzymes involved in lignin biosynthesis, suggesting that they may exist as a complex linked to the plasma membrane, possibly in close proximity to a transporter translocating lignin monomers across the plasma membrane (Nilsson et al. 2010). Another proteomic analysis of the membrane proteins in differentiating secondary vascular tissues of *Populus* (Song et al. 2011) found a total of 226 proteins identified as integral plasma membrane proteins, including receptors, transporters, cell wall formation related or intracellular trafficking proteins. In particular, a group of RLKs were identified in the differentiating xylem and phloem, suggesting their involvement in secondary vascular development. An

endo-1,4- β -mannanase protein identified in the plasma membrane of differentiating xylem tissue potentially produces oligosaccharides that could serve as signaling molecules to suppress cell wall thickening (Zhao et al. 2013).

Culture-based systems for tracheary element differentiation can provide populations of synchronously differentiating cells for proteomics. The secondary cell wall patterning of tracheary elements is guided by underlying microtubules. Using an in vitro system for tracheary element differentiation, microtubule pulldown experiments paired with quantitative proteomic analysis of labeled microtubule interacting proteins identified 605 microtubule interacting proteins associated with specific stages of differentiation. The proteins associated with membrane trafficking, protein synthesis, DNA/RNA binding, and signal transduction peaked during secondary cell wall formation, while proteins associated with stress peaked when approaching tracheary element cell death (Derbyshire et al. 2015). This study thus provided an entire functional microtubule interactome during tracheary element formation, and expanded our understanding of the complexity of microtubule function in xylem development.

Ultimately the integration of proteomics data with other genomic and phenotypic data is required to comprehensively understand a given biological process, requiring the use of systems biology approaches. Bylesho et al (2009) devised a strategy for data generation and integration to model systematic changes in transcript, protein and metabolite profiles associated with lignin biosynthesis in hybrid aspen (Bylesjo et al. 2009). The joint covariation for all profiling platforms was calculated using multiple O2PLS models, and the results quantified genotype-specific perturbations affecting lignin biosynthesis and growth (Bylesjo et al. 2009). A systems biology approach was also applied to analyze effects of oxidative stress in *Populus* after acquiring transcriptomic, proteomic and metabolomic profiles of the cambial region of hybrid aspen plants, and then a multivariate analysis method OnPLS was used to integrate the three types of ‘omics’ data. The results provided a first comprehensive model of multi-level responses to oxidative stress in the vascular cambium (Srivastava et al. 2013).

Comparative Evolutionary Genomics Approaches for Wood Formation

There are a number of established approaches for comparative and evolutionary genomic studies that have been applied in both plants and animals, which could now be applied to wood formation in angiosperm trees. At one end of the spectrum, one common approach is to make detailed comparisons (e.g. of expression pattern) for one or a few genes of interest across species with interesting variation for a trait. At the other extreme, comparative genomic approaches can be used to compare DNA sequence, gene expression, or other comprehensive “omics” data within and across species of interest. Comparative genomic approaches are both more comprehensive as well as more powerful than one-gene-at-a-time approaches, and allow integration

of quantitative genomic, morphological, biochemical, and other wood-related data. This approach also has the advantage of discovering novel genes and mechanisms that have not been previously described.

A critical challenge for most comparative genomics projects is the determination of orthologous relationships of genes across species. For angiosperms, this is especially challenging given the complex evolutionary histories of angiosperm genomes, with many lineages characterized by whole genome duplication events, ploidy variation, hybridization and structural variation (Soltis et al. 2009). Conceptually and practically, the determination of orthologous genes (formally, genes that are homologous or of common decent in the species being compared) varies depending on the phylogenetic distance and evolutionary history of the genomes of the species being compared. At one extreme, closely related species that have not undergone any genome duplications or rearrangements since diverging from a common ancestor have relatively simple and direct orthology (e.g. orthologous genes having one-to-one orthologous relationships). In cases where one species lineage has undergone whole genome duplication since divergence from a common ancestor, homology may be characterized by one-to-two or one-to-many orthologous relationships. Several different computational approaches have been used to define orthologous relationships, and software is available for performing reciprocal BLAST-based approaches, tree-based approaches, and graph-based clustering approaches for estimating orthologous groups, for example see (Huerta-Cepas et al. 2016; Nakaya et al. 2013; Afrasiabi et al. 2013; Ye et al. 2013; Fischer et al. 2011; Li et al. 2003). In the case of related species with fully sequenced genome, syntenic relationships can be used to further infer orthologous relationships in addition to simple sequence data (Lechner et al. 2014). Morphological or experimental data can ultimately be integrated with orthology data to ask questions regarding the genetic evolution of traits, including wood formation.

