Identification of Putative Cell Wall Synthesis Genes in *Betula pendula*

Song Chen
State Key Laboratory of Tree Genetics and Breeding, Northeast Forestry University, Harbin, China
Xin Lin
State Key Laboratory of Tree Genetics and Breeding, Northeast Forestry University, Harbin, China
Xiyang Zhao
State Key Laboratory of Tree Genetics and Breeding, Northeast Forestry University, Harbin, China
Su Chen (chensunefu@163.com)
State Key Laboratory of Tree Genetics and Breeding, Northeast Forestry University, Harbin, China
https://orcid.org/0000-0002-8814-5444

Research article

Keywords: Betula pendula, cell wall, cellulose synthase, RNA-seq, transcription factors

Posted Date: October 27th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-33560/v2

Abstract

Background

Cellulose is an essential structural component of plant cell wall and is an important resource to produce paper, textiles, bioplastics and other biomaterials. The synthesis of cellulose is among the most important but poorly understood biochemical processes, which is precisely regulated by internal and external cues.

Results

Here, we identified 46 gene models in 7 gene families which encoding cellulose synthase and related enzymes of *Betula pendula*, and the transcript abundance of these genes in xylem, root, leaf and flower tissues also be determined. Based on these RNA-seq data, we have identified 8 genes that most likely participate in secondary cell wall synthesis, which include 3 cellulose synthase genes and 5 cellulose synthase-like genes. In parallel, a gene co-expression network was also constructed based on transcriptome sequencing.

Conclusions

In this study, we have identified a total of 46 cell wall synthesis genes in *B. pendula*, which include 8 secondary cell wall synthesis genes. These analyses will help decipher the genetic information of the cell wall synthesis genes, elucidate the molecular mechanism of cellulose synthesis and understand the cell wall structure in *B. pendula*.

1 Background

Silver birch (*Betula pendula*) is a medium-sized deciduous tree that owes its common name to the white peeling bark on the trunk. This species is native to Europe and parts of Asia, and the range extends into Siberia, China, and southwest Asia in the mountains of northern Turkey, the Caucasus, and northern Iran [1]. *B. pendula* is a woody species that play important roles in ecological system due to its strong tolerance to various climates such as salt, cold and drought [2-4]. Flowering at an early age allows *B. pendula* has a faster succession of generations, which together with rapid juvenile development can shorten the breeding cycle. Large-sized logs are produced within relatively short periods with proper silvicultural treatment, and the wood characteristics allow versatile and valuable uses, as shown in Northern Europe. In addition, it can also improve forest resilience by colonizing forest gaps and quickly increasing soil functioning and biodiversity. In the context of societal evolutions and customer perceptions, *B. pendula* will certainly play an increasing role in the building and furniture sectors, and among non-wood forest products [5, 6].

The cell periphery of higher plants is usually surrounded by the cell wall. Plant cell walls are complex networks of polymers that provide protection and structural properties to the cells [7]. The cell wall mainly

includes four chemical polymers: cellulose, hemicellulose, lignin and pectin. Cellulose is the most abundant organic polymer in plant cell walls. It is linear homopolymer of β-1,4-linked glucose residues, and synthesized by rosette terminal complexes (RTCs). The RTCs are hexameric protein structures containing the cellulose synthase enzymes. Annually, plants will produce about 180 billion tons of cellulose making it the largest reservoir of organic carbon on Earth [8]. In addition, plant cells can also control the mechanical properties of the wall by organizing the synthesis and deposition of wall polymers and by modifying the wall architecture according to the needs of the cell [9]. For example, fiber cells and vascular cells always form thick secondary cell walls with high cellulose content to structurally support the plant and to facilitate water transport, so the xylem cell walls always have more higher cellulose content than leaves [10, 11]. Pear et al. [12] isolated and identified the CESA genes encoding cellulose synthase for the first time from cotton in 1996. Subsequent analysis of the *A. thaliana* genome revealed that a total of 10 genes encode CESA proteins with 64% average sequence identity [13, 14]. *Populus trichocarpa* has 18 CESA genes [15], *Hordeum vulgare* has 8 [16] and *Zea mays* has 12 [17]. In *Betula*-related studies, Liu et al. [18] have isolated four full-length CESA cDNAs from *Betula platyphylla* in 2012 by using RT-PCR and RACE method, and calculated the phylogenetic relationship of them. Huang et al. [19] have isolated eight full-length CESA cDNAs from *Betula luminifera* in 2014 based on transcriptome sequencing, and determined their positive influence in tension wood.

As an important tree species in papermaking, understanding the cellulose synthesis pathway of *B. pendula* will greatly contribute to its use in industrial production. Fortunately, the assembled sequences of *B. pendula* have become publicly available, which can help us understand this species at the genome expression level [20]. In this study, we identified the genes that likely encode cellulose synthase and related enzymes during cell wall synthesis in *B. pendula*, which will serve as a basis for further gene functional studies.

