The Plant Cell, Vol. 25: 4324–4341, November 2013, www.plantcell.org

LARGE-SCALE BIOLOGY ARTICLE

SND1 Transcription Factor–Directed Quantitative Functional Hierarchical Genetic Regulatory Network in Wood Formation in *Populus trichocarpa*[C][W]

Ying-Chung Lin,[a,1] Wei Li,[a,1] Ying-Hsuan Sun,[b] Sapna Kumari,[c] Hairong Wei,[c] Quanzi Li,[a,d] Sermsawat Tunlaya-Anukit,[a] Ronald R. Sederoff,[a] and Vincent L. Chiang[a,2]

[a] Forest Biotechnology Group, Department of Forestry and Environmental Resources, North Carolina State University, Raleigh, North Carolina 27695

[b] Department of Forestry, National Chung Hsing University, Taichung 40227, Taiwan

[c] School of Forest Resources and Environmental Science, Michigan Technological University, Houghton, Michigan 49931

[d] College of Forestry, Shandong Agricultural University, Taian, Shandong 271018, China

ORCID IDs: 0000-0001-7120-4690 (Y.-C.L.); 0000-0002-4407-9334 (W.L.); 0000-0002-7152-9601 (V.L.C.).

Wood is an essential renewable raw material for industrial products and energy. However, knowledge of the genetic regulation of wood formation is limited. We developed a genome-wide high-throughput system for the discovery and validation of specific transcription factor (TF)–directed hierarchical gene regulatory networks (hGRNs) in wood formation. This system depends on a new robust procedure for isolation and transfection of *Populus trichocarpa* stem differentiating xylem protoplasts. We overexpressed *Secondary Wall-Associated NAC Domain 1s* (Ptr-*SND1-B1*), a *TF* gene affecting wood formation, in these protoplasts and identified differentially expressed genes by RNA sequencing. Direct Ptr-SND1-B1–DNA interactions were then inferred by integration of time-course RNA sequencing data and top-down Graphical Gaussian Modeling–based algorithms. These Ptr-SND1-B1-DNA interactions were verified to function in differentiating xylem by anti-PtrSND1-B1 antibody-based chromatin immunoprecipitation (97% accuracy) and in stable transgenic *P. trichocarpa* (90% accuracy). In this way, we established a Ptr-SND1-B1–directed quantitative hGRN involving 76 direct targets, including eight *TF* and 61 enzyme-coding genes previously unidentified as targets. The network can be extended to the third layer from the second-layer TFs by computation or by overexpression of a second-layer TF to identify a new group of direct targets (third layer). This approach would allow the sequential establishment, one two-layered hGRN at a time, of all layers involved in a more comprehensive hGRN. Our approach may be particularly useful to study hGRNs in complex processes in plant species resistant to stable genetic transformation and where mutants are unavailable.

INTRODUCTION

Wood formation is a complex developmental process involving differentiation of secondary xylem cells from the vascular cambium, followed by thickening of the cell wall (Evert, 2006). Growth and development in multicellular organisms are regulated at many levels by transacting molecules following well-structured regulatory hierarchies (Riechmann et al., 2000; Davidson, 2001; Wray et al., 2003; Jothi et al., 2009). Understanding the regulatory hierarchy of wood formation will offer novel and more precise genetic approaches to improve the productivity of forest trees. Secondary wall–associated NAC domain (SND) and vascular-related NAC domain (VND) proteins are transcription factors (TFs) known to regulate *TF* and pathway genes affecting secondary cell wall biosynthesis (wood formation) in *Populus* spp (Ohtani et al., 2011; Zhong et al., 2011; Li et al., 2012). However, little is known at the genome-wide level about the regulatory target genes, their quantitative causal relationships, or their regulatory hierarchy.

While TFs typically act cooperatively and combinatorially on their *cis*-regulatory DNA targets for gene expression (Müller, 2001; Levine and Tjian, 2003; Chen and Rajewsky, 2007; Hobert, 2008), a single TF may also target hundreds of genes (Chen and Rajewsky, 2007; Kaufmann et al., 2009; Demura and Ye, 2010; Huang et al., 2012; Li et al., 2012). Many TFs also target their own genes (autoactivation or repression) (Becskei and Serrano, 2000; Wray et al., 2003; Li et al., 2012). Interactions between a TF and its direct targets constitute a regulatory hierarchy. TF–target DNA interactions in vivo can be identified by chromatin immunoprecipitation (ChIP), in which an anti-TF antibody is used to enrich the chromatin that carries a TF and its interacting DNAs (Solomon et al., 1988).

ChIP, together with microarrays (ChIP-chip) or with next-generation sequencing allows genome-wide mapping of TF–DNA interactions (Kim and Ren, 2006; Farnham, 2009; Zhou et al., 2011). However, ChIP does not reveal the regulatory effects of

[1] These authors contributed equally to this work.

[2] Address correspondence to vchiang@ncsu.edu.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantcell.org) is: Vincent L. Chiang (vchiang@ncsu.edu).

[C] Some figures in this article are displayed in color online but in black and white in the print edition.

[W] Online version contains Web-only data.

www.plantcell.org/cgi/doi/10.1105/tpc.113.117697

the TF on its bound targets (Solomon et al., 1988; Farnham, 2009; Bassel et al., 2012). A combination of ChIP sequencing or ChIP-PCR and the TF-induced differential gene expression is needed to reveal the functional network and its regulatory effects. This combination has been used to discover specific functional regulatory networks in animal development, using cell cultures representing different developmental stages where specific sets of *TF*s are induced (Yu and Gerstein, 2006; Farnham, 2009; Bhardwaj et al., 2010; Gerstein et al., 2010; Roy et al., 2010; Cheng et al., 2011). These TF regulatory networks are arranged in well-organized hierarchies. TF functional hierarchical gene regulatory networks (hGRNs), consisting of three to five hierarchical layers, have been described for growth and development of *Caenorhabditis elegans* and *Drosophila melanogaster* (Gerstein et al., 2010; Roy et al., 2010). Subsets of hGRNs with multiple hierarchical layers have also been described for human and mouse (Cheng et al., 2011; Niu et al., 2011).

In plants, ChIP has been applied mainly to *Arabidopsis thaliana* (Kaufmann et al., 2010), focusing on mapping interactions between a single TF and one or a few selected target genes. The regulatory effects of some of these interactions were demonstrated through perturbation or induction of the specific *TF* in transgenics or mutants (Pruneda-Paz et al., 2009; Zheng et al., 2009; Bassel et al., 2012; Huang et al., 2012; Kumar et al., 2012). Knowledge of the TF–DNA interactions in species other than *Arabidopsis* is limited. ChIP techniques have not been reported for any tree species, representing a major challenge to identifying TF–DNA interactions.

Many tree species are recalcitrant to genetic transformation and lack collection of specific mutants (Merkle and Dean, 2000; Song et al., 2006), making studies of the regulatory effects of TF–DNA interactions and hGRNs in these species previously impossible. For tree species that are amenable to genetic transformation, methods are technically demanding and slow, requiring 12 to 18 months of tissue culture (Merkle and Dean, 2000). To reveal a functional hGRN for wood formation, an efficient transgenic system, such as those developed for the cell cultures of yeast (*Saccharomyces cerevisiae*) and animals (Horak et al., 2002; Yu and Gerstein, 2006; Gerstein et al., 2010; Cheng et al., 2011; Niu et al., 2011), is needed where immediate transcriptome responses to *TF* perturbation can be induced, characterized, and quantified.

Plant protoplasts can be cell- or tissue-specific populations of single cells used to study a broad spectrum of processes from physiology to gene function/regulation (Abel and Theologis, 1994; Chiu et al., 1996; Davey et al., 2005; Thorpe, 2007; Yoo et al., 2007). Freshly isolated protoplasts retain cell and transcriptome identity, differentiated state (without dedifferentiation), and original biochemical and regulatory activity (Cocking, 1972; Sheen, 2001; Birnbaum et al., 2003; Yoo et al., 2007; Faraco et al., 2011). These cell properties may be sustained for at least 48 h after isolation (Yoo et al., 2007; Faraco et al., 2011; Chupeau et al., 2013). Therefore, protoplasts are particularly useful for studying early transcriptome responses or the dynamics of such responses to treatments, including perturbation of gene expression.

Mesophyll protoplasts from leaves have been routinely used for transient gene expression (Sheen, 2001; Yoo et al., 2007; Faraco et al., 2011). Such systems have been used extensively to study plant signal transduction. The use of mesophyll protoplasts in these studies is appropriate because some signal transduction pathways highly active in mesophyll cells are conserved in many other meristematic cell types (Inoue et al., 2001; Sheen, 2001; Fujii et al., 2009). At the full transcriptome level, protoplasts from cells representing progressive developmental stages in the *Arabidopsis* root were used to establish a microarray-based global gene expression map linking gene activity to specific cell fates in root development (Birnbaum et al., 2003).

While protoplasts are effective experimental systems, results from one protoplast system cannot be broadly generalized to other cell types because many cellular processes are highly cell or tissue specific. Tissue specificity is particularly important for studying TF–DNA transcriptional regulatory networks because many TFs may require a tissue-specific partner for *trans*-activating or repressing the target DNA expression (Farnham, 2009; Moreno-Risueno et al., 2010; Faraco et al., 2011). Cell- or tissue-specific protoplasts are necessary to study biological processes that are specific to the cells or tissues from which the protoplasts are derived (Birnbaum et al., 2003; Fujii et al., 2009). Leaf mesophyll protoplasts would not be appropriate for studying wood formation.

Protoplasts from wood-forming cells, the immature differentiating xylem cells, are the specific source for information on wood development. The challenge is that woody plant tissues are generally resistant to protoplast isolation (Teulieres et al., 1991; Manders et al., 1992; Gomez-Maldonado et al., 2001; Sun et al., 2009; Tang et al., 2010; Chen et al., 2011; Li et al., 2012). Even for amenable woody plant tissues, protoplast isolation has never before been designed for yield and quality adequate for transgenesis, transcriptome and chromatin enrichment studies.

Using *Populus trichocarpa* protoplasts from stem-differentiating xylem (SDX) as a model, we have begun to describe the SND/VND-directed functional hGRN for wood formation, beginning with the hGRN associated with *Secondary Wall-Associated NAC Domain 1s* (Ptr-*SND1-B1*) (Li et al., 2012). Here, we report the establishment of an efficient SDX protoplast system for monitoring genome-wide Ptr-*SND1-B1*–induced gene transactivation responses and a novel computational method for constructing a hierarchical TF-DNA network. Then, we describe the integration of the SDX protoplast system with computation and modeling for the development of a Ptr-SND1-B1–directed functional hGRN. We then report on the establishment of a robust ChIP assay and ChIP-PCR analysis to demonstrate that the Ptr-SND1-B1–target interactions in the protoplast-inferred hGRN are authentic in intact differentiating xylem tissue. We further describe transcriptome analysis (using RNA sequencing [RNA-seq] and quantitative RT-PCR [qRT-PCR]) to verify that the transactivation effects of such interactions in protoplasts also occur in SDX of stable *P. trichocarpa* transgenics overexpressing Ptr-*SND1-B1*. The developed protoplast system is a simple and dependable genomic tool to study the SND/VND-directed functional hGRN for wood formation.

RESULTS

A Rapid and High-Yield Protoplast Isolation and Efficient Protoplast Transfection System Was Developed for *P. trichocarpa* SDX

We first established an SDX protoplast isolation system that would give high protoplast yields. Greenhouse-grown *P. trichocarpa* at

ages 3 to 9 months were used. We collected SDX tissue strips (see Supplemental Methods 1 online) and used ~2 to 7 g (fresh weight) for each protoplast isolation. Isolation protocols developed for woody tissues (Teulieres et al., 1991; Manders et al., 1992; Gomez-Maldonado et al., 2001; Sun et al., 2009; Tang et al., 2010) were tested and gave very low yields, particularly from plants 3 to 4 months of age.

We then modified the transient expression in *Arabidopsis* mesophyll protoplast protocol (Yoo et al., 2007) focusing on 6- to 9-month-old plants, resulting in high yields of ~2.5 × 10^7 protoplasts/g of fresh weight SDX. For ~2.5 g of SDX that we usually use for each isolation, the modified protocol would allow the generation of ~60 RNA-seq libraries. The total time needed for processing every 2.5 g of SDX to generate protoplasts is ~4 h (~1 h for preparing tissue strips and ~3 h for cell wall digestion). The efficiency of this system is comparable to or better than many others (Manders et al., 1992; He et al., 2007; Riazunnisa et al., 2007; Yoo et al., 2007; Sonntag et al., 2008; Tang et al., 2010; Zhang et al., 2011; Guo et al., 2012; Huddy et al., 2012; Pitzschke and Persak, 2012); however, the throughput (~4 h for every 2.5 g SDX) was still low. We then found that a debarked stem segment (Figure 1; see Supplemental Figure 1 online) with intact SDX can be used directly to release SDX protoplasts by submerging the stem into the enzyme digestion solution. Using this stem dipping approach, further optimization of cell wall digestion conditions led to a major reduction of total processing time from ~3 h down to ~20 min. The trypan blue dye exclusion test demonstrated that, after 20 min or even up to 3 h enzyme digestion under optimized conditions, ~98% of the freshly isolated SDX protoplasts are viable.

Our optimized system (using stem dipping) requires a very short total processing time of ~25 min (from tissue harvest to protoplast recovery) and gives a high yield (~2.5 × 10^7 protoplasts/g SDX), making it one of the most efficient protoplast isolation systems reported (Wu et al., 2009; Tan et al., 2013). Furthermore, our system allows protoplast isolations from milligrams to tens of grams of SDX tissue at a similar efficiency and high protoplast yield.

