Plant Science 181 (2011) 675–687

Contents lists available at ScienceDirect

Plant Science

journal homepage: www.elsevier.com/locate/plantsci

Identification of candidate genes in *Arabidopsis* and *Populus* cell wall biosynthesis using text-mining, co-expression network analysis and comparative genomics

Xiaohan Yang[a,*], Chu-Yu Ye[a], Anjali Bisaria[b], Gerald A. Tuskan[a], Udaya C. Kalluri[a,**]

[a] Biosciences Division and BioEnergy Science Center, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA
[b] Princeton University, Princeton, NJ 08544, USA

ARTICLE INFO

Article history:
Received 28 August 2010
Received in revised form 1 December 2010
Accepted 27 January 2011
Available online 17 February 2011

Keywords:
Gene expression
Cell wall
Biomass
Bioinformatics
Arabidopsis
Populus

ABSTRACT

Populus is an important bioenergy crop for bioethanol production. A greater understanding of cell wall biosynthesis processes is critical in reducing biomass recalcitrance, a major hindrance in efficient generation of biofuels from lignocellulosic biomass. Here, we report the identification of candidate cell wall biosynthesis genes through the development and application of a novel bioinformatics pipeline. As a first step, via text-mining of PubMed publications, we obtained 121 *Arabidopsis* genes that had the experimental evidence supporting their involvement in cell wall biosynthesis or remodeling. The 121 genes were then used as bait genes to query an *Arabidopsis* co-expression database, and additional genes were identified as neighbors of the bait genes in the network, increasing the number of genes to 548. The 548 *Arabidopsis* genes were then used to re-query the *Arabidopsis* co-expression database and re-construct a network that captured additional network neighbors, expanding to a total of 694 genes. The 694 *Arabidopsis* genes were computationally divided into 22 clusters. Queries of the *Populus* genome using the *Arabidopsis* genes revealed 817 *Populus* orthologs. Functional analysis of gene ontology and tissue-specific gene expression indicated that these *Arabidopsis* and *Populus* genes are high likelihood candidates for functional characterization in relation to cell wall biosynthesis.

1. Introduction

The genus *Populus*, representing the most productive temperate tree species in the world, has been selected as an important model woody crop for economic and ecological applications [1–3]. *Populus* species currently supply feedstocks for pulp and paper production, laminated veneer fabrication and an emerging renewable energy industry [1,4]. Since plant cell walls form the basis of this renewable resource [5], understanding the molecular basis of cell wall formation is critical for designing strategies to enhance desirable biomass properties. Based on protein sequence annotation, it is estimated that plants devote approximately 10% of their genome, about 2500 genes in *Arabidopsis*, to biosynthesis and rearrangement of cell walls [6,7]. However, the number of genes that have been experimentally confirmed to be involved in cell wall formation is very limited. The availability of high-quality gene co-expression networks in *Arabidopsis* [8] provides a new opportunity for genome-wide discovery of genes associated with cell wall biosynthesis. In this study, we developed a pipeline for large-scale identification of *Populus* genes involved in cell wall biosynthesis and rearrangement (Fig. 1).

Briefly, the PubMed database was queried to obtain research articles that documented experimental evidence for genes involved in cell wall biosynthesis in *Arabidopsis*, followed by curation of the *Arabidopsis* genes from the selected research articles. The curated *Arabidopsis* genes were then used as bait genes to query an *Arabidopsis* co-expression network to reveal additional *Arabidopsis* cell wall biosynthesis genes, which are associated with the bait genes. The bait genes and their network neighbors were pooled and a co-expression network was reconstructed. The *Arabidopsis* genes in the co-expression network were then computationally divided into clusters. Analyses of gene ontology and tissue-specific expression were performed to determine the functional features of each gene cluster. Candidate *Populus* genes, orthologous to the *Arabidopsis* genes, were identified based on a BlastP search. Finally, the validity of *Populus* candidate genes was tested by comparative analysis of gene expression between *Arabidopsis* and *Populus*. Using this pipeline, we identified several hundred *Populus* genes which are potentially involved in cell wall biosynthesis and rearrangement, providing high-value candidate genes for on-

* Corresponding author at: Biosciences Division, Oak Ridge National Laboratory, P.O. Box 2008 MS6422, Oak Ridge, TN, 37831-6422 USA. Tel.: +1 865 241 6895; fax: +1 865 576 9939.
** Corresponding author at: Biosciences Division, Oak Ridge National Laboratory, P.O. Box 2008 MS6422, Oak Ridge, TN, 37831-6422 USA. Tel.: +1 865 576 9495; fax: +1 865 576 9939.
E-mail addresses: yangx@ornl.gov (X. Yang), kalluriudayc@ornl.gov (U.C. Kalluri).

0168-9452/$ – see front matter © 2011 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.plantsci.2011.01.020

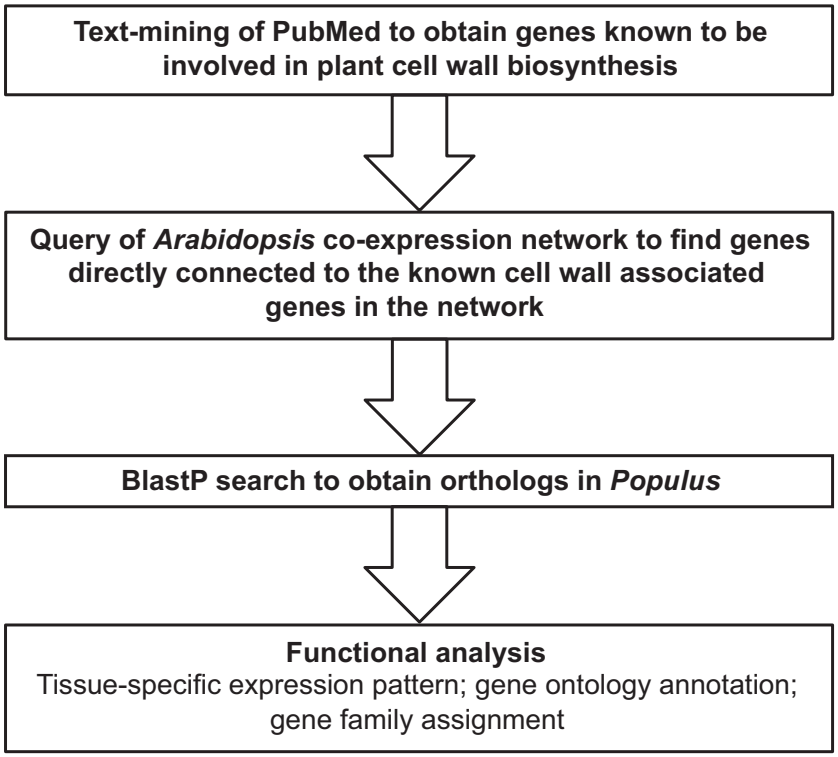

Fig. 1. A computational pipeline based on text mining and co-expression analysis for large-scale discovery of candidate genes involved in cell wall biosynthesis in *Populus*.

going bioenergy research on overcoming biomass recalcitrance in *Populus*.

2. Materials and methods

2.1. Text mining of literature

Abstracts of the articles related to cell wall genes were retrieved from the comprehensive PubMed by query with term "Cell wall AND gene" implemented in the EndNote software (Carlsbad, CA). After manual screening the approximately 8000 articles, the abstracts containing experimental evidence (i.e., native, over-expression or knockdown mutants) for genes involved in plant cell wall biosynthesis and reorganization were retained for extracting gene information to be used in network-querying in Section 2.2.

2.2. Co-expression analysis in Arabidopsis

The *Arabidopsis* genes in the manually curated abstracts/full-text articles that contain experimental evidence for cell wall biosynthesis genes were used as bait genes to query the ATTED-II database [8], and the genes connected directly to the bait genes in the *Arabidopsis* whole-genome co-expression network were obtained. The co-expression relationship between the bait genes and their directly-connected genes were used to re-construct a network of the expanded list of cell wall genes in *Arabidopsis* using ATTED-II [8] and the re-constructed co-expression network was visualized using Cytoscape [9]. The *Arabidopsis* genes in the re-constructed co-expression network were divided into clusters using the ClusterViz Cytoscape plugin [10].

2.3. Identification of orthologs in Populus

The protein sequences of the *Arabidopsis* cell wall associated genes were obtained from TAIR release 9 (http://www.arabidopsis.org/) and used to search the proteins sequences in *Populus* genome annotation v2.0 (http://www.phytozome.net/poplar) using BlastP [11] with e-value cutoff of 1×10^{-4}, the top hits, i.e., with 99–100% of the highest scores, were selected as *Populus* orthologs/co-orthologs. To investigate the conservation of gene expression between *Arabidopsis* and *Populus*, one-to-one *Arabidopsis* to *Populus* ortholog pairs as reciprocal best hits (RBH) in the two-way BlastP searches (i.e., *Arabidopsis* vs. *Populus* and *Populus* vs. *Arabidopsis*) were selected.

2.4. Gene ontology (GO)

GO annotation of the *Arabidopsis* genes was obtained from TAIR (http://www.arabidopsis.org/). GO enrichment was performed in agriGO (http://bioinfo.cau.edu.cn/agriGO/) [12] with default parameters using the whole *Arabidopsis* genome as the background/reference.

2.5. Analysis of gene expression pattern

The *Arabidopsis* microarray data were obtained from AtGenExpress (http://www.weigelworld.org/resources/microarray/AtGenExpress/AtGE_dev_gcRMA.txt.zip/view) [13]. The *Populus* expression data was obtained from PopGenIE (ftp://aspnas.fysbot.umu.se/v1_archive/eFP_data/) [14]. Heat maps were generated using R (http://www.r-project.org/). To detect the conserved gene expression pattern between *Arabidopsis* and *Populus* ortholog gene pairs (i.e., reciprocal best hits in two-way BlastP search), the expression data of 3 tissues (i.e., root, stem/xylem and leaf)

Table 1
Arabidopsis genes known to be involved in cell wall formation.

