Tree Physiology 38, 362–377
doi:10.1093/treephys/tpx121

Research paper

Identification of new protein–protein and protein–DNA interactions linked with wood formation in *Populus trichocarpa*

H. Earl Petzold[1,†], Stephen B. Rigoulot[2,†], Chengsong Zhao[2,4,†], Bidisha Chanda[2,5], Xiaoyan Sheng[1], Mingzhe Zhao[2,6], Xiaoyan Jia[1,7], Allan W. Dickerman[3], Eric P. Beers[2] and Amy M. Brunner[1,8]

[1]Department of Forest Resources and Environmental Conservation, Virginia Tech, Blacksburg, VA 24061, USA; [2]Department of Horticulture, Virginia Tech, Blacksburg, VA 24061, USA; [3]The Biocomplexity Institute at Virginia Tech, Blacksburg, VA 24061, USA; [4]Present address: Department of Plant Pathology, Physiology, and Weed Science, Virginia Tech, Blacksburg, VA 24061, USA; [5]Present address: US Vegetable Laboratory, Charleston, SC 29414, USA; [6]Present address: Agronomy College, Shenyang Agricultural University, Shenyang, Liaoning Province 110866, PR China; [7]Present address: Department of Human Genetics, University of Michigan, Ann Arbor, MI 48109, USA; [8]Corresponding author (abrunner@vt.edu)

Received July 17, 2017; accepted August 30, 2017; published online October 10, 2017; handling Editor Chung-Jui Tsai

Cellular processes, such as signal transduction and cell wall deposition, are organized by macromolecule interactions. Experimentally determined protein–protein interactions (PPIs) and protein–DNA interactions (PDIs) relevant to woody plant development are sparse. To begin to develop a *Populus trichocarpa* Torr. & A. Gray wood interactome, we applied the yeast-two-hybrid (Y2H) assay in different ways to enable the discovery of novel PPIs and connected networks. We first cloned open reading frames (ORFs) for 361 genes markedly upregulated in secondary xylem compared with secondary phloem and performed a binary Y2H screen with these proteins. By screening a xylem cDNA library for interactors of a subset of these proteins and then recapitulating the process by using a subset of the interactors as baits, we ultimately identified 165 PPIs involving 162 different ORFs. Thirty-eight transcription factors (TFs) included in our collection of *P. trichocarpa* wood ORFs were used in a Y1H screen for binding to promoter regions of three genes involved in lignin biosynthesis resulting in 40 PDIs involving 20 different TFs. The network incorporating both the PPIs and PDIs included 14 connected subnetworks, with the largest having 132 members. Protein–protein interactions and PDIs validated previous reports and also identified new candidate wood formation proteins and modules through their interactions with proteins and promoters known to be involved in secondary cell wall synthesis. Selected examples are discussed including a PPI between Mps one binder (MOB1) and a mitogen-activated protein kinase kinase kinase kinase (M4K) that was further characterized by assays confirming the PPI as well as its effect on subcellular localization. Mapping of published transcriptomic data showing developmentally detailed expression patterns across a secondary stem onto the network supported that the PPIs and PDIs are relevant to wood formation, and also illustrated that wood-associated interactions involve gene products that are not upregulated in secondary xylem.

Keywords: interaction network, secondary cell wall, transcription factor, xylem.

Introduction

Populus trichocarpa was the first tree species to undergo genomic sequencing and was selected based upon several attributes, such as economic importance, rapid growth, modest genome size, ease of experimental manipulation and availability of already existing genetic tools (Song et al. 2006, Tuskan et al. 2006). Combined, these attributes make poplar (*Populus* spp.) a logical choice for experiments aimed at further understanding the basic mechanisms involved in woody biomass production and hence improvement of poplar as a biomass feedstock. Towards these goals, reverse genetic experiments that typically select candidate genes based on upregulation in wood-forming tissues and/or homology to genes characterized in Arabidopsis have identified a number of regulatory genes that alter xylem

[†]These authors contributed equally to this work.

development in poplar (e.g., Grant et al. 2010, Robischon et al. 2011, Zawaski et al. 2011, Li et al. 2015).

Recent studies have applied next-generation transcriptomics and genome-wide association studies to identify gene co-expression networks linked to wood formation and candidate genes for various wood traits in poplar (Porth et al. 2013, Gerttula et al. 2015, Fahrenkrog et al. 2017, Shi et al. 2017, Zinkgraf et al. 2017). In between the genotype and phenotype, however, are macromolecule interaction networks; hence, identifying these is crucial for understanding how genetic variation results in trait variation, especially variation in complex traits (Vidal et al. 2011). Whereas computational methods to predict interactions based on homology can be useful, in vivo testing is necessary due to potential biological differences among systems (Vidal and Fields 2014). Of the multiple methods for determining protein–protein interactions (PPIs), yeast-two-hybrid (Y2H) is one of the few that scales well. For example, binary Y2H screens involving ~8000 open reading frames (ORFs) resulted in an interactome map involving 6200 PPIs in Arabidopsis (Arabidopsis Interactome Mapping Consortium 2011) and a screen with ~13,000 ORFs produced a human interactome network of 14,000 PPIs (Rolland et al. 2014). However, most studies have focused on identifying a small number of PPIs, typically involving a few well-characterized proteins. Using various methods, studies have also identified protein–DNA interactions (PDIs), including poplar transcription factors (TFs) involved in wood formation (Zhong et al. 2011, Liu et al. 2015). In Arabidopsis, Taylor-Teeples et al. (2015) used Y1H to screen 50 promoters of genes involved in secondary cell wall (SCW) biosynthesis or xylem formation against 467 TFs expressed in root xylem.

The Arabidopsis PPI and PDI networks (Arabidopsis Interactome Mapping Consortium 2011, Taylor-Teeples et al. 2015) are useful resources, but their transferability to poplar and other trees is limited by lineage-specific genome evolution and differences in physiology and development. Despite the similarities in cell-types observed between secondary growth in the Arabidopsis hypocotyl or root and poplar, there are distinct structural differences in wood formation, e.g., Arabidopsis wood is deficient in rays, parenchyma-like files of cells that comprise the radial component of angiosperm wood (Chaffey et al. 2002). The combination of differences in wood composition and the absence of seasonal cycles of cambial dormancy and activity in annuals highlights the need to study wood formation in a model tree. For this study, our major goals were to discover novel poplar wood-associated PPIs and PDIs and to identify connected networks that included genes/proteins known to be involved wood formation, thereby facilitating new hypotheses and strategies for perturbing networks to manipulate wood traits. To accomplish this, we exploited secondary xylem and phloem transcriptome data from *P. trichocarpa* (Rodgers-Melnick et al. 2012) to select an initial set of poplar wood (PW) proteins and used a combination of binary Y2H screening, reiterative Y2H screening of a secondary xylem cDNA library and Y1H screening of PW TFs against three promoter regions of lignin biosynthesis genes. Selected interactors were independently tested by in vitro pull-down assays and bimolecular fluorescence complementation (BiFC). Finally, we mapped spatially detailed transcriptome data from the secondary stem (Immanen et al. 2016) onto our network to further enhance the utility of this resource.

Materials and methods

Xylem cDNA library preparation and poplar wood ORF cloning

After bark was peeled away from stems (below internode 10) of *P. trichocarpa* Nisqually-1, a razor blade was used to scrape developing secondary xylem directly into liquid nitrogen. Total RNA was isolated using the RNeasy Mini Kit (Qiagen, Hilden, Germany) with a modified protocol (Brunner et al. 2004). mRNA was isolated using the PolyATtract mRNA Isolation System (Promega, Madison, WI, USA). A standard cDNA library was constructed using CloneMiner II cDNA library construction Kit (Invitrogen, Carlsbad, CA, USA). Poplar wood ORFs were cloned from *P. trichocarpa* secondary xylem cDNA using the primers listed in Table S1a available as Supplementary Data at *Tree Physiology* Online into pENTR™/D-TOPO (Invitrogen) according to the Invitrogen manual (25-0434). Subsequent recombination into GATEWAY-compatible destination vectors (pDB-Dest and pAD-Dest) (Walhout et al. 2000) was conducted using LR clonase according to the manufacturer's instructions (Invitrogen).

Yeast two-hybrid analysis

Binary Y2H analysis followed the method of (Walhout and Vidal 2001) as briefly outlined here. Yeast strains MaV103 (MATa) and MaV203 (MATα) (for −His, −Ura selection) or Y8800 (MATa) and Y8930 (MATα) (for −His, −Ade selection) were transformed with vectors pDB-Dest (for MATa) or pAD-Dest (for MATα), containing the Gal4 DNA binding domain (DB) or activation domain (AD) in fusion with PW ORFs (DB-X, AD-Y). Transformants were grown in synthetic complete (Sc) medium minus leucine (Sc −Leu) or tryptophan (Sc −Trp) for DB-X or AD-Y, respectively. DB-X transformed cells were tested for auto-activation, i.e., growth on Sc−Leu, −His, + appropriate level (1 mM or 50 mM) of 3-amino-1,2,4-triazole (3AT), depending on yeast strain and results from titrations with 3AT. Cells from each mating event were resuspended in liquid Sc −Leu, −Trp medium and spotted onto Sc −Leu, −Trp (*LacZ* assay); Sc −Leu, −Trp, −His + 3AT; and Sc −Leu, −Trp, −Ura or Sc −Leu, −Trp, −Ade selective plates for 2 days, after which plates were cloth-cleaned and allowed to re-grow for 2–4 days before scoring for reporter activation. For *LacZ* activity, the nitrocellulose filter overlay assay for colorimetric detection of β-galactosidase activity was used (Deplancke et al. 2006).

Y2H library screening also followed the method of Walhout and Vidal (2001). Yeast cells (MaV203 or Y8930) were co-transformed with vectors for DB-X fusions (i.e., bait proteins)

that did not exhibit auto-activation and the xylem prey library (AD-Y vectors). The resulting co-transformed yeast culture was plated on minimal medium (Sc −Leu, −Trp, −His, + 1 mM 3AT) and allowed to grow for 2–4 days. From these plates, single colonies were selected, diluted and spotted on selection plates: Sc −Leu, −Trp, −His, + 3AT; Sc−Leu, −Trp, −Ura (or Sc −Leu, −Trp, −Ade); and Sc −Leu, −Trp (for *LacZ* assays). Colonies that activated at least two reporters were re-streaked three times, selecting an individual colony each time. PCR was performed to confirm the presence of a single vector, indicated by a single PCR product. Yeast colonies containing a single vector were propagated followed by DNA isolation using Zymoprep™ Yeast Plasmid Miniprep II (http://www.zymoresearch.com) kit. The isolated AD-Y vectors were sequenced and re-tested against the relevant DB-X vector (bait protein).