For the study of the evolution of wood formation in angiosperms, one approach would be to identify genes and pathways regulating wood formation that have been described in model systems (e.g. *Populus* or *Arabidopsis*) and survey their diversification among different lineages. This general approach has been successful in providing fundamental insights into the variation in flower (Becker et al. 2011; Irish and Litt 2005) and leaf (Tsiantis and Hay 2003; Dkhar and Pareek 2014; Tsukaya 2014; Champagne et al. 2007) morphology within the angiosperms. Examples of genes and mechanisms associated with wood formation that could be surveyed include those previously discussed, such as NAC-domain, Class I KNOX and Class III HD ZIP transcription factors, or the mechanisms described by the TDIF/CLE TDR/PXY WOX signaling pathway. However, the selection of candidate genes for study is still relatively subjective, and differences in sequence resources (e.g. genome-level sequence) and experimental tractability (e.g. ability to transform) among species makes this candidate gene approach challenging and non-comprehensive. Indeed, because efficient transformation is difficult for many woody species, performing functional assessment of individual genes in multiple species

would be difficult. However, techniques for visualizing gene transcripts using in situ hybridization or protein epitopes using immunolocalizations within tissue sections have been extended to woody models including *Populus* (Du et al. 2009; Gerttula et al. 2015), and are relatively transferable across species. Examining differences in expression patterns in wood forming tissues during secondary growth could potentially reveal informative differences in the timing or spatial expression of candidate genes among species with unique anatomical, biochemical or other differences in wood development.

A more comprehensive and widely applicable approach would be to use gene expression-based comparative genomics approaches to discover and describe genes and mechanisms underlying phenotypic variation across species. Next generation sequencing (e.g. mRNAsequencing) can now be used to develop the datasets required for such studies from practically any species. This approach has the advantages of discovering genes and mechanisms that may be invariable in population-based approaches or transparent to mutational approaches in model systems. Importantly, this strategy can simultaneously discover mechanisms regulating development and relate them to variation within specific lineages. An example to this approach is given by co-expression gene networks. In this approach, genes that display similar expression across a range of biological conditions (e.g. tissue types or developmental time points) are assigned to coexpression “modules” that often represent genes that work together on common biological processes. For example, a comparative gene regulatory network approach was used to comprehensively describe the transcriptional differences underlying variation in leaf shape for survey of tomato species (Ichihashi et al. 2014). Gene modules were defined based on co-expression relationships, and were then annotated with features including enrichment for genes in specific Gene Ontology classes or correlation with phenotypic traits. To apply such an approach to wood formation, trees from a taxonomic survey spanning some relevant taxonomic range would ideally be grown in a common environment, and then subjected to a variety of experimental perturbations that alter wood development (e.g. hormone treatment). Wood forming tissues would then be harvested and subjected to mRNA sequencing to provide data for comparative co-expression analyses within and among species. Challenges to this approach include collecting together and propagating the required plant material from arboreta or other sources (Groover and Dosmann 2012), and reliably determining orthologous relationships among genes among the different species. However, this approach is currently technically tractable and can utilize analysis approaches and tools developed in other systems. Ideally, such a comparative approach could identify gene modules commonly involved in wood formation across species that might represent ancestral mechanisms, as well as lineage or species-specific modules that could underlie phenotypic differences among species.

New approaches could also be developed for functional genomics within at least some model tree systems that could be scaled to survey larger numbers of

genes. An example of such an approach is given by an irradiation hybrid mutagenesis screen in *Populus*. In this approach, a controlled cross was made between two *Populus* species in which the pollen from the male parent was irradiated to create chromosomal breaks. In a cross between *P. nigra* and *P. deltoides*, over 55% of the 500 F1 progeny produced contained deletions or insertions of chromosomal segments (Henry et al. 2015). The insertions and deletions were mapped with precision in each F1 progeny using low-coverage, whole-genome sequence data. The genomic data can now be used to assist in reverse or forward genetic screens, including association between altered gene dosage in specific regions of the genome and phenotypes of interest. This population also represents a rich source of genetic perturbation useful in ultimately modeling gene regulatory networks (Filkov 2005). A primary aim of this approach is to create novel genetic changes that alter the complex gene dosage relationships that are believed central to the regulation of heterosis and complex quantitative traits (Birchler et al. 2006) including wood development.

Lastly, imaging-based techniques and data are especially useful for all comparative studies of wood formation. Secondary vascular tissues and wood are characterized by complex three dimensional tissues comprised of multiple cell types, and have both radial and longitudinal developmental gradients to consider. An increasing number of options are available to visualize and quantify molecular and anatomical features that can then be integrated with genomics level information. An example of the power of imaging-based techniques is shown in Fig. 1. To determine how woody stems perceived and responded to gravity, an antibody was raised against a peptide from a *Populus* PIN-like auxin transport protein and used in immunolocalizations to reveal the gravity-sensing cells within the stem (Fig. 1a, b). Using an auxin-responsive DR5:GUS reporter, the consequence of radial auxin transport by these PIN-expressing cells was revealed in stems placed horizontally to induce tension wood on the upper side of the stem (Fig. 1c) or opposite wood on the lower side of the stem (Fig. 1d). Differential auxin response in the cambial zone versus the cortex in tension wood and opposite wood, respectively, provides insights into the mechanisms regulating these distinct wood development programs that can now be surveyed in angiosperm trees with differing gravitropic stem responses. Molecular phenotypes including cell wall components can be surveyed using an extensive battery of publically available antibodies developed against cell wall epitopes (<http://www.ccruc.uga.edu/~mao/wallmab/Home/Home.php>). An example is shown in Fig. 2 for an arabinoglactan-recognizing antibody that labels mature gelatinous layers (G-layers) in tension wood fibers in *Populus*. A more generalized technique for high resolution, three dimensional imaging of wood tissues is laser ablation tomography (Fig. 3). In this technique, a high powered laser is used to progressively and precisely remove successive layers of tissue, with each layer imaged at high resolution. The images can then be reconstructed in three dimensions and features quantified, such as number, patterning, length, and interconnectivity among vessels. This technique could presumably be applied to any woody species or sample.