Gene ID	References	Gene ID	References	Gene ID	References
AT1G02730	[30]	AT1G80820	[31]	AT4G12350	[22,32]
AT1G03430	[33]	AT2G01850	[34]	AT4G15900	[35]
AT1G05850	[32]	AT2G03220	[36,37]	AT4G18780	[38–41]
AT1G09540	[32]	AT2G15370	[42]	AT4G22680	[22,43]
AT1G10550	[44]	AT2G15390	[42]	AT4G23920	[45]
AT1G12040	[46]	AT2G19800	[47]	AT4G28500	[22]
AT1G12840	[32]	AT2G20370	[35,36,48]	AT4G32410	[49–52]
AT1G13690	[53]	AT2G21140	[54]	AT4G33450	[22]
AT1G13980	[55]	AT2G21770	[49]	AT4G34230	[56,57]
AT1G14520	[47]	AT2G28110	[58,59]	AT4G36890	[58,60]
AT1G15950	[31,61,62]	AT2G32740	[63]	AT4G38770	[54]
AT1G16490	[64]	AT2G34140	[32]	AT4G39350	[49–51,65]
AT1G17950	[22]	AT2G35100	[66]	AT5G03170	[67,68]
AT1G19300	[58]	AT2G35620	[69]	AT5G05170	[49–51,70,71]
AT1G22620	[72]	AT2G35650	[73]	AT5G09870	[51]
AT1G27440	[21,60]	AT2G37090	[58,60,74]	AT5G10280	[32]
AT1G28470	[22]	AT2G38080	[41,64]	AT5G12870	[22,43,64]
AT1G30620	[35,75]	AT2G39770	[76]	AT5G16600	[22]
AT1G31420	[69]	AT2G40890	[77,78]	AT5G17420	[38–41,70,79–83]
AT1G32770	[22,43,84–87]	AT2G46770	[22,41,64,85,87,88]	AT5G22130	[89]
AT1G45130	[90]	AT3G02230	[91]	AT5G22940	[59]
AT1G56550	[92]	AT3G03050	[93]	AT5G26120	[94]
AT1G58370	[95]	AT3G10740	[94,96,97]	AT5G33290	[98]
AT1G62990	[22,43]	AT3G13870	[99]	AT5G39340	[33]
AT1G63910	[22]	AT3G13890	[41]	AT5G44030	[38–40]
AT1G64440	[45]	AT3G16360	[33,100]	AT5G47820	[101]
AT1G65580	[102]	AT3G19450	[57]	AT5G48930	[103]
AT1G66230	[22]	AT3G21510	[33]	AT5G49360	[104]
AT1G66240	[44]	AT3G25140	[105,106]	AT5G49720	[81,107]
AT1G66340	[69]	AT3G28180	[108]	AT5G54160	[61]
AT1G67490	[52]	AT3G29350	[33]	AT5G54380	[109]
AT1G68560	[97]	AT3G49690	[32]	AT5G54690	[41,58,68,110]
AT1G71930	[22,64]	AT3G51160	[35,111]	AT5G58600	[112]
AT1G73410	[22]	AT3G52840	[90]	AT5G61840	[21,60]
AT1G74380	[113]	AT3G54920	[112,114]	AT5G62220	[63]
AT1G75110	[115]	AT3G61910	[22,41,64,88]	AT5G62380	[22,64]
AT1G75120	[115]	AT3G62160	[53]	AT5G64530	[116]
AT1G78570	[46,117]	AT4G01750	[118]	AT5G64570	[97]
AT1G78580	[119]	AT4G01770	[118]	AT5G64740	[51,120]
AT1G79180	[64]	AT4G08150	[121]	AT5G66680	[122]
AT1G80350	[123,124]				

were selected and a Pearson product-moment correlation was analyzed.

2.6. Gene family assignment

The family assignment of the *Arabidopsis* protein-coding genes were obtained from the plant protein database GreenPhylDB [15].

3. Results

3.1. Genes known to be involved in cell wall biosynthesis

Approximately 8000 publications were retrieved from PubMed using "cell wall + gene" as the query. Among these papers, 159 research articles contained experimental data/evidence for genes involved in plant cell biosynthesis: 98 articles for *Arabidopsis* and 61 for other species (i.e., *Cicer*, *Eucalyptus*, *Gossypium*, *Hordeum*, *Lycopersicon*, *Medicago*, *Nicotiana*, *Oryza*, *Petunia*, *Picea*, *Pinus*, *Populus*, *Solanum*, *Sorghum* and *Zea*). Since the majority of the information was obtained from *Arabidopsis*, and this model plant species has the most abundant genomic resources available, downstream analysis was focused on this species. In the 98 articles related to *Arabidopsis*, 121 genes were characterized with experimental evidence as involved in cell wall biosynthesis (Table 1).

3.2. Co-expression network of Arabidopsis cell wall associated genes

We used the 121 acknowledged cell wall genes as bait genes to query a whole-genome co-expression network [8], and genes directly-connected to the bait genes were obtained, increasing the number of candidate genes associated with cell biosynthesis to 548. This list of 548 *Arabidopsis* genes was then used to reconstruct a co-expression network using ATTED-II [8], expanding the gene list to 694 genes (Supplementary Table 1) (692 genes having the protein sequences in the TAIR release 9). The 694 *Arabidopsis* genes in the network were further divided into 22 clusters using ClusterViz Cytoscape plugin (Fig. 2; Supplementary Fig. 1), with each cluster consisting of at least one bait gene (i.e., *Arabidopsis* genes known to be involved in cell wall formation) (Table 2; Supplementary Fig. 1). Some bait genes are likely to be hub genes, such as CESA7 in Cluster 02, MYB103 in Cluster 04, THE1 in Cluster 07 and SND3 in Cluster 11 (Fig. 2; Supplementary Fig. 1). More than 250 gene families were identified in the 692 protein-coding genes. About one-half of the 692 genes were distributed in 87 families, with seven families (i.e., kinase/LRR superfamily, glycoside hydrolase family, MYB family, glycosyl transferase family, cellulose synthase family, exostosin family and plastocyanin-like family) containing 10 or more gene members (Supplementary Table 2). In addition, 46 and 14 genes were classified as

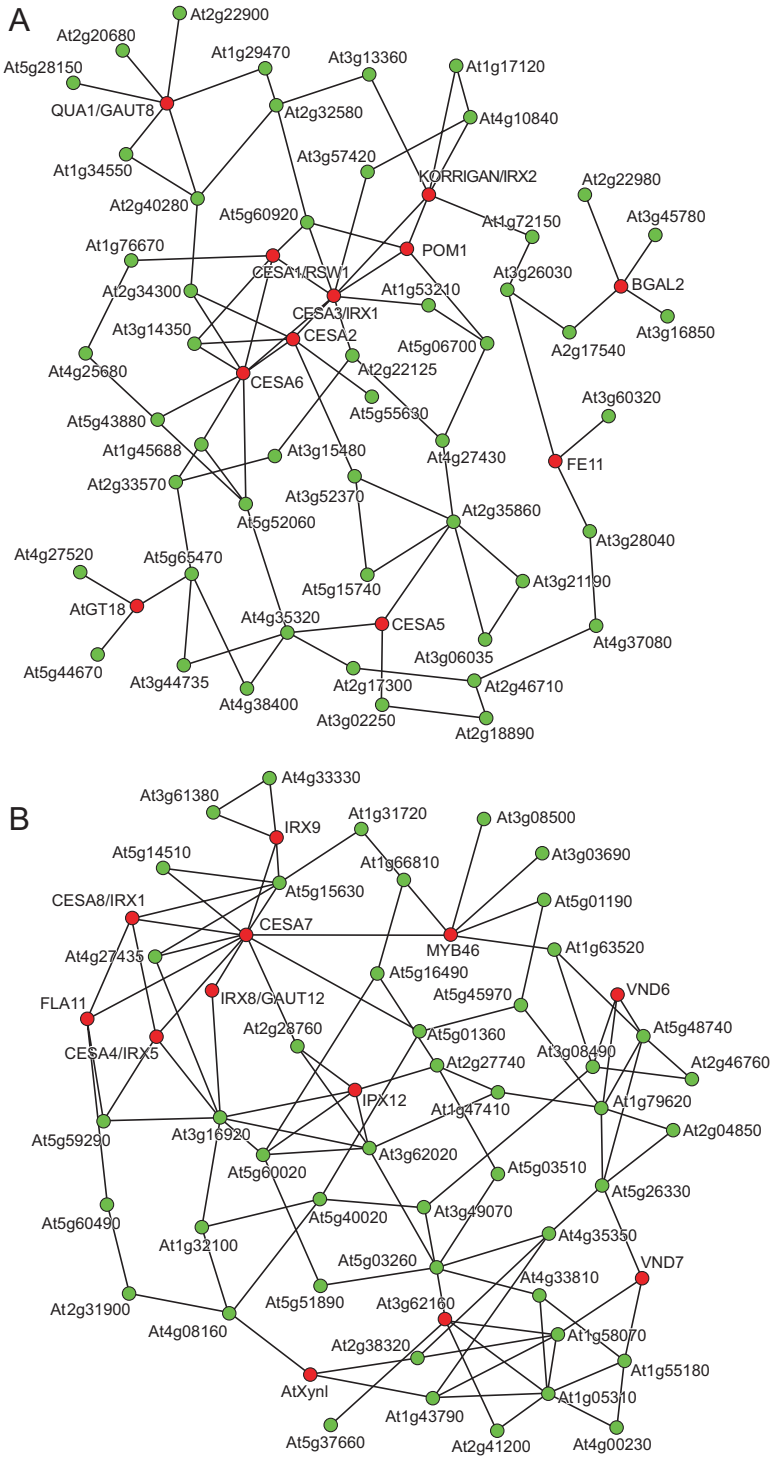

Fig. 2. Clusters C01 (A) and C02 (B) of a co-expression network containing *Arabidopsis* cell wall associated genes. Red dots represent the seed genes (i.e., the genes known to be involved in cell wall formation) listed in Table 1 and green dots represent the genes associated with the seed genes. The lines connecting two genes indicate that the two genes were co-expressed.

domain of unknown function (DUF) and hypothetical protein, respectively.