Yeast one-hybrid

Yeast one-hybrid (Y1H) assays were conducted according to Deplancke et al. (2006), although vector 476-P5E MCS (Addgene plasmid #26,029) was used in place of pDONR-P4-P1R. For promoter cloning, genomic DNA was isolated from *P. trichocarpa* using the DNeasy Plant Mini Kit (Qiagen). The primers used are listed in Table S2f available as Supplementary Data at *Tree Physiology* Online and initial promoter cloning used pGEM-T Easy (Promega) as a shuttle vector. Transcription factors (TFs) previously cloned into pAD-Dest vectors were pooled using 235 ng per TF to give a total mini-library concentration of 10 µg/60 µl. The resulting mini-library was co-transformed with the yeast culture containing the promoter of interest upstream of both *His3* and *LacZ* reporter genes in yeast strain YM4271 (MATa). Yeast growth on Sc −Ura, −Trp, −His, + 60 mM 3AT indicated positive results for PDIs. From these plates, single colonies were selected, diluted and spotted on nitrocellulose filters on Sc −Trp plates to assay for activation of *LacZ*. Individual TFs capable of activating reporters were identified by sequencing and re-tested individually in independent Y1H assays.

RNAseq data analysis

RNAseq data was obtained from the ArrayExpress database, accession number E-MTAB-4631 (http://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-4631/). Files from the database were already processed and mapped according to Immanen et al. (2016). Transcript abundances were calculated using the Trapnell lab RNAseq analysis pipeline (Trapnell et al. 2010) using the Ptrichocarpa_210_v3.0.gene.gtf and Ptrichocarpa_210_v3.0.fa files obtained from the JGI website (phytozome.jgi.doe.gov). Quantification of loci, differential expression, and subsequent visualization and analysis were performed using cuffQuant, cuffDiff and cummerbund R software, respectively. As only mapped read files were available from the ArrayExpress website, additional parameters were included in the cuffQuant analysis to facilitate the generation of correct loci abundances. Based on the library preparation kit used, we inferred that the strandedness of the library was second-stranded and proceeded by specifying this parameter.

Sample clustering was performed on the 12 cryosections and based on clustering as well as section order were renamed Phloem 1 (P1), Phloem 2 (P2), Phloem/Cambium (PC), Cambial Zone 1 (CZ1), Cambial Zone 2 (CZ2), Cambium/Xylem (CX), Young Xylem 1 (YX1), Young Xylem 1 (YX2), Transition Xylem (TX), Mature Xylem 1 (MX1), Mature Xylem 2 (MX2) and Mature Xylem 3 (MX3). These sections were pooled based on sub-clusters into phloem, cambial zone, young xylem, transition xylem and mature xylem for determining the highest tissue specificity scores for genes of interest.

Specificity score calculation

Specificity scores were generated using the csSpecificity() function as part of the cummeRbund analysis software (Goff et al. 2012). A combined specificity profile (CSP) for each interaction was constructed from the individual specificity profiles of the genes encoding the two interacting proteins. The score for each tissue in the CSP is comprised of the lower specificity score from those two profiles. We based the CSP on the gene with the lower specificity score under the assumption that this gene will be the limiting member of the interacting pair and therefore best represent the most conservative prediction of tissue localization for each PPI based on gene expression alone.

Bi-weight mid-correlation score generation

Correlation of expression between genes was scored using 'bi-weight mid-correlation' (bicor), which is a median based correlation measure that is described by Song et al. (2012) as being more robust to outliers than Pearson correlation and more sensitive than Spearman, two alternative correlation measures. Using the WGCNA R package for construction of weighted gene co-expression networks (Langfelder and Horvath 2008) we assigned a correlation score to each of the interactor pairs using the bicorAndPvalue() function. Starting with the genes.fkpm_tracking file from cuffDiff output we generated a queryable matrix from which to obtain the correlations score and confidence *P*-value for each pair of interactors.

Blast2GO description annotation

All DNA sequence data was converted to amino acid sequence using the online translation tool EMBOSS Transeq in the generation of a FASTA file containing the protein sequences for all clones used in this study. For incomplete interaction sequence tags (ISTs), protein sequences were obtained via the bulk data retrieval tool on the phytozome website. To assign descriptions we used Blast2GO basic (Conesa et al. 2005). The parameters for the Blast configuration were as follows, Blast Program blastp-fast, Blast Database Plants/Arabidopsis_thaliana_protein_sequences, Taxonomy filter: flowering plants (taxa: 3398,Magnoliophyta) while limiting the

number of blast hits to 10 and allowing the Blast Description Annotator, which assigns the best description for a sequence based on the top scoring blast results. These descriptions were further curated based on available literature as well as information provided through our interactome.

Network visualization

Cytoscape 3.2.1 was used to visualize Y2H and Y1H interactions (Shannon et al. 2003). File input for network depiction is shown in Tables S2a, S2c and S2e available as Supplementary Data at *Tree Physiology* Online. Table S2a data (all interactors) was used as node attributes and edge attributes were the interactions shown in Table S2c (Y2H results) and Table S2e available as Supplementary Data at *Tree Physiology* Online (Y1H results). Bicor scores were only included as an edge attribute in the cytoscape networks where edge thickness indicated PPI and PDI with significantly (*P*-value ≤0.05) correlated gene expression. The arrangement of PPIs and PDIs shown in Figure 2 was custom sorted by hand to maximize the visibility of nodes and edges involved in the interactome modules.

Transient expression of GFP/YFP/RFP-tagged proteins for localization and BiFC

The vectors used were pK7WGF2 and pK7WGY2 (VIB, http://www.vib.be/en/research/services/Pages/Gateway-Services.aspx) for GFP and YFP, respectively, and pSITE-4CA (Arabidopsis Biological Resource Center, ABRC) for RFP. Gateway™ (Invitrogen) cloning was used for construction of proteins fused to the C-terminus of the fluorescent protein tag (GFP-MOB1, GFP-M4K, GFP-M4KC, GFP-M4KK, YFP-MBP and RFP-CESA8-A). For BiFC, the destination vectors used were pSITE-nEYFP-C1 and pSITE-cEYFP-C1 from ABRC. Gateway cloning was used for fusion of nEYFP (ny) to the N-terminus of MOB1 (ny-MOB1) and cEYFP (cy) to the N-terminus of M4K (cy-M4K) and its deletions. Expression of all fusion proteins in *Nicotiana benthamiana* (tobacco) leaves was directed by the Cauliflower mosaic virus (CaMV) *35S* promoter. For localization and BiFC experiments, Agrobacterium cells were grown o/n and bought to the final concentration of OD_{600} 0.4 before infiltration of tobacco leaves with Agrobacterium, containing vectors for expression of proteins indicated in Figures 3 and 4. The proteins were allowed to express for 48 h followed by observation using a Zeiss LSM 880 confocal laser scanning microscope using a 488-nm argon laser and a 505–550-nm band-pass emission filter (GFP/YFP) and 543-nm HeNe laser and 560-nm band-pass emission filter (RFP).

Pull-downs of GFP-tagged proteins

The GST-MOB1 fusion was cloned using vector pGEX 4T-1 (GE Healthcare, Buckinghamshire, UK). GST-MOB1 fusions were purified using GST SpinTrap™ columns according to the manufacturer's instructions (booklet 28-09,523-59, GE Healthcare). Protein concentrations were determined using the Pierce™ BCA Protein Assay Kit (Thermo Fischer Scientific, Waltham, MA USA). Equal levels of fusion proteins and GST controls were bound to the columns for pull-down assays. Fluorescence of GFP in infiltrated leaves of *N. benthamiana* was confirmed by fluorescence microscopy 72 h post-inoculation. Protein was extracted from GFP-positive leaves by homogenizing 0.1 g leaf tissue in 1 ml of freshly prepared IP buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.5% (v/v) NP-40, 5 mM β-mercaptoethanol, 1% (v/v) Halt protease cocktail (Thermo Fischer Scientific, Waltham, MA, USA). The homogenate was centrifuged at 10,000*g* for 15 min at 4 °C. Clarified supernatant (500 µl) was added to SpinTrap™ columns pre-loaded with either GST-fusion protein or GST only (negative control). Columns were rotated for 2 h at 4 °C, washed with 15 bed volumes of buffer TBST (25 mM Tris-HCl, pH 7.5, 5 mM β-mercaptoethanol, 150 mM NaCl, 0.5% (v/v) Tween-20, 1% (v/v) Halt protease cocktail). A 10 µl slurry of GST-beads, representing 5% of the beads, was boiled in SDS-PAGE sample buffer and analyzed by western blotting using anti-GFP primary antibody (A11122 Invitrogen), and anti-IgG-HRP secondary antibody (4054-05, Southern Biotech, Birmingham, AL, USA). Western blots were developed using ECL™ Prime Western Blotting Detection Reagents (RPN2232, GE Healthcare) according to manufacturer's instructions prior to chemiluminescence detection on ChemiDoc™ plus XRS system (Bio-Rad, Hercules, CA, USA).

Results and discussion

Protein–protein interaction discovery

As a resource for Y2H and Y1H assays and functional analyses of genes associated with poplar wood formation, we cloned ORFs for genes upregulated by at least eightfold based on microarray analysis of gene expression in secondary xylem compared to secondary phloem/cambium of *P. trichocarpa* Nisqually-1 (Rodgers-Melnick et al. 2012). We attempted to clone 455 PW ORFs and succeeded in cloning 361 (see Figure S1 available as Supplementary Data at *Tree Physiology* Online). Open reading frames for 40 additional genes implicated in biomass accumulation and wood formation through results of Y2H assays (discussed below) were also cloned. Together these 401 genes represent the PW gene set available for use during this study (see Table S1a available as Supplementary Data at *Tree Physiology* Online). Recently, RNAseq data from 12 cryosections of the cambial zone and its flanking developing phloem and xylem of *Populus tremula* × *tremuloides* became available (Immanen et al. 2016). After sample clustering and analysis of tissue-specific markers used by Immanen et al. (2016), we classified tissues as phloem (sections P1 and P2), cambial zone (sections CP, C1, C2 and CX), young xylem (sections YX1 and YX2), transition xylem (TX) and mature xylem (sections MX1, MX2, and MX3) and evaluated expression of our 401-member PW gene set using these RNAseq data (see Figure S2, Table S1b available as Supplementary Data at *Tree Physiology* Online). As expected, the vast majority (93%) of PW ORFs cloned

for this study exhibited the highest tissue specificity scores associated with xylem: 66% in mature, 22% in transition and 5% in young xylem. The smallest percentages of PW transcripts had their highest specificity scores associated with phloem and cambial zone samples, at 3% and 4%, respectively (Figure 1A). Together, the microarray and RNAseq analyses support the conclusion that our collection of PW clones is highly relevant to wood formation in poplar (Rodgers-Melnick et al. 2012, Immanen et al. 2016).