2 Results

2.1 Identification of *Betula pendula* cellulose cell wall synthesis genes

A total of 29,439 coding genes in *B. pendula* genome [20] were used to identify putative cell wall synthesis genes. In total, 46 gene models (Table 1) in 7 families were identified as putative cell wall synthesis genes in *B. pendula* genome. The 46 genes encode 10 cellulose synthase proteins (CESAs) and 36 cellulose synthase-like proteins (CSLAs, CSLBs, CSLCs, CSLDs, CSLEs and CSLGs) in 7 families. Among these families, *CESA* was the predominant cellulose synthase gene family and contains ten members. The rest of the gene families all belong to the cellulose synthase-like family, *CSLG* was the largest cellulose synthase-like family containing eleven members, while *CSLA* was the smallest family with only three members. We then applied quantitative criteria to assign the genes likely to be cell wall synthesis genes based on transcript abundance and specificity. The tissue-specific expressional data include xylem, roots, leaves and flowers, and we calculated the expression of the 46 identified genes. A

total of 8 genes showed that expression in the xylem was higher than the expression in both flower and leaf. These genes were identified as the secondary cell wall synthesis genes *BpCESA4*, *BpCESA9*, *BpCESA10*, *BpCSLA2*, *BpCSLA3*, *BpCSLC1*, *BpCSLC4* and *BpCSLD4*.

2.2 Chromosomal location and gene duplication

Cellulose synthase complex mainly includes cellulose synthases (CESAs) and cellulose synthase-like proteins (CSLs), so we investigated the formation of CESAs and CSLs based on the chromosomal location and intra-genome syntenic information. Similar to the *A. thaliana*, the multiple *BpCESAs* were scattered across the *B. pendula* genome and mapped in 13 of the 14 chromosomes (Figure 1). The *BpCESAs* were concentrated on Bpe_Chr6, Bpe_Chr7, Bpe_Chr8, Bpe_Chr9, Bpe_Chr10 and Bpe_Chr11, with one or two genes per chromosome. The *BpCSLs* were scattered on 13 chromosomes except for Bpe_Chr5, and we found that some *BpCSLs* were organized into duplicated blocks, such as *BpCSLB1-7* on Bpe_Chr2, *BpCSLG2-7* on Bpe_Chr14 and *BpCSLG8-10* on Bpe_Chr1. This situation always originated from the duplicative transposition.

2.3 Cellulose synthase (CESA) gene family

Cellulose are the principal ingredient of the cell walls in *B. pendula*, and the small microfibrils are crystallized by 36 tails of H-bonded-β-1,4-Glc chains catalyzed by cellulose synthases [21]. Thus, cellulose synthase (CSEA) was one of the indispensable glycosyltransferases in plants, which plays a crucial role in regulating cell wall cellulose synthesis and plant cell morphogenesis. We have identified 10 *BpCESAs* in the *B. pendula* genome, of which *BpCESA4*, *BpCESA9* and *BpCESA10* were abundant in xylem (Figure 2). *BpCESA4* was the highest expressed gene in the root and xylem of the CESA family. The most similar protein to BpCESA4 was AtCESA4 in *Arabidopsis thaliana*, which confers plant resistance to bacterial and fungal pathogens while encoding a cellulose synthase. The protein most similar to BpCESA9 and BpCESA10 was AtCESA8 in *A. thaliana*.

2.4 Cellulose synthase-like (CSL) gene family

The cellulose synthase-like (CSL) gene family was divided into six families, which were CSLA, CSLB, CSLC, CSLD, CSLE and CSLG. The cellulose synthase-like gene family was divided into six families, which were CSLA, CSLB, CSLC, CSLD, CSLE and CSLG. The functions of the CSL family are still being explored, but a substantial number of studies were published in recent years. Jensen et al. [22] reported that the CSL genes is associated with hemicellulose synthesis, Schreiber et al. [23] and Doblin et al. [24] reported that cellulose synthase-like protein CSLFs and CSLHs mediate the synthesis of cell wall (1,3)(1,4)-β-D-Glucans, but the vast majority of CSL genes functions require further study.

We identified 36 *BpCSLs* in the *B. pendula* genome of which 5 genes were abundant in xylem (Figures 2 and 3). They were *BpCSLA2*, *BpCSLA3*, *BpCSLC1*, *BpCSLC4* and *BpCSLD4*, respectively. Both BpCSLA2 and BpCSLA3 were most similar to AtCSLA9, and BpCSLD4 was most similar to AtCSLD6 in *A. thaliana*. In addition, the most similar protein to BpCSLC1 was AtCSLC4 in *A. thaliana*, which encode a protein similar to cellulose synthase and its mRNA can mobile in cell-to-cell.

2.5 Involvement of transcription factors in cell wall synthesis

Based on transcriptome sequencing data, we performed an extensive analysis between putative cell wall synthesis proteins and 2,816 transcription factors (Table S1) of *B. pendula*. The results showed that a total of 51 transcription factors were co-expressed with 6 cell wall synthesis proteins, which were BpCESA4, BpCESA9, BpCSLA2, BpCSLC1, BpCSLC4 and BpCSLD4 (Figure 4).