3.3. Gene ontology

Gene Ontology (GO) analysis was performed for the 692 *Arabidopsis* protein-coding genes in the co-expression network. Compared with the whole *Arabidopsis* genome annotation, biological processes related to cell wall biosynthesis were significantly enriched in the 692 *Arabidopsis* genes, including processes related to cell wall biogenesis, polysaccharide biosynthesis, secondary cell wall biogenesis, phenylpropanoid metabolism and cell wall organization (Supplementary Table 3). GO cellular component analysis revealed that several protein sub-cellular localizations (i.e., cell wall, membrane, Golgi apparatus and endoplasmic reticulum) were enriched in the co-expression network (Fig. 3). There was differ-

Table 2
Co-expression clusters of *the Arabidopsis* cell wall associated genes (known/bait genes and their nearest-neighbors in the co-expression network).

Cluster	Gene number	Bait gene number	GO enrichment ($P < 0.05$)	GO term level	Gene family number
C01	61	11	Primary cell wall biosynthetic process(GO:0009833) glucan metabolic process(GO:0006073)	8	41
C02	58	12	Secondary cell wall biosynthetic process(GO:0009834) xylem histogenesis(GO:0010089); xylan metabolic process(GO: 0045491)	8	31
C03	54	8	Glucan metabolic process(GO:0006073)	8	35
C04	53	9	NA	NA	30
C05	40	6	Lignin biosynthesis (GO: 0009809)	8	26
C06	37	6	Hormone-mediated signaling(GO:0009755)	6	27
C07	37	3	NA	NA	28
C08	35	4	Purine nucleoside triphosphate biosynthesis(GO:0009145) proton transport(GO:001 5992)	8	28
C09	33	5	NA	NA	25
C10	32	4	Actin cytoskeleton organization and biogenesis(GO: 0030036)	7	23
C11	30	7	Transcription, DNA-dependent(GO:0006351)	7	21
C12	28	5	Lipid metabolic process(GO: 0006629)	4	18
C13	25	5	Regulation of innate immune response(GO:0045088)	7	15
C14	25	4	NA	NA	16
C15	23	5	Cell wall biosynthetic process(GO:0042546)	6	18
C16	23	4	NA	NA	13
C17	22	2	NA	NA	15
C18	17	3	Carbohydrate metabolic process(GO:0005975) generation of precursor metabolites and energy(GO: 0006091)	4	11
C19	16	2	Protein amino acid glycosylation(GO: 0006486) polysaccharide biosynthetic process(GO:0000271) glucan metabolic process(GO:0006073)	8	8
C20	16	2	NA	NA	11
C21	15	1	Cellular morphogenesis (GO: 0000904)	6	11
C22	14	2	NA	NA	8

Table 3
Arabidopsis genes involved in secondary cell wall biogenesis and cell wall organization, which were selected based on GO annotation (biological process).

Secondary Cell wall Biogenesis			Cell wall organization		
Gene	Cluster	Bait gene	Gene	Cluster	Bait gene
AT1G43790	C02		AT1G31420	C01	Yes
AT2G37090	C02	Yes	AT4G38400	C01	
AT2G38080	C02	Yes	AT5G60920	C01	
AT4G18780	C02	Yes	AT1G05310	C02	
AT5G12870	C02	Yes	AT5G16490	C02	
AT5G15630	C02		AT5G17420	C02	Yes
AT5G17420	C02	Yes	AT5G54690	C02	Yes
AT5G44030	C02	Yes	AT1G10550	C03	Yes
AT1G27440	C04	Yes	AT1G19300	C03	Yes
AT1G32770	C04	Yes	AT3G59010	C04	
AT2G35700	C04		AT2G45220	C05	
AT3G61910	C04	Yes	AT3G10720	C07	
AT2G28110	C05	Yes	AT3G45970	C07	
AT5G61840	C06	Yes	AT4G26690	C07	
AT3G13890	C15	Yes	AT1G23200	C09	
			AT1G13980	C10	Yes
			AT5G47820	C10	Yes
			AT2G21140	C12	Yes
			AT3G54920	C12	Yes
			AT1G66340	C14	Yes
			AT5G33290	C15	Yes
			AT2G35620	C16	Yes
			AT5G63800	C18	
			AT3G08550	C19	
			AT5G66680	C19	Yes
			AT3G50410	C20	
			AT4G13390	C21	

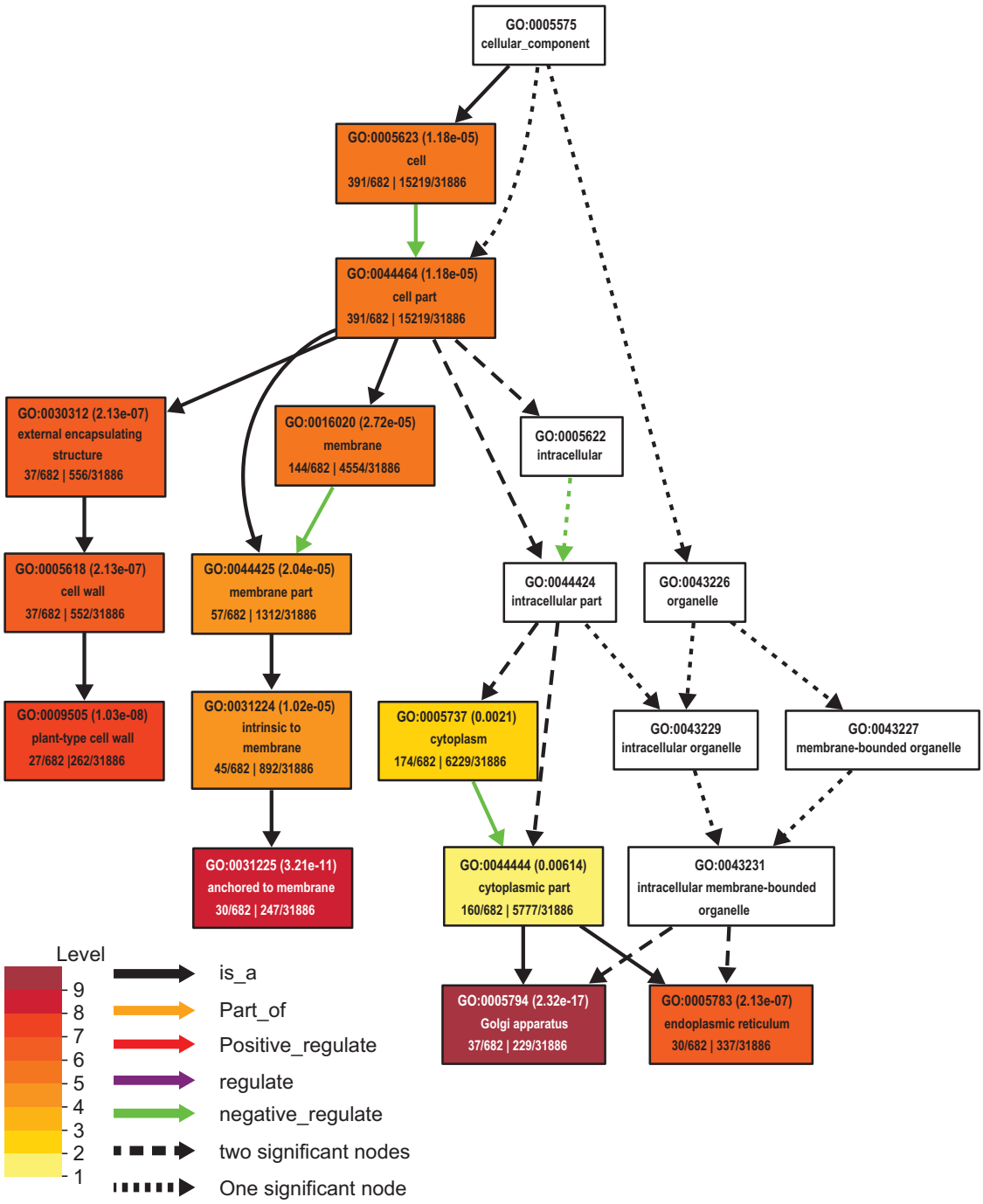

Fig. 3. GO (cellular component) enrichment for all cell wall associated genes detected in the co-expression network in Fig. 2.

ence in the enrichment of biological processes among the gene clusters. For example, the primary cell wall biosynthetic processes were enriched in Cluster01 (Table 2), secondary cell wall biosynthesis in Cluster02 (Table 2; Fig. 4), lignin biosynthesis in Cluster05 (Table 2), hormone-mediated signaling in Cluster06 (Table 2; Fig. 5), regulation of transcription in Cluster11 (Table 2) and regulation of innate immune response in Cluster13 (Table 2). Differences in the tissue-specific expression pattern were also revealed among the alternate GO term groups. For example, the expression patterns of the genes involved in secondary cell wall biogenesis are largely consistent, with most of the genes expressed preferentially in stem (i.e., the second internode counting from the bottom), whereas the genes involved in cell wall organization displayed diverse expression patterns (Table 3; Fig. 6).

3.4. *Populus* orthologs as candidate genes for cell wall formation

In order to obtain the candidate genes associated with cell wall formation in *Populus*, the 692 *Arabidopsis* protein-coding genes identified in this study were used to search the *Populus* genome annotation (V2.0) using BlastP. The top hits, i.e., those with 99–100% of the highest score, were selected as *Populus* orthologs/co-orthologs. 817 *Populus* orthologous genes were obtained as candidate genes involved in cell wall formation (Supplementary Table 1). Also, based on reciprocal BlastP search, using the *Populus* orthologs with the highest score from the 817-gene set as the query to search the *Arabidopsis* genome, 399 pairs of Reciprocal Best Hits (RBHs) were identified (Supplementary Table 1).

Fig. 4. GO (biological process) enrichment for the genes found in Cluster02 in Table 2.

3.5. *Conserved co-expression of cell wall associated genes between Arabidopsis and Populus*

We used similar gene expression patterns among the ortholog gene pairs in two different species to advocate similar functions. To test whether the tissue-specific expression pattern of cell wall associated genes is conserved between *Arabidopsis* and *Populus*, the microarray data were analyzed for 238 *Arabidopsis–Populus* RBH ortholog pairs, for which both *Arabidopsis* and *Populus* genes have expression data in AtGenExpress [13] and PopGenIE [14], respectively (Supplementary Table 4). In general, the tissue-specific expression pattern (leaf, stem and root) of the *Arabidopsis* cell wall associated genes was significantly correlated with that of their *Populus* homologs ($P < 1 \times 10^{-15}$). Furthermore, there was variation in the conservation of co-expression among the gene clusters. Specifically, nine clusters (i.e., C01–05, C08, C11, C16 and C20) showed significant ($P < 0.05$) correlation in gene expression between *Arabidopsis* and *Populus* (Table 4). In particular, there was a strong correlation ($r = 0.65$, $P < 1 \times 10^{-7}$) in expression of the cluster C02 genes between *Arabidopsis* and *Populus* (Table 4; Fig. 7).