We next incorporated the optimized SDX protoplast isolation with a DNA transfection process to develop a transient transgene expression system. We focused on polyethylene glycol (PEG)–directed transfection and tested key reaction parameters. A plasmid DNA (pUC19-35S-sGFP) for expressing a synthetic green fluorescent protein gene (*sGFP*) (Chiu et al., 1996) under the control of a cauliflower mosaic virus (CaMV) 35S promoter was used as the reporter gene, and the green fluorescent protoplasts (see Supplemental Figure 2 online) were counted to estimate the transfection efficiency. We readily achieved at least 30% transfection efficiency using several combinations of parameters (Table 1).

Finally, we tested the entire system (optimized isolation and transfection) with SDX isolated from greenhouse-grown plants at different seasons, ages (6 to 9 months), and developmental stages (internodes 10 to 40). The system is robust, independent of these factors, and gives consistent results: (1) ~25-min total isolation time, (2) high yields (~2.5 × 10^7 protoplasts/g SDX), and (3) 30% transfection efficiency.

We then used this protoplast system as a transient transgenic model to learn the Ptr-SND1-B1–directed functional hGRN for wood formation. We first used qRT-PCR to test the specific transcriptional responses of the protoplast system to the overexpression of Ptr-*SND1-B1*.

Figure 1. *P. trichocarpa* Stem Dipping Approach for SDX Protoplast Isolation.

(A) A 10-cm stem segment.
(B) A debarked stem segment.
(C) Debarked stems in the cell wall digestion enzyme solution in a 50-mL Falcon tube for SDX protoplast isolation.

The *P. trichocarpa* SDX Protoplast System Provides in Vivo Evidence That Ptr-SND1-B1 Is a Transcriptional Activator of Ptr-*MYB002* and Ptr-*MYB021*

We transfected the SDX protoplasts with a plasmid DNA (pUC19-35S-PtrSND1-B1-35S-sGFP) for simultaneous expression of Ptr-*SND1-B1* and *sGFP*, each under the control of a CaMV 35S promoter. A portion of the same batch of protoplasts was also transfected with a pUC19-35S-sGFP plasmid as a control. After 12 h, PCR analysis of total RNAs from the transfected protoplasts using gene-specific primers (Shi and Chiang, 2005) confirmed a high level of Ptr-*SND1-B1* transcript from the genomic DNA of the Ptr-*SND1-B1* transgene. qRT-PCR analysis demonstrated that overexpression of Ptr-*SND1-B1* in SDX protoplasts increased the transcript levels of two R2R3-MYB-type transcription factor genes, Ptr-*MYB002* and Ptr-*MYB021*, by 10- to 20-fold (Figure 2). Ptr-*MYB002* and Ptr-*MYB021* are phylogenetically paired MYB homologs in the *P. trichocarpa* genome (Li et al., 2012). We previously demonstrated that the Ptr-*MYB021* gene is a direct regulatory target of Ptr-SND1-B1

Table 1. Optimization of PEG-Mediated *P. trichocarpa* SDX Protoplast DNA Transfection

Plasmid DNA Purification	DNA/Cell Ratio	PEG		Transfection Time (min)	Transfection Efficiency[a]
		Type	% (v/v)		
Qiagen mini	10 μg/2 $\times 10^4$	PEG4000	40	10	<5% (n > 10)
Qiagen midi	10 μg/2 $\times 10^4$	PEG4000	40	10	15% (n > 10)
CsCl gradient	30 μg/2 $\times 10^4$	PEG4000	40	10	30% (n > 10)
CsCl gradient	20 μg/2 $\times 10^4$	PEG4000	40	10	30% (n > 10)
CsCl gradient	10 μg/2 $\times 10^4$	PEG8000	40	10	<5% (n = 1)
CsCl gradient	10 μg/2 $\times 10^4$	PEG6000	40	10	<15% (n = 1)
CsCl gradient	10 μg/2 $\times 10^4$	PEG4000	10	10	<5% (n = 1)
CsCl gradient	10 μg/2 $\times 10^4$	PEG4000	20	10	<15% (n = 1)
CsCl gradient	10 μg/2 $\times 10^4$	PEG4000	40	0.5	<30% (n = 2)
CsCl gradient	10 μg/2 $\times 10^4$	PEG4000	40	5	30% (n = 2)
CsCl gradient	10 μg/2 $\times 10^4$	PEG4000	40	10	30% (n > 10)[b]
CsCl gradient	10 μg/2 $\times 10^4$	PEG4000	40	15	30% (n > 10)
CsCl gradient	10 μg/2 $\times 10^4$	PEG4000	40	30	<30% (n = 2)
CsCl gradient	10 μg/2 $\times 10^4$	PEG4000	40	60	<15% (n = 2)

[a]*n* represents the number of biological replicates of transfection efficiency test (see Methods) for each set of conditions.
[b]The optimized SDX protoplast transfection condition used for all Ptr-*SND1-B1* overexpression experiments.

(Li et al., 2012). This and current results demonstrated that, in *P. trichocarpa*, Ptr-SND1-B1 is a positive and direct regulator of a pair of sequence-related MYB homologs. We next tested whether the transcriptome of the SDX protoplasts is representative of that of developing xylem.

RNA-Seq Results Reveal That *P. trichocarpa* SDX Protoplasts and SDX Tissue Share High Transcriptome Identity

We performed RNA-seq to profile the genome-wide transcript abundance of SDX protoplasts and of SDX tissue isolated from the same, 6-month-old, greenhouse-grown clonal plants. High-quality RNAs (see Supplemental Figure 3 online) were isolated from these materials for constructing RNA-seq libraries, which gave ~7 million reads per library. After mapping the reads to the *P. trichocarpa* genome (http://www.phytozome.com) using TOPHAT (Trapnell et al., 2009), transcripts of 38,066 and 35,994 genes were detected for SDX protoplasts and SDX tissue, respectively. Approximately 96% of the genes in SDX, or wood-forming tissue, are present in the protoplasts, suggesting that the SDX protoplasts can be a simple and effective model for studying transcriptome responses associated with wood formation. We then used RNA-seq to profile transcriptome changes in SDX protoplasts overexpressed with Ptr-*SND1-B1* to identify the differentially expressed genes (DEGs) and evaluate their causal and hierarchical interactions with Ptr-*SND1-B1*.

Overexpression of Ptr-*SND1-B1* in SDX Protoplasts Induces a Small Set of DEGs

We characterized the transcriptome changes in protoplasts transfected with Ptr-*SND1-B1-sGFP* compared with *sGFP* transfected control. We first used qRT-PCR to quantify alterations in Ptr-*MYB021* transcript levels to estimate the time needed for the transcriptome changes to take place and possible time-dependent variation. SDX protoplasts from a single plant (6 months old) were transfected and incubated at room temperature for 2, 7, 12, 21, 25, 31, and 45 h, with three to five biological replicates (each from a different clonal propagule) and three technical repeats for each biological replicate at each time point. A significant increase (4.39 ± 0.46 fold, mean ± SE) in Ptr-*MYB021* transcript level was first observed at 7 h after transfection. The largest increase (12.43 ± 3.20) was detected at 21 h. All the increases were significant between 7 and 31 h after transfection.

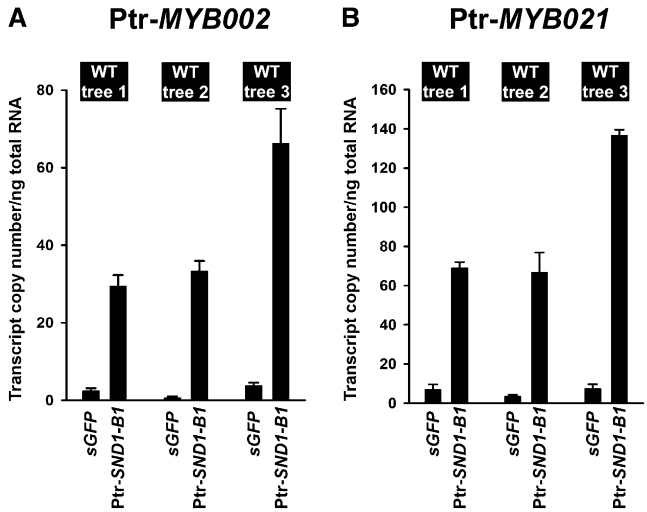

Figure 2. Ptr-SND1-B1 Is a Transcriptional Activator of Ptr-*MYB002* and Ptr-*MYB021*.

qRT-PCR was used to quantify the transcript abundance of Ptr-*MYB002* **(A)** and Ptr-*MYB021* **(B)** in *P. trichocarpa* SDX protoplasts overexpressing Ptr-*SND1-B1* or *sGFP* (control). Three biological replicates (SDX protoplasts from wild-type [WT] trees 1, 2, and 3) were performed. Error bars represent SE of three qRT-PCR technical replicates.

Approximately 95% of the protoplasts remained viable 45 h after transfection. We then focused on full transcriptome responses to Ptr-*SND1-B1* overexpression at 7, 12, and 25 h after transfection. The resulting RNA sequences were mapped to the *P. trichocarpa* genome to obtain read counts as described above.

Scatterplots of transcript abundance (read counts) between Ptr-*SND1-B1-sGFP* and *sGFP* (control) transfected protoplasts were used to evaluate the effect of Ptr-*SND1* overexpression (Figure 3). The scatterplots show strong Pearson correlations between read counts of expressed genes of *PtrSND1-B1-sGFP* and *sGFP*, with coefficients of 0.98 ± 0.002, 0.99 ± 0.002, and 0.99 ± 0.002 for 7, 12, and 25 h, respectively (Figure 3). These high coefficients reflect that the effects of Ptr-*SND1-B1* overexpression are highly specific, causing only a very small number of genes (178) to be differentially expressed ($P < 0.05$; see Supplemental Data Set 1 online). These 178 DEGs, representing the total from the three time points, were identified using edgeR (Robinson et al., 2010) by comparing the transcript abundance of each gene between the *PtrSND1-B1-sGFP* transfection and the *sGFP* transfection control. Of these 178 DEGs, 14 are *TF* genes. We next studied the functional implications of these DEGs in wood formation.

Gene Ontology Functional Enrichment Analysis Indicates Cell Wall Association for Many DEGs Induced by PtrSND1-B1

To explore the functional significance of the 178 DEGs, the g:Profiler Web server (http://biit.cs.ut.ee/gprofiler/; Reimand et al., 2011) was used to analyze for specific functional enrichment. The enrichment analysis is based on the gene ontology (GO) annotation of the Ensembl genome (http://www.ensembl.org). The results showed 603 GO terms for all 178 DEGs identified in this study. Twenty-nine of the 603 GO terms showed highly significant functional enrichment with five major GO hierarchical classes (see Supplemental Table 1 online). They are (1) cellular aromatic compound metabolism, (2) cellular component biogenesis, (3) oxidation reduction, (4) extracellular location, and (5) ion binding. These five major GO hierarchical classes contain 44 DEGs (see Supplemental Table 2 online). The GO subgroups of these five major classes contain significant enriched functional associations with phenylpropanoid and lignin metabolic processes and cell wall biogenesis. Therefore, the Ptr-SND1-B1–induced DEGs in the SDX protoplasts are associated with cell wall formation. We then analyzed the causal interactions between Ptr-*SND1-B1* and these DEGs.

The Time-Dependent Induction of DEGs Suggests a PtrSND1-B1–Directed hGRN

Pairwise comparisons of transcript abundance at different time points (Figure 3) show that the transcriptome responds differentially over time after Ptr-*SND1-B1* transfection. The Ptr-*SND1-B1* transcripts had the highest level of abundance at 7 h, decreasing at 12 and 25 h (Figure 3). Similarly, the total number of DEGs was the highest at 7 h and decreased with increasing time. Among the 178 DEGs, 122 (92 + 13 + 13 + 4 in Venn diagram; Figure 4; see orange columns in Supplemental Data Set 1 online) appeared to respond immediately at 7 h. The remaining 56 (23 + 10 + 23 in

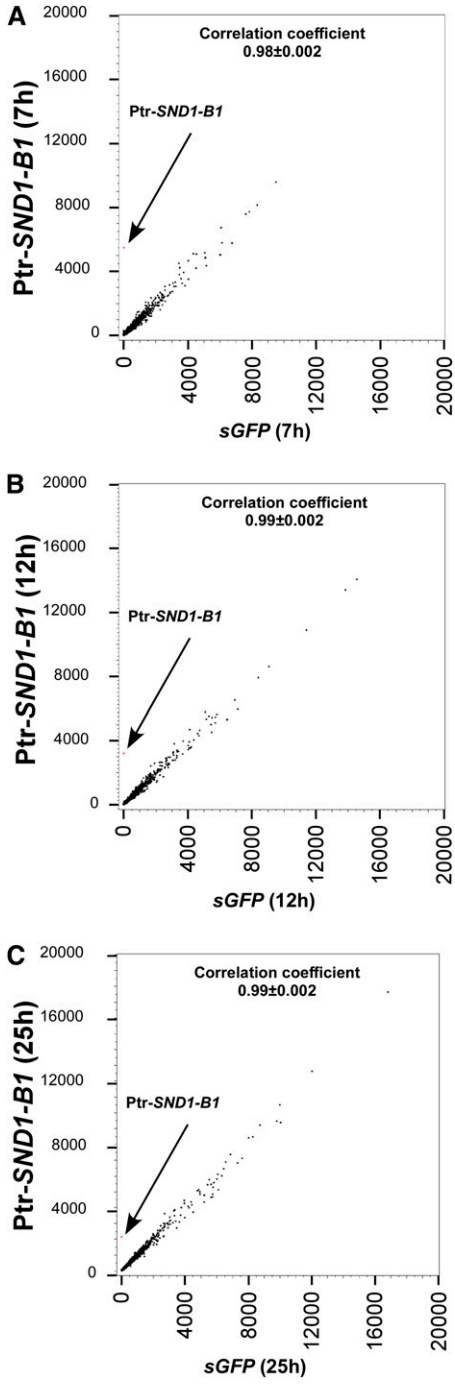

Figure 3. The Effects of Ptr-*SND1-B1* Overexpression in SDX Protoplasts Are Highly Specific.