In preparation for Y2H, DB-fusion PW clones capable of supporting yeast growth on media lacking His and supplemented with 3AT, i.e., auto-activators, were eliminated from further analysis by Y2H, as were selected co-expressed members of multigene families in the PW ORF collection. Ultimately, 290 clones were screened by Y2H as DB and AD fusions of each clone in all pairwise combinations (i.e., a binary Y2H matrix), resulting in the discovery of 14 interactions involving 12 distinct proteins (see Table S2c available as Supplementary Data at *Tree Physiology* Online). We compared our binary Y2H results to those obtained with the relevant subset of homologous proteins screened for the Arabidopsis Interactome Mapping Consortium (2011) Y2H screen, i.e., the AI-1Main dataset. Of the Arabidopsis proteins exhibiting the highest amino acid similarity to the 401 PW proteins, only 122 were represented in the AI-1Main dataset. This 122 × 122 Arabidopsis Y2H binary matrix space yielded a total of six PPIs, i.e., a proportion of PPIs identical to that discovered in our poplar binary Y2H matrix screen (six interactions among 122 Arabidopsis proteins = 5% compared to 14 interactions among 290 poplar proteins = 5%).

Proteins identified as interactors in the binary Y2H matrix and a variety of additional bait proteins (DB fusions) selected from the original 361 PW clones were used in an initial Y2H screen of a library of prey proteins (AD fusions) prepared from PW cDNA. Subsequent Y2H screens of this same library and further testing of PPIs discovered by library screens were carried out using full-length ORFs based on ISTs (see Figure S1, Table S1d available as Supplementary Data at *Tree Physiology* Online). Interaction sequence tags (ISTs) are defined here as any sequence that encodes an in-frame fusion with the Gal4 AD and matches a *P. trichocarpa* locus. In total, 60 proteins were screened against the PW cDNA AD-fusion prey library. Of these, 37 PW ORFs yielded 102 ISTs representing 90 different poplar genes. The majority of ISTs represented N-terminally truncated proteins, with only 11 of 102 ISTs beginning with the predicted initiator Met (see Table S1d available as Supplementary Data at *Tree Physiology* Online). With regard to retesting PPIs discovered by library screens, 61 full-length PW ORFs were used to confirm 64 interactions with other full-length PW ORFs in selected binary Y2H assays. Together, the binary Y2H assays and cDNA library Y2H screens yielded 165 interactions involving 162 distinct poplar full-length ORFs plus ISTs (Figure 2, Table 1; Figure S1, Table S2c available as Supplementary Data at *Tree Physiology* Online).

Tissue specificity of genes for interacting proteins

As noted above, expression of the vast majority (93%) of the PW full-length ORFs cloned for this study exhibited high specificity for xylem (Figure 1A). When specificity scores for expression of all of the interacting proteins (ISTs and full-length clones) are considered, the percentage exhibiting high specificity for all stages of xylem drops to 65%, while scores for phloem and cambial zone increase to 18% and 17%, respectively (Figure 1B). As these Y2H results indicate that PPIs involving PW proteins occur in phloem and cambium as well as xylem, we performed pairwise comparisons of the tissue-specific gene expression specificity scores for all 165 PPIs and generated a CSP for each PPI (Figure 1D; see Table S2c available as Supplementary Data at *Tree Physiology*

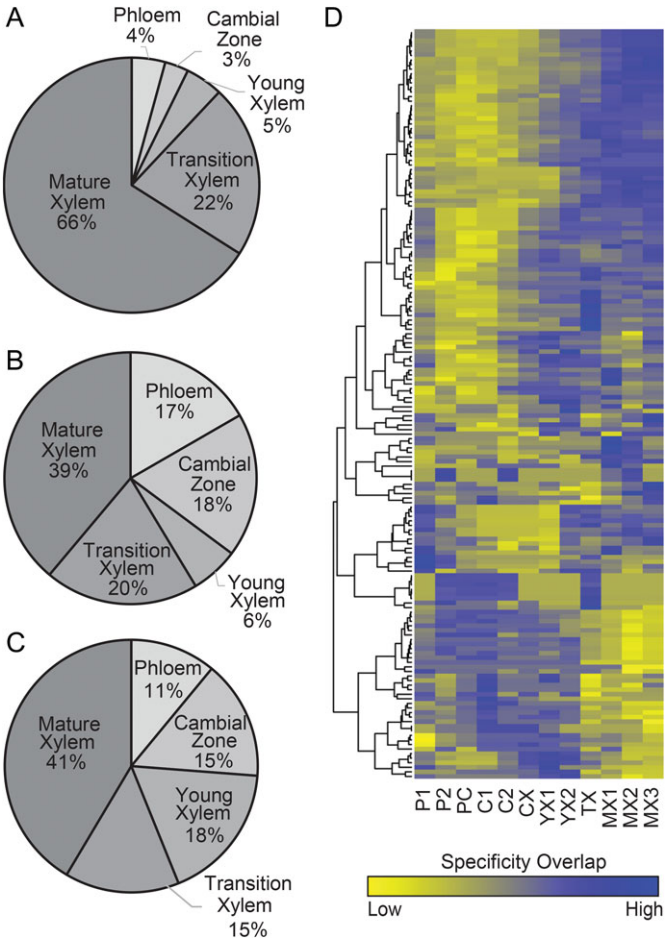

Figure 1. Tissue specificity of single and interacting protein pairs. (A) The members of the PW ORF collection cloned for this study grouped by tissue specificity scores. (B) Individual full-length ORFs and ISTs involved in PPIs grouped by tissue specificity scores. (C) Interacting protein pairs grouped by tissue specificity scores based on their combined specificities profile (CSP). (D) Heatmap and associated dendrogram of the CSPs for all 164 interacting protein pairs across 12 sections of poplar vascular tissue: phloem (P1 and P2); cambial zone (PC, C1, C2 and CX); and young xylem (YX1, YX2), transition xylem (TX) and mature xylem (MX1, MX2 and MX3).

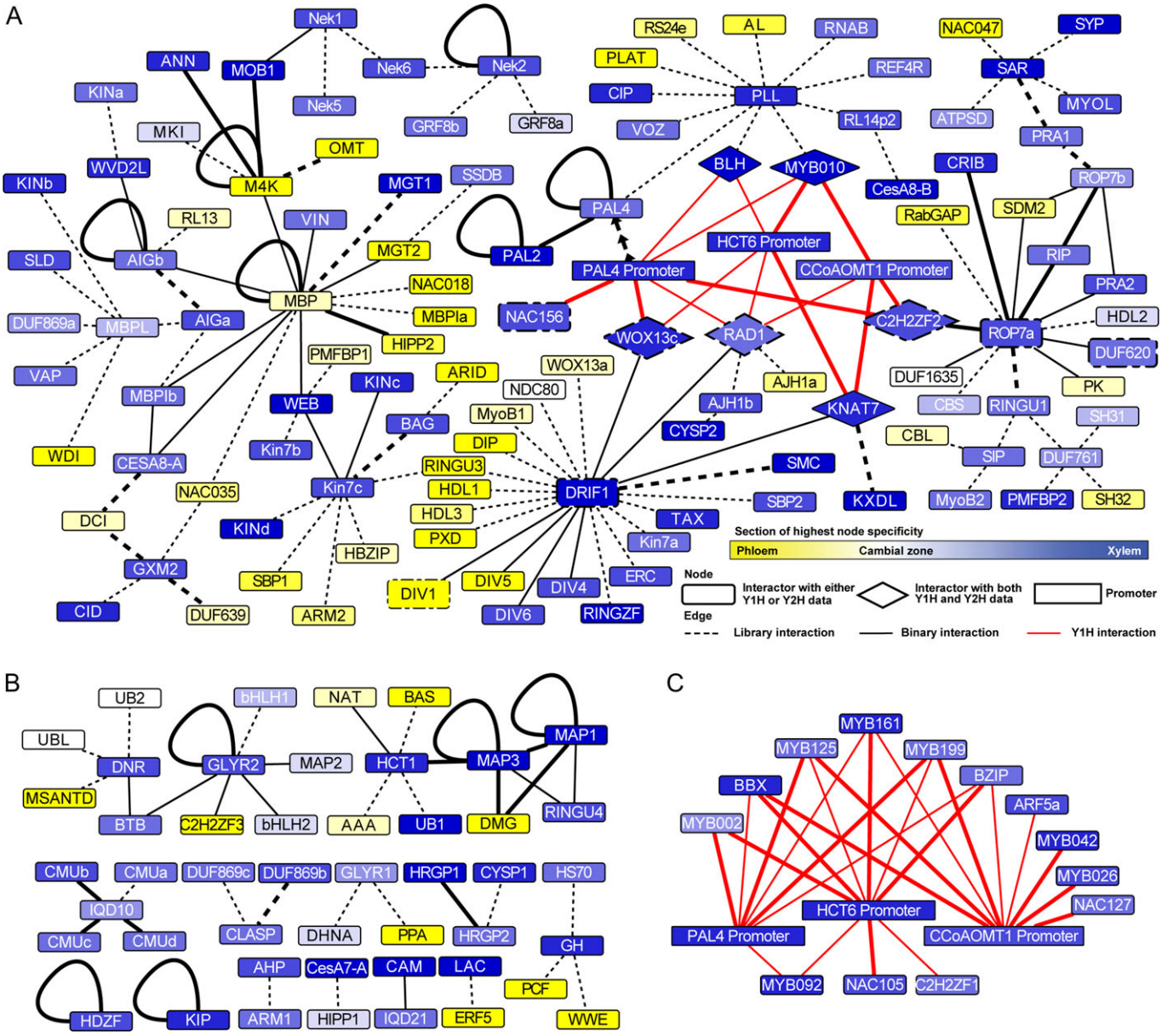

Figure 2. Cytoscape networks of all PPIs and PDIs. (A) Large network presenting Y1H and Y2H data. (B) Smaller Y2H networks. (C) Transcription factor Y1H network. Nodes were color coded based on the section with the highest expression specificity score. Genes targeted for functional studies in transgenic poplar are indicated by a dash dot border line type. Edge thickness indicates Bi-weight mid-correlation scores for interactions which were positively correlated and below significance threshold (*P*-value ≤0.05). For the sole instance where the promoter and gene product were cloned to generate Y1H and Y2H data, PAL4, a separated arrow edge type indicates this relationship.