4. Discussion

4.1. *An efficient approach for discovering candidate genes associated with cell wall biosynthesis*

A limited number of genes have been identified to be involved in cell wall biosynthesis in *Arabidopsis* (Table 1; references

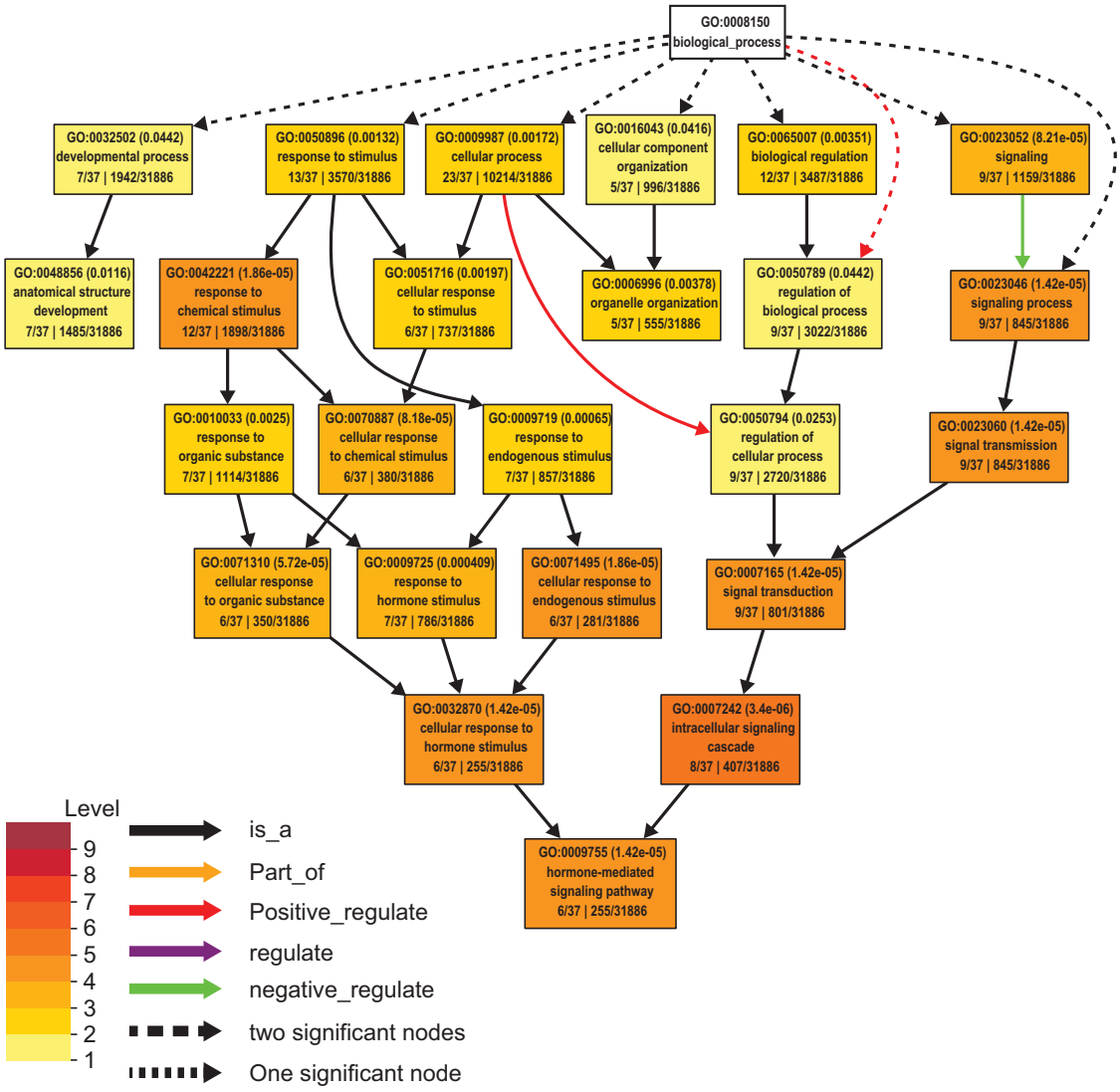

Fig. 5. GO (biological process) enrichment for the genes found in Cluster06 in Table 2.

therein). In this study, we identified more than 550 *Arabidopsis* candidate genes associated with cell wall biosynthesis using the 121 experimentally verified genes to query the *Arabidopsis* co-expression network built from microarray data. The GO annotations revealed that the additional genes are relevant to cell wall biosynthesis and organization (Supplementary Table 3). This demonstrates that querying genome-wide co-expression networks with known cell wall-related genes as bait can be an efficient approach for large-scale identification of additional genes involved in cell wall biosynthesis. In fact, the list of bait genes used for co-expression network query in this study was compiled from papers published through May 2009. Over the past year-and-a-half since then, 22 *Arabidopsis* genes involved in cell wall biosynthesis have been documented in literature with experimental evidence and eight of these genes were already included in our extended list of candidate cell wall biosynthesis genes in *Arabidopsis* (Supplementary Table 5), indicating that the prediction of our bioinformatic pipeline for genome-wide discovery of new genes is robust.

Based on protein sequences in the genome annotation, it was estimated that about 2500 *Arabidopsis* genes are potentially involved biosynthesis and rearrangement of cell walls [6,7]. However, this large list of genes (1) may not be informative regarding the co-ordination of gene subsets, (2) is difficult to characterize gene function in a reverse genetics context, and (3) includes members of large gene families. Moreover, it is well known that the members of large gene families have diverse functions due to subtle but significant variation in both promoter (determining gene expression pattern) and protein-coding (defining protein functions) regions [1,16]. As shown in this study, a number of genes from multiple gene families can work together in a single biological process (Table 2). Therefore, gene co-expression network analysis is complementary to gene family annotation, providing more detailed information about the roles of the candidate genes in biological processes.

The candidate gene discovery strategy used in the study should be considered an on-going dynamic effort. As future transcriptome-sequencing efforts occur, new co-expression network analyses will identify additional genes potentially involved in cell wall biosynthesis. As additional cell wall-related genes are experimentally identified, additional candidate genes could be identified from the next-generation co-expression network analysis.

4.2. Genes with diverse family background involved in cell wall formation

In this study, more than 250 gene families were identified in the *Arabidopsis* protein-coding genes associated with cell wall biosyn-

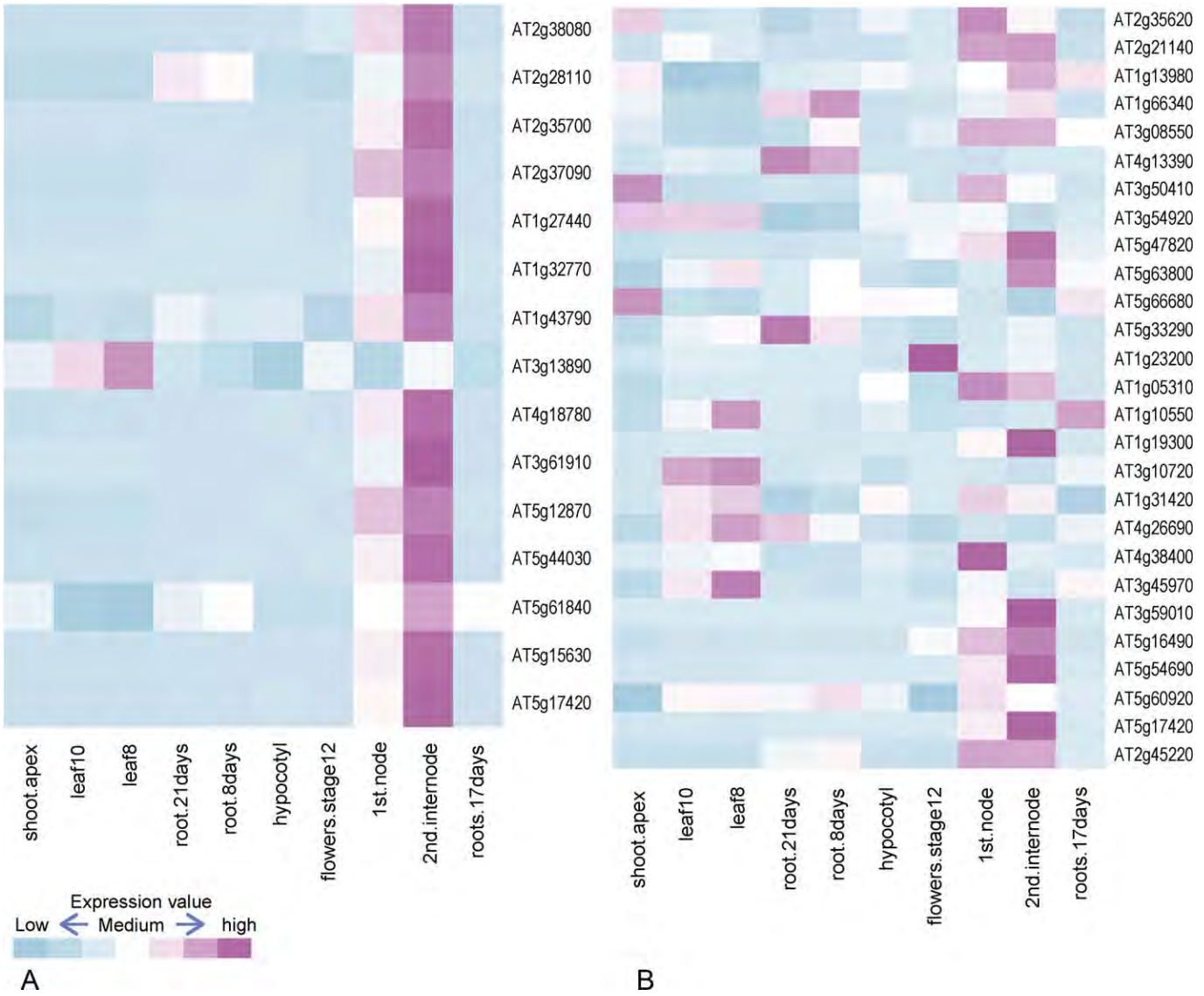

Fig. 6. Tissue-specific expression pattern (based on microarray data) of *Arabidopsis* genes involved in secondary cell wall biogenesis (A) and cell wall organization (B). The genes were selected based on GO annotation (biological process).

thesis and reorganization. Among these gene families, five (i.e., kinase/LRR superfamily, glycoside hydrolase family, MYB family, glycosyl transferase family and cellulose synthase family) were the most abundant, containing 12 or more gene members identified in this study (Supplementary Table 2), and suggesting that members of these gene families play a key role in cell wall biosynthesis and reorganization. This is consistent with previous studies [17–23]. More importantly, 46 and 14 genes, classified as domain of unknown function (DUF) and hypothetical protein, respectively, were discovered as a result of our approach. These genes are candidates for further characterization and experimental validation.