Scatterplots of the average of RNA-seq read counts from three biological replicates of Ptr-*SND1-B1* and *sGFP* (control) transfected SDX protoplasts show high Pearson correlation coefficients after 7- **(A)**, 12- **(B)**, and 25-h **(C)** incubation, demonstrating the high specificity of transactivation effects of Ptr-*SND1-B1*. Red dots indicated by arrows represent the transcript abundance of Ptr-*SND1-B1*.
[See online article for color version of this figure.]

Venn diagram; Figure 4; see green columns in Supplemental Data Set 1 online), found only after 12 h and 25 h incubation, have a delayed response to Ptr-*SND1-B1* overexpression. These differential responses to Ptr-*SND1-B1* overexpression strongly suggest a Ptr-SND1-B1–directed hGRN. The immediate-response group (122 genes, containing Ptr-*MYB002* and Ptr-*MYB021*) likely includes direct regulatory targets of Ptr-SND1-B1, whereas the delayed-response group of 56 genes are candidates for lower layers of the Ptr-SND1-B1 regulatory hierarchy.

In addition to this experimental approach (Method A in Figure 4), we next developed a computational method (Method B in Figure 4) to infer direct targets. We expected that the authentic direct targets can be commonly identified by two independent methods (Figure 4). The computational method is a two-step approach: Step I (Figure 4) is a combination of Fisher's exact test (Fisher, 1922, 1925) and a probability-based method developed in this study (see Methods) to screen the 178 DEGs for genes whose expression is highly concordant with the expression of the Ptr-*SND1-B1* transgene. We call these concordant DEGs responsive genes. Step II (Figure 4), which we developed previously (Lu et al., 2013), uses a partial correlation-based graphical Gaussian model (GGM) (Whittaker, 1990; Edwards, 2000) for a vigorous assessment of direct interactions between Ptr-*SND1-B1* and responsive genes to infer Ptr-SND1-B1's direct targets.

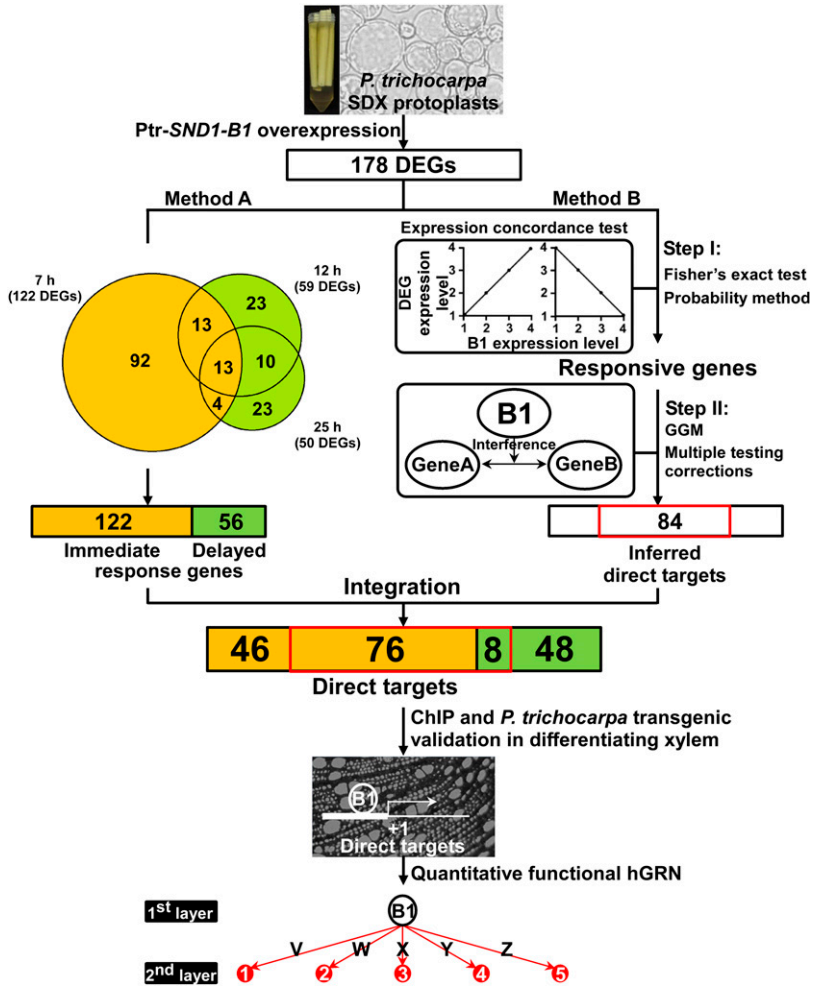

Figure 4. The Workflow for Constructing the Ptr-SND1-B1–Directed Quantitative Functional hGRN in Wood Formation.

SDX protoplasts were isolated using the dipping method and transfected with Ptr-*SND1-B1* to result in 178 DEGs based on RNA-seq analysis. Method A (time-dependent method) was used to identify 122 immediate response genes (orange area). Method B (computational method) was used to infer Ptr-SND1-B1's direct targets. Step I is a combination of Fisher's exact test and a probability-based method to identify responsive genes of Ptr-*SND1-B1* based on expression concordance. Step II is a combination of GGM and multiple testing corrections for assessment of direct interactions between Ptr-*SND1-B1* (B1) and responsive gene pairs (Gene A and Gene B) to infer Ptr-SND1-B1's direct targets (84 candidate direct targets in red box; see Methods). Through the integration of Methods A and B, 76 direct targets of Ptr-SND1-B1 were identified. ChIP-PCR for SDX and stable transgenic *P. trichocarpa* were used to validate Ptr-SND1-B1's direct targets to reveal a Ptr-SND1-B1–directed quantitative functional hGRN. In this hGRN example, B1 indicates Ptr-SND1-B1 at the top and red nodes represent the Ptr-SND1-B1's direct targets (second-layer constituents). The direct targets 1, 2, 3, 4, and 5 are activated by Ptr-SND1-B1 for V, W, X, Y, and Z fold increase (based on RNA-seq), respectively.

A Probability-Based Computational Approach Identifies a Group of DEGs That Are Responsive to the Expression of Ptr-*SND1-B1*

To identify Ptr-*SND1-B1*–responsive genes, we combined Fisher's exact test (at low stringency, $P < 0.25$) with the probability-based method to screen the 178 DEGs for those having high expression concordance with the Ptr-*SND1-B1* transgene (Step I in Figure 4). Because a high expression concordance is indicative of a strong causal relationship, DEGs having such concordances with transgene Ptr-*SND1-B1* are therefore more likely to be regulated directly by this transgene. To define the expression concordance, we first scaled the transcript abundances of all DEGs and Ptr-*SND1-B1* and discretized them arbitrarily into four levels (1, 2, 3, and 4 in *x* and *y* axes in Figure 5; see Methods). Based on pairwise comparisons of the expression levels, we projected possible expression concordances with expression features (DEG:Ptr-*SND1-B1* expression ratios) that suggest strong causal relationships (Figure 5). The values of the expression features of all these concordances are continuous and are bordered by two extremes in two concordance types, Types I and IV (Figure 5). Type I, the concurrent type, specifies that the expression of the DEG is at the same level as that of the transgene Ptr-*SND1-B1* (Figure 5A). Type IV, the inverse concordance, where the expression level of DEG is exactly opposite to that of Ptr-*SND1-B1* (Figure 5D). The expression feature values of the other concordance types can be divided into discrete counterparts for analysis. For simplicity, these concordances were discretized here into two types, Types II and III, each having a unique set of the DEG:Ptr-*SND1-B1* expression ratios (Figures 5B and 5C). We then used our probability-based algorithm (see Methods) to evaluate each DEG:Ptr-*SND1-B1* expression value for the type of expression concordance. Any DEG with the expression that fits with at least one of the four concordance Types with a P value > 0.1 was then classified as a responsive gene. We identified 90 responsive genes (see Supplemental Data Set 2A online) from the 178 DEGs. We then used GGM to model these 90 responsive DEGs to infer the direct targets of Ptr-SND1-B1.

Graphical Gaussian Modeling of Regulatory Interference Allows Inference of Direct Targets of PtrSND1-B1

We applied GGM to subsets of three genes and examined the expression dependence between two of the genes (a pair of the responsive genes) conditional on the third (Ptr-*SND1-B1*) for regulatory inference to identify Ptr-*SND1-B1*'s direct targets (Step II in Figure 4). The underlying idea is that a TF would target multiple genes and that the overexpression of a *TF* would most strongly affect the expression of its target genes in two possible ways: (1) Two such target genes become more significantly coexpressed, and (2) two significantly coexpressed target genes become no longer coexpressed. We then used GGM to examine the three-gene subset for effects of Ptr-*SND1-B1* transgene overexpression on the expression of the 90 responsive genes. If the effect coincides with either one of the two possibilities described above, we then considered that Ptr-*SND1-B1* interferes directly with the pair of the responsive genes.

When all possible combinations of Ptr-*SND1-B1* and two genes from the responsive gene pool were examined for interference, we sorted all the three-gene subsets by P values, as indicators of interference intensity. We then performed multiple testing corrections (Benjamini and Hochberg, 1995) on these P values to obtain the significant three-gene subsets with a false discovery rate < 0.05. The times for each responsive gene that has appeared in the significant subsets were then summed to represent the frequency of interference. Of the 90 Ptr-*SND1-B1*–responsive genes, 84, which include 10 *TF*s, were interfered by Ptr-*SND1-B1* at least one time and were considered the direct targets of Ptr-*SND1-B1* (see Supplemental Data Set 2 online).

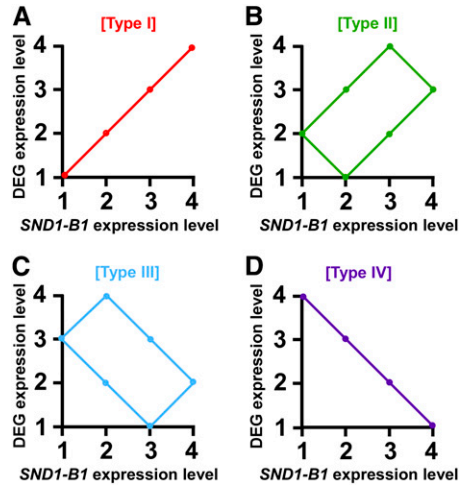

Figure 5. Types of Expression Concordances between the DEG and Ptr-*SND1-B1* That Suggest Strong Causal Relationships.

Each of the four expression concordance types has a unique set of DEG: Ptr-*SND1-B1* transcript level ratios.
(A) Type I, the DEG:Ptr-*SND1-B1* ratios are 1:1, 2:2, 3:3, and 4:4.
(B) Type II, 2:1, 1:2, 2:3, 3:4, 4:3, and 3:2.
(C) Type III, 3:1, 2:2, 1:3, 2:4, 3:3, and 4:2.
(D) Type IV, 4:1, 3:2, 2:3, and 1:4.
[See online article for color version of this figure.]

Integration of Computational and Experimental Approaches Identifies a Unique Set of PtrSND1-B1's Direct Regulatory Targets

We integrated the computational and time-dependent gene expression methods to identify direct regulatory targets of Ptr-SND1-B1. Any targets identified by one method but excluded by the other were disqualified. The time-dependent method classified 122 DEGs in the immediate response group as the putative direct targets (orange area in Method A, Figure 4). Of the 84 computationally inferred direct targets (red framed in Figure 4), 76 belong to this immediate response group and eight to the delayed response group. These eight DEGs were then disqualified because they were excluded by the time-dependent method. Consequently, only the 76 DEGs, inferred by the computational method and commonly included by both methods, were most likely the authentic direct regulatory target genes of Ptr-SND1-B1 (see Supplemental Table 3 online).

A considerable amount of work has been done on predicting secondary wall NAC binding element (Zhong et al., 2010; Wang

et al., 2011) and treachery element regulating binding motifs (Kubo et al., 2005; Pyo et al., 2007) in *Arabidopsis* gene promoters. These putative motif sequences could be used to predict Ptr-SND1-B1-DNA binding for these 76 inferred targets. However, we proceeded from experimentally determined direct TF-DNA binding. We used ChIP to verify whether the 76 inferred targets are bound by Ptr-SND1-B1 in vivo in intact differentiating xylem tissue.

ChIP-PCR Validates That PtrSND1-B1's Direct Target Genes Inferred by the Protoplast System Are the Authentic Targets in the SDX Tissue

We modified the *Arabidopsis* ChIP assay (Kaufmann et al., 2010; Li et al., 2011) to overcome difficulties associated with woody tissues and developed a robust anti-TF antibody-based ChIP protocol for *P. trichocarpa* SDX tissue (see Methods). The TF-specific antibody allows the characterization of the native state of the TF-DNA binding. Positive Ptr-SND1-B1-target binding in chromatin directly from SDX would verify that targets inferred by the protoplast system are the authentic direct targets of Ptr-SND1-B1 in intact differentiating xylem for wood formation.