The highest number of transcription factors were co-expressed with BpCSLC1, up to 27, including ARF, IAA and several other auxin-related transcription factors. BpARF6 was most similar to AtARF17, and BpIAA16 was most similar to AtIAA16, which has transcriptional wiring with cell wall-related genes in *A. thaliana* [25]. In addition to BpCSLC1, there was a co-expression relationship between BpCESA4 and BpCESA9, with 13 transcription factors regulating these two cellulose synthase genes. Among them, BpMYB-HB162 was most similar to AtMYB83, and BpNAM69 was most similar to AtNAC43 (NST1) in *A. thaliana*, which is known to be involved in cellulose synthesis.

3 Discussion

In this study, we identified a total of 10 CESA genes and 36 CSL genes in *B. pendula*, which include 8 genes that most likely involved in secondary cell wall synthesis in *B. pendula*, which should help elucidate the molecular mechanism of cellulose synthesis in *B. pendula*. These genes showed striking consistency compared to the cell wall synthesis genes in *P. trichocarpa*, demonstrating that the cellulose synthesis family is conserved during species evolution.

Cellulose synthesis requires the plant hormones, nitric oxide, cellulose synthase, and a complex transcriptional regulation network. Handakumbura et al. [26] reported that AtCESA4 loss-of-function mutants of *A. thaliana* and *Oryza sativa* have weak stems and thin or irregular cell walls. Glass et al. [27] reported that endo-β-1,4-glucanases AtGH9B5 and AtGH9C2 can impact cellulose crystallization and plant cell wall development by influencing cellulose synthase AtCESA8. In addition, the expression of *CSLs* in *A. thaliana* cells revealed that AtCSLA glycosyltransferases can also encode cell wall glucomannan and intervention the progression of embryogenesis [28, 29]. Galway et al. [30] reported that root hair-specific disruption of cellulose and xyloglucan in AtCSLD3 mutants. The 1,4-beta-glucan synthase AtCSLC4 can form the xylosylated glucan backbone with three xylosyltransferases AtXXT1, AtXXT2 and AtXXT5 in *A. thaliana* [31]. Intriguingly, glucan synthase AtCSLC4 have opposite orientations

in the Golgi membrane [32] with mannan synthase AtCSLA9, which may cause the functional differences between them. As for the related transcription factors, Kim et al. [33, 34] reported that transcription factor AtMYB46 can directly regulate the secondary wall-associated cellulose synthase AtCESA4 and AtCESA8, and the transcription factors AtNAC41, AtbZIP1 and AtMYB46 can directly regulate the expression of *AtCSLA9* in *A. thaliana*. Yang et al. [35] reported that AtARF17 is essential for primexine formation and pollen wall development. Zhong et al. [36] reported that inhibition of the expression of both *AtSND1* and *AtNST1* by RNA interference (RNAi) results in loss of secondary wall formation in stem fibers, and several fiber-associated transcription factor genes will be down-regulation in *A. thaliana*. Ko et al. [37] reported that the AtMYB46/AtMYB83-mediated transcriptional regulatory program is a gatekeeper of secondary wall synthesis.

Given the importance of cellulose synthase's importance to cellulose synthesis, maybe we can limit the rate of cellulose synthesis by directly or indirectly inhibiting the expression of related genes, thereby reducing the cellulose content of *B. pendula*. Oomen et al. [38] reported that reducing of the cellulose content of *Solanum tuberosum* tuber by antisense expression of *StCESA3* clones. Zhong et al. [39] reported that the *AtCesA7* mutant of *A. thaliana* has lower fiber cell wall thickness and cellulose content. However, the process of increasing cellulose content is not as simple as reducing it. Tan et al. [40] reported that overexpressing *HvCESA* showed no increase in cellulose content or stem strength in *Hordeum vulgare*, despite the use of a powerful constitutive promoter. Previous studies [41] have shown that individual CESA and CSL proteins play different roles in the synthase complex and require tightly regulated, so we need more complex strategies in the plant engineering of increasing cellulose content.

4 Conclusion

This study aims to provide information on B. pendula cell wall synthesis genes regarding their potential physiological roles and the molecular mechanism associated. In this study, we have identified a total of 46 cell wall synthesis genes in B. pendula, which include 8 secondary cell wall synthesis genes. And a gene co-expression network was constructed based on synthesis-related genes expression value. These analyses will help decipher the genetic information of the cell wall synthesis genes, elucidate the molecular mechanism of cellulose synthesis and understand the cell wall structure in B. pendula.

5 Methods

5.1 Identification of *B. pendula* cell wall synthesis genes

The *B. pendula* genome [20] and genomic structure information (GFF) were downloaded from the CoGe (Comparative Genomics) platform. The putative cellulose synthase genes were first identified by BLASTP (Basic Local Alignment Search Tool - Protein) v2.9.0 [42] with the *A. thaliana* cellulose synthase genes as queries (E-value ≤ 1E-5). We then further manually examined these putative cell wall synthesis genes using the Conserved Domain Database of NCBI (National Center for Biotechnology Information) [43] to confirm if they were correctly annotated, and divided them into seven subgroups based on their

functional type in *A. thaliana*. In addition, the chromosomal location of the *B. pendula* cell wall synthesis genes was visualized by using TBtools (Toolbox for Biologists) v0.67 [44].

