



# Protoplast isolation and transcriptome analysis of developing xylem in *Pinus massoniana* (Pinaceae)

Tengfei Shen<sup>1,2</sup> · Mengxuan Xu<sup>2</sup> · Haoran Qi<sup>2</sup> · Yuanheng Feng<sup>1</sup> · Zhangqi Yang<sup>1</sup> · Meng Xu<sup>2</sup>

Received: 31 August 2021 / Accepted: 19 November 2021 / Published online: 26 November 2021  
© The Author(s), under exclusive licence to Springer Nature B.V. 2021

## Abstract

**Background** With active physiological and biochemical activities, tissue-specific protoplasts from cambial derivatives, could serve as a specific source for information on xylogenesis for softwood species resistant to stable genetic transformation and lacking available mutants.

**Methods and results** In this study, protoplasts were isolated from developing xylem of the Chinese red pine, *Pinus massoniana*, by enzymolysis. High-quality RNAs were extracted from developing xylem and their protoplasts for constructing transcriptome libraries. Using Illumina HiSeq 2500 PE150 platform, a total of 362,328,426 clean paired-end reads (54.35G) were generated from multiple cDNA libraries and assembled into 146,422 unigenes. The transcriptome data were further analysed to identify 1567 differentially expressed genes (DEGs) between the isolated protoplasts and developing xylem of *P. massoniana* (Masson pine), 1126 DEGs were upregulated in protoplasts relative to developing xylem cells and 441 were downregulated. Most of the differentially expressed genes in biological process terms are related to plant response, which may be due to the response to cell wall removal. Further, the expression pattern of 71 unigenes involved in lignin biosynthesis was verified by RNA-seq.

**Conclusions** This study is the first to report the transcriptome profiles of the developing xylem and its protoplasts of coniferous trees, which provide a new perspective and valuable resource for tracking transcriptional regulatory events in wood formation of Masson pine.

**Keywords** Masson pine · Developing xylem · Protoplasts · Cell wall · Transcriptome

## Introduction

Wood, as a sustainable and renewable resource, plays an increasingly essential role in energy support and industrial production [1]. It is of great significance to better understand the regulatory mechanism of wood formation to improve wood properties and utilization efficiency. Wood formation, namely xylogenesis, is a sequential and complex process that

consists of multiple major steps, i.e., cambial cell mitosis, differentiation, secondary wall biosynthesis and lignification, which are regulated by various environmental factors and endogenous hormonal signals [2]. Although a great deal of work has been devoted to exploring the biosynthesis and assembly of secondary cell walls in trees, our understanding of xylogenesis is still incomplete, especially in gymnosperms [3–6].

Because of its clear genetic background and profound theoretical basis, *Arabidopsis thaliana* gave us a preliminary understanding of secondary cell wall thickening [7]. There are many limitations in *Arabidopsis* as a study system for xylogenesis. Compared with herbaceous plants, perennial forest trees have innately vigorous vascular cambium that produces secondary xylem and phloem, which makes them an indispensable component of the terrestrial ecosystem. Ecologically dominant and economically important hardwood species, such as *Populus* spp. and *Eucalyptus* spp., have been adopted as important systems for studying

Tengfei Shen and Mengxuan Xu contributed equally to this work.

✉ Meng Xu  
xum@njfu.edu.cn

<sup>1</sup> Guangxi Key Laboratory of Superior Timber Trees Resource Cultivation, Guangxi Forestry Research Institute, Nanning 530002, China

<sup>2</sup> Co-Innovation Center for Sustainable Forestry in Southern China, Key Laboratory of Forest Genetics and Biotechnology Ministry of Education, Nanjing Forestry University, Nanjing 210037, China

wood formation, because of their rapid growth, small genome size and mature genetic transformation technology [8]. The wood of broad-leaved trees is known as hardwood while that of conifers is softwood. Along with the progress and application of multi-omics, recent studies on hardwood species are gradually deepening, with a sharpening understanding of xylogenesis in *Populus* and *Eucalyptus* [9–12]. Similarly, the xylogenesis of conifers has also been widely studied [13], and a large number of studies have focused on the anatomical structure and physical mechanical properties of softwood [14, 15]. Softwood species not only have a large and complicated genome but also lack an efficient system for genetic studies. These disadvantages have greatly hindered analysis of their developmental mechanisms in genetic regulatory hierarchies.

Cell- or tissue-specific plant protoplasts have active physiological and biochemical activities, which are sufficient to study various biological processes from physiology to signal transduction [10]. Protoplasts from wood-forming tissues serve as a specific source for information on xylogenesis. Using stem-differentiating xylem of poplar and its protoplasts as materials, protoplast transient transformation and chromatin immunoprecipitation assays were successfully conducted for the discovery and validation of specific regulatory networks in wood formation [11, 12]. These studies have confirmed that protoplasts isolated from cambial derivatives provide valuable materials for biochemical, physiological, and genetic experiments aimed at dissecting miscellaneous regulatory networks of wood formation.

As an ecologically and economically important softwood species, *Pinus massoniana* is widely distributed in Central and South China, with distinguishing features of fast growth, straight trunk shape, and resistance to drought and barren. Its wood, as an important source of woody biomass, is broadly used for construction, pulp and furniture. However, information on the molecular mechanism of softwood formation is still limited [16]. In this study, a simple and rapid protoplast isolation system was established for developing xylem in *P. massoniana*. The viability of isolated protoplasts was more than 90% as assessed by staining with fluorescein diacetate (FDA). Then, the first transcriptome profiles of intact xylem cells (with cell wall) and protoplasts of developing xylem in coniferous trees were revealed. The removal of cell wall triggered a series of stress responses, and most of differentially expressed unigenes were related to the stress response pathway. Even more intriguingly, the transcriptome consistency between intact xylem cells and their protoplasts was more than 90%, and a large number of developing xylem-specific unigenes were identified. These results suggest that protoplasts from developing xylem an ideal material for transient expression assays to track the transcriptional regulatory events associated with xylogenesis. Besides, the application

of protoplasts in signal transduction of plant stress response should be cautious.

## Materials and methods

### Plant materials and protoplasts extraction

Two Masson pines of different genotypes growing on the campus of Nanjing Forestry University, approximately 3 m in height and 10 cm in diameter breast height (DBH), were sources of healthily growing semi-lignified branches. Semi-lignified annual branches of similar size were cut and transferred back to the laboratory (within 30 min), and the bark was peeled to expose the youngest developing xylem. Some of the debarked branches were quickly put into the mixed enzyme solution to isolate protoplasts, and the others were frozen immediately in liquid nitrogen and stored at  $-80^{\circ}\text{C}$ .

Protoplasts of developing xylem in Masson pine were isolated according to a modified method as reported in previous publications [11, 17]. Briefly, the debarked branches exposing developing xylem were cut into 5–10 cm sections, and the sections were immediately transferred into freshly prepared enzyme solution and incubated in the dark at  $28^{\circ}\text{C}$  for 4 h. The new enzyme cocktail consisted of 2% (W/V) cellulase (Sigma, C2730), 1% (W/V) pectinase (Sigma, P2611), 5 M D-mannitol (Sigma, M4125), 20 mM MES (2-(N-morpholino) ethanesulfonic acid), 20 mM KCl, 10 mM  $\text{CaCl}_2$ , and 1% BSA (albumin from bovine serum). After adding W5 solution (2 mM MES at pH 5.7, 154 mM NaCl, 125 mM  $\text{CaCl}_2$ , 5 mM KCl) to stop digestion, protoplasts were released by filtering through nylon meshes, and were collected by centrifugation. The viability of the isolated protoplasts was determined using the fluorescein diacetate (FDA) staining method described previously [18]. In addition, the cross sections of the branches before/after peeling and digestion were observed by scanning electron microscopy (SEM).