We tested the antibody specificity to ensure high levels of specific Ptr-SND1-B1-target complex enrichment over non-specific DNA-protein immunoprecipitation from loci bound by Ptr-SND1-B1 homologs. Ptr-*SND1-B1* has three other family members (*SND1-A1*, *SND1-A2*, and *SND1-B2*) in the genome, and all these share high protein sequence identities at the conserved N-terminal NAC domain (Xie et al., 2000; Ernst et al., 2004; Li et al., 2012). Their sequences at the C-terminal activation domain are divergent (Xie et al., 2000; Ernst et al., 2004; Li et al., 2012). We then identified a C-terminal polypeptide unique to Ptr-SND1-B1 and used it as the immunogen to make polyclonal antibodies. The specificity of this antibody was tested against purified *Escherichia coli* recombinant proteins from the four Ptr-*SND1* members in the form of N-terminal glutathione *S*-transferase (GST)–tagged fusion proteins. These four Ptr-SND1 member proteins could be recognized by the monoclonal anti-GST antibodies, but only Ptr-SND1-B1 could be detected by anti-PtrSND1-B1-peptide antibodies (Figure 6A; see Supplemental Figure 4A online), demonstrating the specificity of this anti-PtrSND1-B1 antibody.

We then used ChIP-PCR to validate the in vivo Ptr-SND1-B1-target binding. We performed PCR amplification of ChIP DNA products focusing on the 2-kb promoter sequence (2000 to 1 bp) upstream of the translation start site of the candidate gene, where TF binding sites are generally located (Thibaud-Nissen et al., 2006; Farnham, 2009; Heisig et al., 2012). PCR was first performed to amplify the enriched DNA products for the promoter sequence (500 to 1 bp) of the candidate gene (Figure 6B). If no amplification could be detected, we then performed additional PCR to amplify the 2000- to 500-bp promoter sequence (see Supplemental Figure 5A online).

From the 76 inferred direct target genes (see Supplemental Table 3 online), we selected 15, which includes all 10 *TF*s (DEG001 to DEG009 and DEG011; Figure 6C) in this category and five enzyme encoding genes (DEG031, DEG053, DEG083, DEG120, and DEG169), for ChIP-PCR validation. One of the *TF*s, Ptr-*MYB021*, a previously validated target of Ptr-SND1-B1 (Li et al., 2012), served as a positive control. Similarly, the five enzyme encoding genes (Figure 6C) were chosen because their *Arabidopsis* homologs were shown to be direct targets of *AtSND1* in *Arabidopsis* leaf protoplasts overexpressing At-*SND1* (Zhong et al., 2010).

ChIP experiments were conducted on chromatin isolated from *P. trichocarpa* SDX of 6-month-old trees. We observed robust enrichment of Ptr-SND1-B1 in the 2-kb promoter region of all these 15 selected targets (Figure 6C), validating that these inferred targets are authentic direct regulatory targets of Ptr-SND1-B in vivo in SDX. These 10 *TF*s include one *NAC*, four *MYB*s, four zinc finger family genes, and a gene encoding an integrase-type DNA binding protein. The *NAC* gene is Ptr-*SND1-L-2*, an SND1-like NAC domain protein that shares 47% sequence identity with Ptr-SND1-B1 (Li et al., 2012). Phylogenetically, the protein sequence of Ptr-SND1-L-2 is in the same clade as the SND1s and VNDs, both involved in the regulation of secondary cell wall thickening (Hu et al., 2010). The four *MYB*s include a pair of paralogs, Ptr-*MYB002* and Ptr-*MYB021* (Li et al., 2012), and two other *MYB* related *TF*s (DEG004 and 005). Ptr-*MYB002* and Ptr-*MYB021* are the two orthologs of *Arabidopsis MYB*46, predicted to directly regulate several laccase genes that are associated with lignin biosynthesis (Berthet et al., 2011; Kim et al., 2012a).

Homologs of only two (Ptr-*MYB002* and an integrase-type *TF*, DEG006; see Supplemental Table 3 online) of these 10 authentic direct target *TF*s in *P. trichocarpa* SDX have previously been reported as the direct targets of At-SND1 in an *Arabidopsis* leaf protoplast system (Zhong et al., 2010). The ChIP-PCR analysis revealed that, based on the number of genes tested, our approach, integrating time-dependent and computational methods, can effectively identify the authentic direct targets with 100% accuracy.

From the 46 DEGs (see Supplemental Data Set 3A online) that were not considered direct targets by our integrated method (Figure 4), we randomly selected six for ChIP-PCR assay. These six included four *TF* (DEG010, DEG012, DEG013, and DEG014; Figure 6C) and two enzyme (DEG045 and DEG172; Figure 6C) coding genes. No enrichment of Ptr-SND1-B1 could be detected in the 2-kb promoter of five of these six genes (Figure 6C; see Supplemental Figure 5B online). We next selected six of the eight DEGs (see Supplemental Data Set 3B online) that were excluded from the direct targets by the time-dependent method and tested by ChIP-PCR (Figure 4). No enrichment of Ptr-SND1-B1 could be detected in the 2-kb promoter of any of these six DEGs (Figure 6C; see Supplemental Figure 5B online). Likewise, six randomly selected DEGs from the group of 48 (see Supplemental Data Set 3C online) that were excluded by both time-dependent and computational methods were also excluded by ChIP-PCR analysis (Figure 6C; see Supplemental Figure 5B online). Overall, ChIP-PCR analysis of 33 Ptr-*SND1-B1*–induced DEGs validated the identities of 32 of them in vivo in intact SDX as either direct or indirect targets of Ptr-SND1-B1, demonstrating the accuracy (32/33, 97%) of our system (SDX protoplasts + computation) in identifying the direct regulatory targets of Ptr-SND1-B1. The results further indicated that SDX protoplasts are representative of intact SDX for revealing hGRNs in xylem differentiation or wood formation.

The transgenic SDX protoplast system together with ChIP-PCR analysis revealed a two-layered regulatory hierarchy directed by

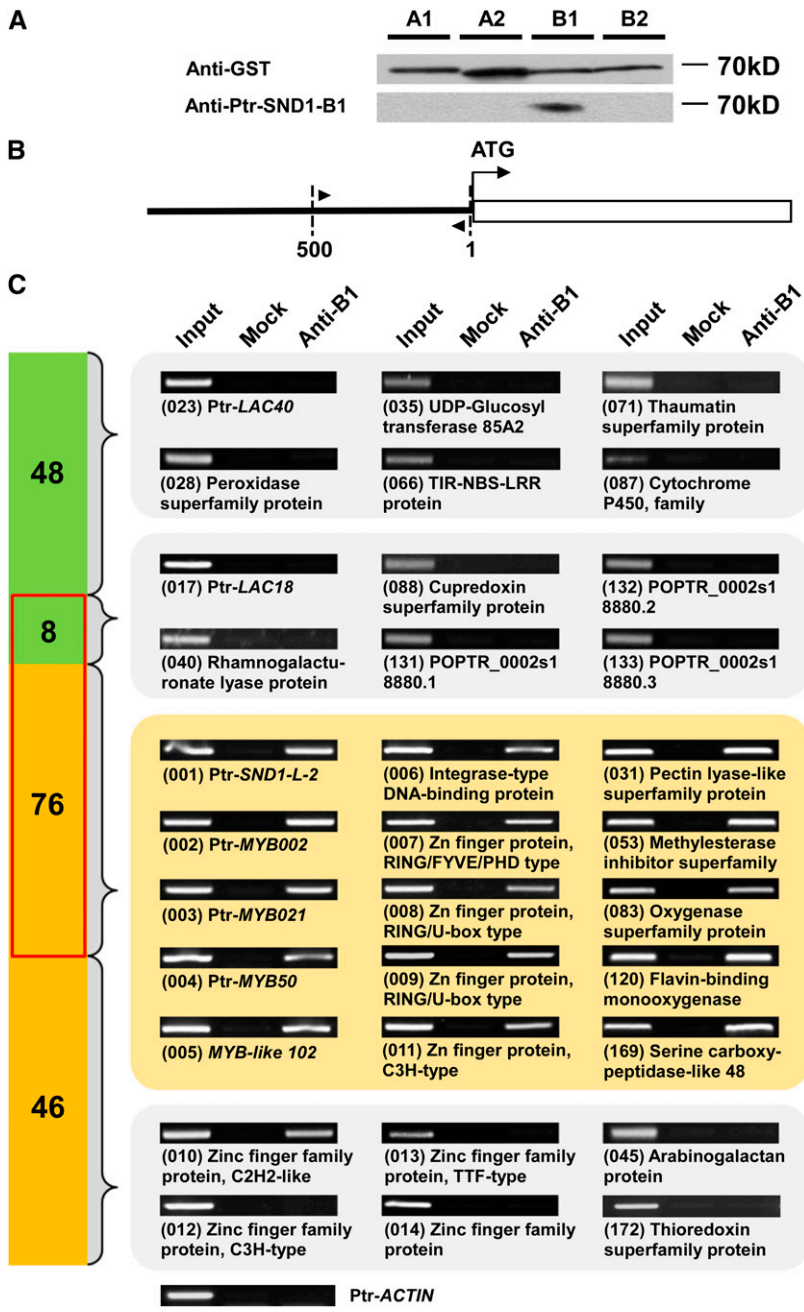

Figure 6. ChIP-PCR Validation of Inferred Direct and Indirect Targets of Ptr-SND1-B1.

(A) Protein gel blot for antibody specificity. Purified Ptr-SND1-A1 (A1), Ptr-SND1-A2 (A2), Ptr-SND1-B1 (B1), and Ptr-SND1-B2 (B2) *E. coli* recombinant proteins fused with GST tag at the N terminus were probed with the anti-GST and anti-PtrSND1-B1 antibodies, respectively. A full-size protein gel blot is shown in Supplemental Figure 4 online.

(B) A simplified gene structure to indicate the locations of the amplified promoter sequences. The thick line corresponds to a gene promoter that drives its gene represented by the rectangle. The arrowheads show the promoter sequence location for primer design.

(C) ChIP-PCR assays of selected genes from each DEG category using chromatin from differentiating xylem and anti-PtrSND1-B1 antibody. The 76 in the orange area and red box is the number of candidate direct targets derived from the computational and time-dependent methods. The 8 in the green area and red box represents the indirect targets excluded by time-dependent method. The 46 in the orange area and 48 in the green area indicate the indirect targets excluded by the computational method and the time-dependent method, respectively. The DEG ID number of selected genes is shown in parentheses (see Supplemental Table 1 online). Input, Mock and Anti-B1 are PCR reactions using the chromatin preparations before immunoprecipitation, immunoprecipitated with preimmune serum and immunoprecipitated with anti-PtrSND1-B1 antibody, respectively. Ptr-*ACTIN* was used as a negative control. Three independent biological replicates of ChIP assays were performed, and the results of one biological replicate are shown.

Ptr-SND1-B1 (at the top; Figure 7). In this hierarchy, Ptr-SND1-B1 directly regulates 76 target genes (the second layer), including 10 *TF*s. We next built a part of the third layer of the hierarchy from two second-layer *TF*s, Ptr-*MYB002* and Ptr-*MYB021*, using the computational approach.

The SDX Protoplast System Reveals a PtrSND1-B1–Directed Regulatory Hierarchy

As described above, 84 (Figure 4) of the 178 DEGs are the computationally inferred direct targets of Ptr-SND1-B1. Likely, the remaining 94 DEGs consist of direct targets of the TFs in the lower (second and third) layers of the Ptr-SND1-B1–directed hGRN. Therefore, these remaining 94 DEGs can be used as input for computationally inferring the direct targets of the second-layer TFs. We selected Ptr-*MYB002* and Ptr-*MYB021* to test this approach. We selected these two *MYB*s because the direct targets of *Arabidopsis MYB46* (the ortholog of Ptr-*MYB002* and Ptr-*MYB021*) have been proposed or validated (Kim et al., 2012a, 2012b), providing a basis to assess the robustness of the approach. The probability-based algorithm identified 12 and 14 Ptr-*MYB002*– and Ptr-*MYB021*–responsive genes, respectively (see Supplemental Data Set 4 online). From the Ptr-*MYB*–specific responsive genes, GGM inferred nine direct targets for Ptr-*MYB002* and 11 for Ptr-*MYB021* (Figure 7; see Supplemental Data Set 5 online).

The two Ptr-*MYB*s target a total of 12 unique genes, of which eight are common targets, including five homologs of laccase genes (Ptr-*LAC14*, Ptr-*LAC15*, Ptr-*LAC40*, Ptr-*LAC41*, and Ptr-*LAC49*; Figure 7), suggesting redundant or combinatorial regulatory roles for these two sequence related Ptr-*MYB*s. The five Ptr-*LAC*s and *Arabidopsis LAC4* are homologs and were shown to control lignin quantity in *P. trichocarpa* (Lu et al., 2013) and *Arabidopsis* (Berthet et al., 2011), respectively. At-*LAC4* is a predicted direct target of At-MYB46 based on the MYB46-responsive *cis*-acting binding site (RKTWGGTR) in the At-*LAC4* gene promoter. All these 12 inferred direct targets of the two Ptr-*MYB*s have at least one exact RKTWGGTR binding site in the 2-kb proximal promoter region (see Supplemental Data Set 5 online). These results strongly suggest that the 12 inferred genes are authentic direct targets of Ptr-*MYB002* and Ptr-*MYB021* and further demonstrate the accuracy of our integrated approach (Figure 4) in identifying direct TF–DNA interactions.