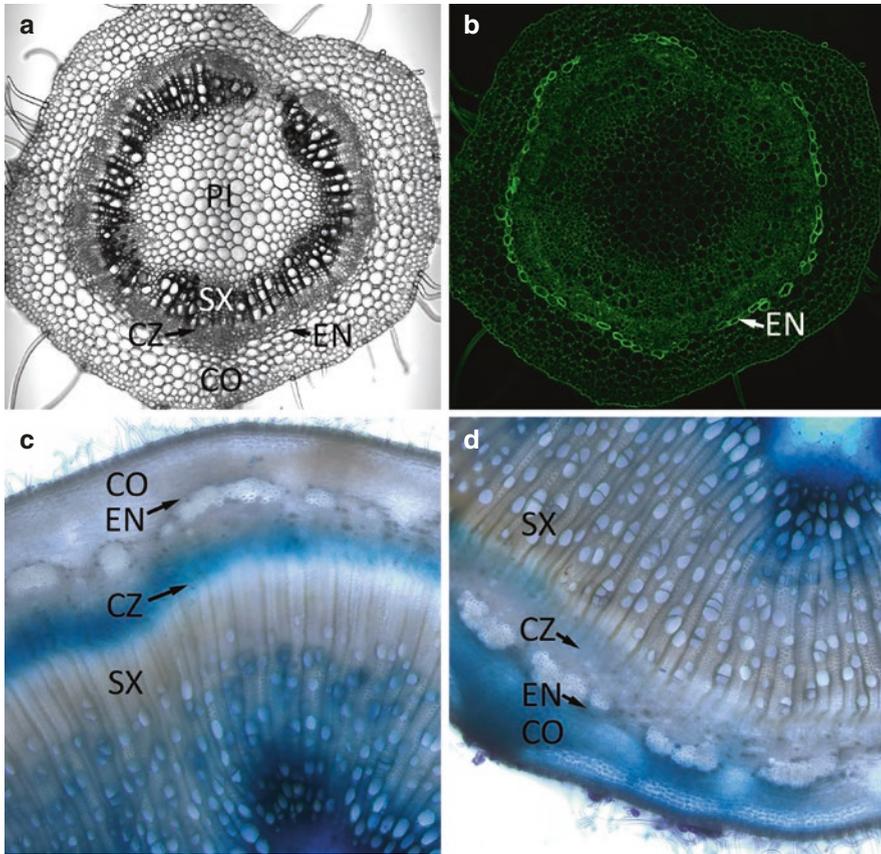


Fig. 1 Imaging of molecular events associated with tension wood formation in *Populus*. **(a)** Bright field image of stem cross section. **(b)** Confocal image of immunolocalization of a *Populus* PIN3-like protein in the same tissue section. The strong green signal corresponds to the endodermis, which is the innermost layer of the cortex. **(c)** Tension wood from DR5:GUS*Populus* stem of a GA treated tree. **(d)** Opposite wood from DR5:GUS*Populus* stem of a GA treated tree. Blue signal corresponds to auxin response (Note response is strong in the cambial zone of tension wood, versus the cortex of opposite wood. CO cortex, CZ cambial zone, EN endodermis, SX secondary xylem, PI pith)

Conclusions and Future Perspectives

The quickly advancing fields of genomics and computational biology make comparative and evolutionary genomic studies increasingly attractive. The large number of diverse angiosperm tree species of ecological or economic interest makes it infeasible to produce the range of resources for each to become a model species. However, sequencing-based experimental approaches such as mRNAsequencing paired with computational approaches such as comparative gene co-expression

Fig. 2 Imaging of arabinogalactan protein epitopes in *Populus* tension wood. (a) A cross section of a *Populus* stem that has transitioned to tension wood formation. Strong labeling by the antibody JIM14 (red signals) of the gelatinous cell wall layer in the lumen of fibers within tension wood takes time to mature and be labeled by the antibody, providing a molecular marker of fiber type and differentiation. (b) Higher magnification of tension wood fibers with labeled gelatinous layers. Blue signal is UV autofluorescence. *CO* cortex, *CZ* cambial zone, *DF* differentiating fibers, *MF* mature fibers, *NW* normal wood, *TW* tension wood