### RNA-seq library construction and sequencing

Total RNA was extracted from frozen samples by using an RNeasy Plant Mini Kit (Qiagen, Hilden, Germany), and the quality of total RNA was checked by agarose gel electrophoresis and using an Agilent 2100 Bioanalyzer. The protoplasts suspended in W5 solution (1–2 ml,  $10^6$ – $10^7$ /ml) were centrifuged to remove the supernatant, and then the RNA extraction solution was added according to the manufacturer's protocol without grinding. A total of 4 ml protoplast suspensions ( $10^6$ – $10^7$ /ml) for RNA extraction were prepared for each sample. After quantification and qualification, a total of 1.5  $\mu\text{g}$  RNA per sample of high-quality RNA was used for cDNA library construction. Two cDNA libraries

of developing xylem were named PM1\_X and PM2\_2, and those of protoplasts isolated from developing xylem were named PM1\_XP and PM2\_XP. Following the manufacturer's recommendations, four libraries were generated using the NEBNext® Ultra™ RNA Library Prep Kit for Illumina® (NEB, Ipswich, MA, USA), and library quality was assessed on the Agilent 2100 Bioanalyzer system. The qualified libraries were normalized and merged with equal volume (1.5 ng/μl) for the subsequent cluster generation. The indexed clusters were proceeded on a cBot Cluster Generation System using TruSeq PE Cluster Kit v3-cBot-HS (Illumina) following the manufacturer's instructions. The prepared libraries were sequenced using Illumina the HiSeq 2500 platform with 150 bp paired-end reads.

### RNA-seq data processing

Raw reads in FASTQ format obtained by CASAVA base-calling were filtered by in-house Perl scripts to remove reads containing adaptors and poly-N, and low-quality reads. The cutoffs were as follows: (1) removed the reads with adapter; (2) removed the reads with the percentage of N more than 0.1% (N means that the base cannot be determined); (3) removed the low quality reads (the base number of Qphred  $\leq 20$  accounts for more than 50% of the total reads). While high-quality clean reads were obtained, the Q value and GC-content level of clean reads were calculated, and the next analyses were all based on the clean data. Because there was no reference genome in Masson pine, de novo assembly was accomplished by Trinity v2.4.0 with min\_kmer\_cov set to 3 and other parameters set to their defaults [19]. Unigenes, as the longest transcripts of assembled transcripts, were usually selected for the gene functional annotation and other analyses. Unigenes were blasted against Nr (NCBI non-redundant protein sequences), Swiss-Prot (a manually annotated and reviewed protein sequence database) and KOG/COG (Clusters of Orthologous Groups of proteins) databases by diamond v0.8.22 with an e-value of  $1e-5$  and sensitivity mode of more sensitivity [20]. Unigenes were matched against the Nt (NCBI non-redundant nucleotide sequence) database by NCBI BLAST 2.2.28 + with an e-value of  $1e-5$  [21]. Annotation of unigenes in KEGG (Kyoto Encyclopedia of Genes and Genomes), Pfam (Protein family) and GO (Gene Ontology) databases was performed by KAAS r140224 [22], hmmscan [23] and Blast2GO [24] with e-values of  $1e-10$ , 0.01 and  $1e-6$ , respectively.

### Identification of differentially expressed unigenes

Read counts per unigene were obtained by mapping back clean data onto the assembled transcriptome, and the gene expression level was estimated by RSEM v1.2.15 with the parameters of bowtie2 at a mismatch of 0 [25]. All the

unigenes were used for PCA analysis. Based on FPKM (expected number of fragments per kilobase of transcript sequence per millions base pairs sequenced) normalized from read counts, differentially expressed genes (DEGs) between protoplasts and developing xylem were obtained by the DESeq2 R package (1.12.1) with  $\text{padj} < 0.05$  and  $\text{llog}_2(\text{foldchange}) > 1$  [26]. GO enrichment analysis of DEGs was performed by clusterprofiler v3.14.3 with the corrected p-value  $< 0.05$  [27]. Enriched GO terms and hierarchical relationships were visualized by topGO 2.10.0 [28]. KOBAS v2.0.12 was used to calculate the statistical enrichment of DEGs and classify them into terms of KEGG pathways [29]. Enrichment analysis of KEGG pathway was performed by clusterprofiler v3.14.3, and the cut-off threshold was the same as that for GO enrichment.

### qRT-PCR validation

Total RNA from various tissues was normalized and reverse transcribed into cDNA using Prime Script™ RT Master Mix Kit (TaKaRa). Specific primers for quantitative real-time RT-PCR (qRT-PCR) were designed by an online biological tool (<https://www.genscript.com/tools/pcr-prime-rs-designer/>). In accordance with the manufacturer's protocol, qRT-PCR for ten unigenes and one reference gene was performed on an ABI ViiA 7 Real-Time PCR platform (ABI, Carlsbad, USA) with FastStart Universal SYBR Green Master with ROX, and we did three technical replicates for each gene. Relative expression abundance was estimated by the  $2^{-\Delta\Delta C_t}$  method with *UBI4* as the endogenous control [30, 31].

## Results

### Protoplast extraction and RNA-seq

By incubating freshly debarked branches with young developing xylem in an enzyme cocktail (Fig. 1), protoplasts were isolated from developing xylem exposed by peeling the bark of actively growing trees. The sampling of developing xylem must occur in the vigorous growing stage of Masson pine, that is, during the active period of the vascular cambium. Developing xylem consists of ray and fusiform cells in softwood species, and ray cells are not only smaller than fusiform cells, but also more numerous [17]. Here, the isolated protoplasts represented two distinct cell types, the small diameter protoplasts (10–28 μm) from ray cells and the larger diameter protoplasts (30–55 μm) from fusiform cambial derivatives (Supplementary Figure S1), and the former were obviously more abundant than the latter. By staining with FDA [18], under fluorescence microscope, the active protoplasts showed green fluorescence. Statistics of cells

**Fig. 1** Isolation of protoplasts from developing xylem in Masson pine. **A** A semi-lignified branch. **B** A debarked branch. **C** Debarked branches in the freshly prepared enzyme solution. **D** The cross sections of the branches before (a)/after (b) peeling and digestion (c) observed by scanning electron microscope micrograph



with fluorescence signals in different visual fields revealed that the vitality of the protoplasts from developing xylem was more than 90%. High-quality RNAs with RIN greater than 7 were extracted from these developing xylem and protoplasts for constructing transcriptome libraries (Supplementary Figure S2). By Illumina paired-end sequencing, a total of 367,768,550 raw paired-end reads were generated from the four libraries, and all raw reads have been deposited to the NCBI Sequence Read Archive (SRR12596930, SRR12596931, SRR12596932 and SRR1259933). After filtering out low quality reads, 362,328,426 clean reads

(54.35G) were obtained for subsequent analysis (Table 1). Due to the lack of a high-quality reference genome in Masson pine, de novo assembly of the transcriptome was conducted. De novo assembly of these clean reads generated 185,829 transcripts with a mean length of 1096 bp and an N50 of 1892, and 146,422 unigenes with a mean length of 1319 and an N50 of 1990, respectively (Supplementary Table S1). Of the assembled transcripts, 108,509 (58.15%) were 500 bp or longer. Among the unigenes, 26.53%, 26.60%, 26.26% and 20.61% showed lengths of 200–500 bp, 501–1000 bp, 1001–2000 bp and > 2000 bp, respectively.