5.2 Phylogenetic analyses of *B. pendula* cell wall synthesis genes

To investigate the phylogenetic relationships of the cellulose synthases (CESAs) and cellulose synthase-like proteins (CSLs), the phylogenetic tree was constructed for every subgroup. The multiple sequence comparison was performed by MUSCLE (Multiple Sequence Alignment with High Accuracy and High Throughput) v3.8.1551 [45] with default parameters, and the constraint maximum likelihood phylogenetic trees of each subgroup were then be generated by RAxML (Randomized Axelerated Maximum Likelihood) v8.2.12 [46] with 1,000 bootstrap trials. The model was selected for the GAMMA model and visualized by iTOL (Interactive Tree of Life) v5 [47].

5.3 RNA-seq expression analysis of *B. pendula* cell wall synthesis genes

We downloaded the transcriptome data (PRJNA535361) [48] from the NCBI SRA database to investigate the expressional patterns of *B. pendula* cellulose synthase genes in different tissues. The clean reads of three replicates per tissues were aligned to the *B. pendula* transcriptome by using STAR (Spliced Trans Alignment to a Reference) v2.7.3a [49], and the accurate transcript quantification was estimated by using RSEM (RNA-seq by Expectation-Maximization) v1.3.3 pipeline [50] with paired-end sequencing mode. The normalized expression value was all selected as TMM (trimmed mean of M-values).

5.4 Transcription factor regulatory networks in *B. pendula* cell wall synthesis

The transcription factors of *B. pendula* were identified by PlantTFcat [51], and the Conserved Domain Database of NCBI [43] was be used to determine whether they are correctly annotated. To perform the weighted correlation network analysis between cell wall synthesis genes and transcription factors, we used the WGCNA (Weighted Correlation Network Analysis) R package v1.69 [52] to construct the co-expression network. The TMM value from different tissues of *B. pendula* was as input expression data for this software, and only genes with TMM values larger than 10 for all samples were kept. The threshold power (β) value was determined to be 13 from pickSoftThreshold output, and the Pearson algorithm is then be used to calculate the correlation coefficient. Finally, the co-expression network was generated block wise using the WGCNA function blockwiseModules with the following settings: TOM-type, unsigned; mergeCutHeight, 0.15; deepSplit, 2; minModuleSize, 30; and eventually visualized by the Cytoscape v3.8.0 (http://cytoscape.org/).

6 Abbreviations

ARF: Auxin Response Factor

bZIP: Basic Leucine-Zipper

BLASTP: Basic Local Alignment Search Tool – Protein

CoGe: Comparative Genomics

CESA: Cellulose Synthase

CSL: Cellulose Synthase-like

CSLA: Cellulose Synthase-like A

CSLB: Cellulose Synthase-like B

CSLC: Cellulose Synthase-like C

CSLD: Cellulose Synthase-like D

CSLE: Cellulose Synthase-like E

CSLG: Cellulose Synthase-like G

GH9B5: Glycosyl Hydrolase 9B5

GH9C2: Glycosyl Hydrolase 9C2

iTOL: Interactive Tree of Life

IAA: Auxin Induced

MUSCLE: Multiple Sequence Alignment with High Accuracy and High Throughput

MYB: V-Myb Avian Myeloblastosis Viral Oncogene Homolog

NAC: No Apical Meristem (NAM), *Arabidopsis thaliana* Transcription Activation Factor (ATAF1/2) and Cup Shaped Cotyledon (CUC2)

NAM: No Apical Meristem

NCBI: National Center for Biotechnology Information

NST: NAC Secondary Wall Thickening Promoting Factor

RAxML: Randomized Axelerated Maximum Likelihood

RTC: Rosette Terminal Complex

SND: Secondary Wall-Associated NAC Domain

STAR: Spliced Trans Alignment to a Reference

TBtools: Toolbox for Biologists

TMM: Trimmed Mean of M-values

WGCNA: Weighted Correlation Network Analysis

XXT: Xyloglucan Xylosyltransferase

7 Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

The RNA datasets used in the current study are available in the NCBI SRA (Sequence Read Archive) database (Accession No. PRJNA535361). The leaves, roots, xylem, and flowers of the two-year-old *B. pendula* were sampled and sequenced by Illumina HiSeq 2500.

Competing interests

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Funding

Study design and data collection were supported by the National Natural Science Foundation of China, grant number 31870659. Data analysis and data interpretation were supported by the Fundamental Research Funds for the Central Universities, grant number 2572019CG08. The labour service and travel

expense were supported by Heilongjiang Touyan Innovation Team Program (Tree Genetics and Breeding Innovation Team).

Authors' contributions

All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by SoC and XL. Conceived and supervised were performed by XZ and SuC.

Acknowledgements

Not applicable.