Finally, we used stable transgenic *P. trichocarpa* to verify the adequacy of using the protoplast/computation system to study gene regulation that occurs in intact wood forming tissue at the whole-plant level. We tested whether the SDX protoplast-inferred TF–DNA interactions and the regulatory effects of these interactions also take place in SDX of transgenic *P. trichocarpa* plants overexpressing Ptr-*SND1-B1*.

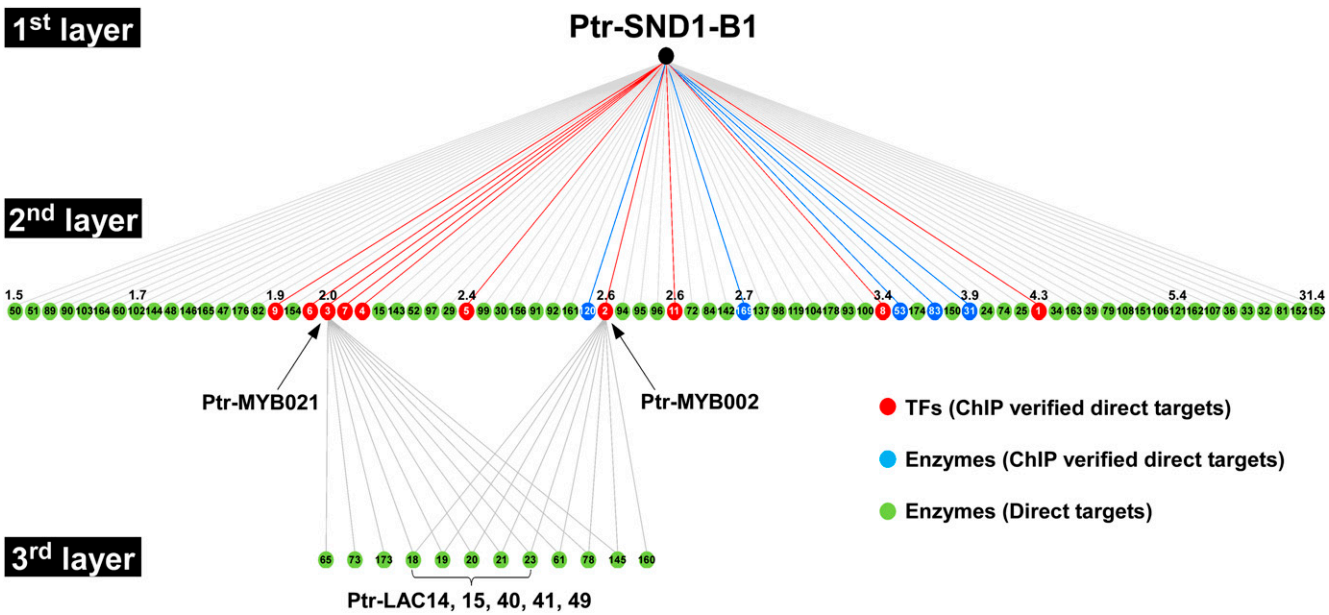

Figure 7. Ptr-SND1-B1–Directed Quantitative Functional hGRN in Wood Formation.

Ptr-SND1-B1 is at the top (first layer) of this hGRN. The second layer of the hGRN consists of 76 Ptr-SND1-B1 direct targets (second layer) inferred from the integration of the time-dependent and computational methods. Among these 76 direct targets, 10 TFs (red nodes) and five enzymes (blue nodes) were validated by ChIP-PCR in SDX. Two of the 10 TFs are PtrMYB021 and PtrMYB002 as indicated. On the third layer, 11 direct targets for PtrMYB021 and nine for PtrMYB002 (green nodes) were inferred using the computational approach. PtrMYB002 and 021 share eight common direct targets, of which five are laccase genes (Ptr-*LAC14*, Ptr-*LAC15*, Ptr-*LAC40*, Ptr-*LAC41*, and Ptr-*LAC49*). The number in the nodes indicates the DEG ID number. The number above the Ptr-SND1-B1's direct targets represents the $\log_2$ fold change of the direct targets in SDX protoplasts induced by Ptr-*SND1-B1* overexpression. The Ptr-SND1-B1's direct targets are displayed based on the induction level by Ptr-*SND1-B1* overexpression in an increasing order from left to right.

Transgenic *P. trichocarpa* Overexpressing Ptr-*SND1-B1* Further Proves That PtrSND1-B1's Direct Target Genes Are Overexpressed in the SDX Tissue

Thirteen transgenic *P. trichocarpa* lines overexpressing Ptr-*SND1-B1* driven by a CaMV 35S promoter were generated, and three with the highest transgene transcript levels (Figure 8) were selected and maintained in a greenhouse. We quantified the transcript levels of the 10 ChIP-PCR positive *TF* targets in the SDX of these three transgenic lines and found that the expression of all these targets was upregulated by 2- to 25-fold (Figure 9B). The expression of four of the five ChIP-PCR–verified enzyme encoding targets was also induced by ~2- to 50-fold in the transgenics (Figure 9C). Six additional genes were randomly selected from the inferred direct targets for qRT-PCR, and five of them had increased transcript levels (approximately two- to eightfold increases) in the transgenics (Figure 9D). Overall, ~90% (19/21) of the protoplast-inferred TF–DNA interactions tested were validated for their regulatory effects in intact transgenic SDX at the whole-plant level. The validation suggests that our protoplast/computation system is sufficient to reveal tissue- or cell-specific TF-DNA regulatory networks, without using stable and whole-plant transgenics.

In summary, the results and insights demonstrate the value of the SDX system (SDX protoplast expression + computation) for understanding the hierarchical structure of transacting regulation in xylem differentiation for wood formation.

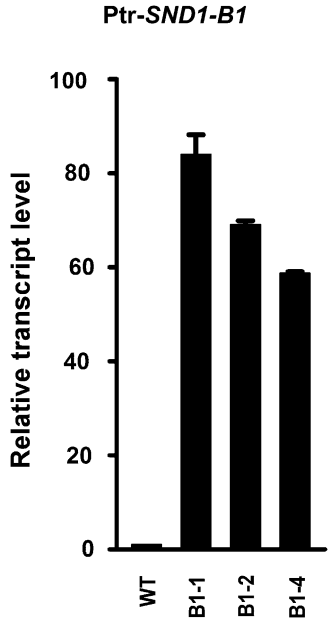

Figure 8. Overexpression of Ptr-*SND1-B1* in Transgenic *P. trichocarpa* Plants.

Ptr-*SND1-B1*, driven by a 35S promoter, was overexpressed in *P. trichocarpa*. The transcript abundance of Ptr-*SND1-B1* in three wild-type (WT) plants and three selected transgenic lines (B1-1, B1-2, and B1-4) was estimated by qRT-PCR. The average of three biological replicates of wild-type plants was set as 1. Error bars in three transgenic lines represent the SE of three qRT-PCR technical replicates.

DISCUSSION

In this study, we used *P. trichocarpa* SDX protoplasts as a model system to begin to reveal the hierarchy of transcriptional regulation of xylogenesis (wood formation). Plant cell protoplasts are known to maintain their differentiated state without dedifferentiation in isotonic solutions and are simple experimental systems, much like mammalian cell lines, for studying specific cellular activities (Davey et al., 2005). SDX protoplasts are expected to retain specific cellular and transcriptomic potentials for wood formation.

We first developed a robust SDX protoplast isolation and PEG-mediated DNA transfection system suitable for genome-wide high-throughput transient gene expression and transactivation analysis. The system provides information on gene perturbation responses in only a few days, and many experiments can be performed in parallel. Using RNA-seq, we demonstrated that the transcriptome of the SDX protoplast is highly (~96%) representative of that of the intact SDX tissue. To study the hGRN for wood formation, we perturbed Ptr-*SND1-B1* expression in SDX protoplasts. In all protoplast perturbation experiments reported here, we overexpressed Ptr-*SND1-B1* using a vector containing a *sGFP* gene and compared that to transfection of the same vector lacking this *TF* gene. The Pearson correlation coefficients of transcript abundance between these two constructs are very high (0.98 to 0.99; Figure 3) and provide efficient detection of DEGs resulting from the overexpression of Ptr-*SND1-B1*. These high coefficients also strongly indicate high specificity of Ptr-SND1-B1–directed regulatory effects, consistent with the small number of genes (178 DEGs) that were differentially expressed. Therefore, the protoplasts are an efficient system for genome-wide quantification of transregulation of TF–target DNA interactions. To map direct TF–DNA interactions during wood formation, we established a ChIP assay system for the differentiating xylem tissue of *P. trichocarpa*.

A common ChIP-based approach to map in vivo TF–DNA interactions in *Arabidopsis* is the use of a transgenic tagged TF to engineer these interactions, which can then be immunoprecipitated through the tag. The results of this approach are insightful, if the transgenic tagged TF is functionally identical to the native TF. This functional identity must be validated through demonstration that the tagged TF can functionally complement the specific loss-of-function phenotype. This tagged TF approach is not applicable to woody plants because no TF mutants are known for these species. Here, we developed an anti-TF antibody-based ChIP approach to map in vivo TF–DNA interactions in intact secondary differentiating xylem. This TF-specific antibody approach allows more exclusive enrichment of the genomic regions that are bound to the specific native TF of interest.

The SDX protoplast-based system (SDX protoplast expression + computation) allowed us to infer a hierarchical layer of 76 genes immediately downstream of Ptr-*SND1-B1* (Figure 7). We used ChIP-PCR and transgenic *P. trichocarpa* overexpressing Ptr-*SND1-B1* to validate this hierarchical network. We selected 15 genes (including all 10 inferred direct *TF* genes) of the second hierarchical layer (Figure 6C) for ChIP-PCR using chromatin from

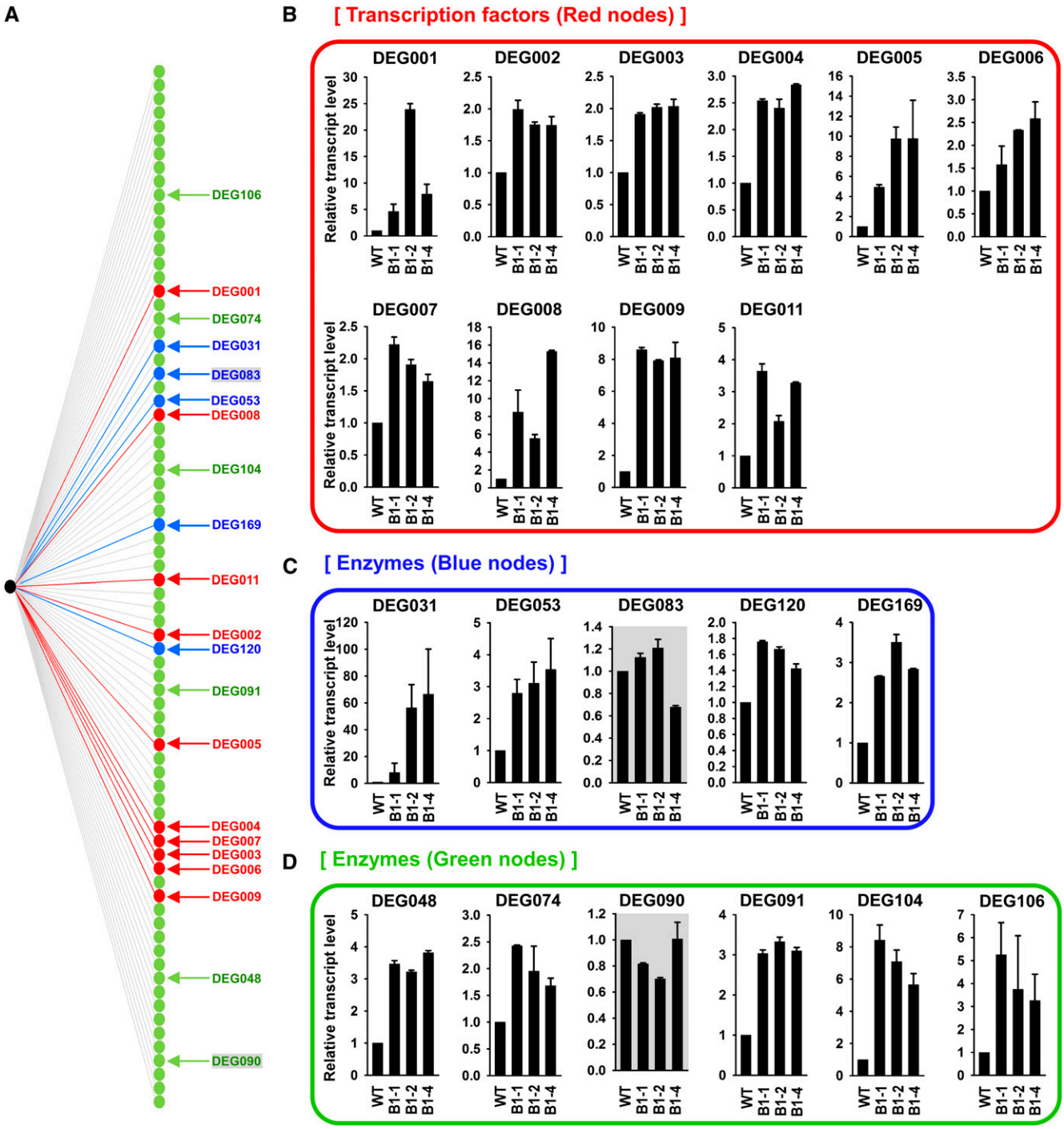

Figure 9. Validation of Ptr-SND1-B1's Direct Targets in Stable Transgenic *P. trichocarpa*.

Ptr-SND1-B1's direct target genes derived from the SDX protoplast system were verified by qRT-PCR for their induced expression in differentiating xylem of stable transgenic *P. trichocarpa* overexpressing Ptr-*SND1-B1*.