**Table 1** Summary statistics for Illumina HiSeq transcriptome data of Masson pine

Library	Raw reads	Clean reads	Clean Bases	Q30 (%)	Mapped reads (rate %)
PM1_X	84,134,844	82,767,314	12.42G	94.19	72,926,930 (88.11%)
PM2_X	96,511,700	95,290,124	14.29G	92.67	84,759,182 (88.95%)
PM1_XP	101,367,068	99,694,660	14.95G	93.94	90,667,510 (90.95%)
PM2_XP	85,754,938	84,576,328	12.69G	92.24	75,617,754 (89.41%)
Total	367,768,550	362,328,426	54.35G	–	–

## Annotation of unigenes

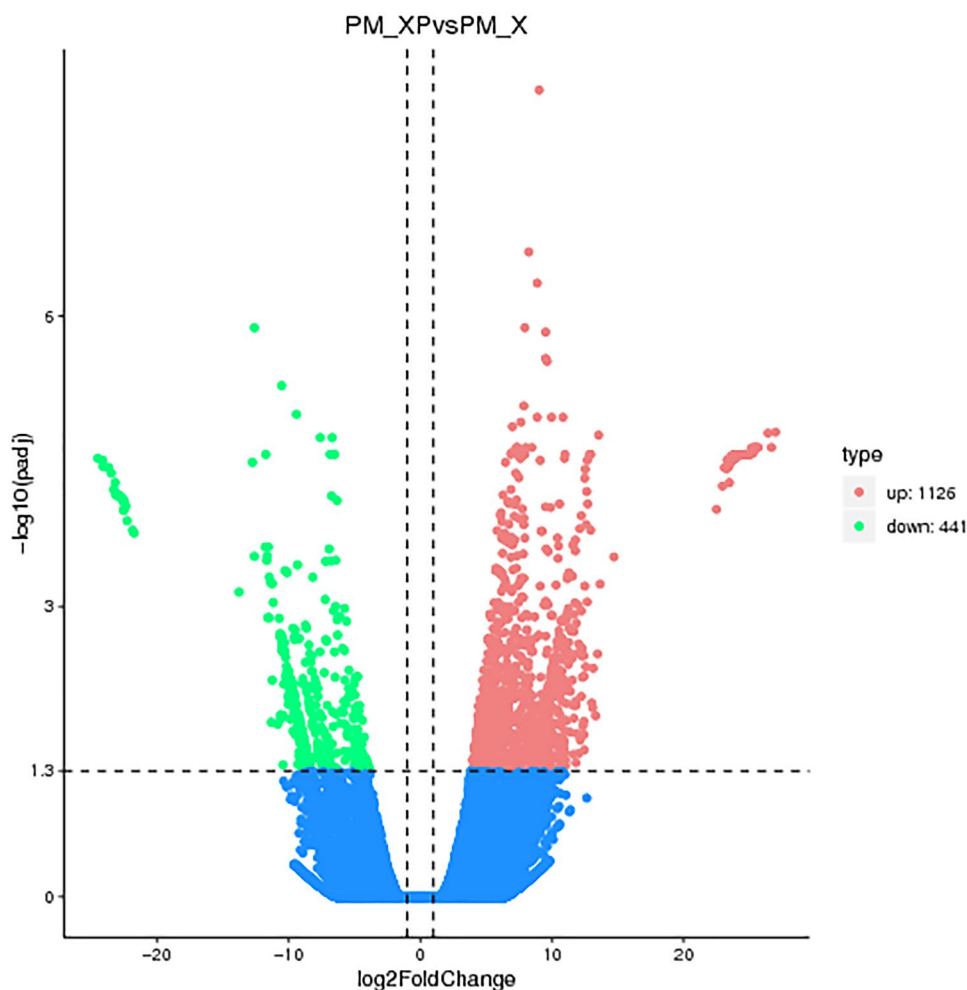
Unigenes were used as queries to identify homologous sequences in protein and nucleotide databases. Of the 146,422 unigenes, 117,455 (80.21%) were annotated in at least one database, and 7973 (5.44%) were successfully annotated in all seven databases. The results of annotating unigenes to the seven public databases show that 54,733 (37.38%), 95,643 (65.32%), 37,031 (25.29%), 80,247 (54.80%), 78,892 (53.87%), 78,892 (53.87%) and 23,545 (16.08) of the unigenes had significant BLAST matches in the Nr, Nt, KO, Swissport, Pfam, GO and KOG databases, respectively (Supplementary Figure S3A). To better show the numbers of annotations in each database and the overlap of their representation, we constructed a seven-way Venn diagram (Supplementary Figure S3B), which indicated that 7244 unigenes matched in all seven databases. There were 78,892 unigenes annotated by GO were hierarchically classified into 56 terms, which consisted of 26 terms of biological process (BP) terms, 20 cellular component (CC) terms, and 10 molecular function (MF) terms. Additionally, binding (MF), cellular process (BP), metabolic process (BP), catalytic activity (MF) and single-organism process (BP) were the top-five enriched terms in GO classification (Supplementary Figure S4A). For KOG analysis, there were 23,545 unigenes divided into 25 KOG groups, and the top three groups were [O] posttranslational modification, protein turnover, and chaperones; [R] general function prediction only; and [J] translation, ribosomal structure and biogenesis (Supplementary Figure S4B). A total of 37,031 unigenes were assigned to 19 KEGG metabolic pathways. The 19 pathways were subordinate to five branches (pathway hierarchy 1), including cellular processes (A, 1775 unigenes), environmental information processing (B, 1235 unigenes), genetic information processing (C, 7341 unigenes), metabolism (D, 17,942 unigenes), and organismal systems (E, 1970 unigenes). Carbohydrate metabolism (4129), translation (3078), folding, sorting and degradation (2429), overview (2329), and amino acid metabolism (2233) were the top-five enriched pathways (pathway hierarchy 2) in KEGG classification (Supplementary Figure S4C).

## Identification of differentially expressed unigenes and enrichment

For further analysis of gene expression levels, there were 72,926,930 (88.11%) clean reads in library PM1\_X, 84,759,182 (88.95%) in PM2\_X, 90,667,510 (90.95%) in PM1\_XP, and 75,617,754 (89.41%) PM2\_XP that were successfully mapped back onto the transcripts assembled (Table 1). After statistically analysing the mapping results, FPKM values were calculated for results under different experimental conditions, and the overall values and differences between samples were also evaluated through FPKM boxplots and FPKM interval distributions (Supplementary Figure S5). PCA analysis results show that the two protoplasts are well clustered together (Supplementary Figure S6). Based on the thresholds  $p < 0.05$  and  $|\log_2(\text{foldchange})| > 1$ , 1567 DEGs of PM\_XP vs PM\_X were obtained by DESeq2, including 1,126 upregulated DEGs and 441 downregulated DEGs (Fig. 2, supplementary Table S2). According to a standard FPKM value greater than 0.3, 75,623 unigenes were coexpressed in the two groups, 9,547 unigenes were specifically expressed in group PM\_X, and 5660 unigenes were specifically expressed in group PM\_XP (Supplementary Table S3).

The DEGs obviously enriched in each GO term are shown in the column chart. The upregulated DEGs were enriched in 50 terms, including 20 BP terms, 20 MF terms and 10 CC terms (Supplementary Figure S7A). The downregulated DEGs were enriched in 51 terms, including 20 BP terms, 13 CC terms and 18 MF terms (Supplementary Figure S7B). The top three enriched terms were metabolic process (GO:0008152), catalytic activity (GO:0003824), and single-organism process (GO:0044699), which were associated with 661, 586 and 508 DEGs, respectively. For directed acyclic graphs (DAGs) of DEGs, upregulated DEGs were significantly enriched in glutathione metabolic process (GO: 0006749, BP), glutathione peroxidase activity (GO: 0004602, MF) and Holliday junction (GO: 009,379, CC) terms, and downregulated DEGs were significantly enriched in lyase activity (GO: 0016829, MF), photosynthesis (GO: 0015979, BP) and photosystem (GO: 0009521, CC) related terms. DEGs were enriched in 90 KEGG pathways. The

**Fig. 2** Volcano plot of DEGs in masson pine



top five significantly enriched pathways were glutathione metabolism (57 DEGs), arachidonic acid metabolism (26 DEGs), protein processing in endoplasmic reticulum (72 DEGs), photosynthesis (23 DEGs) and photosynthesis-antenna proteins (15 DEGs), with  $p.adjust$  of  $7.75e-22$ ,  $3.22e-17$ ,  $4.53e-15$ ,  $6.13e-09$  and  $2.17e-08$ , respectively. Among them, the upregulated DEGs were significantly enriched in top five pathways of glutathione metabolism, protein processing in endoplasmic reticulum, arachidonic acid metabolism, selenocompound metabolism, and pyrimidine metabolism, while the downregulated DEGs were significantly enriched in pathways of glyoxylate and dicarboxylate metabolism; photosynthesis; photosynthesis-antenna proteins; carbon fixation in photosynthetic organisms; and glycine, serine and threonine metabolism (Supplementary Table S4).