4.3. Conservation of gene expression is associated with functional categories

It is hypothesized that genes with essential functions (e.g., transcription factors) are disproportionately retained following speciation [24–28]. This hypothesis is supported by our comparative analysis of gene expression between *Arabidopsis* and *Populus* (Tables 2 and 4; Fig. 7). For example, the co-expression patterns of genes in nine clusters (C01–05, C08, C11, C16 and C20) are significantly conserved between *Arabidopsis* and *Populus* (Table 4). The clusters showing conserved co-expression between the two species are involved in biological processes essential for cell wall formation, including primary cell wall biosynthesis (Cluster01), secondary cell wall biosynthesis (Cluster02), glucan metabolic process (Cluster03), lignin biosynthesis (Cluster05), and DNA-dependent transcription (Cluster11) (Table 2). The *Populus* orthologs corresponding to the conserved co-expression clusters should be considered high confidence candidates involved in cell wall biosynthesis. One possible evolutionary mechanism leading to non-conserved co-expression in the other 12 clusters is that orthologs were lost via small deletions or pseudonization in one species [16]. Alternatively, some of the genes may have experienced rapid lineage-specific expansion, as demonstrated in the F-box gene family in herbaceous annual plants, which would not be found in woody perennial plants [29].

Fig. 7. Expression pattern (based on microarray data) of *Arabidopsis* Cluster02 genes and the corresponding *Populus* orthologs (Supplementary Table 1) which were selected based on reciprocal best hits (RBH) in two-way BlastP searches (i.e., *Arabidopsis* vs. *Populus* and *Populus* vs. *Arabidopsis*).

4.4. Differential expression between cell wall biosynthesis and cell wall organization

Both cell wall biosynthesis and cell wall organization are important biological processes contributing to the cell wall formation. This study revealed differences in the tissue-specific expression pattern between the genes involved in secondary cell wall biogenesis and the genes involved in cell wall organization, with the cell wall biogenesis gene category showing consistent preferential expression in the second internode whereas the cell wall organization gene category displaying diverse expression patterns (Table 3; Fig. 6). The second internode is the principal location of secondary cell wall formation and maturation in the stems of *Arabidopsis*. Despite the fact that most plant cell types share a common pathway to generate basic polymers (e.g., cellulose, lignin, etc.) for building the cell wall, our analysis detected significantly higher expression of the majority of the cell wall biogenesis genes in the second internode. More expectedly, the genes identified by our network analysis associated with cell wall organization varied among the different tissues. It can be hypothesized then that manipulating of genes involved in cell wall organization may hold great promise for reducing the biomass recalcitrance in targeted

Table 4
Correlation in gene expression pattern between *Arabidopsis* (leaf, stem and root) and *Populus* (leaf, xylem and root).

Cluster	Pearson correlation (*r*)	*P*-value	Gene number
C01–C22**	0.40	2.20E−16	238
C01*	0.28	2.42E−02	22
C02**	0.65	8.66E−08	18
C03**	0.35	3.55E−03	22
C04**	0.62	9.99E−06	14
C05**	0.49	4.58E−05	21
C06	0.17	2.36E−01	17
C07	0.10	5.49E−01	12
C08*	0.40	1.43E−02	12
C09	0.31	2.15E−01	6
C10	0.19	4.58E−01	6
C11**	0.58	4.64E−04	11
C12	0.20	1.81E−01	16
C13	0.07	7.42E−01	8
C14	0.08	7.40E−01	6
C15	0.59	9.22E−02	3
C16*	0.43	2.60E−02	9
C17	−0.33	7.50E−02	10
C18	0.18	3.97E−01	8
C19	0.11	5.70E−01	9
C20**	0.62	6.32E−03	6
C21	−0.86	3.47E−01	1
C22	−0.43	7.17E−01	1

* Significance at $P < 0.05$.
** Significance at $P < 0.01$.

tissues, independent of modifications in cell wall biosynthesis genes.

Acknowledgments

We would like to thank T.J. Tschaplinski and T. Li for thoughtful and insightful comments on the manuscript. This research was supported by the U.S. DOE BioEnergy Science Center. The BioEnergy Science Center is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science. Oak Ridge National Laboratory is managed by UT-Battelle, LLC for the U.S. Department of Energy under Contract Number DE-AC05-00OR22725.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.plantsci.2011.01.020.

References

[1] X. Yang, U.C. Kalluri, S.P. DiFazio, S.D. Wullschleger, T.J. Tschaplinski, Z.M. Cheng, G.A. Tuskan, Poplar genomics: state of the science, Crit. Rev. Plant Sci. 28 (2009) 285–308.
[2] T.G. Whitham, S.P. DiFazio, J.A. Schweitzer, S.M. Shuster, G.J. Allan, J.K. Bailey, S.A. Woolbright, Perspective—extending genomics to natural communities and ecosystems, Science 320 (2008) 492–495.
[3] G.A. Tuskan, S.P. DiFazio, T. Teichmann, Poplar genomics is getting popular: the impact of the poplar genome project on tree research, Plant Biol. 6 (2004) 2–4.
[4] G. Tuskan, Short-rotation woody crop supply systems in the United States: what do we know and what do we need to know? Biomass Bioenergy 14 (1998) 307–315.
[5] A.H. Liepman, R. Wightman, N. Geshi, S.R. Turner, H.V. Scheller, *Arabidopsis*—a powerful model system for plant cell wall research, Plant J. 61 (2010) 1107–1121.
[6] M.C. McCann, N.C. Carpita, Designing the deconstruction of plant cell walls, Curr. Opin. Plant Biol. 11 (2008) 314–320.
[7] W.D. Yong, B. Link, R. O'Malley, J. Tewari, C.T. Hunter, C.A. Lu, X.M. Li, A.B. Bleecker, K.E. Koch, M.C. McCann, et al., Genomics of plant cell wall biogenesis, Planta 221 (2005) 747–751.
[8] T. Obayashi, S. Hayashi, M. Saeki, H. Ohta, K. Kinoshita, ATTED-II provides coexpressed gene networks for *Arabidopsis*, Nucleic Acids Res. 37 (2009) D987–991.
[9] P. Shannon, A. Markiel, O. Ozier, N.S. Baliga, J.T. Wang, D. Ramage, N. Amin, B. Schwikowski, T. Ideker, Cytoscape: a software environment for integrated models of biomolecular interaction networks, Genome Res. 13 (2003) 2498–2504.
[10] J. Cai, G. Chen, J. Wang, ClusterViz: A Cytoscape Plugin for Graph Clustering and Visualization, School of Information Science and Engineering, Central South University, Changsha, China, 2010.
[11] S.F. Altschul, T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, W. Miller, D.J. Lipman, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, Nucleic Acids Res. 25 (1997) 3389–3402.
[12] Z. Du, X. Zhou, Y. Ling, Z. Zhang, Z. Su, agriGO: a GO analysis toolkit for the agricultural community, Nucleic Acids Res. 38 (Suppl.) (2010) W64–70.
[13] M. Schmid, T.S. Davison, S.R. Henz, U.J. Pape, M. Demar, M. Vingron, B. Scholkopf, D. Weigel, J.U. Lohmann, A gene expression map of *Arabidopsis thaliana* development, Nat. Genet. 37 (2005) 501–506.
[14] A. Sjodin, N.R. Street, G. Sandberg, P. Gustafsson, S. Jansson, The *Populus* genome integrative explorer (PopGenIE): a new resource for exploring the *Populus* genome, New Phytol. (2009).
[15] M.G. Conte, S. Gaillard, N. Lanau, M. Rouard, C. Perin, GreenPhylDB: a database for plant comparative genomics, Nucleic Acids Res. 36 (2008) D991–998.
[16] X. Yang, G.A. Tuskan, Z.M. Cheng, Divergence of the Dof gene families in poplar, *Arabidopsis*, and rice suggests multiple modes of gene evolution after duplication, Plant Physiol. 142 (2006) 820–830.
[17] B.J Steinwand, J.J. Kieber, The role of receptor-like kinases in regulating cell wall function, Plant Physiol. 153 (2010) 479–484.
[18] C. Dubos, R. Stracke, E. Grotewold, B. Weisshaar, C. Martin, L. Lepiniec, MYB transcription factors in *Arabidopsis*, Trends Plant Sci. (2010).
[19] H.J. Gilbert, The biochemistry and structural biology of plant cell wall deconstruction, Plant Physiol. 153 (2010) 444–455.
[20] Y. Yin, H. Chen, M.G. Hahn, D. Mohnen, Y. Xu, Evolution and function of the plant cell wall synthesis-related glycosyltransferase family 8, Plant Physiol. 153 (2010) 1729–1746.
[21] A.M. Wu, C. Rihouey, M. Seveno, E. Hornblad, S.K. Singh, T. Matsunaga, T. Ishii, P. Lerouge, A. Marchant, The *Arabidopsis* IRX10 and IRX10-LIKE glycosyltransferases are critical for glucuronoxylan biosynthesis during secondary cell wall formation, Plant J. 57 (2009) 718–731.
[22] R. Zhong, C. Lee, J. Zhou, R.L. McCarthy, Z.H. Ye, A battery of transcription factors involved in the regulation of secondary cell wall biosynthesis in *Arabidopsis*, Plant Cell 20 (2008) 2763–2782.
[23] R. Wightman, S. Turner, Trafficking of the plant cellulose synthase complex, Plant Physiol. 153 (2010) 427–432.
[24] J.A. Birchler, R.A. Veitia, The gene balance hypothesis: implications for gene regulation, quantitative traits and evolution, New Phytol. 186 (2010) 54–62.
[25] G Blanc, K.H. Wolfe, Functional divergence of duplicated genes formed by polyploidy during *Arabidopsis* evolution, Plant Cell 16 (2004) 1679–1691.
[26] P.P. Edger, J.C. Pires, Gene and genome duplications: the impact of dosage-sensitivity on the fate of nuclear genes, Chromosome Res. 17 (2009) 699–717.
[27] U.C. Kalluri, S.P. Difazio, A.M. Brunner, G.A. Tuskan, Genome-wide analysis of Aux/IAA and ARF gene families in *Populus trichocarpa*, BMC Plant Biol. 7 (2007) 59.
[28] G.A. Tuskan, S. Difazio, S. Jansson, J. Bohlmann, I. Grigoriev, U. Hellsten, N. Putnam, S. Ralph, S. Rombauts, A. Salamov, et al., The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray), Science 313 (2006) 1596–1604.
[29] X. Yang, U.C. Kalluri, S. Jawdy, L.E. Gunter, T.M. Yin, T.J. Tschaplinski, D.J. Weston, P. Ranjan, G.A. Tuskan, The F-box gene family is expanded in herbaceous annual plants relative to woody perennial plants, Plant Physiol. 148 (2008) 1189–1200.
[30] A.J. Bernal, J.K. Jensen, J. Harholt, S. Sorensen, I. Moller, C. Blaukopf, B. Johansen, R. de Lotto, M. Pauly, H.V. Scheller, et al., Disruption of *ATCSLD5* results in reduced growth, reduced xylan and homogalacturonan synthase activity and altered xylan occurrence in *Arabidopsis*, Plant J. 52 (2007) 791–802.
[31] M. Mir Derikvand, J.B. Sierra, K. Ruel, B. Pollet, C.T. Do, J. Thevenin, D. Buffard, L. Jouanin, C. Lapierre, Redirection of the phenylpropanoid pathway to feruloyl malate in *Arabidopsis* mutants deficient for cinnamoyl-CoA reductase 1, Planta 227 (2008) 943–956.
[32] L.A. Rogers, C. Dubos, C. Surman, J. Willment, I.F. Cullis, S.D. Mansfield, M.M. Campbell, Comparison of lignin deposition in three ectopic lignification mutants, New Phytol. 168 (2005) 123–140.
[33] C.E. Hutchison, J. Li, C. Argueso, M. Gonzalez, E. Lee, M.W. Lewis, B.B. Maxwell, T.D. Perdue, G.E. Schaller, J.M. Alonso, et al., The *Arabidopsis* histidine phosphotransfer proteins are redundant positive regulators of cytokinin signaling, Plant Cell 18 (2006) 3073–3087.
[34] A. Matsui, R. Yokoyama, M. Seki, T. Ito, K. Shinozaki, T. Takahashi, Y. Komeda, K. Nishitani, *AtXTH27* plays an essential role in cell wall modification during the development of tracheary elements, Plant J. 42 (2005) 525–534.
[35] Y. Li, C. Smith, F. Corke, L. Zheng, Z. Merali, P. Ryden, P. Derbyshire, K. Waldron, M.W. Bevan, Signaling from an altered cell wall to the nucleus mediates sugar-responsive growth and development in *Arabidopsis thaliana*, Plant Cell 19 (2007) 2500–2515.
[36] P. Ryden, K. Sugimoto-Shirasu, A.C. Smith, K. Findlay, W.D. Reiter, M.C. McCann, Tensile properties of *Arabidopsis* cell walls depend on both a xyloglucan cross-linked microfibrillar network and rhamnogalacturonan II-borate complexes, Plant Physiol. 132 (2003) 1033–1040.
[37] G.F. Vanzin, M. Madson, N.C. Carpita, N.V. Raikhel, K. Keegstra, W.D. Reiter, The *mur2* mutant of *Arabidopsis thaliana* lacks fucosylated xyloglucan because