Online) for the purpose of predicting tissue localization. As was discovered by analysis of the specificity profile for each of the 401 PW genes cloned for the project, the tissue for which most interactions have their highest CSP score is xylem at 74%, comprised of young, transition and mature xylem at 18%, 15% and 42%, respectively. Cambial zone follows at 15% and phloem at 10% (Figure 1C). Gene expression levels do not always correlate with protein levels. Hence, the predictions of tissue localization for PPIs derived from the CSPs are tentative and require further investigation. Subcellular co-localization is another essential feature of physiologically relevant PPIs. Predicted localizations are shown in Table S2d available as Supplementary Data at *Tree Physiology* Online and selected PW proteins (discussed below) have been experimentally localized for this report.

Integrating PPIs and PDIs identifies connected networks

The cytoscape network representing the PW PPIs and PDIs reported here includes 179 nodes. There are 11 self-loops, and 14 connected subnetworks, the largest being comprised of 132 members (Table 1, Figure 2). The other subnetworks are considerably smaller and range from 10- to 2-member subnetworks and two isolated

Table 1. List of poplar wood-associated proteins involved in PPIs or PDIs. Interactors are grouped by subnetworks shown in Figure 2 in descending order of subnetwork membership number.

Potri ID	Arabidopsis	Name	Description
132-member subnetwork			
Y1H interactions			
Potri.009G099800	AT4G34050	CCoAOMT1p	S-adenosyl-L-methionine-dependent methyltransferase
Potri.001G042900	AT5G48930	HCT6p	Hydroxycinnamoyl transferase family protein
Potri.010G224100	AT2G37040	PAL4p	Phenylalanine ammonia-lyase
Potri.002G024700	AT1G19850	ARF5a	B3 family auxin-responsive factor
Potri.007G121100	AT4G15248	BBX	B-box type zinc finger with CCT domain-containing
Potri.008G018400	AT5G04840	BZIP	Basic-leucine zipper (bZIP) transcription factor family
Potri.010G209400	AT2G28200	C2H2ZF1	C2H2-type zinc finger family
Potri.001G258700	AT5G12870	MYB002	Myb domain
Potri.005G063200	AT4G33450	MYB026	Myb domain
Potri.010G004300	AT1G57560	MYB042	Myb domain
Potri.001G118800	AT4G12350	MYB092	Myb domain
Potri.003G114100	AT4G12350	MYB125	Myb domain
Potri.007G134500	AT1G17950	MYB161	Myb domain
Potri.012G127700	AT4G22680	MYB199	Myb domain
Potri.011G058400	AT4G28500	NAC105	NAC (no apical meristem) domain superfamily
Potri.018G068700	AT4G29230	NAC127	NAC (no apical meristem) domain superfamily
Potri.007G135300	AT4G28500	NAC156	NAC (no apical meristem) domain superfamily
Y1H and Y2H interactions			
Potri.010G197300	AT2G16400	BLH	BEL1-like homeodomain
Potri.014G066200	AT3G60580	C2H2ZF2	C2H2-type zinc finger family
Potri.001G112200	AT5G25220	KNAT7	Homeobox knotted
Potri.001G099800	AT1G63910	MYB010	Myb domain
Potri.002G260000	AT1G75250	RAD1	Duplicated homeodomain-like superfamily
Potri.002G008800	AT4G35550	WOX13c	WUSCHEL related homeobox 13 C
Y2H interactions			
Potri.019G077300	AT1G33970	AIGa	P-loop containing NTP hydrolases superfamily
Potri.013G104800	AT1G33970	AIGb	P-loop containing NTP hydrolase superfamily
Potri.006G275100	AT1G22920	AJH1a	Mov34 MPN PAD-1 family
Potri.018G006100	AT1G22920	AJH1b	Mov34 MPN PAD-1 family
Potri.006G145300	AT5G26210	AL	Alfin-like
Potri.002G095600	AT1G35720	ANN	Annexin
Potri.004G208900	AT2G17410	ARID	ARID BRIGHT DNA-binding domain-containing
Potri.012G074400	AT5G62580	ARM2	ARM repeat superfamily
Potri.016G139000	AT2G38580	ATPSD	Mitochondrial ATP synthase D chain-related
Potri.015G135500	AT5G52060	BAG	BCL-2-associated athanogene
Potri.001G371700	AT5G55990	CBL	Calcineurin B 10
Potri.001G303900	AT4G36910	CBS	Cystathionine beta-synthase (CBS) family
Potri.011G069600	AT4G18780	CESA8-A	Cellulose synthase family
Potri.004G059600	AT4G18780	CesA8-B	Cellulose synthase family
Potri.011G152700	AT4G10610	CID	CTC-interacting domain
Potri.001G164900	AT3G14172	CIP	COP1-interacting
Potri.T058500	AT5G16490	CRIB	ROP-interactive CRIB motif-containing
Potri.005G256000	AT1G20850	CYSP2	Granulin repeat cysteine protease family
Potri.014G097800	AT3G61700	DIP	helicase with zinc finger
Potri.016G112300	AT2G38090	DIV1	Duplicated homeodomain-like superfamily
Potri.009G042600	AT5G58900	DIV4	Duplicated homeodomain-like superfamily
Potri.010G193000	AT5G05790	DIV5	Duplicated homeodomain-like superfamily
Potri.008G064200	AT5G05790	DIV6	Duplicated homeodomain-like superfamily
Potri.007G132700	AT3G07565	DRIF1	DIV and RAD INTERACTING FACTOR1
Potri.009G006000	AT5G22930	DUF1635	DUF1635
Potri.004G087600	AT5G16720	MyoB1	Myosin-binding protein1 (DUF593)
Potri.002G242900	AT4G13630	MyoB2	Myosin-binding protein2 (DUF593)
Potri.010G174200	AT1G79420	DUF620	DUF620

(*Continued*)

Table 1. (Continued)

Potri ID	Arabidopsis	Name	Description
Potri.012G055800	AT1G48840	DUF639	Heat-inducible transcription repressor (DUF639)
Potri.001G406600	AT5G54300	DUF761	Cotton fiber (DUF761)
Potri.005G140300	AT2G23360	DUF869a	Filament (DUF869)
Potri.014G139000	AT1G03290	ERC	ELKS Rab6-interacting CAST family
Potri.010G056200	AT1G27850	DCI	Disorded CesA interactor
Potri.002G103800	AT5G65430	GRF8a	General regulatory factor 8a
Potri.005G157700	AT5G65430	GRF8b	General regulatory factor 8b
Potri.013G102200	AT1G09610	GXM2	Glucuronoxylan 4-O-methyltransferase (DUF579)
Potri.016G059000	AT5G06710	HBZIP	Homeobox-leucine zipper family
Potri.001G314800	AT4G13640	HDL1	Homeodomain-like superfamily
Potri.008G025600	AT5G03680	HDL2	Homeodomain-like superfamily
Potri.017G054800	AT4G13640	HDL3	Homeodomain-like superfamily
Potri.016G006600	AT5G50740	HIPP2	Heavy metal transport detoxification superfamily
Potri.001G416300	AT5G54670	KINa	Kinesin-like
Potri.008G221400	AT3G16630	KINb	Kinesin-like
Potri.008G226200	AT3G16630	KINc	Kinesin-like
Potri.T115700	AT3G16630	KINd	Kinesin-like
Potri.011G098200	AT3G29130	KXDL	KXDL motif
Potri.011G116300	AT3G15220	M4K	Kinase superfamily
Potri.015G099100	NA	MBP	Multiple binding partners
Potri.007G022100	AT3G50910	MBPla	Multiple binding partners interacting a
Potri.005G196000	NA	MBPlb	Multiple binding partners interacting b
Potri.012G100900	NA	MBPL	Multiple binding partners-like
Potri.001G043200	AT1G16010	MGT1	Magnesium transporter
Potri.008G161400	AT5G64560	MGT2	Magnesium transporter
Potri.015G029300	AT4G27750	MKI	MAPK-interacting protein
Potri.006G194900	AT3G12020	Kin7a	ATP binding microtubule motor family
Potri.016G060400	AT3G12020	Kin7b	ATP binding microtubule motor family
Potri.004G161100	AT4G39050	Kin7c	ATP binding microtubule motor family
Potri.001G132700	AT4G19045	MOB1	Mob1 phocein family
Potri.006G021200	AT2G32240	MYOL	Myosin heavy chain-like
Potri.001G452700	AT1G33060	NAC018	NAC (no apical meristem) domain superfamily
Potri.005G200100	AT1G34190	NAC035	NAC (no apical meristem) domain superfamily
Potri.008G031800	AT3G10500	NAC047	NAC (no apical meristem) domain superfamily
Potri.001G208700	AT3G54630	NDC80	Kinetochore protein
Potri.005G051600	AT1G54510	Nek1	NIMA-related kinase
Potri.002G049400	AT3G04810	Nek2	NIMA-related kinase
Potri.001G218100	AT3G44200	Nek5	NIMA-related kinase
Potri.006G056300	AT3G12200	Nek6	NIMA-related kinase
Potri.012G006400	AT5G54160	OMT	O-Methyltransferase family
Potri.008G038200	AT3G53260	PAL2	phenylalanine ammonia-lyase
Potri.010G224100	AT2G37040	PAL4	phenylalanine ammonia-lyase
Potri.010G080700	AT3G22960	PK	Pyruvate kinase family
Potri.001G167700	AT1G72520	PLAT	PLAT LH2 domain-containing lipoxygenase family
Potri.003G131700	AT4G23500	PLL	Pectin lyase-like superfamily
Potri.005G115500	AT2G17990	PMFBP1	Polyamine-modulated factor 1-binding
Potri.007G012600	AT2G17990	PMFBP2	Polyamine-modulated factor 1-binding
Potri.016G126400	AT2G38360	PRA1	Prenylated RAB acceptor
Potri.019G124100	AT2G40380	PRA2	Prenylated RAB acceptor
Potri.018G026000	AT4G32160	PXD	Phox (PX) domain-containing
Potri.006G133400	AT2G37290	RabGAP	Ypt Rab-GAP domain of gyp1p superfamily
Potri.008G201600	AT3G23590	REF4R	REF4-related
Potri.004G217700	AT2G27950	RINGU1	RING U-box superfamily
Potri.014G055800	AT2G44950	RINGU3	RING U-box superfamily
Potri.012G055100	AT3G18290	RINGZF	C3HC4-type RING Zinc finger
Potri.016G085900	AT2G37080	RIP	ROP interactive partner

(*Continued*)

Table 1. (Continued)