To verify the accuracy of transcriptional differences between developing xylem (intact cells) and its protoplasts, the expression profiles of ten DEGs identified based on Illumina RNA-Seq data were analysed by qRT-PCR using specific primers (Supplementary Table S5). The ten DEGs

encoded galactinol synthase, glutathione transferase, pectin methylesterase, caffeic acid 3-O-methyltransferase, raffinose synthetase, ferulate 5-hydroxylase, trehalose 6-phosphate phosphatase, glycosyltransferases, glutathione peroxidase, and sucrose synthetase, which involved in secondary cell wall biosynthesis (Supplementary Table S5). The qRT-PCR results showed that the expression trends of nine genes (those other than Cluster-18410.55705) were basically the same in the two platforms (Fig. 3), indicating that the transcriptome data were accurate and effective for detecting transcriptome differences between the intact xylem cells and their protoplasts of *P. massoniana*.

### Analysis of PM\_X-specific genes

According to the FPKM value, 9,547 PM\_X-specific expressed unigenes were used to perform GO and KEGG enrichment analysis, and the results found those unigenes were significantly enriched in 20 GO terms (Supplementary Figure S8, supplementary Table S6,  $p.adjust < 0.05$ ). Under BP, unigenes were significantly

enriched in proteolysis involved in cellular protein catabolic process (GO:0051603,  $p.adjust < 0.01$ ), amino sugar metabolic process (GO:0006040,  $p.adjust < 0.01$ ), and N-acetylmannosamine metabolic process (GO:0006051,  $p.adjust = 0.01$ ). Under CC, unigenes were significantly enriched in photosystem II oxygen evolving complex (GO:0009654,  $p.adjust < 0.01$ ), extrinsic component of membrane (GO:0019898,  $p.adjust < 0.01$ ), and chloroplast (GO:0009507,  $p.adjust < 0.01$ ). Under MF, the top three significantly enriched terms were lyase activity (GO:0016829,  $p.adjust < 0.01$ ), threonine-type endopeptidase activity (GO:0004298,  $p.adjust < 0.01$ ), and heat shock protein binding (GO:0031072,  $p.adjust = 0.01$ ). The study found that xylem-specific genes in poplar were significantly enriched in the single organism process, localization, and growth of the BP subclass; membrane, membrane part, supramolecular fiber, extracellular region, cell junction, and symplast of the CC subclass; and transporter activity and electron carrier activity of the MF subclass. The xylem-specific genes in *Larix kaempferi* were significantly enriched in single organism process of the BP subclass; membrane part, membrane, and nucleoid of the CC subclass; and catalytic activity, nucleic acid binding transcription factor activity, and structural molecule activity of the MF subclass [32]. According to the results, the PM\_X-specific unigenes were involved in the biomolecular metabolism process. In contrast, genes in poplar and *L. kaempferi* are involved in plant growth and structural integrity of a complex, respectively.

9547 PM\_X-specific expressed unigenes of *P. massoniana* were significantly enriched in 20 KEGG terms (Supplementary Figure S9), such as Glyoxylate and dicarboxylate metabolism (ko00630,  $p.adjust < 0.01$ ), Pyruvate metabolism (ko00620,  $p.adjust < 0.01$ ), Carotenoid biosynthesis (ko00906,  $p.adjust < 0.01$ ), Aminoacyl-tRNA biosynthesis (ko00970,  $p.adjust < 0.01$ ), Zeatin biosynthesis (ko00908,  $p.adjust < 0.01$ ), Glycerolipid metabolism (ko00561,  $p.adjust < 0.01$ ), Glycine, serine and threonine metabolism (ko00260,  $p.adjust < 0.01$ ), and Diterpenoid biosynthesis (ko00904,  $p.adjust < 0.01$ ). 13,907 xylem-specific expressed genes of poplar were enriched in photosynthesis (ko00195), metabolic pathways (ko01100), and photosynthesis antenna proteins (ko00196). 2596 xylem-specific expressed unigenes of *L. kaempferi* were enriched in starch and sucrose metabolism (ko00500), flavonoid biosynthesis (ko00941), and metabolic pathways (ko01100). 52 upregulated genes in the xylem cells of *A. thaliana* are involved in signal transduction, transcriptional regulation, and cell wall metabolism [33]. The PM\_X-specific unigenes were found to be involved in biomolecular biosynthesis, whereas genes in poplar, *L. kaempferi*, and *A. thaliana* were found to be involved in photosynthesis, starch and sucrose metabolism, and cell wall biosynthesis, respectively.