(A) The same two-layered hGRN in Figure 8 is shown here. Arrows indicate the selected targets for qRT-PCR tests in transgenic *P. trichocarpa*.

(B) to **(D)** The transcript abundance ChIP-PCR verified TFs (red nodes) **(B)**, ChIP-PCR verified enzyme genes (blue nodes) **(C)**, and inferred enzyme gene targets (green nodes) **(D)** were quantified by qRT-PCR in three wild-type (WT) and three Ptr-*SND1-B1* transgenic lines (B1-1, B1-2, and B1-4). DEG083 (shaded in **[C]**) and DEG090 (shaded in **[D]**) were two genes not affected by Ptr-*SND1-B1* overexpression in stable transgenic *P. trichocarpa*. The average of three biological replicates of wild-type plants was set as 1. Error bars in three transgenic lines represent the SE of three qRT-PCR technical replicates.

intact SDX (not from SDX protoplasts). We confirmed that all 15 are Ptr-SND1-B1's direct targets in vivo in differentiating xylem (Figure 6C). Fourteen of these 15 genes and five of six additional genes from the second hierarchical layer (19 out of 21) were further found to have significantly increased transcript levels in SDX of the transgenic *P. trichocarpa* (Figure 9). These verifications demonstrated that the protoplast-inferred Ptr-*SND1-B1* regulation hierarchy is functional in SDX, the wood-forming tissue.

Our use of the SDX protoplast system revealed a novel Ptr-SND1-B1–directed hGRN for wood formation. Ptr-SND1-B1 is the immediate transactivator of 10 *TF*s (Figure 7), of which two (Ptr-*MYB002* and an integrase-type *TF*; see Supplemental Table 3 online) have homologs in *Arabidopsis* identified as *SND1* targets, and the remaining eight (DEG001, DEG003, DEG004, DEG005, DEG007, DEG008, DEG009, and DEG011; see Supplemental Table 3 online) are novel *SND1* targets. One of these eight regulatory targets is a Ptr-*SND1* family member, Ptr-*SND1-L-2*, and such transactivation suggests autoregulation among NAC family members in the *SND/VND* regulatory hierarchy. The involvement of another novel target, a zinc finger family protein (DEG007 in Supplemental Table 3 online), in this wood formation hGRN is consistent with the association of its homolog in tension wood formation in *Populus tremula* × *Populus tremuloides* (Andersson-Gunneråas et al., 2006). The precise functions of these 10 *TF*s in wood formation need to be further characterized. The transgenic *P. trichocarpa* plants overexpressing Ptr-*SND1-B1* may help provide clues to these functions. This work strongly suggests a Ptr-SND1-B1–directed regulatory structure consisting of 10 direct *TF* and 64 enzyme coding genes in the second hierarchical layer of a two-layered hGRN (Figure 7) in vivo in wood forming tissue of *P. trichocarpa*.

Our protoplast system also accurately identified genes that are indirect targets of Ptr-SND1-B1, as verified by ChIP-PCR analysis of chromatin from intact differentiating xylem (Figure 6C). These indirect targets are the candidate targets (the third hierarchical layer) of the second-layer *TF* genes. For example, 12 of these indirect targets were inferred as direct targets of two second-layer *TF* genes, Ptr-*MYB002* and Ptr-*MYB021* (Figure 7; see Supplemental Data Set 5 online). These two sequence-related Ptr-*MYB*s share eight common targets, of which five are laccase genes (Ptr-*LAC14*, Ptr-*LAC15*, Ptr-*LAC40*, Ptr-*LAC41*, and Ptr-*LAC49*). Using transgenic *P. trichocarpa* overexpressing a microRNA (miRNA), ptr-miR397a, we recently demonstrated that this miRNA and many *TF* genes, including Ptr-*MYB021*, regulate directly 13 laccase (including Ptr-*LAC14*, Ptr-*LAC15*, Ptr-*LAC40*, Ptr-*LAC41*, and Ptr-*LAC49*) and four peroxidase (Ptr-*PO*) genes in a transcriptional regulatory network controlling lignin content during wood formation (Lu et al., 2013). These miRNA transgenic results strongly supported the accuracy of developed top-down algorithm (Figure 4) for predicting the direct TF-DNA relationships. We only selected two second-layer *TF*s (Ptr-*MYB002* and Ptr-*MYB021*) to predict their direct targets because many of the TF-DNA interactions are known for their homologs (At-*MYB46*) (Kim et al., 2012a) and therefore are evidence for the validity of the prediction. Our purpose here is to further demonstrate the robustness of our computational algorithms. The completion of the third layer of the hGRN is a substantial subject of its own for later investigation.

The SND1-related hGRN for wood formation revealed in this study is different from an SND1 network derived from *Arabidopsis* leaves (Zhong et al., 2010). A total of 138 direct targets were found for At-*SND1*, the *Arabidopsis* homolog of Ptr-*SND1-B1*. Of these 138 genes, only seven (two *TF* and five enzyme coding genes; i.e., homologs of Ptr-*MYB002* and Integrase-type DNA binding superfamily protein; see Supplemental Table 3 online) were found in our 76 direct targets of Ptr-SND1-B1 in *P. trichocarpa* differentiating xylem, demonstrating that leaf and xylem cells behave quite differently in the transcriptome response to *SND* regulation. In fact, the biological function of At-*SND1* in regulating the 138 genes in leaf cells is unclear because, in the leaf tissue, At-*SND1* was not detected (Zhong et al., 2006). By contrast, Ptr-*SND1-B1* is abundantly and specifically expressed in *P. trichocarpa* differentiating xylem (Li et al., 2012).

The SDX protoplast system is most effective for studying a complex regulatory hierarchy of multiple layers. We discovered that all direct regulatory targets of Ptr-SND1-B1 are activated within 7 h after Ptr-*SND1-B1* transfection for overexpression (Figure 4). Therefore, to identify the direct regulatory targets of a *TF*, the SDX protoplast expression system can focus on the immediate response group of DEGs from 7 h of incubation for computational analysis (Figure 4). This simple approach (7 h of *TF* overexpression in SDX protoplasts + computation) will then allow the sequential establishment, one two-layered hGRN at a time in a top-down manner, of all regulatory layers involved in a complete hGRN. For example, in this study, 10 *TF* genes were identified as the direct regulatory targets in the hierarchical layer immediately below Ptr-*SND1-B1* (Figure 7). From each of these 10 *TF*s in this second hierarchical layer, the same approach (7 h of *TF* overexpression in SDX protoplasts + computation) could then lead to the third hierarchical layer, encompassing ultimately the direct targets of all these 10 *TF*s. Likewise, subsequent layers can be built progressively to reveal a complete Ptr-SND1-B1–directed hGRN. A comprehensive and interactive transcriptional regulatory network integrating the sub-networks for all 20 *NAC TF*s can then be revealed for genetic regulation in wood formation. All such sub-hGRN and the complete interactive hGRN are functional networks depicting not only connectivity but quantitative information of interference frequencies or transregulation effects. We believe that our approach can be readily extended to other cell or tissue types to study functional regulatory hierarchies in complex developmental processes specific in these cells/tissues in plants, particularly in species that lack mutants or are resistant to stable genetic transformation.

METHODS

Plant Materials

Populus trichocarpa plants (genotype Nisqually-1) were maintained in a greenhouse as described (Song et al., 2006). Stem internodes of healthy 3- to 9-month-old plants were used to collect xylem tissue and isolate SDX protoplasts.

SDX Protoplast Isolation, DNA Transfection, and Transfection Efficiency Determination

All detailed methods for these experiments are described in Supplemental Methods 1 online. The cellulolytic enzyme solution and buffers were adopted

from the transient expression in *Arabidopsis thaliana* mesophyll protoplast system (Yoo et al., 2007) with modifications, and two forms of differentiating xylem tissues, tissue strips and debarked stem segments, were used to isolate protoplasts. The xylem tissues were placed in the freshly prepared enzyme solution for ~3 h (tissue strips) and ~20 min (debarked stem segments) at room temperature, and the digested xylem tissues were transferred to W5 solution (2 mM MES, pH 5.7, 125 mM CaCl2, 154 mM NaCl, 0.1 M Glc, and 5 mM KCl) to release SDX protoplasts. The protoplasts were then filtered, resuspended in W5 solution, chilled on ice, and finally resuspended in MMG solution (4 mM MES, pH 5.7, 0.5 M mannitol, and 15 mM $MgCl_2$, room temperature) to a concentration of 2×10^5 cells/mL for immediate transfection. We constructed pUC19-35S-PtrSND1-B1-35S-sGFP to overexpress Ptr-*SND1-B1* in SDX protoplasts and used pUC19-35S-sGFP (Li et al., 2012) as a control transgene. For transfection, several conditions were tested: (1) plasmid DNA purification methods, (2) DNA/cell ratios, (3) the concentrations and types of PEG solutions, and (4) time periods of transfection. Three to five biological replicates of each DNA transfection for different time points were performed, with three technical repeats for each biological replicate. The viability of protoplasts after isolation and incubation was tested by the 0.4% trypan blue dye (Invitrogen) exclusion method (Larkin, 1976). Protoplast suspension was visualized by a Zeiss Axioskop 40 microscope to calculate the transfection efficiency.

Total RNA Extraction

Total RNA from SDX or SDX protoplasts was isolated according to Li et al. (2012). The RNA quality was examined by an Agilent 2100 Bioanalyzer using Agilent RNA 6000 Pico Assay chips (see Supplemental Figure 3 online). Total RNA (~800 ng) could be extracted from ~8×10^5 protoplasts. The RNA was used for RT-PCR, qRT-PCR, and RNA-seq.

RT-PCR

For testing the overexpression of Ptr-*SND1-B1* in transfected SDX protoplasts, reverse transcription was performed as described (Shi and Chiang, 2005) to synthesize cDNA using RNA from Ptr-*SND1-B1-sGFP* and *sGFP* (control) transfected SDX protoplasts and poly(T) adapter as the primer. The PtrSND1-B1-2F/R (see Supplemental Data Set 6 online) PCR primer set was used to detect Ptr-*SND1-B1*, using the *P. trichocarpa* actin gene, Ptr-*Actin*, as the internal control and detected by the PtrActin-F/R primer set (see Supplemental Data Set 6 online).

qRT-PCR

qRT-PCR was performed as described (Li et al., 2012) to detect the transcript abundance of Ptr-*MYB002* and Ptr-*MYB021* in the *PtrSND1-B1-sGFP* and *sGFP* (control) transfected SDX protoplasts, using primers listed in Supplemental Data Set 6 online, with three biological replicates for each transfection and three technical repeats for each biological replicate.

Full-Transcriptome RNA-Seq Analysis of SDX and SDX Protoplasts and Identification of DEGs

RNA-seq was performed with three biological replicates each for SDX, SDX protoplasts, and SDX protoplasts transfected with Ptr-*SND1-B1-sGFP* and *sGFP* for 7, 12, and 25 h. Total RNA of each sample (750 ng) was used for library construction using Illumina TruSeq RNA sample preparation kit. Each library was constructed with different index sequences in the adaptors. The quality and concentration of libraries was examined by the Agilent 2100 bioanalyzer with Agilent high-sensitivity DNA assay chips. Six libraries with different index numbers were pooled by mixing equal quantities of DNA from each library and applied as one lane for sequencing; 72-bp average read lengths were obtained. After removing the 4-bp library sequence index sequences from each read, the remaining 68 bp were mapped to the reference *P. trichocarpa* genome release v2.2 (Phytozome V7.0; http://www.phytozome.com) using the program TOPHAT (Trapnell et al., 2009).

The frequency of raw counts was determined and normalized as described (Lu et al., 2013). DEGs were identified using edgeR (Robinson et al., 2010) by comparing the relative transcript abundance for each gene between the Ptr-*SND1-B1* and the *sGFP* (control) transfection at each incubation time point (7, 12, and 25 h). The global false discovery rate of the differential gene identification was set at 0.05 level.

GO Functional Enrichment Analysis

This analysis was performed using the g:Profiler Web server (http://biit.cs.ut.ee/gprofiler/; Reimand et al., 2011). The *P. trichocarpa* GO functional enrichment analysis in the g:Profiler Web server is based on the Ensembl Genome (http://www.ensembl.org) annotation for *P. trichocarpa*. The annotation contains 3892 GO terms for 29,042 genes. The statistical significances of functional enrichment are calculated (g:Profiler) for the 178 DEGs considering all known *P. trichocarpa* genes as the background control.

Probability-Based Identification of Ptr-*SND1-B1*–Responsive Genes

Details of this computational approach are described in Supplemental Methods 2 online. Briefly, to identify which DEGs are Ptr-*SND1-B1*–responsive genes, we first scaled the normalized gene expression abundances (see the full-transcriptome RNA-seq section above) into values between 0 and 1 and then discretized the scaled values into four levels. Fisher's exact test was then applied to the discretized data to identify those DEGs that are dependent on Ptr-*SND1-B1* in expression. We then calculated a conditional probability table (four levels [Ptr-*SND1-B1*] $\times$ four levels [Ptr-*SND1-B1*–dependent DEG] =16) between Ptr-*SND1-B1* and each Ptr-*SND1-B1*–dependent DEG. The conditional probabilities in each such table were then classified into four types of concordance: Type I, II, III, and IV (Figure 5). The statistical significance of each of these four types was tested with Pearson's χ^2. If any of these four types is statistically significant, the Ptr-*SND1-B1*–dependent DEG is defined as a Ptr-*SND1-B1*–responsive gene.