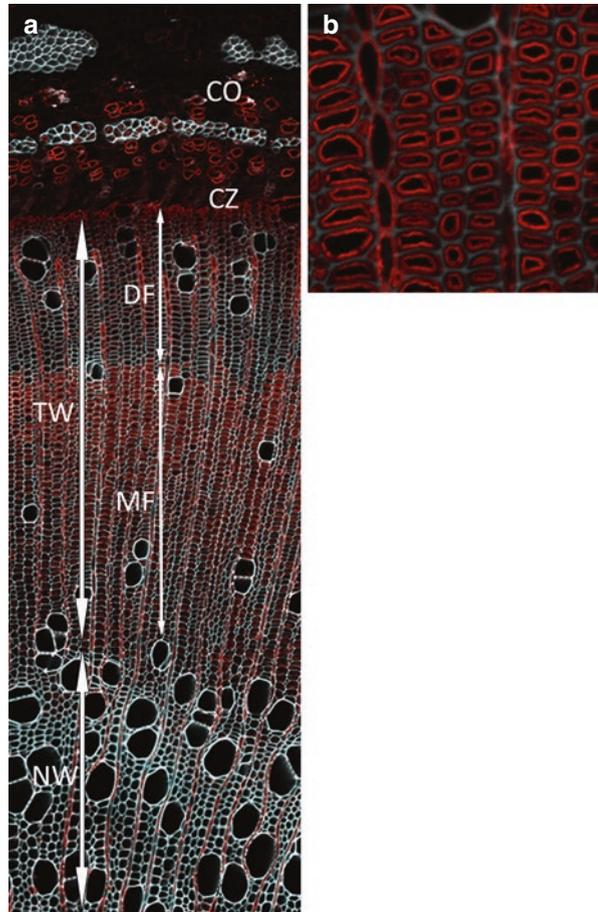
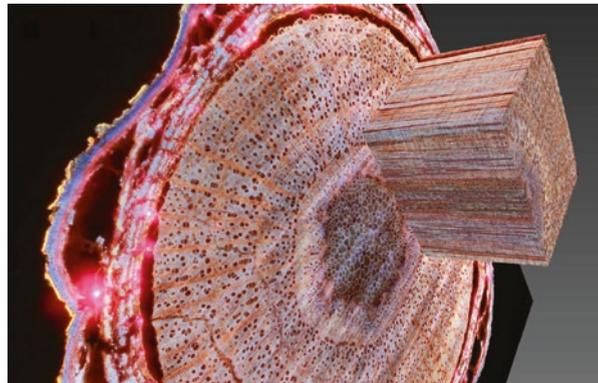


Fig. 3 Three dimensional rendering of wood sample using laser ablation tomography. Wood block rendered in three dimensions by integration of multiple z-stacked images. For information about this technology see <http://l4is.com/Image> courtesy of Lasers for Innovative Solutions, LLC



network analysis could be extended to most species. Importantly, comparative approaches can actually be much more powerful and address questions that one-species-a-time approaches cannot. The power of looking beyond a handful of model tree species would enable answers to such questions as, “what are the core set of genes and mechanisms that make a tree a tree.”

Acknowledgements This work was supported by grant 2015-67013-22891 from USDA AFRI, DE-SC0007183 from DOE Office of Science, Office of Biological and Environmental Research (BER) to AG, and grant 31270219 and 31570581 from National Natural Science Foundation of China to XH. We thank Keith Woeste for helpful comments on the manuscript.

References

- Abraham P, Adams R, Giannone RJ, Kalluri U, Rnajan P, Erickson B, Shah MTGA. Defining the boundaries and characterizing the landscape of functional genome expression in vascular tissues of populus using shotgun proteomics. *J Proteome Res.* 2012;11:449–60.
- Afrasiabi C, Samad B, Dineen D, Meacham C, Sjölander K. The PhyloFacts FAT-CAT web server: ortholog identification and function prediction using fast approximate tree classification. *Nucleic Acids Res.* 2013;41(W1):W242–8. doi:[10.1093/nar/gkt399](https://doi.org/10.1093/nar/gkt399).
- Arnold DH, Mauseth JD. Effects of environmental factors on development of wood. *Am J Bot.* 1999;86(3):367–71.
- Baas P. Systematic, phylogenetic and ecological wood anatomy – history and perspectives. In: Baas P, editor. *New perspectives in wood anatomy*. Springer Science Business Media, Springer Netherlands B.V; 1982.
- Bailey IW. The development of vessels in angiosperms and its significance in morphological research. *Am J Bot.* 1944;31(7):421–8.
- Becker A, Alix K, Damerval C. The evolution of flower development: current understanding and future challenges. *Ann Bot.* 2011;107(9):1427–31. doi:[10.1093/aob/mcr122](https://doi.org/10.1093/aob/mcr122).
- Birchler JA, Yao H, Chudalayandi S. Unraveling the genetic basis of hybrid vigor. *Proc Natl Acad Sci.* 2006;103(35):12957–8. doi:[10.1073/pnas.0605627103](https://doi.org/10.1073/pnas.0605627103).
- Bliss MC. The vessel in seed plants. *Bot Gaz.* 1921;71(4):314–26.
- Brady SM, Zhang L, Megraw M, Martinez NJ, Jiang E, Yi CS, Liu W, Zeng A, Taylor-Teeples M, Kim D, Ahnert S, Ohler U, Ware D, Walhout AJM, Benfey PN. A stele-enriched gene regulatory network in the Arabidopsis root. *Mol Syst Biol.* 2011;7:459. doi:[10.1038/msb.2010.114](https://doi.org/10.1038/msb.2010.114).
- Bylesjo M, Nilsson R, Srivastava V, Gronlund A, Johansson A, Jansson S, Karlsson J, Moritz T, Wingsle G, Trygg J. Integrated analysis of transcript, protein and metabolite data to study lignin biosynthesis in hybrid aspen. *J Proteome Res.* 2009;8:199–210.
- Carlquist S. *Comparative wood anatomy. Systematic, ecological, and evolutionary aspects of dicotyledon wood*. Berlin: Spinger-Verlag; 2001.
- Carlquist S. Successive cambia revisited: ontogeny, histology, diversity, and functional significance. *J Torrey Bot Soc.* 2007;134(2):301–32.
- Carlquist S. Xylem heterochrony: an unappreciated key to angiosperm origin and diversifications. *Bot J Linn Soc.* 2009;161(1):26–65.
- Carlquist S, Schneider EL. The tracheid–vessel element transition in angiosperms involves multiple independent features: cladistic consequences. *Am J Bot.* 2002;89(2):185–95. doi:[10.3732/ajb.89.2.185](https://doi.org/10.3732/ajb.89.2.185).
- Carlsbecker A, Lee J, Roberts C, Dettmer J, Lehesranta S, Zhou J, Lindgren O, Moreno-Risueno M, Vaten A, Thitamadee S, Campilho A, Sebastian J, Bowman J, Helariutta Y, Benfey P. Cell