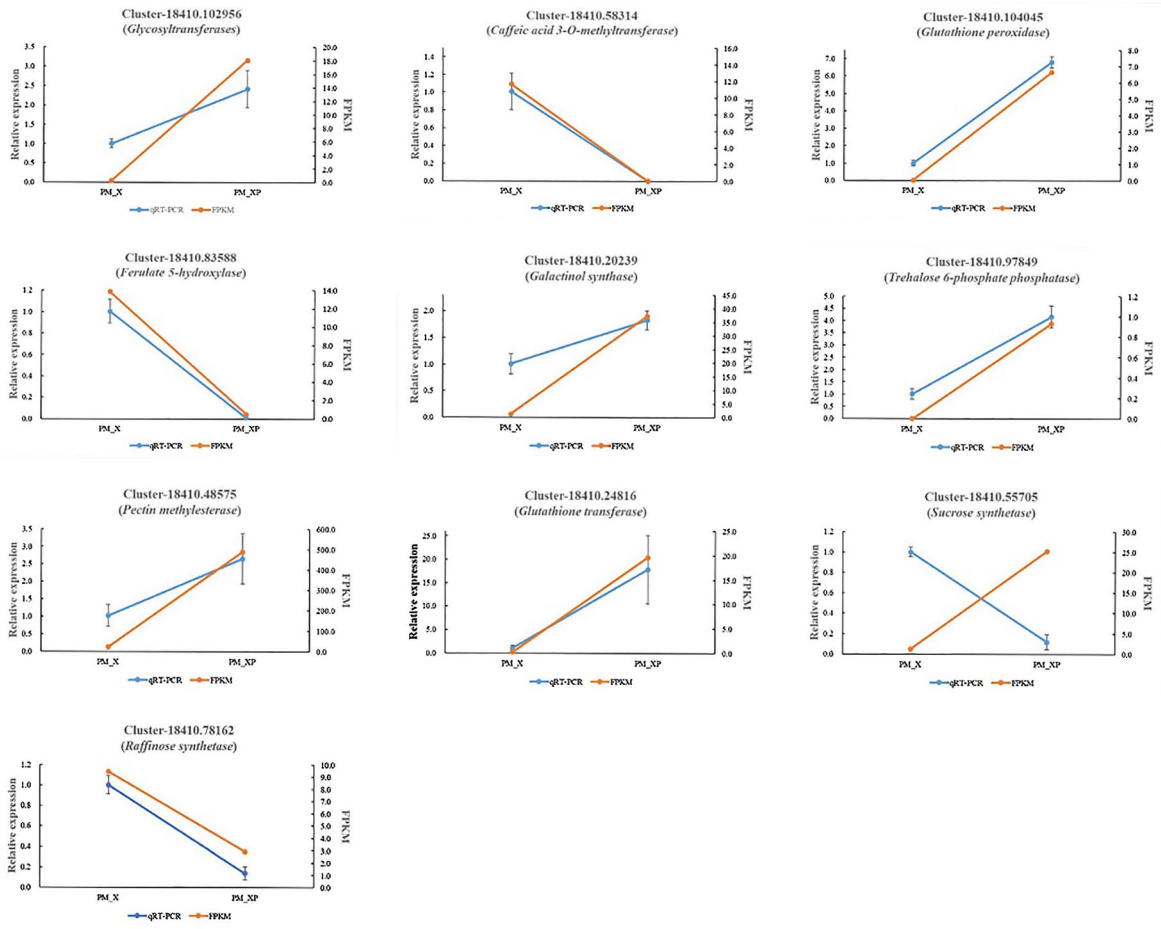
## Gene involved in lignin biosynthesis

Secondary cell walls (SCW) in plants accumulate the majority of photosynthesis products and are the main raw material for wood. Lignin is an important component of vascular plants' SCW and performs a major biological function in that it not only provides mechanical support and allows long-distance transport of water and nutrients, enabling trees to grow to great heights, but it also protects plants from pathogen invasion. In this study, we identified 71 unigenes in developing xylem of *P. massoniana* that may be involved in lignin biosynthesis, including ten *phenylalanine ammonia lyase* (*PAL*), one *cinnamate 4-hydroxylase* (*C4H*), two *coumarate 3-hydroxylase* (*C3H*), two *hydroxycinnamoyl transferase* (*HCT*), one *caffeoyl shikimate esterase* (*CSE*), one *ferulate 5-hydroxylase* (*F5H*), two *catechol-O-methyltransferase* (*COMT*), four *caffeoyl-CoA O-methyltransferase* (*CCoAOMT*), three *cinnamoyl CoA reductase* (*CCR*), five *cinnamyl alcohol dehydrogenase* (*CAD*), 28 *laccase* (*LAC*) and 12 *peroxidase* (*POD*) genes (Fig. 4). Using FPKM value to measure the expressions pattern of these unigenes in PM\_XP and PM\_X, we found that the expressions of these unigenes were very different, such as *CSE01* (Cluster-18410.58248) had 387.48 and 415.705 FPKM in PM\_X and PM\_XP, respectively, in contrast to *CCoAOMT04* (Cluster-18410.58641) had 0.685 and 0.385 FPKM in PM\_X and PM\_XP, respectively, which was much smaller than *CSE01* (Fig. 4). A similar situation occurred between different members of the same gene family. The FPKM value of *LAC01* (Cluster-18410.65339) in PM\_X and PM\_XP were 307.035 and 64.47, respectively, while *LAC19* (Cluster-18410.63178) were 20.62 and 10.17, respectively (Fig. 4). Those results indicate that different genes play different roles in the lignin biosynthesis pathway.

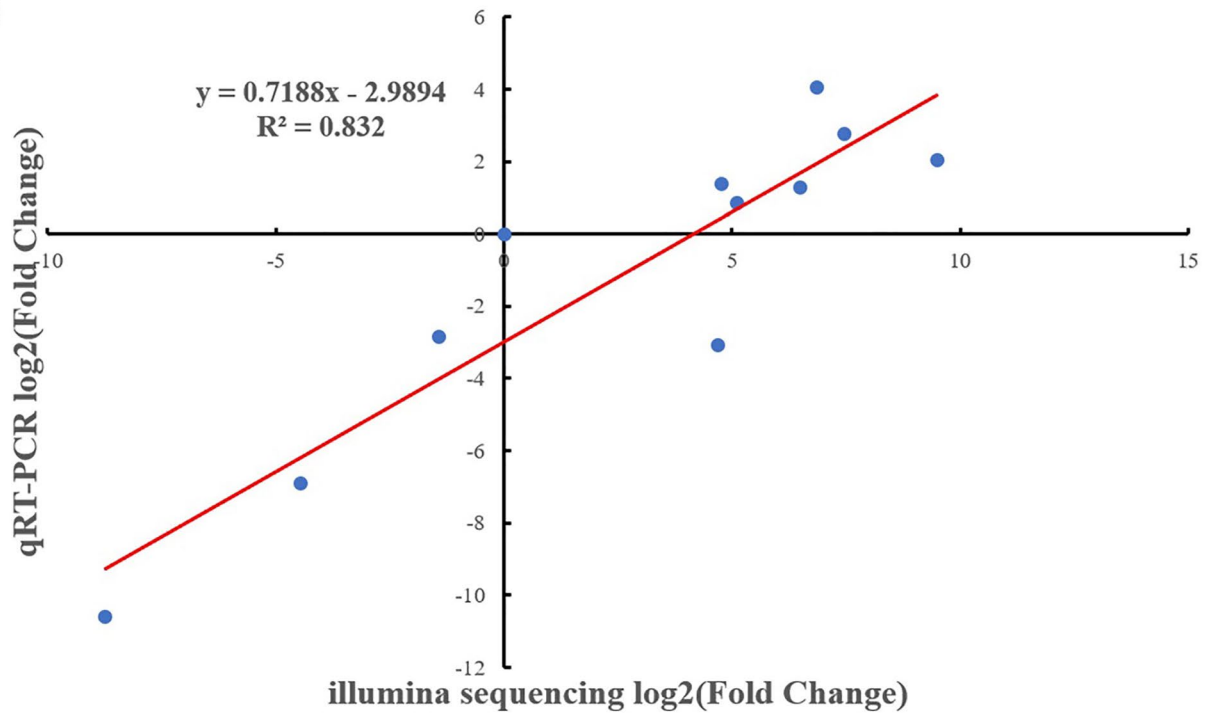
## Discussion

Previous studies have successfully isolated high-quality protoplasts from developing xylem in poplar, and the RNA-seq results revealed that approximately 96% of the genes in poplar developing xylem were present in the protoplasts [11]. The threshold of FPKM for gene expression detection and the number of DEGs were not mentioned in the literature; thus, it is difficult to accurately evaluate transcriptome identity between intact cells and their protoplasts from poplar developing xylem. In this study, on the basis of previous studies [11, 17, 18], an enzymatic hydrolysis method of plant cell wall was effectively improved and successfully applied to the isolation of protoplasts from developing xylem in *P. massoniana*, and the isolated protoplasts retained high vitality and integrity (Fig. 1), which ensured the acquisition of high-quality RNA and RNA-seq library construction

A



B





**Fig. 3** DEG expression patterns. **A** qRT-PCR validation of DEG expression patterns between developing xylem and its protoplasts. The blue lines represent the relative expression level determined with qRT-PCR (left y-axis) and the orange lines represent the FPKM value of selected unigenes (right y-axis). The error bars indicate the standard errors from biological replicates. **B** The correlation between RNA-seq and qPCR

(Table 1, supplementary Figure S2). More than 90% of the unigenes of pine developing xylem were detected in the protoplasts, indicating that the two shared a high transcriptome identity. Subsequent analysis of DEGs revealed more valuable information.

The genus *Pinus* includes more than 110 species and is the world's major tree species for timber and turpentine production. A number of studies have been reported on wood formation in *P. massoniana*. A study on wood biomass and the developing xylem transcriptome found that downregulation genes of the lignin biosynthetic pathway may be a cause of growth retardation in some *P. massoniana* individuals [34]. MicroRNAs (miRNAs) and other non-coding small RNAs also regulate lignin biosynthesis, and a study found that the target genes of pma-novel\_16, pma-novel\_92, and pma-novel\_57 and pma-novel\_104 encode MYB, ARF16, and LAC, respectively. These pathways are also present in angiosperms, indicating that some conserved regulatory modules are shared in wood formation in gymnosperms and angiosperms [35]. These studies have deepened our understanding of the wood formation process and regulatory mechanisms in *P. massoniana*, and this study found that the mechanism of wood formation can be better studied using protoplasts for species lacking genetic transformation systems.

The upregulated DEGs were significantly enriched in glutathione metabolism-related terms, and genes encoding glutathione peroxidase (GPX, EC: 1.11.1.9) and glutathione transferase (GST, EC 2.5.1.18) were significantly upregulated in protoplasts. GST has been considered to be involved in responses of plants to stress [36], and increases in GST transcripts are important indicators of plant responses [37]. GPXs are the main enzymes that convert glutathione (GSH) into glutathione disulfide (GSSG), which can prevent the cell membrane from being damaged by peroxidation. Thioredoxin reductase may change the redox state to activate GPXs [38]. On the other hand, the enrichment of upregulated DEGs in oxidative stress also indicated that protoplasts responded to the environmental stimulus cell wall removal to protect the cell membrane from oxidative damage. In galactose metabolism, galactinol synthase (GS, EC: 2.4.1.123) and raffinose synthetase (RS, EC: 2.4.4.82) are key enzymes in the synthesis of inositol galactoside and raffinose, which regulate the distribution of sucrose and oligosaccharides in plants as osmotic pressure protectants [39]. Trehalose plays an important

role in regulating the homeostasis and stress tolerance of plant cells via starch and sucrose metabolism [40]. The DEGs encoding GS, RS, sucrose synthetase (SuSy, EC: 2.4.1.13) and trehalose 6-phosphate phosphatase (otsb, EC: 3.1.3.12) were significantly upregulated in protoplasts (Supplementary Table S2). The osmotic pressure of protoplasts changed after cell wall enzymolysis, and the morphology and physiological function of cells were maintained by this osmotic pressure adjustment. The enrichment of oligosaccharide synthesis-related genes may indicate that the response of protoplasts to osmotic pressure changes activates an anti-stress response of cells. According to the upregulated expression of DEGs in the biosynthesis pathway of stress resistance and osmotic pressure balance-related substances in protoplasts, we speculated that the upregulated expression of related genes in protoplasts might be associated with the osmotic pressure caused by the response to cell wall enzymolysis.