Potri ID	Arabidopsis	Name	Description
Potri.017G054600	AT5G48760	RL13	Ribosomal L13 family
Potri.010G066400	AT3G04400	RL14p2	Ribosomal L14p L23e family
Potri.003G204300	AT3G51950	RNAB	RNA-binding (RRM RBD RNP motifs) family
Potri.011G061500	AT1G75840	ROP7a	RAC-like
Potri.019G092300	AT2G17800	ROP7b	RAC-like
Potri.005G049400	AT3G04920	RS24e	Ribosomal S24e family
Potri.002G240900	AT2G32670	SAR	Synaptobrevin-related
Potri.011G142700	AT3G12920	SBP1	SBP (S-ribonuclease binding) family
Potri.012G119200	AT5G45100	SBP2	SBP (S-ribonuclease binding) family
Potri.006G023400	AT2G32170	SDM2	S-adenosyl-L-methionine-dependent methyltransferase
Potri.001G354700	AT4G18060	SH31	SH3 domain-containing
Potri.003G105200	AT1G31440	SH32	SH3 domain-containing
Potri.013G155600	AT2G30360	SIP	SOS3-interacting
Potri.017G054300	AT2G32720	SLD	Fatty acid sphingolipid desaturase
Potri.008G048100	AT5G13560	SMC	Structural maintenance of chromosomes
Potri.003G118600	AT3G58630	SSDB	Sequence-specific DNA binding transcription factor
Potri.019G036700	AT3G03800	SYP	Syntaxin of plants
Potri.010G219700	AT5G50840	TAX	Alpha-taxilin-like protein
Potri.005G044900	AT1G08820	VAP	Vesicle associated
Potri.018G091500	AT4G30200	VIN	Vernalization insensitive
Potri.011G060000	AT1G28520	VOZ	Vascular plant one zinc finger
Potri.017G106400	AT5G11390	WDI	WPP domain-interacting
Potri.004G222800	NA	WEB	WEB family
Potri.005G101800	AT4G35550	WOX13a	WUSCHEL related homeobox 13 A
Potri.008G162800	AT3G23090	WVD2L	Wave dampened 2-like
Ten-member subnetwork			
Y2H interactions			
Potri.015G048000	AT5G08130	bHLH1	Basic helix-loop-helix (bHLH) DNA-binding superfamily
Potri.012G065000	AT5G08130	bHLH2	Basic helix-loop-helix (bHLH) DNA-binding superfamily
Potri.014G093700	AT3G61600	BTB	BTB POZ domain-containing
Potri.005G027200	AT1G03840	C2H2ZF3	C2H2-like zinc finger
Potri.011G057500	AT1G28400	DNR	GATA zinc finger
Potri.017G087200	AT5G39570	GLYR2	Glycine-rich protein
Potri.008G135100	AT1G68060	MAP2	Microtubule-associated protein
Potri.015G137800	AT5G05800	MSANTD	Myb SANT-like DNA-binding domain
Potri.005G100200	AT2G17200	UB2	Ubiquitin family
Potri.005G247900	AT5G42220	UBL	Ubiquitin-like superfamily
Nine-member subnetwork			
Y2H interactions			
Potri.016G028000	AT1G14840	AAA	AAA-type ATPase family
Potri.007G002500	AT1G78950	BAS	Terpenoid cyclases family
Potri.010G137300	AT1G13635	DMG	DNA glycosylase superfamily
Potri.003G183900	AT5G48930	HCT1	Hydroxycinnamoyl transferase family protein
Potri.016G006900	AT1G14840	MAP1	Microtubule-associated protein
Potri.006G018000	AT2G01750	MAP3	Microtubule-associated protein
Potri.001G279400	AT2G30090	NAT	Acyl-*N*-acyltransferases (NAT) superfamily
Potri.001G127100	AT5G45290	RINGU4	RING U-box superfamily
Potri.004G205100	AT2G17200	UB1	Ubiquitin family
Five-member subnetwork			
Y2H interactions			
Potri.001G087800	AT4G10840	CMUa	Cellulose synthase-microtubule uncoupling protein
Potri.008G094700	AT3G27960	CMUb	Cellulose synthase-microtubule uncoupling protein
Potri.014G100400	AT4G10840	CMUc	Cellulose synthase-microtubule uncoupling protein
Potri.003G143200	AT4G10840	CMUd	Cellulose synthase-microtubule uncoupling protein
Potri.011G096500	AT3G15050	IQD10	IQ-domain 10

(*Continued*)

Table 1. (Continued)

Potri ID	Arabidopsis	Name	Description
Four-member subnetwork			
Y2H interactions			
Potri.T167100	AT4G16260	GH	Glycosyl hydrolase superfamily
Potri.008G054700	AT5G02500	HS70	Heat shock 70 (Hsp 70) family
Potri.007G134000	AT2G36480	PCF	Pre-mRNA cleavage complex II
Potri.001G137200	AT1G32230	WWE	WWE-interaction domain family
Three-member subnetworks			
Y2H interactions			
Potri.016G067200	AT3G11750	DHNA	Dihydroneopterin aldolase
Potri.015G111800	NA	GLYR1	Glycine-rich protein
Potri.007G022700	AT3G53620	PPA	Pyrophosphorylase
Potri.004G160300	AT4G39090	CYSP1	Papain family cysteine protease
Potri.009G073600	AT1G07120	HRGP1	Hydroxyproline-rich glycoprotein family
Potri.001G279000	AT1G07120	HRGP2	Hydroxyproline-rich glycoprotein family
Potri.002G253200	AT2G20190	CLASP	CLIP-associated protein
Potri.007G098500	AT3G05270.	DUF869b	DUF869
Potri.005G070400	AT3G05270	DUF869c	DUF869
Two-member subnetworks			
Y2H interactions			
Potri.008G197600	AT3G21510	AHP	Histidine-containing phosphotransmitter
Potri.010G090900	AT5G19330	ARM1	ARM repeat interacting with ABF
Potri.006G181900	AT4G39350	CesA7-A	Cellulose synthase family
Potri.004G175400	AT4G38580	HIPP1	Heavy metal transport detoxification superfamily
Potri.010G146900	AT5G49480	CAM	Calmodulin
Potri.015G012500	AT3G49260	IQD21	IQ-domain 21
Potri.003G150800	AT5G51190	ERF5	Integrase-type DNA-binding superfamily
Potri.008G064000	AT2G38080	LAC	Laccase Diphenol oxidase family
Y2H self-interactions			
Potri.017G082900	AT3G28920	HDZF	Homeobox 31
Potri.003G164200	AT1G03080	KIP	Kinase interacting (KIP1-like) family

self-interactors. The network indicates new candidate regulators of wood development and enables the phenotypic effects of different network perturbations (e.g., manipulated hub proteins versus less connected proteins) to be studied. Accordingly, we selected genes (see Figure 2) for functional characterization via transgenic manipulation, including targeted expression of altered genes in the cambium/xylem and downregulation. These transgenics are currently being evaluated in field and greenhouse studies. Below we discuss examples of previously identified PPIs and PDIs (discussed in next section) confirmed during this work, and highlight selected novel wood-formation-associated proteins implicated via interactions with proteins or promoters for genes that have well-established roles in secondary cell wall synthesis and wood formation.

Protein–protein networks include new wood-associated hubs and interactions

Two of the poplar proteins involved in binary Y2H matrix screen interactions, Rho of Plants 7a (ROP7a) and IQ-domain 21 (IQD21) (see Table S2c available as Supplementary Data at *Tree Physiology* Online) belong to families previously reported in connection with trafficking or functioning of cellulose-synthase complexes and their association with the cytoskeleton (Oda and Fukuda 2012, Burstenbinder et al. 2013, Liu et al. 2016). In a screen against a xylem cDNA prey library, IQD21 interacted with calmodulin, while IQD10 interacted with four tetratricopeptide repeat (TPR)-like proteins (Table 1, Figure 2). The TPR-like proteins are putatively orthologous to the Arabidopsis cellulose synthase-microtubule uncoupling proteins (CMUa/b/c/d). Together these interactions with IQD10 and IQD21 confirm previous findings with homologs from Arabidopsis (Burstenbinder et al. 2013, Liu et al. 2016). In Arabidopsis, the small GTPase ROP11 in association with microtubule depletion domain 1 (MIDD1) is required for secondary cell wall pitting (Oda and Fukuda 2012). In contrast, past efforts to identify xylem-specific roles for ROP7 in Arabidopsis were inconclusive (Brembu et al. 2005). From our cDNA library screen and binary Y2H assays with selected full-length clones, we were able to demonstrate that three of the five proteins shown to interact with ROP11 by the Arabidopsis Interactome Mapping Consortium (2011) also interacted with ROP7a, i.e., DUF620, RIP (a MIDD1 homolog), and Cystathionine beta-synthase-like (CBS), as did additional well-known GTPase-interacting proteins Prenylated Rab acceptor

2 (PRA2), Rab-GTPase activating protein (RabGAP) and Cdc42/Rac interactive binding protein (CRIB) (Wu et al. 2001, Alvim Kamei et al. 2008, Frasa et al. 2012).

Cellulose synthase complexes ultimately localize to the plasma membrane (Haigler and Brown 1985) and thus may not be visible in an assay based on nuclear-localized Y2H interactions. However, a complete picture of trafficking and assembly of cellulose synthase complexes has not yet been presented (Bashline et al. 2014). Consequently, it is possible that PPIs relevant to discrete aspects of the cell biology of cellulose synthase can be discovered by the Y2H system used here. We used the PW cellulose synthase subunit CESA8-A as bait in our Y2H screens of the PW cDNA library. CESA8-A interacted with three uncharacterized proteins, Disordered CESA Interactor (DCI), Multiple binding partners (MBP) and MBP interacting b (MBPIb), all of which lacked known conserved domains. MBP homologs are present in many plant species, but apparently absent from Arabidopsis. However, identification of homologs by BLAST may be confounded by the characteristics of MBP, a relatively small (216 amino-acid), low complexity, disordered protein, i.e., 62% of MBP amino acids are in disordered regions (PONDR® www.pondr.com). Attempts to independently confirm the MBP-CESA8-A interaction by BiFC were unsuccessful. However, during transient expression in tobacco, MBP co-localized with mCherry-ER in the endoplasmic reticulum (ER) (Figure 3). Likewise, MBP and CESA8-A co-localized during transient expression (Figure 4), possibly to the ER, although CESA subunits have not yet been found in the ER (Haigler and Brown 1985, Wightman and Turner 2010). MBP screened against the PW cDNA library interacted with a wide range of proteins, including HEAVY METAL ASSOCIATED ISOPRENYLATED PLANT PROTEIN 2 (HIPP2), Magnesium transporters (MGT1, MGT2), a GTP-binding protein similar to the avirulence-induced gene product (AIGb), transcription factors (NAC018, NAC035) and M4K (Figure 2). Several of these interactions with MBP were confirmed by Y2H using full-length clones, i.e., for MGT2, HIPP2, AIGb and M4K. MBP has a paralog in poplar, MBPL. As with MBP, MBPL is predicted to be disordered and interacted with a variety of prey proteins in the PW cDNA library, including those with links to the cytoskeleton (Kinesin b (KINb), VESICLE ASSOCIATED PROTEIN (VAP)) and nuclear movement or shape (WPP DOMAIN-INTERACTING PROTEIN (WDI)) (Oda and Fukuda 2013, Tamura et al. 2013, Wang et al. 2016). The ability of MBP to interact with multiple proteins may be linked with its high degree of disorder and thus further investigations of MBP as a potential disordered hub protein are warranted by these PPI results (Dunker et al. 2005).