8 References

1. Hynynen J, Niemistö P, Viherä-Aarnio A, Brunner A, Hein S, Velling P: **Silviculture of birch (Betula pendula Roth and Betula pubescens Ehrh.) in northern Europe**. *Forestry* 2010, **83**(1):103-119.
2. Ritonga FN, Chen S: **Physiological and Molecular Mechanism Involved in Cold Stress Tolerance in Plants**. *Plants* 2020, **9**(5):560.
3. Aspelmeier S, Leuschner C: **Genotypic variation in drought response of silver birch (Betula pendula Roth): leaf and root morphology and carbon partitioning**. *Trees* 2006, **20**(1):42-52.
4. Zhao X, Zheng T, Shao L, Xiao Z, Wang F, Li S, Zang L, Zheng M, Li Y, Qu G-Z: **Variation analysis of physiological traits in betula platyphylla overexpressing TaLEA-ThbZIP gene under salt stress**. *PloS one* 2016, **11**(11):e0164820.
5. Dubois H, Verkasalo E, Claessens H: **Potential of Birch (Betula pendula Roth and B. pubescens Ehrh.) for Forestry and Forest-Based Industry Sector within the Changing Climatic and Socio-Economic Context of Western Europe**. *Forests* 2020, **11**(3):336.
6. Zhao X, Xia H, Wang X, Wang C, Liang D, Li K, Liu G: **Variance and stability analyses of growth characters in half-sib Betula platyphylla families at three different sites in China**. *Euphytica* 2016, **208**(1):173-186.
7. Buchanan BB, Gruissem W, Jones RL: **Biochemistry and molecular biology of plants**: John Wiley & Sons; 2015.
8. Festucci-Buselli RA, Otoni WC, Joshi CP: **Structure, organization, and functions of cellulose synthase complexes in higher plants**. *Brazilian Journal of Plant Physiology* 2007, **19**(1):1-13.
9. Li S, Bashline L, Zheng Y, Xin X, Huang S, Kong Z, Kim SH, Cosgrove DJ, Gu Y: **Cellulose synthase complexes act in a concerted fashion to synthesize highly aggregated cellulose in secondary cell walls of plants**. *Proceedings of the National Academy of Sciences* 2016, **113**(40):11348-11353.
10. Kimura S, Laosinchai W, Itoh T, Cui X, Linder CR, Brown RM: **Immunogold labeling of rosette terminal cellulose-synthesizing complexes in the vascular plant Vigna angularis**. *The Plant Cell* 1999,

11(11):2075-2085.

11. Chen S, Zhao Y, Zhao X, Chen S: **Identification of putative lignin biosynthesis genes in Betula pendula**.
12. Pear JR, Kawagoe Y, Schreckengost WE, Delmer DP, Stalker DM: **Higher plants contain homologs of the bacterial celA genes encoding the catalytic subunit of cellulose synthase**. *Proceedings of the National Academy of Sciences* 1996, **93**(22):12637-12642.
13. Holland N, Holland D, Helentjaris T, Dhugga KS, Xoconostle-Cazares B, Delmer DP: **A comparative analysis of the plant cellulose synthase (CesA) gene family**. *Plant Physiol* 2000, **123**(4):1313-1324.
14. Richmond T: **Higher plant cellulose synthases**. *Genome biology* 2000, **1**(4):reviews3001. 3001.
15. Djerbi S, Lindskog M, Arvestad L, Sterky F, Teeri TT: **The genome sequence of black cottonwood (Populus trichocarpa) reveals 18 conserved cellulose synthase (CesA) genes**. *Planta* 2005, **221**(5):739-746.
16. Burton RA, Shirley NJ, King BJ, Harvey AJ, Fincher GB: **The CesA gene family of barley. Quantitative analysis of transcripts reveals two groups of co-expressed genes**. *Plant Physiol* 2004, **134**(1):224-236.
17. Appenzeller L, Doblin M, Barreiro R, Wang H, Niu X, Kollipara K, Carrigan L, Tomes D, Chapman M, Dhugga KS: **Cellulose synthesis in maize: isolation and expression analysis of the cellulose synthase (CesA) gene family**. *Cellulose* 2004, **11**(3-4):287-299.
18. Liu X, Wang Q, Chen P, Song F, Guan M, Jin L, Wang Y, Yang C: **Four novel cellulose synthase (CESA) genes from birch (Betula platyphylla Suk.) involved in primary and secondary cell wall biosynthesis**. *International journal of molecular sciences* 2012, **13**(10):12195-12212.
19. Huang H, Jiang C, Tong Z, Cheng L, Zhu M, Lin E: **Eight distinct cellulose synthase catalytic subunit genes from Betula luminifera are associated with primary and secondary cell wall biosynthesis**. *Cellulose* 2014, **21**(4):2183-2198.
20. Salojärvi J, Smolander O-P, Nieminen K, Rajaraman S, Safronov O, Safdari P, Lamminmäki A, Immanen J, Lan T, Tanskanen J: **Genome sequencing and population genomic analyses provide insights into the adaptive landscape of silver birch**. *Nature genetics* 2017, **49**(6):904.
21. Joshi CP: **Xylem-specific and tension stress—responsive expression of cellulose synthase genes from aspen trees**. In: *Biotechnology for Fuels and Chemicals.* Springer; 2003: 17-25.
22. Jensen JK, Schultink A, Keegstra K, Wilkerson CG, Pauly M: **RNA-Seq analysis of developing nasturtium seeds (Tropaeolum majus): identification and characterization of an additional galactosyltransferase involved in xyloglucan biosynthesis**. *Molecular plant* 2012, **5**(5):984-992.
23. Schreiber M, Wright F, MacKenzie K, Hedley PE, Schwerdt JG, Little A, Burton RA, Fincher GB, Marshall D, Waugh R: **The barley genome sequence assembly reveals three additional members of the CslF (1, 3; 1, 4)-β-glucan synthase gene family**. *Plos One* 2014, **9**(3).
24. Doblin MS, Pettolino FA, Wilson SM, Campbell R, Burton RA, Fincher GB, Newbigin E, Bacic A: **A barley cellulose synthase-like CSLH gene mediates (1, 3; 1, 4)-β-D-glucan synthesis in transgenic Arabidopsis**. *Proceedings of the National Academy of Sciences* 2009, **106**(14):5996-6001.