The down-regulated DEGs were significantly enriched in photosystem-related items, indicating that the photosynthetic activity of intact cells in developing xylem was higher than that of protoplasts. DEGs involved in photosynthesis-antenna proteins, photosynthesis, and glyoxylate and dicarboxylate metabolism pathways were also significantly upregulated in developing xylem tissues (Supplementary Table S2). Glucose is produced by the carbon fixation reaction of photosynthesis, and then transformed into the basic material of cell growth and development. Therefore, photosynthetic biochemical reactions in the developing xylem tissues may be more active than those in protoplasts because of the latter's inhibition by wall removal [41]. The cell wall can maintain a relatively stable environment for organisms and provides space for biochemical reactions. At the same time, there are many enzymes in the cell wall, especially the hydrolases and synthetases related to carbohydrate synthesis and metabolism. For example, hydrogen peroxide is produced in the cell wall and is involved in lignin synthesis [42]. The genes related to HAO (EC: 1.1.3.15, (S)-2-hydroxyacid oxidase) and katE (EC: 1.11.1.6) were upregulated in intact cells, which may affect the synthesis of lignin in the cell wall.

In general, due to wall removal, protoplasts may be damaged in terms of the biological functions of cells, such as photosynthesis, and responses to environmental stimuli [43]. Using the developing xylem and its protoplasts as materials not only is beneficial to the study of vascular cambium differentiation but also provides a new clue for understanding the biosynthesis of the xylem cell wall. The plant cell wall is a complex network structure that includes cellulose, hemicellulose, and pectin. The effect of wall enzymolysis on cell biochemical activity can be revealed through DEG analyses of both intact cells and protoplasts of developing xylem, which also provides

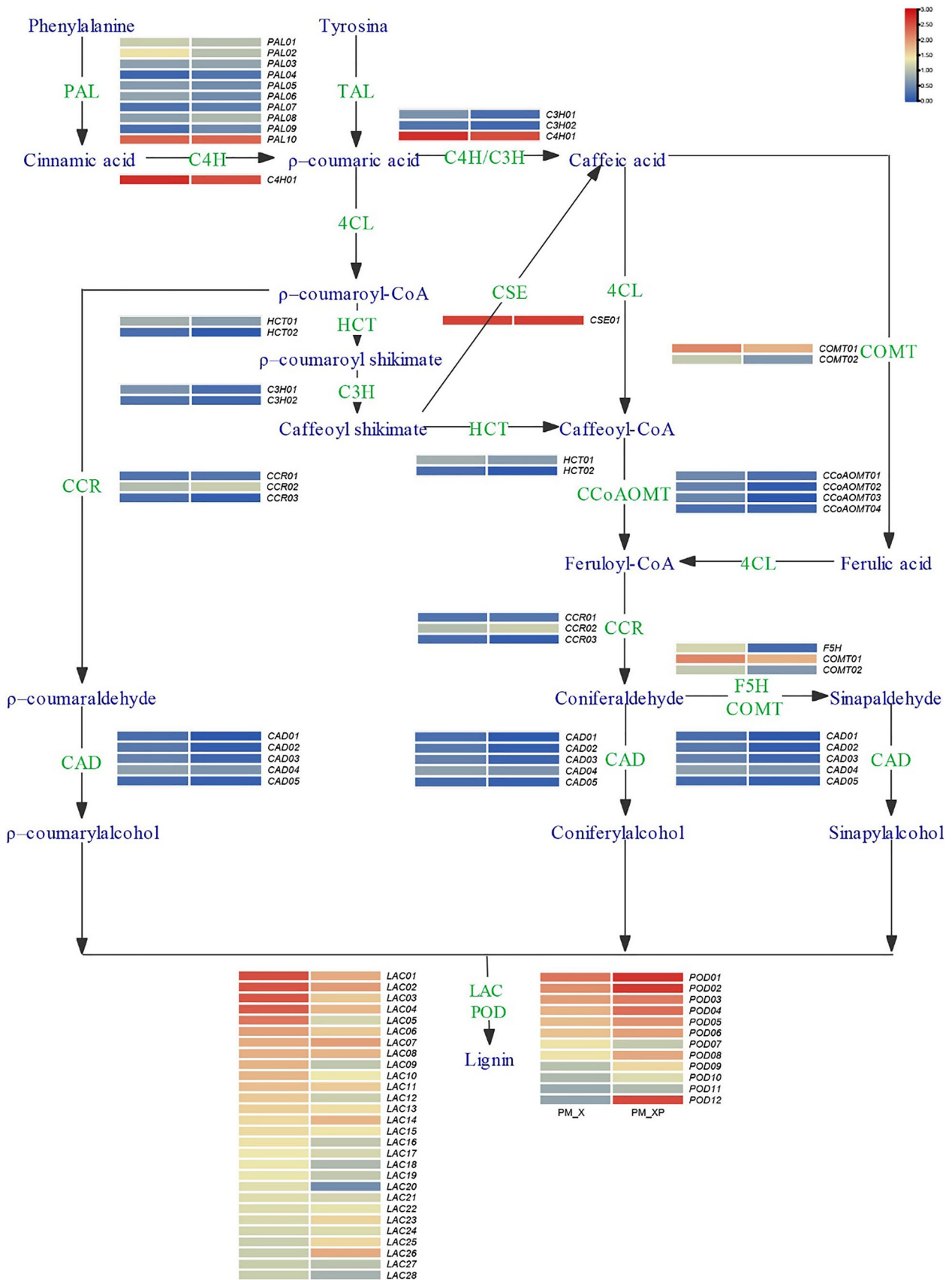


Fig. 4 Genes involved in lignin biosynthesis

a new perspective for tracking transcriptional regulatory events related to xylogenesis of Masson pine.

Cellulose is the main component of the plant cell wall, and the synthesis of cellulose is catalysed by cellulose synthase and proteins related to microfibril synthesis on cell membrane, with the substrate UDP glucose being produced by sucrose synthase (SuSy) [44]. THESEUS 1 (plasma-membrane-bound receptor-like kinase) belongs to the receptor-like kinase CrRLK1L (*Catharanthus roseus* receptor-like kinase 1-like) subfamily, and THESEUS 1, as a cell-wall-integrity sensor, can sense the state of the cell wall and coordinate cell wall synthesis and cell elongation [45]. In pine protoplasts, the expression of sucrose synthase and RLK genes was upregulated, and the expression of cellulose synthesis-related genes was downregulated. The downregulated DEGs were also significantly enriched in the membrane of cellular components. Genes associated with microfibril structure, such as COBRA [46], were highly expressed, however, there was no significant difference in the expression of these genes encoding cellulase and microfibril synthesis protein between protoplasts and intact cells of developing xylem. Protoplasts may promote the recovery of cell wall, but enzymatic hydrolysis of the cell wall may also lead to membrane damage and affect the regeneration of the cell wall.

Pectin is synthesized in the Golgi apparatus and transported to the cell wall in membrane vesicles [47]. The methylation and demethylation levels of pectin in the cell wall are controlled by methylsterases and pectin methylsterase (PME) [48], and several dozens of enzymes including glycosyltransferase (GT), methyltransferase and acetyltransferase, are required for pectin synthesis [49]. PME- and GT-related genes were significantly upregulated in the protoplasts from developing xylem in *P. massoniana*, suggesting that the hydrolysis of the cell wall may affect the pectin synthesis of protoplasts [50].