- signalling by microRNA165/6 directs gene dose-dependent root cell fate. *Nature*. 2010;465:316–21.
- Champagne CEM, Goliber TE, Wojciechowski MF, Mei RW, Townsley BT, Wang K, Paz MM, Geeta R, Sinha NR. Compound leaf development and evolution in the legumes. *Plant Cell*. 2007;19(11):3369–78. doi:10.1105/tpc.107.052886.
- Consortium TEP. An integrated encyclopedia of DNA elements in the human genome. *Nature*. 2012;489(7414):57–74. doi:http://www.nature.com/nature/journal/v489/n7414/abs/nature11247.html#supplementary-information.
- Dafoe N, Constabel C. Proteomic analysis of hybrid poplar xylem sap. *Phytochemistry*. 2009;70:856–63.
- Derbyshire P, Menard D, Green P, Saalbach G, Buschmann H, Lloyd C, Pesquet E. Proteomic analysis of microtubule interacting proteins over the course of xylem tracheary element formation in *Arabidopsis*. *Plant Cell*. 2015;27:2709–26.
- Ding Q, Zeng J, He X. Deep sequencing on a genome-wide scale reveals diverse stage-specific microRNAs in cambium during dormancy-release induced by chilling in poplar. *BMC Plant Biol*. 2014;14:267.
- Dkhar J, Pareek A. What determines a leaf's shape? *EvoDevo*. 2014;5(1):1–19. doi:10.1186/2041-9139-5-47.
- Du J, Xie HL, Zhang DQ, He XQ, Wang MJ, Li YZ, Cui KM, Lu MZ. Regeneration of the secondary vascular system in poplar as a novel system to investigate gene expression by a proteomic approach. *Proteomics*. 2006;6(3):881–95. doi:10.1002/pmic.200401348.
- Du J, Mansfield SD, Groover AT. The *Populus* homeobox gene *ARBORKNOX2* regulates cell differentiation during secondary growth. *Plant J*. 2009;60(6):1000–14.
- Du J, Robischon M, Miura E, Martinez C, Groover AT. The *Populus* Class III HD ZIP transcription factor *POPCORONA* affects patterning and cell differentiation during secondary growth of woody stems. *PLoS One*. 2011;6(2), e17458.
- Emery JF, Floyd SK, Alvarez J, Eshed Y, Hawker NP, Izhaki A, Baum SF, Bowman JL. Radial patterning of *Arabidopsis* shoots by class III HD-ZIP and *KANADI* genes. *Curr Biol*. 2003;13(20):1768–74.
- Esau K. Anatomy of seed plants. 2nd ed. New York: Wiley; 1977.
- Escamez S, Tuominen H. Programmes of cell death and autolysis in tracheary elements: when a suicidal cell arranges its own corpse removal. *J Exp Bot*. 2014;65(5):1313–21. doi:10.1093/jxb/eru057.
- Etchells JP, Turner SR. The PXY-CLE41 receptor ligand pair defines a multifunctional pathway that controls the rate and orientation of vascular cell division. *Development*. 2010;137(5):767–74. doi:10.1242/dev.044941.
- Etienne Paux VC, Marques C, Mendes de Sousa A, Borrallho N, Sivadon P, Grima-Pettenati J. Transcript profiling of *Eucalyptus* xylem genes during tension wood formation. *New Phytol*. 2005;167(1):89–100.
- Filkov V. Identifying gene regulatory networks from gene expression data. In: Aluru S, editor. Handbook of computational molecular biology. Boca Raton: Chapman & Hall/CRC Press; 2005.
- Fiorani C, de Andrade A, Meireles K, Gallo D, Caldas D, Moon D, Carneiro R, Franceschini L, Oda S, Labate C. Proteomic analysis of the cambial region in juvenile *eucalyptus grandis* at three ages. *Proteomics*. 2007;7:2258–74.
- Fischer S, Brunk BP, Chen F, Gao X, Harb OS, Iodice JB, Shanmugam D, Roos DS, Stoeckert CJ (2011) Using OrthoMCL to assign proteins to OrthoMCL-DB groups or to cluster proteomes into new ortholog groups. *Curr Protoc Bioinformatics/editorial board*, Andreas D Baxevanis [et al] Chapter: Unit-6.1219. doi:10.1002/0471250953.bi0612s35.
- Floyd SK, Zalewski CS, Bowman JL. Evolution of class III homeodomain-leucine zipper genes in streptophytes. *Genetics*. 2006;173(1):373–88. doi:10.1534/genetics.105.054239.
- Gerttula S, Zinkgraf M, Muday GK, Lewis DR, Ibatullin FM, Brumer H, Hart F, Mansfield SD, Filkov V, Groover A. Transcriptional and hormonal regulation of gravitropism of woody stems in populus. *Plant Cell*. 2015;27(10):2800–13. doi:10.1105/tpc.15.00531.