of a lesion in fucosyltransferase AtFUT1, Proc. Natl. Acad. Sci. U.S.A. 99 (2002) 3340–3345.
[38] I.I. Atanassov, J.K. Pittman, S.R. Turner, Elucidating the mechanisms of assembly and subunit interaction of the cellulose synthase complex of *Arabidopsis* secondary cell walls, J. Biol. Chem. 284 (2009) 3833–3841.
[39] J.C. Gardiner, N.G. Taylor, S.R. Turner, Control of cellulose synthase complex localization in developing xylem, Plant Cell 15 (2003) 1740–1748.
[40] N.G. Taylor, R.M. Howells, A.K. Huttly, K. Vickers, S.R. Turner, Interactions among three distinct CesA proteins essential for cellulose synthesis, Proc. Natl. Acad. Sci. U.S.A. 100 (2003) 1450–1455.
[41] C. Yang, Z. Xu, J. Song, K. Conner, G. Vizcay Barrena, Z.A. Wilson, *Arabidopsis MYB26/MALE STERILE35* regulates secondary thickening in the endothecium and is essential for anther dehiscence, Plant Cell 19 (2007) 534–548.
[42] R. Sarria, T.A. Wagner, M.A. O'Neill, A. Faik, C.G. Wilkerson, K. Keegstra, N.V. Raikhel, Characterization of a family of *Arabidopsis* genes related to xyloglucan fucosyltransferase1, Plant Physiol. 127 (2001) 1595–1606.
[43] R. Zhong, E.A. Richardson, Z.H. Ye, The MYB46 transcription factor is a direct target of SND1 and regulates secondary wall biosynthesis in *Arabidopsis*, Plant Cell 19 (2007) 2776–2792.
[44] I. Ndamukong, A. Chetram, A. Saleh, Z. Avramova, Wall-modifying genes regulated by the *Arabidopsis* homolog of trithorax, ATX1: repression of the *XTH33* gene as a test case, Plant J. (2009).
[45] J. Rosti, C.J. Barton, S. Albrecht, P. Dupree, M. Pauly, K. Findlay, K. Roberts, G.J. Seifert, UDP-glucose 4-epimerase isoforms UGE2 and UGE4 cooperate in providing UDP-galactose for cell wall biosynthesis and growth of *Arabidopsis thaliana*, Plant Cell 19 (2007) 1565–1579.
[46] A. Diet, B. Link, G.J. Seifert, B. Schellenberg, U. Wagner, M. Pauly, W.D. Reiter, C. Ringli, The *Arabidopsis* root hair cell wall formation mutant *lrx1* is suppressed by mutations in the *RHM1* gene encoding a UDP-L-rhamnose synthase, Plant Cell 18 (2006) 1630–1641.
[47] U. Kanter, B. Usadel, F. Guerineau, Y. Li, M. Pauly, R. Tenhaken, The inositol oxygenase gene family of *Arabidopsis* is involved in the biosynthesis of nucleotide sugar precursors for cell-wall matrix polysaccharides, Planta 221 (2005) 243–254.
[48] M. Madson, C. Dunand, X. Li, R. Verma, G.F. Vanzin, J. Caplan, D.A. Shoue, N.C. Carpita, W.D. Reiter, The *MUR3* gene of *Arabidopsis* encodes a xyloglucan galactosyltransferase that is evolutionarily related to animal exostosins, Plant Cell 15 (2003) 1662–1670.
[49] T. Beeckman, G.K. Przemeck, G. Stamatiou, R. Lau, N. Terryn, R. De Rycke, D. Inze, T. Berleth, Genetic complexity of cellulose synthase a gene function in *Arabidopsis* embryogenesis, Plant Physiol. 130 (2002) 1883–1893.
[50] J.E. Burn, C.H. Hocart, R.J. Birch, A.C. Cork, R.E. Williamson, Functional analysis of the cellulose synthase genes *CesA1*, *CesA2*, and *CesA3* in *Arabidopsis*, Plant Physiol. 129 (2002) 797–807.
[51] T. Desprez, M. Juraniec, E.F. Crowell, H. Jouy, Z. Pochylova, F. Parcy, H. Hofte, M. Gonneau, S. Vernhettes, Organization of cellulose synthase complexes involved in primary cell wall synthesis in *Arabidopsis thaliana*, Proc. Natl. Acad. Sci. U.S.A. 104 (2007) 15572–15577.
[52] C.S. Gillmor, P. Poindexter, J. Lorieau, M.M. Palcic, C. Somerville, Alpha-glucosidase I is required for cellulose biosynthesis and morphogenesis in *Arabidopsis*, J. Cell Biol. 156 (2002) 1003–1013.
[53] S. Sawa, T. Demura, G. Horiguchi, M. Kubo, H. Fukuda, The *ATE* genes are responsible for repression of transdifferentiation into xylem cells in *Arabidopsis*, Plant Physiol. 137 (2005) 141–148.
[54] T.J. Fowler, C. Bernhardt, M.L. Tierney, Characterization and expression of four proline-rich cell wall protein genes in *Arabidopsis* encoding two distinct subsets of multiple domain proteins, Plant Physiol. 121 (1999) 1081–1092.
[55] D.E. Shevell, T. Kunkel, N.H. Chua, Cell wall alterations in the *Arabidopsis emb30* mutant, Plant Cell 12 (2000) 2047–2060.
[56] A.P. Chen, N.Q. Zhong, Z.L. Qu, F. Wang, N. Liu, G.X. Xia, Root and vascular tissue-specific expression of glycine-rich protein AtGRP9 and its interaction with AtCAD5, a cinnamyl alcohol dehydrogenase, in *Arabidopsis thaliana*, J. Plant Res. 120 (2007) 337–343.
[57] R. Sibout, A. Eudes, G. Mouille, B. Pollet, C. Lapierre, L. Jouanin, A. Seguin, *CINNAMYL ALCOHOL DEHYDROGENASE-C* and *-D* are the primary genes involved in lignin biosynthesis in the floral stem of *Arabidopsis*, Plant Cell 17 (2005) 2059–2076.
[58] D.M. Brown, F. Goubet, V.W. Wong, R. Goodacre, E. Stephens, P. Dupree, S.R. Turner, Comparison of five xylan synthesis mutants reveals new insight into the mechanisms of xylan synthesis, Plant J. 52 (2007) 1154–1168.
[59] C. Lee, Q. Teng, W. Huang, R. Zhong, Z.H. Ye, The F8H glycosyltransferase is a functional paralog of FRA8 involved in glucuronoxylan biosynthesis in *Arabidopsis*, Plant Cell Physiol. (2009).
[60] D.M. Brown, Z. Zhang, E. Stephens, P. Dupree, S.R. Turner, Characterization of IRX10 and IRX10-like reveals an essential role in glucuronoxylan biosynthesis in *Arabidopsis*, Plant J. 57 (2009) 732–746.
[61] T. Goujon, V. Ferret, I. Mila, B. Pollet, K. Ruel, V. Burlat, J.P. Joseleau, Y. Barriere, C. Lapierre, L. Jouanin, Down-regulation of the *AtCCR1* gene in *Arabidopsis thaliana*: effects on phenotype, lignins and cell wall degradability, Planta 217 (2003) 218–228.
[62] L. Jones, A.R. Ennos, S.R. Turner, Cloning and characterization of *irregular xylem4* (*irx4*): a severely lignin-deficient mutant of *Arabidopsis*, Plant J. 26 (2001) 205–216.
[63] X. Li, I. Cordero, J. Caplan, M. Molhoj, W.D. Reiter, Molecular analysis of 10 coding regions from *Arabidopsis* that are homologous to the MUR3 xyloglucan galactosyltransferase, Plant Physiol. 134 (2004) 940–950.