The phenylpropanoid pathway, serving as a starting point for the production of many important compounds and metabolites, is required for lignin biosynthesis [51, 52]. Interestingly, the two key genes related to caffeic acid 3-O-methyltransferase (COMT, EC: 2.1.1.68) and ferulate 5-hydroxylase (F5H), were significantly downregulated in the protoplasts. Lignin consists of syringyl (S), guaiacyl (G) and hydroxyphenyl (H) units produced through the phenylpropanoid pathway. Lignin component varies across species and genera, and guaiacyl lignin is the main composition of softwood lignin. However, F5H catalyses the formation of S-lignin monomers [53], while COMT participates in the synthesis of the S-lignin precursor [54]. The reason for the specific downregulation of these two genes between protoplasts and intact cells of pine developing xylem is an interesting problem worthy of further study.

In conclusion, a new enzyme cocktail was formulated to isolate protoplasts with high vitality and integrity from pine developing xylem that ensured the acquisition of high-quality RNA and the reliability of transcriptome data. At the transcriptome level, the hydrolysis of the cell wall can stimulate protoplasts to respond to environmental changes, and protoplasts may undergo a series of physiological reactions to achieve internal balance. The removal of the cell wall leads to a loss of some enzymes in the cell wall, which disturbs the physiological activities of cells and affects the biosynthesis of cellulose, pectin and lignin in protoplasts. Transcriptome data are valuable in further revealing gene functions and regulatory pathways related to xylogenesis in softwood trees.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s11033-021-06995-6>.

**Author contributions** MX participated in the conception and design of the experiments and reviewed a draft of the manuscript. TS and MXX participated in conducting the experiments and analyzing the data and drafted the manuscript. HQ participated in conducting the experiments and analyzing the data. FY and ZY participated in the conception and design of the experiments. All authors read and approved the final version of the manuscript.

**Funding** This work was supported by Grants from the Open Research Fund of Guangxi Key Laboratory of Superior Timber Trees Resource Cultivation (17-B-03-01), the Science and Technology Major Program of Guangxi (AA17204087-1) and the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD).

**Data availability** The raw data has been submitted to SRA database with accession numbers SRR12596930 (PM\_XP2), SRR12596931 (PM\_XP1), SRR12596932 (PM\_X2) and SRR12596933 (PM\_X1).

## Declarations

**Conflict of interest** No potential conflicts of interest were disclosed.

## References

1. Chen H, Wang JP, Liu H, Li H, Lin Y-CJ, Shi R, Yang C, Gao J, Zhou C, Li Q et al (2019) Hierarchical transcription factor and chromatin binding network for wood formation in black cottonwood (*Populus trichocarpa*). *Plant Cell* 31(3):602–626. <https://doi.org/10.1105/tpc.18.00620>
2. Plomion C, Leprovost G, Stokes A (2001) Wood formation in trees. *Plant Physiol* 127(4):1513–1523
3. Moyle R, Schrader J, Stenberg A, Olsson O, Saxena S, Sandberg G, Bhalerao RP (2002) Environmental and auxin regulation of wood formation involves members of the Aux/IAA gene family in hybrid aspen. *Plant J* 31(6):675–685
4. Lin Y-C, Li W, Sun Y-H, Kumari S, Wei H, Li Q, Tunlaya-Anukit S, Sederoff RR, Chiang VL (2013) SND1 transcription factor-directed quantitative functional hierarchical genetic regulatory network in wood formation in *Populus trichocarpa*. *Plant Cell* 25(11):4324–4341. <https://doi.org/10.1105/tpc.113.117697>

5. Schuetz M, Smith R, Ellis B (2013) Xylem tissue specification, patterning, and differentiation mechanisms. *J Exp Bot* 64(1):11–31. <https://doi.org/10.1093/jxb/ers287>
6. Shu W, Zhou H, Jiang C, Zhao S, Wang L, Li Q, Yang Z, Groover A, Lu M-Z (2019) The auxin receptor TIR1 homolog (PagFBL1) regulates adventitious rooting through interactions with Aux/IAA28 in *Populus*. *Plant Biotechnol J* 17(2):338–349. <https://doi.org/10.1111/pbi.12980>
7. Chaffey N, Cholewa E, Regan S, Sundberg B (2002) Secondary xylem development in *Arabidopsis*: a model for wood formation. *Physiol Plant* 114(4):594–600
8. Jansson S, Douglas CJ (2007) *Populus*: a model system for plant biology. *Annu Rev Plant Biol* 58:435–458
9. Wang JP, Matthews ML, Williams CM, Shi R, Yang C, Tunlaya-Anukit S, Chen H-C, Li Q, Liu J, Lin C-Y et al (2018) Improving wood properties for wood utilization through multi-omics integration in lignin biosynthesis. *Nat Commun* 9(1):1579. <https://doi.org/10.1038/s41467-018-03863-z>
10. Eeckhout T, Lakshmanan PS, Deryckere D, Van Bockstaele E, Van Huylenbroeck J (2013) Progress in plant protoplast research. *Planta* 238(6):991. <https://doi.org/10.1007/s00425-013-1936-7>
11. Lin Y-C, Li W, Chen H, Li Q, Sun Y-H, Shi R, Lin C-Y, Wang JP, Chen H-C, Chuang L et al (2014) A simple improved-throughput xylem protoplast system for studying wood formation. *Nat Protoc* 9(9):2194–2205. <https://doi.org/10.1038/nprot.2014.147>
12. Li W, Lin Y-C, Li Q, Shi R, Lin C-Y, Chen H, Chuang L, Qu G-Z, Sederoff RR, Chiang VL (2014) A robust chromatin immunoprecipitation protocol for studying transcription factor-DNA interactions and histone modifications in wood-forming tissue. *Nat Protoc* 9(9):2180–2193. <https://doi.org/10.1038/nprot.2014.146>
13. Pascual MB, Llebrés M-T, Craven-Bartle B, Cañas RA, Cánovas FM, Ávila C (2018) PpNAC1, a main regulator of phenylalanine biosynthesis and utilization in maritime pine. *Plant Biotechnol J* 16(5):1094–1104. <https://doi.org/10.1111/pbi.12854>
14. Neale DB, Wegrzyn JL, Stevens KA, Zimin AV, Puiu D, Crepeau MW, Cardeno C, Koriabine M, Holtz-Morris AE, Liechty JD et al (2014) Decoding the massive genome of loblolly pine using haploid DNA and novel assembly strategies. *Genome Biol* 15(3):R59. <https://doi.org/10.1186/gb-2014-15-3-r59>
15. Nystedt B, Street NR, Wetterbom A, Zuccolo A, Lin Y-C, Scofield DG, Vezzi F, Delhomme N, Giacomello S, Alexeyenko A et al (2013) The Norway spruce genome sequence and conifer genome evolution. *Nature* 497(7451):579–584. <https://doi.org/10.1038/nature12211>
16. Ni Z, Bai T, Chen Y, Huang Y, Xu L-A (2019) Effects of parental genetic distance on offspring growth performance in *Pinus massoniana*: significance of parental-selection in a clonal seed orchard. *Euphytica* 215(12):195. <https://doi.org/10.1007/s10681-019-2525-z>
17. Leinhos V, Savidge RA (1993) Isolation of protoplasts from developing xylem of *Pinus banksiana* and *Pinus strobus*. *Can J For Res* 23:343–348
18. Tan B, Xu M, Chen Y, Huang M (2013) Transient expression for functional gene analysis using *Populus* protoplasts. *Plant Cell Tissue Organ Cult (PCTOC)* 114(1):11–18. <https://doi.org/10.1007/s11240-013-0299-x>
19. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q et al (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol* 29(7):644–652. <https://doi.org/10.1038/nbt.1883>
20. Buchfink B, Xie C, Huson DH (2015) Fast and sensitive protein alignment using DIAMOND. *Nat Methods* 12(1):59–60. <https://doi.org/10.1038/nmeth.3176>
21. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25(17):3389–3402
22. Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M (2007) KAAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res* 35:W182–W185
23. Potter SC, Luciani A, Eddy SR, Park Y, Lopez R, Finn RD (2018) HMMER web server: 2018 update. *Nucleic Acids Res* 46(W1):W200–W204. <https://doi.org/10.1093/nar/gky448>
24. Götz S, García-Gómez JM, Terol J, Williams TD, Nagaraj SH, Nueda MJ, Robles M, Talón M, Dopazo J, Conesa A (2008) High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res* 36(10):3420–3435. <https://doi.org/10.1093/nar/gkn176>
25. Li B, Dewey CN (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinform* 12:323. <https://doi.org/10.1186/1471-2105-12-323>
26. Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15(12):550
27. Yu G, Wang L-G, Han Y, He Q-Y (2012) clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 16(5):284–287. <https://doi.org/10.1089/omi.2011.0118>
28. Alexa A, Rahnenführer J, Lengauer T (2006) Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics* 22(13):1600–1607
29. Mao X, Cai T, Olyarchuk JG, Wei L (2005) Automated genome annotation and pathway identification using the KEGG Orthology (KO) as a controlled vocabulary. *Bioinformatics* 21(19):3787–3793
30. Chen H, Yang Z, Hu Y, Tan J, Jia J, Xu H, Chen X (2016) Reference genes selection for quantitative gene expression studies in *Pinus massoniana* L. *Trees* 30(3):685–696. <https://doi.org/10.1007/s00468-015-1311-3>
31. Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the  $2^{-\Delta\Delta CT}$  method. *Methods* 25(4):402–408. <https://doi.org/10.1006/meth.2001.1262>
32. Li H, Chen G, Pang H, Wang Q, Dai X (2021) Investigation into different wood formation mechanisms between angiosperm and gymnosperm tree species at the transcriptional and post-transcriptional level. *Front Plant Sci* 12:1273
33. Ko J-H, Beers EP, Han K-H (2006) Global comparative transcriptome analysis identifies gene network regulating secondary xylem development in *Arabidopsis thaliana*. *Mol Genet Genomics* 276(6):517–531. <https://doi.org/10.1007/s00438-006-0157-1>
34. Ni Z, Han X, Yang Z, Xu M, Feng Y, Chen Y, Xu L-A (2020) Integrative analysis of wood biomass and developing xylem transcriptome provide insights into mechanisms of lignin biosynthesis in wood formation of *Pinus massoniana*. *Int J Biol Macromol* 163:1926–1937. <https://doi.org/10.1016/j.ijbiomac.2020.08.253>
35. Shen T, Xu M, Qi H, Feng Y, Yang Z, Xu M (2021) Uncovering miRNA-mRNA regulatory modules in developing xylem of *Pinus massoniana* via small RNA and degradome sequencing. *Int J Mol Sci* 22(18):10154. <https://doi.org/10.3390/ijms221810154>
36. Alscher RG (1989) Biosynthesis and antioxidant function of glutathione in plants. *Physiol Plant* 77(3):457–464. <https://doi.org/10.1111/j.1399-3054.1989.tb05667.x>
37. Edwards R, Dixon DP, Walbot V (2000) Plant glutathione S-transferases: enzymes with multiple functions in sickness and in health. *Trends Plant Sci* 5(5):193–198
38. Herbet S, Lenne C, Leblanc N, Julien J-L, Drevet JR, Roeckel-Drevet P (2002) Two GPX-like proteins from *Lycopersicon esculentum* and *Helianthus annuus* are antioxidant enzymes with