- Groover A, Cronk Q. From Nehemiah Grew to Genomics: the emerging field of evo-devo research for woody plants. *Int J Plant Sci.* 2013;174(7):959–63. doi:10.1086/671569.
- Groover A, Dosmann M. The importance of living botanical collections for plant biology and the “next generation” of evo-devo research. *Front Plant Sci.* 2012;3. doi:10.3389/fpls.2012.00137.
- Groover A, Jones A. Tracheary element differentiation uses a novel mechanism coordinating programmed cell death and secondary cell wall synthesis. *Plant Physiol.* 1999;119:375–84.
- Groover A, Mansfield S, DiFazio S, Dupper G, Fontana J, Millar R, Wang Y. The *Populus* homeobox gene *ARBORKNOX1* reveals overlapping mechanisms regulating the shoot apical meristem and the vascular cambium. *Plant Mol Biol.* 2006;61(6):917–32.
- Henry IM, Zinkgraf MS, Groover AT, Comai L. A system for dosage-based functional genomics in poplar. *Plant Cell.* 2015;27(9):2370–83. doi:10.1105/tpc.15.00349.
- Hirakawa Y, Bowman JL. A role of TDIF peptide signaling in vascular cell differentiation is conserved among euphylllophytes. *Front Plant Sci.* 2015;6:1048. doi:10.3389/fpls.2015.01048.
- Hirakawa Y, Shinohara H, Kondo Y, Inoue A, Nakanomyo I, Ogawa M, Sawa S, Ohashi-Ito K, Matsubayashi Y, Fukuda H. Non-cell-autonomous control of vascular stem cell fate by a CLE peptide/receptor system. *Proc Natl Acad Sci.* 2008;105(39):15208–13. doi:10.1073/pnas.0808444105.
- Hirakawa Y, Kondo Y, Fukuda H. TDIF peptide signaling regulates vascular stem cell proliferation via the *WOX4* homeobox gene in *Arabidopsis*. *Plant Cell.* 2010;22(8):2618–29. doi:10.1105/tpc.110.076083.
- Huerta-Cepas J, Szklarczyk D, Forslund K, Cook H, Heller D, Walter MC, Rattei T, Mende DR, Sunagawa S, Kuhn M, Jensen LJ, von Mering C, Bork P. eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res.* 2016;44(D1):D286–93. doi:10.1093/nar/gkv1248.
- Ichihashi Y, Aguilar-Martínez JA, Farhi M, Chitwood DH, Kumar R, Millon LV, Peng J, Maloof JN, Sinha NR. Evolutionary developmental transcriptomics reveals a gene network module regulating interspecific diversity in plant leaf shape. *Proc Natl Acad Sci.* 2014;111(25):E2616–21. doi:10.1073/pnas.1402835111.
- Irish VF, Litt A. Flower development and evolution: gene duplication, diversification and redeployment. *Curr Opin Genet Dev.* 2005;15(4):454–60. doi:http://dx.doi.org/10.1016/j.gde.2005.06.001.
- Ito Y, Nakanomyo I, Motose H, Iwamoto K, Sawa S, Dohmae N, Fukuda H. Dodeca-CLE peptides as suppressors of plant stem cell differentiation. *Science.* 2006;313(5788):842–5. doi:10.1126/science.1128436.
- Kaku T, Serada S, Baba K, Tanaka F, Hayash T. Proteomic analysis of the G-layer in poplar tension wood. *J Wood Sci.* 2009;55:250–7.
- Ko JH, Prassinis C, Han KH. Developmental and seasonal expression of PtaHB1, a *Populus* gene encoding a class III HD-Zip protein, is closely associated with secondary growth and inversely correlated with the level of microRNA (miR166). *New Phytol.* 2006;169(3):469–78. doi:10.1111/j.1469-8137.2005.01623.x. NPH1623 [pii].
- Kubo M, Udagawa M, Nishikubo N, Horiguchi G, Yamaguchi M, Ito J, Mimura T, Fukuda H, Demura T. Transcription switches for protoxylem and metaxylem vessel formation. *Genes Dev.* 2005;19(16):1855–60. doi:10.1101/gad.1331305.
- Larson PR. The vascular cambium. Berlin: Springer; 1994.
- Lechner M, Hernandez-Rosales M, Doerr D, Wieseke N, Thévenin A, Stoye J, Hartmann RK, Prohaska SJ, Stadler PF. Orthology detection combining clustering and synteny for very large datasets. *PLoS One.* 2014;9(8), e105015. doi:10.1371/journal.pone.0105015.
- Li L, Stoeckert CJ, Roos DS. OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res.* 2003;13(9):2178–89. doi:10.1101/gr.1224503.
- Liu L, Ramsay T, Zinkgraf M, Sundell D, Street NR, Filkov V, Groover A. A resource for characterizing genome-wide binding and putative target genes of transcription factors expressed during secondary growth and wood formation in *Populus*. *Plant J.* 2015a. doi:10.1111/tjpi.12850.
- Liu L, Zinkgraf M, Petzold HE, Beers EP, Filkov V, Groover A. The *Populus* ARBORKNOX1 homeodomain transcription factor regulates woody growth through binding to evolutionarily conserved target genes of diverse function. *New Phytol.* 2015b;205(2):682–94. doi:10.1111/nph.13151.