The observed interaction between MBP and M4K is notable in light of recent work with another MAP4K, serine/threonine kinase 1 (SIK1), from Arabidopsis. SIK1 interacted with Mps one binder (MOB1) (Xiong et al. 2016). Functional analyses implicated both members of this interacting pair in the control of cell proliferation and expansion, mirroring some of the roles attributed to the mammalian Mst/Hippo kinase cascade, wherein MOB1 is a core component (Rawat and Chernoff 2015). We used M4K as bait in a Y2H screen of the PW cDNA library and found that M4K also interacted with MOB1.

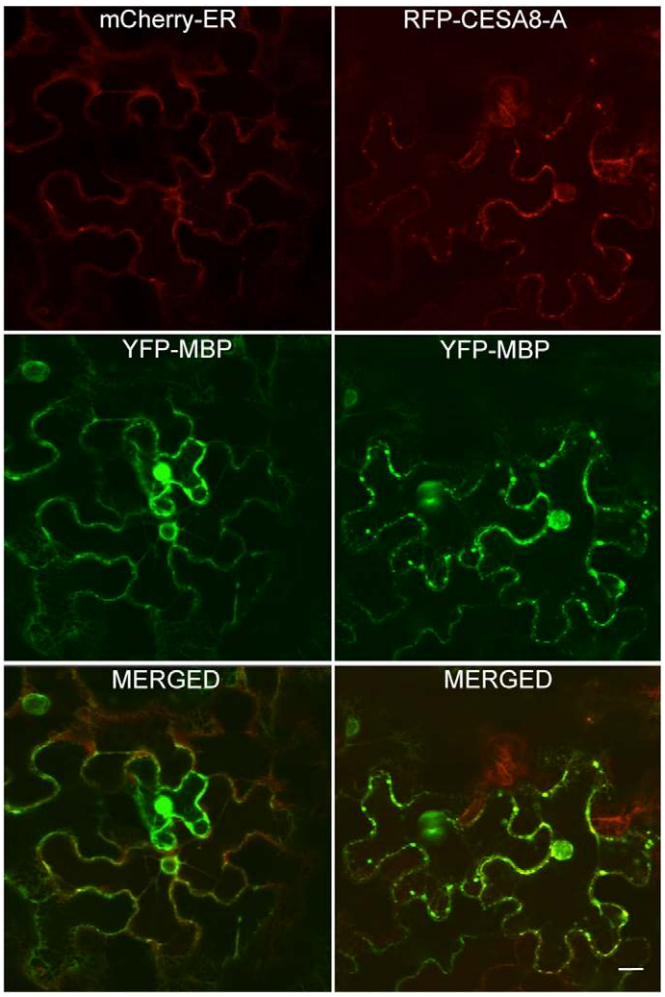

Figure 3. Multiple binding partners (MBP) co-localizes with the endoplasmic reticulum marker mCherry-ER and CESA8-A. YFP-MBP and mCherry-ER (left panels) or YFP-MBP and RFP-CESA8-A (right panels) were transiently co-expressed in tobacco and fluorescence was observed after 48 h under the confocal microscope. Merged images show sites of YFP-MBP co-localized with mCherry-ER and YFP-MBP co-localized with RFP-CESA8-8, visible as yellow where mCherry-ER or RFP and YFP signals co-localized. Scale bar in bottom right panel is 20 μm for all panels.

We performed additional analyses of the M4K–MOB1 interaction using Y2H, pull-down assays and BiFC. Y2H assays testing reciprocal interactions between M4K, the M4K C-terminal domain (M4KC, amino acids 271–665), or the N-terminal kinase domain (K4KK, amino acids 1–270) and MOB1 indicated that MOB1 interacted with the C-terminal domain but not the N-terminal kinase domain. The Y2H results were consistent with those observed with BiFC using M4KC or M4KK co-expressed with MOB1, i.e., the combination of full-length M4K or M4KC with

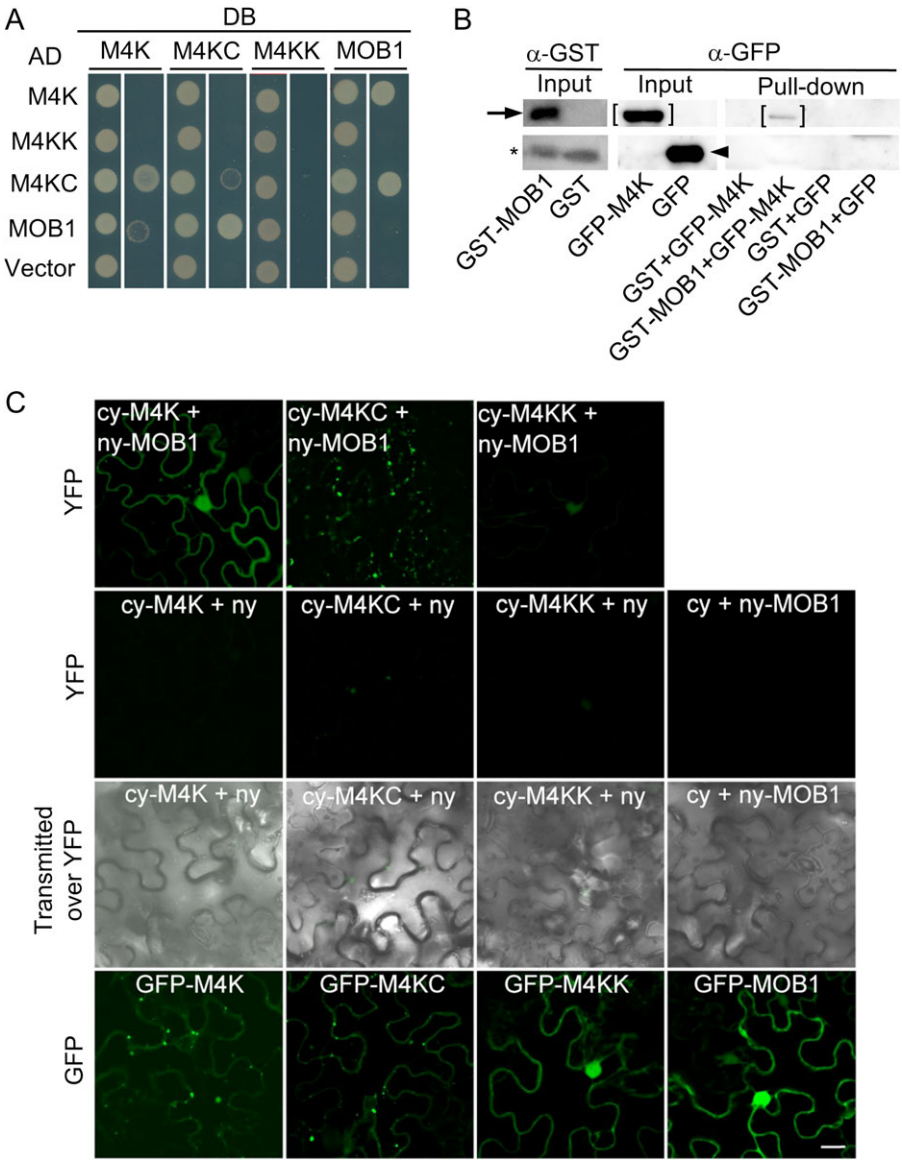

Figure 4. Yeast two-hybrid, pull-down and bimolecular fluorescence complementation (BiFC) assays showing the interaction of M4K and MOB1. (A) A yeast two-hybrid binary interaction matrix consisting of M4K, the M4K C-terminal domain (M4KC), M4K kinase domain (M4KK), or MOB1 fused with the Gal4 DNA binding domain (DB; above yeast colonies) or the Gal4 activation domain (AD; at left of yeast colonies) was tested in all pairwise combinations. For each pair of lanes showing yeast growth beneath each DB fusion protein, the left-hand lane shows growth on synthetic dropout medium lacking leucine and tryptophan (SD) and the right-hand lane shows yeast growth on SD without adenine, indicating protein–protein interactions where yeast growth is observed. The AD vector-only negative control shows absence of auto-activation by each DB fusion. (B) For pull-downs, GST-MOB1 fusion protein (arrow) bound to glutathione beads was incubated with GFP-M4K fusion protein (brackets) for interaction or with GFP only (arrowhead) as a negative control. Additional negative controls include GST (asterisk) incubated with GFP-M4K and GST incubated with GFP. GST in GST-MOB1 lane is a breakdown product that co-purified with GST-MOB1. (C) For BiFC, M4K, M4KC or M4KK, fused to the C-terminal portion of YFP (cy), was transiently co-expressed in tobacco with MOB1 fused to the N-terminal portion of YFP (ny). Negative controls consisted of each M4K-based cy fusion co-expressed with the ny-only vector and ny-MOB1 with the cy-only vector. Subcellular localizations of GFP-M4K, GFP-M4KC, GFP-M4KK and GFP-MOB1 fusion proteins are shown (bottom row) for comparison with localizations of BiFC-dependent YFP. Scale bar in GFP-MOB1 panel (bottom right) is 20 μm for all panels in (C).

MOB1 produced a strong YFP signal, while the M4KK/MOB1 combination yielded barely detectable YFP that was only slightly above background levels observed for negative controls (Figure 4). Localization of GFP-MK4 was punctate, suggesting association with organelles or vesicles currently of unknown identity, but similar in appearance to localization of SIK1 to *trans*-Golgi/early endosome (Xiong et al. 2016). MOB1 localized to the nucleus and other compartments following the cell perimeter, possibly the plasma membrane or tonoplast, as reported by Xiong et al. (2016) for MOB1 from Arabidopsis. A comparison of GFP-fusion localization results for full-length and truncated M4K revealed that the C-terminal domain is both necessary and sufficient for the

observed punctate localization of GFP. Moreover, the interaction between MOB1 and M4K in BiFC experiments resulted in localization resembling that of MOB1 alone rather than M4K alone, suggesting that MOB1 determines localization of interacting proteins. Alternatively, masking of motifs by MOB1 binding to the C-terminal portion of M4K or other interaction-dependent conformational changes or posttranslational modifications may allow motifs in the N-terminal region of M4K to determine localization. Finally, GFP-M4K was shown to specifically interact with GST-MOB1 fusion protein bound to glutathione sepharose beads. MOB1 interaction characteristics of SIK1 and M4K were similar, in that the kinase domain was not necessary for the interaction with MOB1, yet distinct, in that the interaction between SIK1 and MOB1 in plant cells resulted in a SIK1-like localization pattern, while the M4K–MOB1 interaction in plant cells led to a MOB1-like localization pattern. It will be interesting to discover through future research whether the M4K–MOB1 interaction plays an important role in regulating wood formation possibly through its reported ability to influence cell division (Xiong et al. 2016).