25. Mutwil M, Ruprecht C, Giorgi FM, Bringmann M, Usadel B, Persson S: **Transcriptional wiring of cell wall-related genes in Arabidopsis**. *Molecular plant* 2009, **2**(5):1015-1024.
26. Handakumbura PP, Matos DA, Osmont KS, Harrington MJ, Heo K, Kafle K, Kim SH, Baskin TI, Hazen SP: **Perturbation ofBrachypodium distachyon CELLULOSE SYNTHASE A4or7results in abnormal cell walls**. *Bmc Plant Biol* 2013, **13**(1):131.
27. Glass M, Barkwill S, Unda F, Mansfield SD: **Endo-β-1, 4-glucanases impact plant cell wall development by influencing cellulose crystallization**. *Journal of integrative plant biology* 2015, **57**(4):396-410.
28. Liepman AH, Wilkerson CG, Keegstra K: **Expression of cellulose synthase-like (Csl) genes in insect cells reveals that CslA family members encode mannan synthases**. *Proceedings of the National Academy of Sciences* 2005, **102**(6):2221-2226.
29. Goubet F, Barton CJ, Mortimer JC, Yu X, Zhang Z, Miles GP, Richens J, Liepman AH, Seffen K, Dupree P: **Cell wall glucomannan in Arabidopsis is synthesised by CSLA glycosyltransferases, and influences the progression of embryogenesis**. *The Plant Journal* 2009, **60**(3):527-538.
30. Galway ME, Eng RC, Schiefelbein JW, Wasteneys GO: **Root hair-specific disruption of cellulose and xyloglucan in AtCSLD3 mutants, and factors affecting the post-rupture resumption of mutant root hair growth**. *Planta* 2011, **233**(5):985-999.
31. Chou Y-H, Pogorelko G, Zabotina OA: **Xyloglucan xylosyltransferases XXT1, XXT2, and XXT5 and the glucan synthase CSLC4 form Golgi-localized multiprotein complexes**. *Plant Physiol* 2012, **159**(4):1355-1366.
32. Davis J, Brandizzi F, Liepman AH, Keegstra K: **Arabidopsis mannan synthase CSLA9 and glucan synthase CSLC4 have opposite orientations in the Golgi membrane**. *The Plant Journal* 2010, **64**(6):1028-1037.
33. Kim W-C, Kim J-Y, Ko J-H, Kang H, Han K-H: **Identification of direct targets of transcription factor MYB46 provides insights into the transcriptional regulation of secondary wall biosynthesis**. *Plant molecular biology* 2014, **85**(6):589-599.
34. Kim W-C, Reca I-B, Kim Y, Park S, Thomashow MF, Keegstra K, Han K-H: **Transcription factors that directly regulate the expression of CSLA9 encoding mannan synthase in Arabidopsis thaliana**. *Plant molecular biology* 2014, **84**(4-5):577-587.
35. Yang J, Tian L, Sun M-X, Huang X-Y, Zhu J, Guan Y-F, Jia Q-S, Yang Z-N: **AUXIN RESPONSE FACTOR17 is essential for pollen wall pattern formation in Arabidopsis**. *Plant Physiol* 2013, **162**(2):720-731.
36. Zhong R, Richardson EA, Ye Z-H: **Two NAC domain transcription factors, SND1 and NST1, function redundantly in regulation of secondary wall synthesis in fibers of Arabidopsis**. *Planta* 2007, **225**(6):1603-1611.
37. Ko J-H, Jeon H-W, Kim W-C, Kim J-Y, Han K-H: **The MYB46/MYB83-mediated transcriptional regulatory programme is a gatekeeper of secondary wall biosynthesis**. *Ann Bot-London* 2014, **114**(6):1099-1107.