Inference of Direct Target Genes of Ptr-*SND1-B1* Using GGM

Detailed mathematical procedures for inferring Ptr-SND1-B1's direct targets are described in Supplemental Methods 3 online. Briefly, to identify which Ptr-*SND1-B1*–responsive genes are likely to be the direct targets of Ptr-*SND1-B1*, we employed GGM model to evaluate which Ptr-*SND1-B1*–responsive genes have causal relationships with Ptr-*SND1-B1*. The GGM-based algorithm evaluated one at a time a subset of three genes: Ptr-*SND1-B1* and two of the Ptr-*SND1-B1*–responsive genes. The algorithm examined if the presence of Ptr-*SND1-B1* would significantly interfere the expression relationships of the two Ptr-*SND1-B1*–responsive genes in the subset. After the Ptr-*SND1-B1* and all possible pairwise combinations of two Ptr-*SND1-B1*–responsive genes were examined for interference, we calculated the times that Ptr-*SND1-B1* interfered with each responsive gene. In this study, the responsive genes interfered at least once by Ptr-SND1-B1 were classified as the inferred direct targets of Ptr-SND1-B1.

ChIP Assay of *P. trichocarpa* Differentiating Xylem

ChIP assays were performed using anti-PtrSND1-B1 antibody and chromatin from the SDX of 6-month-old *P. trichocarpa* plants. A peptide, TQDYNNEIDLWNFTTRSSPD, located in the C terminus of Ptr-SND1-B1 was synthesized and used for polyclonal antibody production as described

(Li et al., 2012). Ptr-SND1-A1, Ptr-SND1-A2, Ptr-SND1-B1, and Ptr-SND1-B2 N-terminal GST-tagged fusion proteins were expressed in *Escherichia coli* and purified as described (Li et al., 2012). These purified Ptr-SND1 member proteins were used to verify the specificity of the peptide-based anti-PtrSND1-B1 antibody.

We developed ChIP assay for *P. trichocarpa* differentiating xylem by modifying protocols used for *Arabidopsis* (Kaufmann et al., 2010; Li et al., 2011). All ChIP assay–related details are described in Supplemental Methods 4 online. Briefly, 5 g of SDX was immersed in cross-link buffer (1% [w/w] formaldehyde) for cross-linking protein-DNA complexes (vacuum, 30 min, stopped by 0.125 M Gly). The SDX was washed (double-distilled water ×3), dried, ground in liquid nitrogen, suspended in Buffer 1 (0.4 M Suc, 10 mM Tris-HCl, pH 8.0, 5 mM β-mercaptoethanol, and protease inhibitors [1 mM phenylmethylsulfonyl fluoride, 1 μg/mL pepstatin A, and 1 μg/mL leupeptin]), agitated at 4°C, filtered, and pelleted (1800*g*, 10 min, 4°C). The pellet was resuspended in Buffer 2 (0.25 M Suc, 10 mM Tris-HCl, pH 8.0, 5 mM β-mercaptoethanol, 10 mM $MgCl_2$, 1% Triton X-100, and protease inhibitors), centrifuged (16,000*g*, 10 min, 4°C), resuspended in Buffer 3 (1.7 M Suc, 10 mM Tris-HCl, pH 8.0, 5 mM β-mercaptoethanol, 2 mM $MgCl_2$, 0.15% Triton X-100, and protease inhibitors), and centrifuged (16,000*g*, 1 h, 4°C). The pellet was resuspended in lysis buffer, sonicated to shear the DNA, and centrifuged (16,000*g*, 10 min, 4°C). A small aliquot of the supernatant was used as input control, and the remaining supernatant was diluted (×10) by dilution buffer and divided equally into two tubes for reactions (4°C overnight) with anti-PtrSND1-B1 antibody and preimmune serum (as mock control), respectively. The solution was treated (4°C, 2 h) with Dynabeads protein G (Invitrogen), and the beads were washed (×2) with dilution buffer, high-salt wash buffer (×1), LiCl buffer (×1), and Tris-EDTA buffer (×2). Freshly prepared elution buffer (65°C) was added to elute (65°C, 15 min, ×2) the protein-DNA complexes, of which cross-linking was reversed (5 M NaCl, 65°C overnight) and treated with protease/RNase buffer (45°C for 1 h). The DNA was extracted (phenol/chloroform, ethanol precipitation with glycogen) and centrifuged (13,800*g*, 15 min, 4°C), and the pellet was resuspended in double-distilled water, from which a 2-μL aliquot was used for PCR (25-μL volumes, 32 to 35 cycles). Three independent biological replicates of ChIP assays were performed. The primers for ChIP-PCR are shown in Supplemental Data Set 6 online.

Stable *P. trichocarpa* Transgenic Plant Production and Verification of PtrSND1-B1's Direct Targets in Transgenic Differentiating Xylem

pBI121-35S-PtrSND1-B1 was constructed for the *P. trichocarpa* stable transformation. Briefly, the Ptr-*SND1-B1* cDNA sequence was amplified from *P. trichocarpa* xylem cDNA with primers Ptr-SND1-B1-1F/R (see Supplemental Data Set 6 online) and then inserted into pBI121 using *Bam*HI-*Sac*I sites. The construct was sequence confirmed and mobilized into *Agrobacterium tumefaciens* strain C58 to transform *P. trichocarpa* as described (Song et al., 2006; Lu et al., 2013). qRT-PCR was conducted as described above to quantify the transcript abundance of Ptr-*SND1-B1* and the selected direct targets of Ptr-SND1-B1 (Figure 9). Total RNA of SDX was extracted from three wild-type trees and transgenics (6 months old). The primers for qRT-PCR are listed in Supplemental Data Set 6 online.

Accession Numbers

The RNA-seq data discussed in this study can be found in National Center for Biotechnology Information's Gene Expression Omnibus through GEO Series accession number GSE49911 (nonpermanent URL for reviewer access: http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?token=xfwthkmwciwigvkandacc=GSE49911). Gene model names (*P. trichocarpa* genome release v2.2; Phytozome V7.0; http://www.phytozome.com) of the genes used in this work are listed in the supplemental tables and data sets online, and the gene model name of Ptr-*SND1-B1* is POPTR_0014s10060.1.

Supplemental Data

The following materials are available in the online version of this article.

Supplemental Figure 1. Procedures of Xylem Protoplast Isolation.

Supplemental Figure 2. Transient Expression of *sGFP* in *P. trichocarpa* SDX Protoplasts.

Supplemental Figure 3. RNA Quality of SDX Protoplasts.

Supplemental Figure 4. Protein Gel Blots for Antibody Specificity.

Supplemental Figure 5. ChIP-PCR Assays on the Promoter Sequence (2000 to 500 bp) of the Selected Genes.

Supplemental Table 1. Significant Gene Ontology Functional Classes of the Ptr-SND1-B1–Induced DEGs.

Supplemental Table 2. Genes of the Major GO Classes in the g: Profiler Functional Enrichment Analysis.

Supplemental Table 3. Seventy-Six Direct Targets of Ptr-SND1-B1.

Supplemental Methods 1. SDX Protoplast Isolation, Transfection, and Transfection Efficiency.

Supplemental Methods 2. Probability-Based Identification of Ptr-*SND1-B1*–Responsive Genes.

Supplemental Methods 3. Inference of Direct Target Genes of Ptr-*SND1-B1* Using GGM.

Supplemental Methods 4. ChIP Assay of *P. trichocarpa* Differentiating Xylem.

Supplemental Data Set 1. 178 Differentially Expressed Genes List.

Supplemental Data Set 2. Lists of Ptr-*SND1-B1*–Responsive Genes and Interfered Genes.

Supplemental Data Set 3. Indirect Targets Excluded by Computational and Time-Dependent Methods (Alone and in Combination; Three Sheets).

Supplemental Data Set 4. List of Ptr-*MYB002*– and Ptr-*MYB021*–Responsive Genes.

Supplemental Data Set 5. List of Interfered Genes List of Ptr-*MYB002* and Ptr-*MYB021*.

Supplemental Data Set 6. Primer List.

ACKNOWLEDGMENTS

We thank Jung-Ying Tzeng (North Carolina State University) for suggestions with statistical analysis and Choun-Sea Lin and Fu-Hui Wu (both from Academia Sinica, Taiwan) for suggestions with protoplast isolation. This work was supported by the Office of Science (Biological and Environmental Research), Department of Energy Grant DE-SC000691 (to V.L.C.). We also thank the support of the North Carolina State University Jordan Family Distinguished Professor Endowment.

AUTHOR CONTRIBUTIONS

Y.-C.L., W.L., Y.-H.S., H.W., R.R.S., and V.L.C. designed research. Y.-C.L., W.L., Y.-H.S., S.K., H.W., and Q.L. performed research. Y.-C.L., W.L., Y.-H.S., S.K., H.W., S.T.-A., and V.L.C. analyzed data. Y.-C.L., W.L., Y.-H.S., H.W., R.R.S., and V.L.C. wrote the article.

Received August 21, 2013; revised October 13, 2013; accepted October 23, 2013; published November 26, 2013.

REFERENCES

Abel, S., and Theologis, A. (1994). Transient transformation of *Arabidopsis* leaf protoplasts: A versatile experimental system to study gene expression. Plant J. **5:** 421–427.

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

Bassel, G.W., Gaudinier, A., Brady, S.M., Hennig, L., Rhee, S.Y., and De Smet, I. (2012). Systems analysis of plant functional, transcriptional, physical interaction, and metabolic networks. Plant Cell **24:** 3859–3875.

Becskei, A., and Serrano, L. (2000). Engineering stability in gene networks by autoregulation. Nature **405:** 590–593.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. J. R. Statist. Soc. B **57:** 289–300.

Berthet, S., Demont-Caulet, N., Pollet, B., Bidzinski, P., Cézard, L., Le Bris, P., Borrega, N., Hervé, J., Blondet, E., Balzergue, S., Lapierre, C., and Jouanin, L. (2011). Disruption of *LACCASE4* and *17* results in tissue-specific alterations to lignification of *Arabidopsis thaliana* stems. Plant Cell **23:** 1124–1137.

Bhardwaj, N., Yan, K.K., and Gerstein, M.B. (2010). Analysis of diverse regulatory networks in a hierarchical context shows consistent tendencies for collaboration in the middle levels. Proc. Natl. Acad. Sci. USA **107:** 6841–6846.

Birnbaum, K., Shasha, D.E., Wang, J.Y., Jung, J.W., Lambert, G.M., Galbraith, D.W., and Benfey, P.N. (2003). A gene expression map of the *Arabidopsis* root. Science **302:** 1956–1960.

Chen, H.C., Li, Q., Shuford, C.M., Liu, J., Muddiman, D.C., Sederoff, R.R., and Chiang, V.L. (2011). Membrane protein complexes catalyze both 4- and 3-hydroxylation of cinnamic acid derivatives in monolignol biosynthesis. Proc. Natl. Acad. Sci. USA **108:** 21253–21258.

Chen, K., and Rajewsky, N. (2007). The evolution of gene regulation by transcription factors and microRNAs. Nat. Rev. Genet. **8:** 93–103.

Cheng, C., Yan, K.K., Hwang, W., Qian, J., Bhardwaj, N., Rozowsky, J., Lu, Z.J., Niu, W., Alves, P., Kato, M., Snyder, M., and Gerstein, M. (2011). Construction and analysis of an integrated regulatory network derived from high-throughput sequencing data. PLoS Comput. Biol. **7:** e1002190.

Chiu, W., Niwa, Y., Zeng, W., Hirano, T., Kobayashi, H., and Sheen, J. (1996). Engineered GFP as a vital reporter in plants. Curr. Biol. **6:** 325–330.

Chupeau, M.C., Granier, F., Pichon, O., Renou, J.P., Gaudin, V., and Chupeau, Y. (2013). Characterization of the early events leading to totipotency in an *Arabidopsis* protoplast liquid culture by temporal transcript profiling. Plant Cell **25:** 2444–2463.

Cocking, E.C. (1972). Plant cell protoplasts-isolation and development. Annu. Rev. Plant Physiol. **23:** 29–50.

Davey, M.R., Anthony, P., Power, J.B., and Lowe, K.C. (2005). Plant protoplasts: Status and biotechnological perspectives. Biotechnol. Adv. **23:** 131–171.

Davidson, E.H. (2001). Genomic Regulatory Systems: Development and Evolution. (San Diego, CA: Academic Press).

Demura, T., and Ye, Z.H. (2010). Regulation of plant biomass production. Curr. Opin. Plant Biol. **13:** 299–304.

Edwards, D. (2000). Introduction to Graphical Modelling, 2nd ed. (New York: Springer Verlag).

Ernst, H.A., Olsen, A.N., Larsen, S., and Lo Leggio, L. (2004). Structure of the conserved domain of ANAC, a member of the NAC family of transcription factors. EMBO Rep. **5:** 297–303.

Evert, R.R. (2006). Esau's Plant Anatomy: Meristems, Cells, and Tissues of the Plant Body: Their Structure, Function, and Development, 3rd ed. (Hoboken, NJ: John Wiley & Sons).

Faraco, M., Di Sansebastiano, G.P., Spelt, K., Koes, R.E., and Quattrocchio, F.M. (2011). One protoplast is not the other! Plant Physiol. **156:** 474–478.

Farnham, P.J. (2009). Insights from genomic profiling of transcription factors. Nat. Rev. Genet. **10:** 605–616.

Fisher, R.A. (1922). On the interpretation of $\chi 2$ from contingency tables, and the calculation of P. J. R. Statist. Soc. **85:** 87–94.

Fisher, R.A. (1925). Statistical Methods for Research Workers. (Edinburgh, UK: Oliver & Boyd).