[64] J. Zhou, C. Lee, R. Zhong, Z.H. Ye, MYB58 and MYB63 are transcriptional activators of the lignin biosynthetic pathway during secondary cell wall formation in *Arabidopsis*, Plant Cell 21 (2009) 248–266.
[65] Z. Chu, H. Chen, Y. Zhang, Z. Zhang, N. Zheng, B. Yin, H. Yan, L. Zhu, X. Zhao, M. Yuan, et al., Knockout of the *AtCESA2* gene affects microtubule orientation and causes abnormal cell expansion in *Arabidopsis*, Plant Physiol. 143 (2007) 213–224.
[66] J. Harholt, J.K. Jensen, S.O. Sorensen, C. Orfila, M. Pauly, H.V. Scheller, ARABINAN DEFICIENT 1 is a putative arabinosyltransferase involved in biosynthesis of pectic arabinan in *Arabidopsis*, Plant Physiol. 140 (2006) 49–58.
[67] S. Ito, Y. Suzuki, K. Miyamoto, J. Ueda, I. Yamaguchi, AtFLA11, a fasciclin-like arabinogalactan–protein, specifically localized in sclerenchyma cells, Biosci. Biotechnol. Biochem. 69 (2005) 1963–1969.
[68] S Persson, H. Wei, J. Milne, G.P. Page, C.R. Somerville, Identification of genes required for cellulose synthesis by regression analysis of public microarray data sets, Proc. Natl. Acad. Sci. U.S.A. 102 (2005) 8633–8638.
[69] S.L. Xu, A. Rahman, T.I. Baskin, J.J. Kieber, Two leucine-rich repeat receptor kinases mediate signaling, linking cell wall biosynthesis and ACC synthase in *Arabidopsis*, Plant Cell 20 (2008) 3065–3079.
[70] D.M. Brown, L.A. Zeef, J. Ellis, R. Goodacre, S.R. Turner, Identification of novel genes in *Arabidopsis* involved in secondary cell wall formation using expression profiling and reverse genetics, Plant Cell 17 (2005) 2281–2295.
[71] A. Cano-Delgado, S. Penfield, C. Smith, M. Catley, M. Bevan, Reduced cellulose synthesis invokes lignification and defense responses in *Arabidopsis thaliana*, Plant J. 34 (2003) 351–362.
[72] R. Zhong, D.H. Burk, C.J. Nairn, A. Wood-Jones, W.H. Morrison 3rd, Z.H. Ye, Mutation of SAC1, an *Arabidopsis* SAC domain phosphoinositide phosphatase, causes alterations in cell morphogenesis, cell wall synthesis, and actin organization, Plant Cell 17 (2005) 1449–1466.
[73] F. Goubet, A. Misrahi, S.K. Park, Z. Zhang, D. Twell, P. Dupree, AtCSLA7, a cellulose synthase-like putative glycosyltransferase, is important for pollen tube growth and embryogenesis in *Arabidopsis*, Plant Physiol. 131 (2003) 547–557.
[74] S. Bauer, P. Vasu, S. Persson, A.J. Mort, C.R. Somerville, Development and application of a suite of polysaccharide-degrading enzymes for analyzing plant cell walls, Proc. Natl. Acad. Sci. U.S.A. 103 (2006) 11417–11422.
[75] E.G. Burget, R. Verma, M. Molhoj, W.D. Reiter, The biosynthesis of L-arabinose in plants: molecular cloning and characterization of a golgi-localized UDP-D-xylose 4-epimerase encoded by the *MUR4* gene of *Arabidopsis*, Plant Cell 15 (2003) 523–531.
[76] W. Lukowitz, T.C. Nickle, D.W. Meinke, R.L. Last, P.L. Conklin, C.R. Somerville, *Arabidopsis cyt1* mutants are deficient in a mannose-1-phosphate guanylyltransferase and point to a requirement of N-linked glycosylation for cellulose biosynthesis, Proc. Natl. Acad. Sci. U.S.A. 98 (2001) 2262–2267.
[77] N. Abdulrazzak, B. Pollet, J. Ehlting, K. Larsen, C. Asnaghi, S. Ronseau, C. Proux, M. Erhardt, V. Seltzer, J.P. Renou, et al., A coumaroyl-ester-3-hydroxylase insertion mutant reveals the existence of nonredundant meta-hydroxylation pathways and essential roles for phenolic precursors in cell expansion and plant growth, Plant Physiol. 140 (2006) 30–48.
[78] R. Franke, M.R. Hemm, J.W. Denault, M.O. Ruegger, J.M. Humphreys, C. Chapple, Changes in secondary metabolism and deposition of an unusual lignin in the ref8 mutant of *Arabidopsis*, Plant J. 30 (2002) 47–59.
[79] S. Bosca, C.J. Barton, N.G. Taylor, P. Ryden, L. Neumetzler, M. Pauly, K. Roberts, G.J. Seifert, Interactions between *MUR10*/*CesA7*-dependent secondary cellulose biosynthesis and primary cell wall structure, Plant Physiol. 142 (2006) 1353–1363.
[80] M.A. Ha, I.M. MacKinnon, A. Sturcova, D.C. Apperley, M.C. McCann, S.R. Turner, M.C. Jarvis, Structure of cellulose-deficient secondary cell walls from the *irx3* mutant of *Arabidopsis thaliana*, Phytochemistry 61 (2002) 7–14.
[81] P.M. Szyjanowicz, I. McKinnon, N.G. Taylor, J. Gardiner, M.C. Jarvis, S.R. Turner, The *irregular xylem 2* mutant is an allele of korrigan that affects the secondary cell wall of *Arabidopsis thaliana*, Plant J. 37 (2004) 730–740.
[82] N.G. Taylor, W.R. Scheible, S. Cutler, C.R. Somerville, S.R. Turner, The *irregular xylem3* locus of *Arabidopsis* encodes a cellulose synthase required for secondary cell wall synthesis, Plant Cell 11 (1999) 769–780.
[83] R. Zhong, W.H. Morrison 3rd, G.D. Freshour, M.G. Hahn, Z.H. Ye, Expression of a mutant form of cellulose synthase AtCesA7 causes dominant negative effect on cellulose biosynthesis, Plant Physiol. 132 (2003) 786–795.
[84] J.H. Ko, S.H. Yang, A.H. Park, O. Lerouxel, K.H. Han, ANAC012, a member of the plant-specific NAC transcription factor family, negatively regulates xylary fiber development in *Arabidopsis thaliana*, Plant J. 50 (2007) 1035–1048.
[85] N. Mitsuda, A. Iwase, H. Yamamoto, M. Yoshida, M. Seki, K. Shinozaki, M. Ohme-Takagi, NAC transcription factors NST1 and NST3, are key regulators of the formation of secondary walls in woody tissues of *Arabidopsis*, Plant Cell 19 (2007) 270–280.
[86] R. Zhong, T. Demura, Z.H. Ye, SND1 a NAC domain transcription factor, is a key regulator of secondary wall synthesis in fibers of *Arabidopsis*, Plant Cell 18 (2006) 3158–3170.
[87] R. Zhong, E.A. Richardson, Z.H. Ye, Two NAC domain transcription factors, SND1 and NST1, function redundantly in regulation of secondary wall synthesis in fibers of *Arabidopsis*, Planta 225 (2007) 1603–1611.
[88] N. Mitsuda, M. Seki, K. Shinozaki, M. Ohme-Takagi, The NAC transcription factors NST1 and NST2 of Arabidopsis regulate secondary wall thickenings and are required for anther dehiscence, Plant Cell 17 (2005) 2993–3006.
[89] C.S. Gillmor, W. Lukowitz, G. Brininstool, J.C. Sedbrook, T. Hamann, P. Poindexter, C. Somerville, Glycosylphosphatidylinositol-anchored proteins