38. Oomen RJ, Tzitzikas EN, Bakx EJ, Straatman-Engelen I, Bush MS, McCann MC, Schols HA, Visser RG, Vincken J-P: **Modulation of the cellulose content of tuber cell walls by antisense expression of different potato (Solanum tuberosum L.) CesA clones**. *Phytochemistry* 2004, **65**(5):535-546.
39. Zhong R, Morrison WH, Freshour GD, Hahn MG, Ye Z-H: **Expression of a mutant form of cellulose synthase AtCesA7 causes dominant negative effect on cellulose biosynthesis**. *Plant Physiol* 2003, **132**(2):786-795.
40. Tan H-T, Shirley NJ, Singh RR, Henderson M, Dhugga KS, Mayo GM, Fincher GB, Burton RA: **Powerful regulatory systems and post-transcriptional gene silencing resist increases in cellulose content in cell walls of barley**. *Bmc Plant Biol* 2015, **15**(1):62.
41. Doblin MS, Kurek I, Jacob-Wilk D, Delmer DP: **Cellulose biosynthesis in plants: from genes to rosettes**. *Plant Cell Physiol* 2002, **43**(12):1407-1420.
42. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL: **BLAST+: architecture and applications**. *Bmc Bioinformatics* 2009, **10**(1):421.
43. Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, Geer RC, He J, Gwadz M, Hurwitz DI: **CDD: NCBI's conserved domain database**. *Nucleic acids research* 2015, **43**(D1):D222-D226.
44. Chen C, Xia R, Chen H, He Y: **TBtools, a Toolkit for Biologists integrating various biological data handling tools with a user-friendly interface**. *BioRxiv* 2018:289660.
45. Edgar RC: **MUSCLE: multiple sequence alignment with high accuracy and high throughput**. *Nucleic Acids Res* 2004, **32**(5):1792-1797.
46. Stamatakis A: **RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies**. *Bioinformatics* 2014, **30**(9):1312-1313.
47. Letunic I, Bork P: **Interactive Tree Of Life (iTOL) v4: recent updates and new developments**. *Nucleic Acids Res* 2019, **47**(W1):W256-W259.
48. Chen S, Lin X, Zhang D, Li Q, Zhao X, Chen S: **Genome-Wide Analysis of NAC Gene Family in Betula pendula**. *Forests* 2019, **10**(9):741.
49. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR: **STAR: ultrafast universal RNA-seq aligner**. *Bioinformatics* 2013, **29**(1):15-21.
50. Li B, Dewey CN: **RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome**. *Bmc Bioinformatics* 2011, **12**.
51. Dai X, Sinharoy S, Udvardi M, Zhao PX: **PlantTFcat: an online plant transcription factor and transcriptional regulator categorization and analysis tool**. *Bmc Bioinformatics* 2013, **14**(1):321.
52. Langfelder P, Horvath S: **WGCNA: an R package for weighted correlation network analysis**. *Bmc Bioinformatics* 2008, **9**.

9 Table

Table 1 Putative *B. pendula* cellulose synthase genes in 7 gene families

Gene family	Gene name	Gene ID (*B. pendula*)
CESA	*BpCESA1*	*Bpev01.c0196.g0006*
	BpCESA2	*Bpev01.c0205.g0006*
	BpCESA3	*Bpev01.c0777.g0012*
	BpCESA4	***Bpev01.c0000.g0006***
	BpCESA5	*Bpev01.c0402.g0034*
	BpCESA6	*Bpev01.c0480.g0087*
	BpCESA7	*Bpev01.c0598.g0015*
	BpCESA8	*Bpev01.c0603.g0003*
	BpCESA9	***Bpev01.c0374.g0017***
	BpCESA10	***Bpev01.c0374.g0018***
CSLA	*BpCSLA1*	*Bpev01.c0902.g0015*
	BpCSLA2	***Bpev01.c0169.g0024***
	BpCSLA3	***Bpev01.c2286.g0004***
CSLB	*BpCSLB1*	*Bpev01.c1000.g0017*
	BpCSLB2	*Bpev01.c1000.g0013*
	BpCSLB3	*Bpev01.c1000.g0018*
	BpCSLB4	*Bpev01.c1193.g0003*
	BpCSLB5	*Bpev01.c1193.g0012*
	BpCSLB6	*Bpev01.c1193.g0006*
	BpCSLB7	*Bpev01.c1000.g0016*
CSLC	***BpCSLC1***	***Bpev01.c0094.g0029***
	BpCSLC2	*Bpev01.c0515.g0003*
	BpCSLC3	*Bpev01.c0058.g0002*
	BpCSLC4	***Bpev01.c0018.g0093***
CSLD	*BpCSLD1*	*Bpev01.c0016.g0057*
	BpCSLD2	*Bpev01.c0016.g0055*
	BpCSLD3	*Bpev01.c0423.g0009*
	BpCSLD4	***Bpev01.c0949.g0008***
	BpCSLD5	*Bpev01.c1082.g0006*
	BpCSLD6	*Bpev01.c1484.g0010*
	BpCSLD7	*Bpev01.c0364.g0008*
CSLE	*BpCSLE1*	*Bpev01.c1469.g0001*
	BpCSLE2	*Bpev01.c1782.g0020*
	BpCSLE3	*Bpev01.c1782.g0018*
	BpCSLE4	*Bpev01.c2470.g0006*
CSLG	*BpCSLG1*	*Bpev01.c1225.g0008*
	BpCSLG2	*Bpev01.c1739.g0002*
	BpCSLG3	*Bpev01.c1739.g0001*
	BpCSLG4	*Bpev01.c0188.g0037*
	BpCSLG5	*Bpev01.c2210.g0001*
	BpCSLG6	*Bpev01.c2469.g0001*
	BpCSLG7	*Bpev01.c0995.g0003*
	BpCSLG8	*Bpev01.c1270.g0001*
	BpCSLG9	*Bpev01.c0774.g0001*
	BpCSLG10	*Bpev01.c0774.g0003*
	BpCSLG11	*Bpev01.c0774.g0002*