DRIF1, the protein with the most PPIs, had not been previously implicated in wood development. Interactions involving the SANT-MYB domain containing proteins RADIALIS (RAD) and DIVARICATA (DIV) with DIV- and RAD-Interacting Factor (DRIF) have been associated with dorsoventral asymmetry in Antirrhinum flowers (Raimundo et al. 2013) and homologous PPIs were linked to regulation of cell expansion during fruit development in tomato (Machemer et al. 2011). In consideration of the xylem-biased expression pattern of RAD1 (Rodgers-Melnick et al. 2012), we performed a Y2H screen with RAD1 as bait and identified a PPI with DRIF1. We then used DRIF1 as bait in a subsequent Y2H library screen and identified numerous interaction partners including RAD1 and multiple DIV homologs, suggesting that these conserved PPIs play various roles in plants depending on their developmental or tissue context. It is also possible that the DRIF-RAD-DIV module has been repeatedly co-opted to have taxon-specific roles. Moreover, DRIF1 formed novel interactions with numerous proteins showing highest expression values in xylem, phloem or the cambial zone (Figure 2). Among these new PPIs are three cytoskeletal-related proteins, including a member of the myosin binding protein (MyoB) family that controls cytoplasmic streaming, cell growth and possibly nucleocytoplasmic trafficking (Peremyslov et al. 2013, Kurth et al. 2017). DRIF1 also interacted with a putative Nuclear Division Cycle 80 (NDC80) protein, a component of the kinetochore (Du and Dawe 2007), and Kin7a, a member of the kinesin-7 family that regulates microtubule polymerization, cell polarity and chromosome separation (Moschou et al. 2016). Protein–protein interactions involving members of the homeodomain superfamily were also identified, including KNAT7, which regulates SCW biosynthesis in Arabidopsis and tobacco (Li et al. 2012, Pandey et al. 2016). Considered together, the various DRIF1 interactors suggest that this hub protein could be involved in protein complexes regulating cell division, cell expansion/shape and cell differentiation.

Wood transcription factors connect PPIs with activation of lignin biosynthesis genes

A transcription factor (TF) mini-library comprised of 38 PW TFs was used in a Y1H experiment to identify TFs that regulate SCW gene expression (see Figure S3 available as Supplementary Data at *Tree Physiology* Online; Figure 2). Upstream regions representing promoters for three genes encoding enzymes involved in lignin biosynthesis, phenylalanine ammonia-lyase 4 (PAL4), hydroxycinnamoyl transferase 6 (HCT6) and Caffeoyl CoA O-methyltransferase 1 (CCoAOMT1), were used as bait (see Table S2f available as Supplementary Data at *Tree Physiology* Online). Screening the TF mini-library against these promoters yielded 40 interactions for 20 different TF-AD fusion proteins. Of the 20 transcription factors that activated yeast reporters via *PAL4*, *HCT6* or *CCoAOMT1* promoters, six were also involved in PPIs (Figure 2). Similar to previous studies of Arabidopsis TF interactions with promoters of SCW biosynthesis genes (Zhong and Ye 2012, Taylor-Teeples et al. 2015), most of the PW TFs interacted with more than one promoter. We determined Bicor scores for co-expression of TFs yielding positive Y1H results and their target promoter gene products. All but two of the TF-promoter target pairs were positively correlated and Bicor scores for 55% of Y1H interactions were significant (P-value ≤ 0.05; only positive correlations were significant) (see Figure S3, Table S2e available as Supplementary Data at *Tree Physiology* Online).

Some PDIs were consistent with previously reported PDIs for poplar, such as MYB026 with the *CCoAOMT1* promoter (Zhong et al. 2011), and orthologous interactions in Arabidopsis, such as MYB46 (poplar MYB002 homolog) that directly activates *CCoAOMT1* and *PAL4* (Zhong and Ye 2012), and MYB103 (poplar MYB010 ortholog) that interacts with the *HCT* promoter (Taylor-Teeples et al. 2015). Although their PDIs identified here are new, poplar MYB092, MYB161 and MYB199 have previously been implicated in the regulation of SCW biosynthesis (Zhong et al. 2011). Specifically, they were upregulated in leaves of poplar overexpressing *WOOD-ASSOCIATED NAC DOMAIN 2B* (*WND2B*) that show ectopic SCW deposition. KNAT7 was reported to be a negative regulator of SCW biosynthesis in Arabidopsis (Li et al. 2012), but its tobacco ortholog was a positive regulator of SCW formation in tobacco (Pandey et al. 2016). The ability of poplar KNAT7 to activate reporters in Y1H assays is not necessarily inconsistent with a role as a negative regulator. KNAT7 alone is not an autoactivator in Y2H assays, indicating that its ability to activate reporters in Y1H depends on its fusion with the Gal4 AD. Protein–DNA interactions were also identified for two other homeodomain proteins, BEL1-like homeodomain (BLH)

and WUSCHEL related homeobox 13C (WOX13C). Consistent with the WOX13C PDIs, a genome-wide association study in *Populus deltoides* recently linked genetic variation in *WOX13C* with lignin percentage in wood (Fahrenkrog et al. 2017).

NAC TFs constitute one of the largest TF families in plants and function in a diverse array of processes; several NAC TFs are master regulators of lignin biosynthesis and secondary cell wall development (Petricka et al. 2012, Nakano et al. 2015, Jung Kim et al. 2016). NAC105 and NAC156 are members of the Secondary wall-associated NAC domain protein 2, 3 (SND2, 3) subgroup, while NAC127 is orthologous to Arabidopsis NAC075 (see Figure S3 available as Supplementary Data at *Tree Physiology* Online). In Arabidopsis, SND2 was shown to activate the *CESA8* promoter (Zhong et al. 2013) and affect fiber cell wall thickness (Zhong et al. 2008, Hussey et al. 2011), whereas NAC075 induced ectopic vessel-like cells (Endo et al. 2015). Three of the poplar NAC TFs homologous to SND2, SND3 or NAC075 interacted with the promoters tested here. Other TFs we found to interact with lignin pathway promoters that have not been previously linked to SCW biosynthesis, include RAD1, a B-box microprotein (BBX), a basic-leucine zipper (bZIP) protein and two C2H2-type zinc finger (C2H2ZF) proteins. C2H2ZF2 activates both the *PAL4* and *CCoAOMT1* promoters and *C2H2ZF2* expression is significantly and positively correlated with *PAL4* and *CCoAOMT1* expression (see Figure S3 available as Supplementary Data at *Tree Physiology* Online). C2H2ZF2 also has a PPI with ROP7a, suggesting that in addition to regulating SCW patterning (Oda and Fukuda 2012), ROP-GTPases could also have roles in regulating synthesis of SCW components. The aforementioned interactions support the key roles of MYB and NAC TFs in regulating SCW biosynthesis and identify new candidate regulators.

Conclusions

Most studies have approached interactome mapping via cloning of thousands of ORFs and performing binary Y2H screens (e.g., Arabidopsis Interactome Mapping Consortium 2011, Rolland et al. 2014). Here, we showed that a hybrid and more targeted approach can identify connected networks with new interactions associated with wood formation. Key to achieving this was a modest symmetric binary matrix guided by xylem-biased expression patterns followed by reiterative screening of a xylem library (interactome walking) for discovery of PPIs and using Y1H assays to link PW TFs to SCW biosynthesis and connect PPI networks to SCW PDIs. The wood interaction network can be expanded using the same approach and with advances in de novo DNA synthesis and decreasing costs (Kosuri and Church 2014), it may become more feasible to approach a wood proteome-scale PPI network via binary Y2H screens.

Supplementary Data

Supplementary Data for this article are available at *Tree Physiology* Online.

Acknowledgments

We thank Dr Marc Vidal for providing the pAD-Dest and pDB-Dest vectors, Dr Nathan Lawson for providing the 476-P5E MCS vector, and Dr. Aureliano Bombarely for advice on RNAseq analyses.

Conflict of interest

None declared.

Funding

This work was supported by the US Department of Energy, Office of Science, Office of Biological and Environmental Research (BER), Grant No. DE-FG02-07ER64449 to E.P.B. and A.M.B. Network assembly and RNAseq data analysis by S.B.R. was supported by US Department of Agriculture (USDA), National Institute of Food and Agriculture (NIFA), Grant No. 2014-67013-21580 to E.P.B. and A.M.B. Support was also provided by the Virginia Agricultural Experiment Station and the McIntire Stennis and Hatch Programs of USDA-NIFA, and the Virginia Tech Institute for Critical Technology and Applied Science.

References

Alvim Kamei CL, Boruc J, Vandepoele K, Van den Daele H, Maes S, Russinova E, Inze D, De Veylder L (2008) The PRA1 gene family in Arabidopsis. Plant Physiol 147:1735–1749.

Arabidopsis Interactome Mapping Consortium (2011) Evidence for network evolution in an *Arabidopsis* interactome map. Science 333: 601–607.

Bashline L, Li SD, Gu Y (2014) The trafficking of the cellulose synthase complex in higher plants. Ann Bot 114:1059–1067.

Brembu T, Winge P, Bones AM (2005) The small GTPase AtRAC2/ROP7 is specifically expressed during late stages of xylem differentiation in *Arabidopsis*. J Exp Bot 56:2465–2476.

Brunner AM, Yakovlev IA, Strauss SH (2004) Validating internal controls for quantitative plant gene expression studies. BMC Plant Biol 4:14.

Burstenbinder K, Savchenko T, Muller J, Adamson AW, Stamm G, Kwong R, Zipp BJ, Dinesh DC, Abel S (2013) *Arabidopsis* calmodulin-binding protein IQ67-domain 1 localizes to microtubules and interacts with kinesin light chain-related protein-1. J Biol Chem 288:1871–1882.

Chaffey N, Cholewa E, Regan S, Sundberg B (2002) Secondary xylem development in *Arabidopsis*: a model for wood formation. Physiol Plant 114:594–600.

Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, Robles M (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21:3674–3676.