are required for cell wall synthesis and morphogenesis in *Arabidopsis*, Plant Cell 17 (2005) 1128–1140.
[90] D. Gantulga, Y. Turan, D.R. Bevan, A. Esen, The *Arabidopsis* At1g45130 and At3g52840 genes encode beta-galactosidases with activity toward cell wall polysaccharides, Phytochemistry 69 (2008) 1661–1670.
[91] I.J. Delgado, Z. Wang, A. de Rocher, K. Keegstra, N.V. Raikhel, Cloning and characterization of *AtRGP1*. A reversibly autoglycosylated *Arabidopsis* protein implicated in cell wall biosynthesis, Plant Physiol. 116 (1998) 1339–1350.
[92] J. Egelund, I. Damager, K. Faber, C.E. Olsen, P. Ulvskov, B.L. Petersen, Functional characterisation of a putative rhamnogalacturonan II specific xylosyltransferase, FEBS Lett. 582 (2008) 3217–3222.
[93] X. Wang, G. Cnops, R. Vanderhaeghen, S. De Block, M. Van Montagu, M. Van Lijsebettens, *AtCSLD3*, a cellulose synthase-like gene important for root hair growth in *Arabidopsis*, Plant Physiol. 126 (2001) 575–586.
[94] L.M. Fulton, C.S. Cobbett, Two alpha-L-arabinofuranosidase genes in *Arabidopsis thaliana* are differentially expressed during vegetative growth and flower development, J. Exp. Bot. 54 (2003) 2467–2477.
[95] M. Suzuki, A. Kato, N. Nagata, Y. Komeda, A xylanase, AtXyn1, is predominantly expressed in vascular bundles, and four putative xylanase genes were identified in the *Arabidopsis thaliana* genome, Plant Cell Physiol. 43 (2002) 759–767.
[96] R.A. Chavez Montes, P. Ranocha, Y. Martinez, Z. Minic, L. Jouanin, M. Marquis, L. Saulnier, L.M. Fulton, C.S. Cobbett, F. Bitton, et al., Cell wall modifications in *Arabidopsis* plants with altered alpha-L-arabinofuranosidase activity, Plant Physiol. 147 (2008) 63–77.
[97] Z. Minic, C. Rihouey, C.T. Do, P. Lerouge, L. Jouanin, Purification and characterization of enzymes exhibiting beta-D-xylosidase activities in stem tissues of *Arabidopsis*, Plant Physiol. 135 (2004) 867–878.
[98] J.K. Jensen, S.O. Sorensen, J. Harholt, N. Geshi, Y. Sakuragi, I. Moller, J. Zandleven, A.J. Bernal, N.B. Jensen, C. Sorensen, et al., Identification of a xylogalacturonan xylosyltransferase involved in pectin biosynthesis in *Arabidopsis*, Plant Cell 20 (2008) 1289–1302.
[99] Y. Hu, R. Zhong, W.H. Morrison 3rd, Z.H. Ye, The *Arabidopsis RHD3* gene is required for cell wall biosynthesis and actin organization, Planta 217 (2003) 912–921.
[100] K.W. Jung, S.I. Oh, Y.Y. Kim, K.S. Yoo, M.H. Cui, J.S. Shin, *Arabidopsis* histidine-containing phosphotransfer factor 4 (AHP4) negatively regulates secondary wall thickening of the anther endothecium during flowering, Mol. Cells 25 (2008) 294–300.
[101] R. Zhong, D.H. Burk, W.H. Morrison 3rd, Z.H. Ye, A kinesin-like protein is essential for oriented deposition of cellulose microfibrils and cell wall strength, Plant Cell 14 (2002) 3101–3117.
[102] R. Zhong, D.H. Burk, W.H. Morrison 3rd, Z.H. Ye, *FRAGILE FIBER3*, an *Arabidopsis* gene encoding a type II inositol polyphosphate 5-phosphatase, is required for secondary wall synthesis and actin organization in fiber cells, Plant Cell 16 (2004) 3242–3259.
[103] L. Hoffmann, S. Besseau, P. Geoffroy, C. Ritzenthaler, D. Meyer, C. Lapierre, B. Pollet, M. Legrand, Silencing of hydroxycinnamoyl-coenzyme A shikimate/quinate hydroxycinnamoyltransferase affects phenylpropanoid biosynthesis, Plant Cell 16 (2004) 1446–1465.
[104] T. Goujon, Z. Minic, A. El Amrani, O. Lerouxel, E. Aletti, C. Lapierre, J.P. Joseleau, L. Jouanin, *AtBXL1*, a novel higher plant (*Arabidopsis thaliana*) putative beta-xylosidase gene, is involved in secondary cell wall metabolism and plant development, Plant J. 33 (2003) 677–690.
[105] S. Bouton, E. Leboeuf, G. Mouille, M.T. Leydecker, J. Talbotec, F. Granier, M. Lahaye, H. Hofte, H.N. Truong, *QUASIMODO1* encodes a putative membrane-bound glycosyltransferase required for normal pectin synthesis and cell adhesion in *Arabidopsis*, Plant Cell 14 (2002) 2577–2590.
[106] C. Orfila, S.O. Sorensen, J. Harholt, N. Geshi, H. Crombie, H.N. Truong, J.S. Reid, J.P. Knox, H.V. Scheller, *QUASIMODO1* is expressed in vascular tissue of *Arabidopsis thaliana* inflorescence stems, and affects homogalacturonan and xylan biosynthesis, Planta 222 (2005) 613–622.
[107] S. Sato, T. Kato, K. Kakegawa, T. Ishii, Y.G. Liu, T. Awano, K. Takabe, Y. Nishiyama, S. Kuga, Y. Nakamura, et al., Role of the putative membrane-bound endo-1,4-beta-glucanase KORRIGAN in cell elongation and cellulose synthesis in *Arabidopsis thaliana*, Plant Cell Physiol. 42 (2001) 251–263.
[108] J.C. Cocuron, O. Lerouxel, G. Drakakaki, A.P. Alonso, A.H. Liepman, K. Keegstra, N. Raikhel, C.G. Wilkerson, A gene from the cellulose synthase-like C family encodes a beta-1,4 glucan synthase, Proc. Natl. Acad. Sci. U.S.A. 104 (2007) 8550–8555.
[109] K. Hematy, P.E. Sado, A. Van Tuinen, S. Rochange, T. Desnos, S. Balzergue, S. Pelletier, J.P. Renou, H. Hofte, A receptor-like kinase mediates the response of *Arabidopsis* cells to the inhibition of cellulose synthesis, Curr. Biol. 17 (2007) 922–931.
[110] S. Persson, K.H. Caffall, G. Freshour, M.T. Hilley, S. Bauer, P. Poindexter, M.G. Hahn, D. Mohnen, C. Somerville, The *Arabidopsis irregular xylem8* mutant is deficient in glucuronoxylan and homogalacturonan, which are essential for secondary cell wall integrity, Plant Cell 19 (2007) 237–255.
[111] C.P. Bonin, G. Freshour, M.G. Hahn, G.F. Vanzin, W.D. Reiter, The *GMD1* and *GMD2* genes of *Arabidopsis* encode isoforms of GDP-D-mannose 4,6-dehydratase with cell type-specific expression patterns, Plant Physiol. 132 (2003) 883–892.
[112] J.P. Vogel, T.K. Raab, C.R. Somerville, S.C. Somerville, Mutations in *PMR5* result in powdery mildew resistance and altered cell wall composition, Plant J. 40 (2004) 968–978.
[113] O.A. Zabotina, W.T. van de Ven, G. Freshour, G. Drakakaki, D. Cavalier, G. Mouille, M.G. Hahn, K. Keegstra, N.V. Raikhel, *Arabidopsis XXT5* gene encodes a putative alpha-1,6-xylosyltransferase that is involved in xyloglucan biosynthesis, Plant J. 56 (2008) 101–115.
[114] J.P. Vogel, T.K. Raab, C. Schiff, S.C. Somerville, *PMR6*, a pectate lyase-like gene required for powdery mildew susceptibility in *Arabidopsis*, Plant Cell 14 (2002) 2095–2106.
[115] J. Egelund, N. Obel, P. Ulvskov, N. Geshi, M. Pauly, A. Bacic, B.L. Petersen, Molecular characterization of two *Arabidopsis thaliana* glycosyltransferase mutants, *rra1* and *rra2*, which have a reduced residual arabinose content in a polymer tightly associated with the cellulosic wall residue, Plant Mol. Biol. 64 (2007) 439–451.
[116] C. Zhao, U. Avci, E.H. Grant, C.H. Haigler, E.P. Beers, XND1, a member of the NAC domain family in *Arabidopsis thaliana*, negatively regulates lignocellulose synthesis and programmed cell death in xylem, Plant J. 53 (2008) 425–436.
[117] J. Wang, Q. Ji, L. Jiang, S. Shen, Y. Fan, C. Zhang, Overexpression of a cytosol-localized rhamnose biosynthesis protein encoded by *Arabidopsis RHM1* gene increases rhamnose content in cell wall, Plant Physiol. Biochem. 47 (2009) 86–93.
[118] J. Egelund, B.L. Petersen, M.S. Motawia, I. Damager, A. Faik, C.E. Olsen, T. Ishii, H. Clausen, P. Ulvskov, N. Geshi, *Arabidopsis thaliana RGXT1* and *RGXT2* encode golgi-localized (1,3)-alpha-D-xylosyltransferases involved in the synthesis of pectic rhamnogalacturonan-II, Plant Cell 18 (2006) 2593–2607.
[119] L.D. Gomez, S. Baud, A. Gilday, Y. Li, I.A. Graham, Delayed embryo development in the *ARABIDOPSIS TREHALOSE-6-PHOSPHATE SYNTHASE 1* mutant is associated with altered cell wall structure, decreased cell division and starch accumulation, Plant J. 46 (2006) 69–84.
[120] I.M. MacKinnon, A. Sturcova, K. Sugimoto-Shirasu, I. His, M.C. McCann, M.C. Jarvis, Cell-wall structure and anisotropy in procuste, a cellulose synthase mutant of *Arabidopsis thaliana*, Planta 224 (2006) 438–448.
[121] G. Mele, N. Ori, Y. Sato, S. Hake, The *knotted1-like* homeobox gene *BREVIPEDICELLUS* regulates cell differentiation by modulating metabolic pathways, Genes Dev. 17 (2003) 2088–2093.
[122] O. Lerouxel, G. Mouille, C. Andeme-Onzighi, M.P. Bruyant, M. Seveno, C. Loutelier-Bourhis, A. Driouich, H. Hofte, P. Lerouge, Mutants in DEFECTIVE GLYCOSYLATION, an *Arabidopsis* homolog of an oligosaccharyltransferase complex subunit, show protein underglycosylation and defects in cell differentiation and growth, Plant J. 42 (2005) 455–468.
[123] D.H. Burk, B. Liu, R. Zhong, W.H. Morrison, Z.H. Ye, A katanin-like protein regulates normal cell wall biosynthesis and cell elongation, Plant Cell 13 (2001) 807–827.
[124] D.H. Burk, Z.H. Ye, Alteration of oriented deposition of cellulose microfibrils by mutation of a katanin-like microtubule-severing protein, Plant Cell 14 (2002) 2145–2160.