- Long J, Moan E, Medford J, Barton M. A member of the KNOTTED class of homeodomain proteins encoded by the SHOOTMERISTEMLESS gene of Arabidopsis. *Nature*. 1996;379:66–9.
- Lu S, Sun Y, Shi R, Clark C, Li L, Chiang V. Novel and mechanical stress-responsive MicroRNAs in populus trichocarpa that are absent from Arabidopsis. *Plant Cell*. 2005;17:2186–203.
- Lu S, Li Q, Wei H, Chang M, Tunlaya-Anukit S, Kim H, Liu J, Song J, Sun Y, Yuan L, Yeh T, Peszlen I, Ralph J, Sederoff R, Chiang V. Ptr-miR397a is a negative regulator of laccase genes affecting lignin content in populus trichocarpa. *Proc Natl Acad Sci U S A*. 2013;110:10848–53.
- Mauriat M, Leple J, Claverol S, Bartholome J, Negroni L, Richet N, Lalanne C, Bonneau M, Coutand C, Plomion C. Quantitative proteomic and phosphoproteomic approaches for deciphering the signaling pathway for tension wood formation in poplar. *J Proteome Res*. 2015;14:3188–203.
- Mauseth J, Plemons-Rodriguez B. Evolution of extreme xeromorphic characters in wood: a study of nine evolutionary lines in Cactaceae. *Am J Bot*. 1998;85(2):209.
- Meng Y, Shao C, Wang H, Chen M. The regulatory activities of plant MicroRNAs: a more dynamic perspective. *Plant Physiol*. 2011;157(4):1583–95. doi:10.1104/pp.111.187088.
- Mijnsbrugge K, Meyermans H, Van Montagu M, Bauw G, Boerjan W. Wood formation in poplar: identification, characterization, and seasonal variation of xylem proteins. *Planta*. 2000;210:589–98.
- Nakaya A, Katayama T, Itoh M, Hiranuka K, Kawashima S, Moriya Y, Okuda S, Tanaka M, Tokimatsu T, Yamanishi Y, Yoshizawa AC, Kanehisa M, Goto S. KEGG OC: a large-scale automatic construction of taxonomy-based ortholog clusters. *Nucleic Acids Res*. 2013;41(D1):D353–7. doi:10.1093/nar/gks1239.
- Nilsson R, Bernfur K, Gustavsson N, Bygdell J, Wingsle G, Larsson C. Proteomics of plasma membranes from poplar trees reveals tissue distribution of transporters, receptors, and proteins in cell wall formation. *Mol Cell Proteomics*. 2010;9:368–87.
- Plomion C, Pionneau C, Baillères H. Identification of tension-wood responsive proteins in the developing xylem of eucalyptus. *Holzforschung*. 2003;57:353–8.
- Porth I, Klápště J, Skyba O, Friedmann MC, Hannemann J, Ehling J, El-Kassaby YA, Mansfield SD, Douglas CJ. Network analysis reveals the relationship among wood properties, gene expression levels and genotypes of natural Populus trichocarpa accessions. *New Phytol*. 2013a;200(3):727–42. doi:10.1111/nph.12419.
- Porth I, Klápště J, Skyba O, Hannemann J, McKown AD, Guy RD, DiFazio SP, Muchero W, Ranjan P, Tuskan GA, Friedmann MC, Ehling J, Cronk QCB, El-Kassaby YA, Douglas CJ, Mansfield SD. Genome-wide association mapping for wood characteristics in Populus identifies an array of candidate single nucleotide polymorphisms. *New Phytol*. 2013b;200(3):710–26. doi:10.1111/nph.12422.
- Robischon M, Du J, Miura E, Groover A. The Populus Class III HD ZIP, *popREVOLUTA*, influences cambium initiation and patterning of woody stems. *Plant Physiol*. 2011;155:1214–25.
- Schrader J, Nilsson J, Mellerowicz E, Berglund A, Nilsson P, Hertzberg M, Sandberg G. A high-resolution transcript profile across the wood-forming meristem of poplar identifies potential regulators of cambial stem cell identity. *Plant Cell*. 2004;16(9):2278–92. doi:10.1105/tpc.104.024190. tpc.104.024190 [pii].
- Smith SA, Beaulieu JM. Life history influences rates of climatic niche evolution in flowering plants. *Pro R Soc Lond B Biol Sci*. 2009. doi:10.1098/rspb.2009.1176.
- Smith SA, Donoghue MJ. Rates of molecular evolution are linked to life history in flowering plants. *Science*. 2008;322(5898):86–9. doi:10.1126/science.1163197.
- Soltis DE, Albert VA, Leebens-Mack J, Bell CD, Paterson AH, Zheng C, Sankoff D, dePamphilis CW, Wall PK, Soltis PS. Polyploidy and angiosperm diversification. *Am J Bot*. 2009;96(1):336–48. doi:10.3732/ajb.0800079.
- Song D, Xi W, Shen J, Bi T, Li L. Characterization of the plasma membrane proteins and receptor-like kinases associated with secondary vascular differentiation in poplar. *Plant Mol Biol*. 2011;76(97–115).
- Spicer R, Groover A. The evolution of development of the vascular cambium and secondary growth. *New Phytol*. 2010;186:577–92.