- phospholipid hydroperoxide glutathione peroxidase and thioredoxin peroxidase activities. *Eur J Biochem* 269(9):2414–2420
39. Castillo EM, De Lumen BO, Reyes PS, De Lumen HZ (1990) Raffinose synthase and galactinol synthase in developing seeds and leaves of legumes. *J Agric Food Chem* 38(2):351–355. <https://doi.org/10.1021/jf00092a003>
  40. Garg AK, Kim J-K, Owens TG, Ranwala AP, Choi YD, Kochian LV, Wu RJ (2002) Trehalose accumulation in rice plants confers high tolerance levels to different abiotic stresses. *Proc Natl Acad Sci USA* 99(25):15898–15903
  41. Jouanneau JP, Lalous D, Guern J (1991) In plant protoplasts, the spontaneous expression of defense reactions and the responsiveness to exogenous elicitors are under auxin control. *Plant Physiol* 96(2):459–466
  42. Mäder M, Ungemach J, Schloß P (1980) The role of peroxidase isoenzyme groups of *Nicotiana tabacum* in hydrogen peroxide formation. *Planta* 147(5):467–470. <https://doi.org/10.1007/BF00380189>
  43. Steponkus PL, Dowgert MF, Gordon-Kamm WJ (1983) Destabilization of the plasma membrane of isolated plant protoplasts during a freeze-thaw cycle: the influence of cold acclimation. *Cryobiology* 20(4):448–465
  44. Coleman HD, Yan J, Mansfield SD (2009) Sucrose synthase affects carbon partitioning to increase cellulose production and altered cell wall ultrastructure. *Proc Natl Acad Sci USA* 106(31):13118–13123. <https://doi.org/10.1073/pnas.0900188106>
  45. Hématy K, Sado P-E, Van Tuinen A, Rochange S, Desnos T, Balzergue S, Pelletier S, Renou J-P, Höfte H (2007) A receptor-like kinase mediates the response of Arabidopsis cells to the inhibition of cellulose synthesis. *Curr Biol* 17(11):922–931
  46. Roudier F, Fernandez AG, Fujita M, Himmelspach R, Borner GHH, Schindelman G, Song S, Baskin TI, Dupree P, Wasteneys GO et al (2005) COBRA, an Arabidopsis extracellular glycosylphosphatidyl inositol-anchored protein, specifically controls highly anisotropic expansion through its involvement in cellulose microfibril orientation. *Plant Cell* 17(6):1749–1763
  47. Ibar C, Orellana A (2007) The import of S-adenosylmethionine into the Golgi apparatus is required for the methylation of homogalacturonan. *Plant Physiol* 145(2):504–512
  48. Sénéchal F, Graff L, Surcouf O, Marcelo P, Rayon C, Bouton S, Mareck A, Mouille G, Stintzi A, Höfte H et al (2014) Arabidopsis PECTIN METHYLESTERASE17 is co-expressed with and processed by SBT3.5, a subtilisin-like serine protease. *Ann Bot* 114(6):1161–1175. <https://doi.org/10.1093/aob/mcu035>
  49. Mohnen D, Bar-Peled M, Somerville C (2008) Cell wall polysaccharide synthesis. *Biomass Recalcitrance*. <https://doi.org/10.1002/9781444305418.ch5>
  50. Liu H, Chen H, Ding G, Li K, Ren Q (2020) Identification of candidate genes conferring tolerance to aluminum stress in *Pinus massoniana* inoculated with ectomycorrhizal fungus. *BMC Plant Biol* 20(1):521. <https://doi.org/10.1186/s12870-020-02719-3>
  51. Zubieta C, Kota P, Ferrer J-L, Dixon RA, Noel JP (2002) Structural basis for the modulation of lignin monomer methylation by caffeic acid/5-hydroxyferulic acid 3/5-O-methyltransferase. *Plant Cell* 14(6):1265–1277
  52. Ruegger M, Meyer K, Cusumano JC, Chapple C (1999) Regulation of ferulate-5-hydroxylase expression in Arabidopsis in the context of sinapate ester biosynthesis. *Plant Physiol* 119(1):101–110
  53. Sibout R, Baucher M, Gatineau M, Van Doorsselaere J, Mila I, Pollet B, Maba B, Pilate G, Lapierre C, Boerjan W et al (2002) Expression of a poplar cDNA encoding a ferulate-5-hydroxylase/coniferaldehyde 5-hydroxylase increases S lignin deposition in *Arabidopsis thaliana*. *Plant Physiol Biochem* 40(12):1087–1096. [https://doi.org/10.1016/S0981-9428\(02\)01474-2](https://doi.org/10.1016/S0981-9428(02)01474-2)
  54. Guo D, Chen F, Inoue K, Blount JW, Dixon RA (2001) Down-regulation of caffeic acid 3-O-methyltransferase and caffeoyl CoA 3-O-methyltransferase in transgenic alfalfa. Impacts on lignin structure and implications for the biosynthesis of G and S lignin. *Plant Cell* 13(1):73–88

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.