Fujii, H., Chinnusamy, V., Rodrigues, A., Rubio, S., Antoni, R., Park, S.Y., Cutler, S.R., Sheen, J., Rodriguez, P.L., and Zhu, J.K. (2009). In vitro reconstitution of an abscisic acid signalling pathway. Nature **462:** 660–664.

Gerstein, M.B., et al. (2010). Integrative analysis of the *Caenorhabditis elegans* genome by the modENCODE project. Science **330:** 1775–1787.

Gomez-Maldonado, J., Crespillo, R., Avila, C., and Canovas, F.M. (2001). Efficient preparation of maritime pine (*Pinus pinaster*) protoplasts suitable for transgene expression analysis. Plant Mol. Biol. Rep. **19:** 361–366.

Guo, J., Morrell-Falvey, J.L., Labbé, J.L., Muchero, W., Kalluri, U.C., Tuskan, G.A., and Chen, J.G. (2012). Highly efficient isolation of *Populus* mesophyll protoplasts and its application in transient expression assays. PLoS ONE **7:** e44908.

He, P., Shan, L., and Sheen, J. (2007). The use of protoplasts to study innate immune responses. Methods Mol. Biol. **354:** 1–9.

Heisig, J., Weber, D., Englberger, E., Winkler, A., Kneitz, S., Sung, W.K., Wolf, E., Eilers, M., Wei, C.L., and Gessler, M. (2012). Target gene analysis by microarrays and chromatin immunoprecipitation identifies HEY proteins as highly redundant bHLH repressors. PLoS Genet. **8:** e1002728.

Hobert, O. (2008). Gene regulation by transcription factors and microRNAs. Science **319:** 1785–1786.

Horak, C.E., Luscombe, N.M., Qian, J., Bertone, P., Piccirrillo, S., Gerstein, M., and Snyder, M. (2002). Complex transcriptional circuitry at the G1/S transition in *Saccharomyces cerevisiae*. Genes Dev. **16:** 3017–3033.

Hu, R., Qi, G., Kong, Y., Kong, D., Gao, Q., and Zhou, G. (2010). Comprehensive analysis of NAC domain transcription factor gene family in *Populus trichocarpa*. BMC Plant Biol. **10:** 145.

Huang, W., Pérez-García, P., Pokhilko, A., Millar, A.J., Antoshechkin, I., Riechmann, J.L., and Mas, P. (2012). Mapping the core of the *Arabidopsis* circadian clock defines the network structure of the oscillator. Science **336:** 75–79.

Huddy, S.M., Meyers, A.E., and Coyne, V.E. (2012). Protoplast isolation optimization and regeneration of cell wall in *Gracilaria gracilis* (Gracilariales, Rhodophyta). J. Appl. Phycol. **25:** 433–443.

Inoue, T., Higuchi, M., Hashimoto, Y., Seki, M., Kobayashi, M., Kato, T., Tabata, S., Shinozaki, K., and Kakimoto, T. (2001). Identification of CRE1 as a cytokinin receptor from *Arabidopsis.* Nature **409:** 1060–1063.

Jothi, R., Balaji, S., Wuster, A., Grochow, J.A., Gsponer, J., Przytycka, T.M., Aravind, L., and Babu, M.M. (2009). Genomic analysis reveals a tight link between transcription factor dynamics and regulatory network architecture. Mol. Syst. Biol. **5:** 294.

Kaufmann, K., Muiño, J.M., Jauregui, R., Airoldi, C.A., Smaczniak, C., Krajewski, P., and Angenent, G.C. (2009). Target genes of the MADS transcription factor SEPALLATA3: Integration of developmental and hormonal pathways in the *Arabidopsis* flower. PLoS Biol. **7:** e1000090.

Kaufmann, K., Muiño, J.M., Østerås, M., Farinelli, L., Krajewski, P., and Angenent, G.C. (2010). Chromatin immunoprecipitation (ChIP)

of plant transcription factors followed by sequencing (ChIP-SEQ) or hybridization to whole genome arrays (ChIP-CHIP). Nat. Protoc. **5:** 457–472.

Kim, T.H., and Ren, B. (2006). Genome-wide analysis of protein-DNA interactions. Annu. Rev. Genomics Hum. Genet. **7:** 81–102.

Kim, W.C., Ko, J.H., and Han, K.H. (2012a). Identification of a *cis*-acting regulatory motif recognized by MYB46, a master transcriptional regulator of secondary wall biosynthesis. Plant Mol. Biol. **78:** 489–501.

Kim, W.C., Ko, J.H., Kim, J.Y., Kim, J.M., Bae, H.J., and Han, K.H. (2012b). MYB46 directly regulates the gene expression of secondary wall-associated cellulose synthases in *Arabidopsis*. Plant J. **73:** 26–36.

Kubo, M., Udagawa, M., Nishikubo, N., Horiguchi, G., Yamaguchi, M., Ito, J., Mimura, T., Fukuda, H., and Demura, T. (2005). Transcription switches for protoxylem and metaxylem vessel formation. Genes Dev. **19:** 1855–1860.

Kumar, S.V., Lucyshyn, D., Jaeger, K.E., Alós, E., Alvey, E., Harberd, N.P., and Wigge, P.A. (2012). Transcription factor PIF4 controls the thermosensory activation of flowering. Nature **484:** 242–245.

Larkin, P.J. (1976). Purification and viability determinations of plant protoplasts. Planta **128:** 213–216.

Levine, M., and Tjian, R. (2003). Transcription regulation and animal diversity. Nature **424:** 147–151.

Li, Q., Lin, Y.C., Sun, Y.H., Song, J., Chen, H., Zhang, X.H., Sederoff, R.R., and Chiang, V.L. (2012). Splice variant of the SND1 transcription factor is a dominant negative of SND1 members and their regulation in *Populus trichocarpa*. Proc. Natl. Acad. Sci. USA **109:** 14699–14704.

Li, W., Liu, H., Cheng, Z.J., Su, Y.H., Han, H.N., Zhang, Y., and Zhang, X.S. (2011). DNA methylation and histone modifications regulate *de novo* shoot regeneration in *Arabidopsis* by modulating *WUSCHEL* expression and auxin signaling. PLoS Genet. **7:** e1002243.

Lu, S., et al. (2013). Ptr-miR397a is a negative regulator of laccase genes affecting lignin content in *Populus trichocarpa*. Proc. Natl. Acad. Sci. USA **110:** 10848–10853.

Manders, G., Dossantos, A.V.P., Vaz, F.B.D., Davey, M.R., and Power, J.B. (1992). Transient gene-expression in electroporated protoplasts of *Eucalyptus citriodora* Hook. Plant Cell Tissue Organ Cult. **30:** 69–75.

Merkle, S.A., and Dean, J.F. (2000). Forest tree biotechnology. Curr. Opin. Biotechnol. **11:** 298–302.

Moreno-Risueno, M.A., Busch, W., and Benfey, P.N. (2010). Omics meet networks - Using systems approaches to infer regulatory networks in plants. Curr. Opin. Plant Biol. **13:** 126–131.

Müller, C.W. (2001). Transcription factors: Global and detailed views. Curr. Opin. Struct. Biol. **11:** 26–32.

Niu, W., et al. (2011). Diverse transcription factor binding features revealed by genome-wide ChIP-seq in *C. elegans*. Genome Res. **21:** 245–254.

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

Pitzschke, A., and Persak, H. (2012). Poinsettia protoplasts - A simple, robust and efficient system for transient gene expression studies. Plant Methods **8:** 14.

Pruneda-Paz, J.L., Breton, G., Para, A., and Kay, S.A. (2009). A functional genomics approach reveals CHE as a component of the *Arabidopsis* circadian clock. Science **323:** 1481–1485.

Pyo, H., Demura, T., and Fukuda, H. (2007). TERE; a novel *cis*-element responsible for a coordinated expression of genes related to programmed cell death and secondary wall formation during differentiation of tracheary elements. Plant J. **51:** 955–965.

Reimand, J., Arak, T., and Vilo, J. (2011). g:Profiler—A web server for functional interpretation of gene lists (2011 update). Nucleic Acids Res. **39** (Web Server issue)**:** W307–W315.

Riazunnisa, K., Padmavathi, L., Scheibe, R., and Raghavendra, A.S. (2007). Preparation of *Arabidopsis* mesophyll protoplasts with high rates of photosynthesis. Physiol. Plant. **129:** 879–886.

Riechmann, J.L., et al. (2000). *Arabidopsis* transcription factors: Genome-wide comparative analysis among eukaryotes. Science **290:** 2105–2110.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics **26:** 139–140.

Roy, S., et al; modENCODE Consortium (2010). Identification of functional elements and regulatory circuits by *Drosophila* modENCODE. Science **330:** 1787–1797.

Sheen, J. (2001). Signal transduction in maize and Arabidopsis mesophyll protoplasts. Plant Physiol. **127:** 1466–1475.

Shi, R., and Chiang, V.L. (2005). Facile means for quantifying microRNA expression by real-time PCR. Biotechniques **39:** 519–525.

Solomon, M.J., Larsen, P.L., and Varshavsky, A. (1988). Mapping protein-DNA interactions in vivo with formaldehyde: evidence that histone H4 is retained on a highly transcribed gene. Cell **53:** 937–947.

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

Sonntag, K., Ruge-Wehling, B., and Wehling, P. (2008). Protoplast isolation and culture for somatic hybridization of *Lupinus angustifolius* and *L. subcarnosus*. Plant Cell Tissue Organ Cult. **96:** 297–305.

Sun, J., et al. (2009). NaCl-induced alternations of cellular and tissue ion fluxes in roots of salt-resistant and salt-sensitive poplar species. Plant Physiol. **149:** 1141–1153.

Tan, B., Xu, M., Chen, Y., and Huang, M. (2013). Transient expression for functional gene analysis using *Populus* protoplasts. Plant Cell Tissue Organ Cult. **114:** 11–18.

Tang, R.J., Liu, H., Bao, Y., Lv, Q.D., Yang, L., and Zhang, H.X. (2010). The woody plant poplar has a functionally conserved salt overly sensitive pathway in response to salinity stress. Plant Mol. Biol. **74:** 367–380.

Teulieres, C., Grima-Pettenati, J., Curie, C., Teissie, J., and Boudet, A.M. (1991). Transient foreign gene expression in polyethylene/glycol treated or electropulsated *Eucalyptus gunnii* protoplasts. Plant Cell Tissue Organ Cult. **25:** 125–132.

Thibaud-Nissen, F., Wu, H., Richmond, T., Redman, J.C., Johnson, C., Green, R., Arias, J., and Town, C.D. (2006). Development of *Arabidopsis* whole-genome microarrays and their application to the discovery of binding sites for the TGA2 transcription factor in salicylic acid-treated plants. Plant J. **47:** 152–162.

Thorpe, T.A. (2007). History of plant tissue culture. Mol. Biotechnol. **37:** 169–180.

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: Discovering splice junctions with RNA-Seq. Bioinformatics **25:** 1105–1111.

Wang, H.Z., Zhao, Q., Chen, F., Wang, M.Y., and Dixon, R.A. (2011). NAC domain function and transcriptional control of a secondary cell wall master switch. Plant J. **68:** 1104–1114.

Whittaker, J. (1990). Graphical Models in Applied Multivariate Statistics. (New York: John Wiley and Sons).

Wray, G.A., Hahn, M.W., Abouheif, E., Balhoff, J.P., Pizer, M., Rockman, M.V., and Romano, L.A. (2003). The evolution of transcriptional regulation in eukaryotes. Mol. Biol. Evol. **20:** 1377–1419.

Wu, F.H., Shen, S.C., Lee, L.Y., Lee, S.H., Chan, M.T., and Lin, C.S. (2009). Tape-*Arabidopsis* Sandwich: A simpler *Arabidopsis* protoplast isolation method. Plant Methods **5:** 16.

Xie, Q., Frugis, G., Colgan, D., and Chua, N.H. (2000). *Arabidopsis* NAC1 transduces auxin signal downstream of TIR1 to promote lateral root development. Genes Dev. **14:** 3024–3036.

Yoo, S.D., Cho, Y.H., and Sheen, J. (2007). *Arabidopsis* mesophyll protoplasts: A versatile cell system for transient gene expression analysis. Nat. Protoc. **2:** 1565–1572.

Yu, H., and Gerstein, M. (2006). Genomic analysis of the hierarchical structure of regulatory networks. Proc. Natl. Acad. Sci. USA **103:** 14724–14731.

Zhang, Y., Su, J., Duan, S., Ao, Y., Dai, J., Liu, J., Wang, P., Li, Y., Liu, B., Feng, D., Wang, J., and Wang, H. (2011). A highly efficient rice green tissue protoplast system for transient gene expression and studying light/chloroplast-related processes. Plant Methods **7:** 30.

Zheng, Y., Ren, N., Wang, H., Stromberg, A.J., and Perry, S.E. (2009). Global identification of targets of the *Arabidopsis* MADS domain protein AGAMOUS-Like15. Plant Cell **21:** 2563–2577.

Zhong, R., Demura, T., and Ye, Z.H. (2006). SND1, a NAC domain transcription factor, is a key regulator of secondary wall synthesis in fibers of *Arabidopsis*. Plant Cell **18:** 3158–3170.

Zhong, R., Lee, C., and Ye, Z.H. (2010). Global analysis of direct targets of secondary wall NAC master switches in *Arabidopsis*. Mol. Plant **3:** 1087–1103.

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

Zhou, V.W., Goren, A., and Bernstein, B.E. (2011). Charting histone modifications and the functional organization of mammalian genomes. Nat. Rev. Genet. **12:** 7–18.