Deplancke B, Vermeirssen V, Arda HE, Martinez NJ, Walhout AJ (2006) Gateway-compatible yeast one-hybrid screens. CSH Protocols doi: 10.1101/pdb.prot4590.

Du Y, Dawe RK (2007) Maize NDC80 is a constitutive feature of the central kinetochore. Chromosome Res 15:767–775.

Dunker AK, Cortese MS, Romero P, Iakoucheva LM, Uversky VN (2005) Flexible nets – the roles of intrinsic disorder in protein interaction networks. FEBS J 272:5129–5148.

Endo H, Yamaguchi M, Tamura T et al. (2015) Multiple classes of transcription factors regulate the expression of VASCULAR-RELATED NAC-DOMAIN7, a master switch of xylem vessel differentiation. Plant Cell Physiol 56:242–254.

Fahrenkrog AM, Neves LG, Resende MF Jr et al. (2017) Genome-wide association study reveals putative regulators of bioenergy traits in *Populus deltoides*. New Phytol 213:799–811.

Frasa MA, Koessmeier KT, Ahmadian MR, Braga VM (2012) Illuminating the functional and structural repertoire of human TBC/RABGAPs. Nat Rev Mol Cell Biol 13:67–73.

Gerttula S, Zinkgraf M, Muday GK et al. (2015) Transcriptional and hormonal regulation of gravitropism of woody stems in *Populus*. Plant Cell 27:2800–2813.

Goff LA, Trapnell C, Kelley D (2012) CummeRbund: visualization and exploration of Cufflinks high-throughput sequencing data R Package Version 2.2. http://compbio.mit.edu/cummeRbund/

Grant EH, Fujino T, Beers EP, Brunner AM (2010) Characterization of NAC domain transcription factors implicated in control of vascular cell differentiation in Arabidopsis and Populus. Planta 232:337–352.

Haigler CH, Brown RM (1985) The mechanisms of cellulose biosynthesis. Am J Bot 72:881–881.

Hussey SG, Mizrachi E, Spokevicius AV, Bossinger G, Berger DK, Myburg AA (2011) SND2, a NAC transcription factor gene, regulates genes involved in secondary cell wall development in *Arabidopsis* fibres and increases fibre cell area in *Eucalyptus*. BMC Plant Biol 11:173.

Immanen J, Nieminen K, Smolander OP et al. (2016) Cytokinin and auxin display distinct but interconnected distribution and signaling profiles to stimulate cambial activity. Curr Biol 26:1990–1997.

Jung Kim H, Gil Nam H, Ok Lim P (2016) Regulatory network of NAC transcription factors in leaf senescence. Curr Opin Plant Biol 33:9.

Kosuri S, Church GM (2014) Large-scale de novo DNA synthesis: technologies and applications. Nat Methods 11:499–507.

Kurth EG, Peremyslov VV, Turner HL, Makarova KS, Iranzo J, Mekhedov SL, Koonin EV, Dolja VV (2017) Myosin-driven transport network in plants. Proc Natl Acad Sci USA 114:E1385–E1394.

Langfelder P, Horvath S (2008) WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics 9:559.

Li C, Wang X, Ran L, Tian Q, Fan D, Luo K (2015) PtoMYB92 is a transcriptional activator of the lignin biosynthetic pathway during secondary cell wall formation in *Populus tomentos*a. Plant Cell Physiol 56:2436–2446.

Li E, Bhargava A, Qiang W et al. (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.

Liu L, Zinkgraf M, Petzold HE, Beers EP, Filkov V, Groover A (2015) The Populus ARBORKNOX1 homeodomain transcription factor regulates woody growth through binding to evolutionarily conserved target genes of diverse function. New Phytol 205:682–694.

Liu Z, Schneider R, Kesten C, Zhang Y, Somssich M, Zhang Y, Fernie AR, Persson S (2016) Cellulose-microtubule uncoupling proteins prevent lateral displacement of microtubules during cellulose synthesis in arabidopsis. Dev Cell 38:305–315.

Machemer K, Shaiman O, Salts Y, Shabtai S, Sobolev I, Belausov E, Grotewold E, Barg R (2011) Interplay of MYB factors in differential cell expansion, and consequences for tomato fruit development. Plant J 68:337–350.

Moschou PN, Gutierrez-Beltran E, Bozhkov PV, Smertenko A (2016) Separase promotes microtubule polymerization by activating CENP-E-related kinesin Kin7. Dev Cell 37:350–361.

Nakano Y, Yamaguchi M, Endo H, Rejab NA, Ohtani M (2015) NAC-MYB-based transcriptional regulation of secondary cell wall biosynthesis in land plants. Front Plant Sci 6:288.

Oda Y, Fukuda H (2012) Initiation of cell wall pattern by a Rho- and microtubule-driven symmetry breaking. Science 337:1333–1336.

Oda Y, Fukuda H (2013) Rho of plant GTPase signaling regulates the behavior of *Arabidopsis* kinesin-13A to establish secondary cell wall patterns. Plant Cell 25:4439–4450.

Pandey SK, Nookaraju A, Fujino T, Pattathil S, Joshi CP (2016) Virus-induced gene silencing (VIGS)-mediated functional characterization of two genes involved in lignocellulosic secondary cell wall formation. Plant Cell Rep 35:2353–2367.

Peremyslov VV, Morgun EA, Kurth EG, Makarova KS, Koonin EV, Dolja VV (2013) Identification of myosin XI receptors in *Arabidopsis* defines a distinct class of transport vesicles. Plant Cell 25:3022–3038.

Petricka J, Winter C, Benfey P (2012) Control of *Arabidopsis* root development. Annu Rev Plant Biol 63:30.

Porth I, Klapste J, Skyba O et al. (2013) Genome-wide association mapping for wood characteristics in *Populus* identifies an array of candidate single nucleotide polymorphisms. New Phytol 200:710–726.

Raimundo J, Sobral R, Bailey P, Azevedo H, Galego L, Almeida J, Coen E, Costa MM (2013) A subcellular tug of war involving three MYB-like proteins underlies a molecular antagonism in Antirrhinum flower asymmetry. Plant J 75:527–538.

Rawat SJ, Chernoff J (2015) Regulation of mammalian Ste20 (Mst) kinases. Trends Biochem Sci 40:149–156.

Robischon M, Du J, Miura E, Groover A (2011) The *Populus* class III HD ZIP, *popREVOLUTA*, influences cambium initiation and patterning of woody stems. Plant Physiol 155:1214–1225.

Rodgers-Melnick E, Mane SP, Dharmawardhana P, Slavov GT, Crasta OR, Strauss SH, Brunner AM, Difazio SP (2012) Contrasting patterns of evolution following whole genome versus tandem duplication events in *Populus*. Genome Res 22:95–105.

Rolland T, Tasan M, Charloteaux B et al. (2014) A proteome-scale map of the human interactome network. Cell 159:1212–1226.

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13:2498–2504.

Shi R, Wang JP, Lin YC, Li Q, Sun YH, Chen H, Sederoff RR, Chiang VL (2017) Tissue and cell-type co-expression networks of transcription factors and wood component genes in *Populus trichocarpa*. Planta 245:927–938.

Song J, Lu S, Chen ZZ, Lourenco R, Chiang VL (2006) Genetic transformation of *Populus trichocarpa* genotype Nisqually-1: a functional genomic tool for woody plants. Plant Cell Physiol 47:1582–1589.

Song L, Langfelder P, Horvath S (2012) Comparison of co-expression measures: mutual information, correlation, and model based indices. BMC Bioinformatics 13:21.

Tamura K, Iwabuchi K, Fukao Y, Kondo M, Okamoto K, Ueda H, Nishimura M, Hara-Nishimura I (2013) Myosin XI-i links the nuclear membrane to the cytoskeleton to control nuclear movement and shape in arabidopsis. Curr Biol 23:1776–1781.

Taylor-Teeples M, Lin L, de Lucas M et al. (2015) An *Arabidopsis* gene regulatory network for secondary cell wall synthesis. Nature 517:571–575.

Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol 28:511–515.

Tuskan GA, DiFazio S, Jansson S et al. (2006) The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). Science 313:1596–1604.

Vidal M, Fields S (2014) The yeast two-hybrid assay: still finding connections after 25 years. Nat Methods 11:1203–1206.

Vidal M, Cusick ME, Barabasi AL (2011) Interactome networks and human disease. Cell 144:986–998.

Walhout AJ, Vidal M (2001) High-throughput yeast two-hybrid assays for large-scale protein interaction mapping. Methods 24:297–306.

Walhout AJ, Temple GF, Brasch MA, Hartley JL, Lorson MA, van den Heuvel S, Vidal M (2000) GATEWAY recombinational cloning: application to the cloning of large numbers of open reading frames or ORFeomes. Methods Enzymol 328:575–592.

Wang PW, Richardson C, Hawkins TJ, Sparkes I, Hawes C, Hussey PJ (2016) Plant VAP27 proteins: domain characterization, intracellular localization and role in plant development. New Phytol 210:1311–1326.

Wightman R, Turner S (2010) Trafficking of the plant cellulose synthase complex. Plant Physiol 153:427–432.

Wu G, Gu Y, Li S, Yang Z (2001) A genome-wide analysis of Arabidopsis Rop-interactive CRIB motif-containing proteins that act as Rop GTPase targets. Plant Cell 13:2841–2856.

Xiong J, Cui X, Yuan X, Yu X, Sun J, Gong Q (2016) The Hippo/STE20 homolog SIK1 interacts with MOB1 to regulate cell proliferation and cell expansion in Arabidopsis. J Exp Bot 67:1461–1475.

Zawaski C, Kadmiel M, Ma C, Gai Y, Jiang X, Strauss SH, Busov VB (2011) SHORT INTERNODES-like genes regulate shoot growth and xylem proliferation in *Populus*. New Phytol 191:678–691.

Zhong R, Ye ZH (2012) MYB46 and MYB83 bind to the SMRE sites and directly activate a suite of transcription factors and secondary wall biosynthetic genes. Plant Cell Physiol 53:368–380.

Zhong R, Lee C, Zhou J, McCarthy RL, Ye ZH (2008) A battery of transcription factors involved in the regulation of secondary cell wall biosynthesis in *Arabidopsis*. Plant Cell 20:2763–2782.

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.

Zhong R, McCarthy RL, Haghighat M, Ye ZH (2013) The poplar MYB master switches bind to the SMRE site and activate the secondary wall biosynthetic program during wood formation. PLoS One 8: e69219.

Zinkgraf M, Liu L, Groover A, Filkov V (2017) Identifying gene coexpression networks underlying the dynamic regulation of wood-forming tissues in *Populus* under diverse environmental conditions. New Phytol 214:1464–1478.