- Srivastava V, Obudulu O, Bygdell J, Lofstedt T, Ryden P, Nilsson R, Ahnlund M, Johansson A, Jonsson P, Freyhult E, Qvarnstrom J, Karlsson J, Melzer M, Moritz T, Trygg J, Hvidsten T, Wingsle G. OnPLS integration of transcriptomic, proteomic and metabolomic data shows multi-level oxidative stress responses in the cambium of transgenic hiPl- superoxide dismutase populus plants. *BMC Genomics*. 2013;14:893.
- Street N, Jansson S, Hvidsten T. A systems biology model of the regulatory network in *Populus* leaves reveals interacting regulators and conserved regulation. *BMC Plant Biol*. 2011;11(1):13.
- Sun Y, Shi R, Zhang X, Chiang V, Sederoff R. MicroRNAs in trees. *Plant Mol Biol*. 2012;80:37–53.
- Taylor-Teeples M, Lin L, de Lucas M, Turco G, Toal TW, Gaudinier A, Young NF, Trabucco GM, Veling MT, Lamothe R, Handakumbura PP, Xiong G, Wang C, Corwin J, Tsoukalas A, Zhang L, Ware D, Pauly M, Kliebenstein DJ, Dehesh K, Tagkopoulos I, Breton G, Pruneda-Paz JL, Ahnert SE, Kay SA, Hazen SP, Brady SM. An Arabidopsis gene regulatory network for secondary cell wall synthesis. *Nature*. 2015;517(7536):571–5. doi:10.1038/nature14099. <http://www.nature.com/nature/journal/v517/n7536/abs/nature14099.html#supplementary-information>.
- The Angiosperm Phylogeny G. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Bot J Linn Soc*. 2009;161(2):105–21. doi:10.1111/j.1095-8339.2009.00996.x.
- Tsiantis M, Hay A. Comparative plant development: the time of the leaf? *Nat Rev Genet*. 2003;4(3):169–80.
- Tsukaya H. Comparative leaf development in angiosperms. *Curr Opin Plant Biol*. 2014;17:103–9. doi:<http://dx.doi.org/10.1016/j.pbi.2013.11.012>.
- Victor M. MicroRNAs in differentiating tissues of populus and eucalyptus trees. Thesis University of Pretoria; 2006.
- Ye Y, Wei B, Wen L, Rayner S. BlastGraph: a comparative genomics tool based on BLAST and graph algorithms. *Bioinformatics*. 2013;29(24):3222–4. doi:10.1093/bioinformatics/btt553.
- Zanne AE, Tank DC, Cornwell WK, Eastman JM, Smith SA, FitzJohn RG, McGlenn DJ, O'Meara BC, Moles AT, Reich PB, Royer DL, Soltis DE, Stevens PF, Westoby M, Wright IJ, Aarssen L, Bertin RI, Calaminus A, Govaerts R, Hemmings F, Leishman MR, Oleksyn J, Soltis PS, Swenson NG, Warman L, Beaulieu JM. Three keys to the radiation of angiosperms into freezing environments. *Nature*. 2014;506(7486):89–92. doi:10.1038/nature12872. <http://www.nature.com/nature/journal/v506/n7486/abs/nature12872.html#supplementary-information>.
- Zhao C, Craig JC, Petzold HE, Dickerman AW, Beers EP. The xylem and phloem transcriptomes from secondary tissues of the Arabidopsis root-hypocotyl. *Plant Physiol*. 2005;138(2):803–18. doi:10.1104/pp.105.060202.
- Zhao Y, Song D, Sun J, Li L. 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*. 2013;74:473–85.
- Zhong R, Ye Z-H. Secondary cell walls: biosynthesis, patterned deposition and transcriptional regulation. *Plant Cell Physiol*. 2015;56(2):195–214. doi:10.1093/pcp/pcu140.
- Zhong R, Demura T, Ye Z-H. SND1, a NAC domain transcription factor, is a Key regulator of secondary wall synthesis in fibers of Arabidopsis. *Plant Cell*. 2006;18(11):3158–70. doi:10.1105/tpc.106.047399.
- Zhong R, Lee C, Zhou J, McCarthy RL, Ye Z-H. A battery of transcription factors involved in the regulation of secondary cell wall biosynthesis in Arabidopsis. *Plant Cell*. 2008;20(10):2763–82. doi:10.1105/tpc.108.061325.