[a] Gene information in bold is for the genes most probably encode secondary cell wall synthesis enzymes

Figures

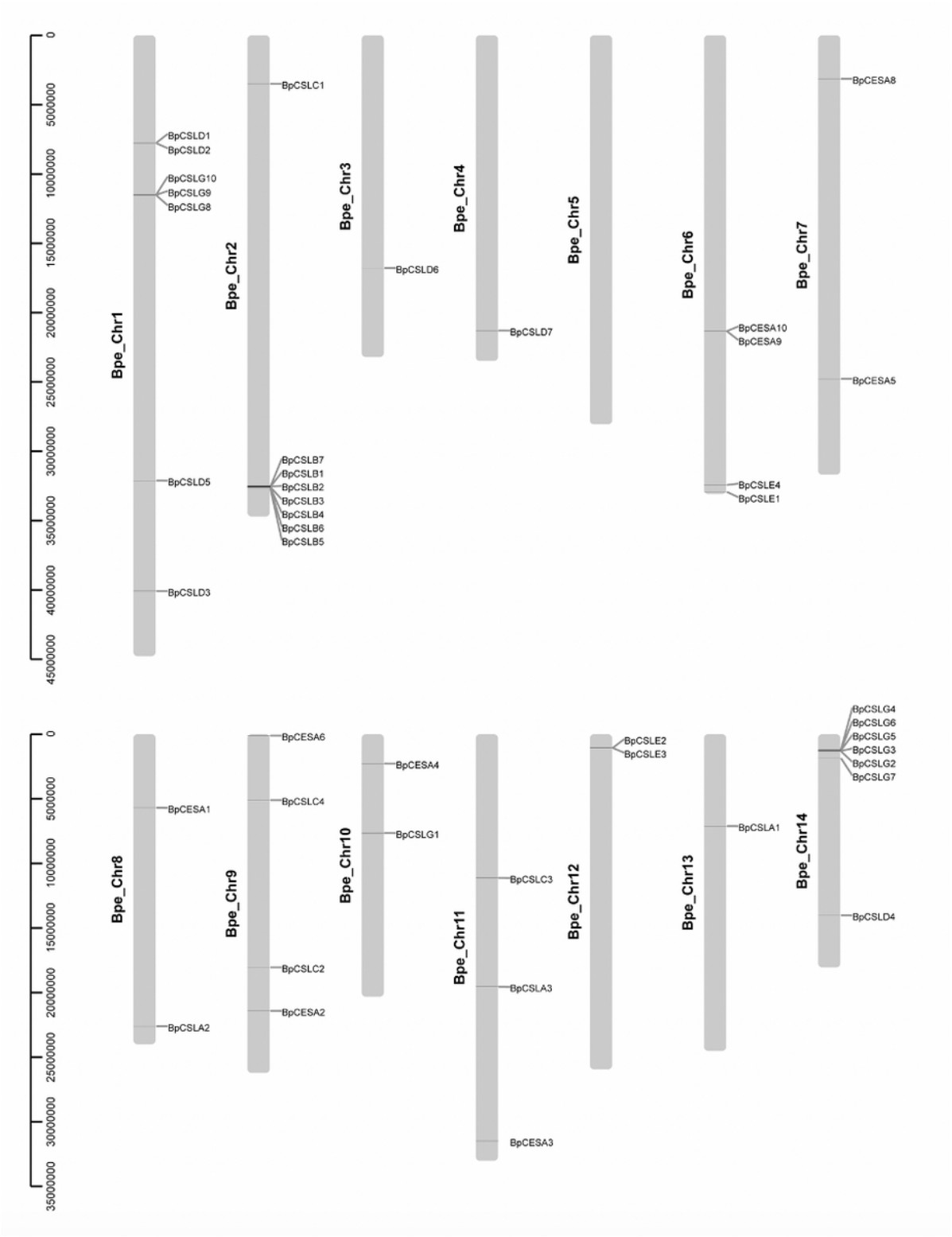

Figure 1

The chromosomal location of B. pendula cell wall synthesis genes. The silver line represents the chromosome of B. pendula, and the black line represents the relative location of CESA and CSL genes on the chromosome.

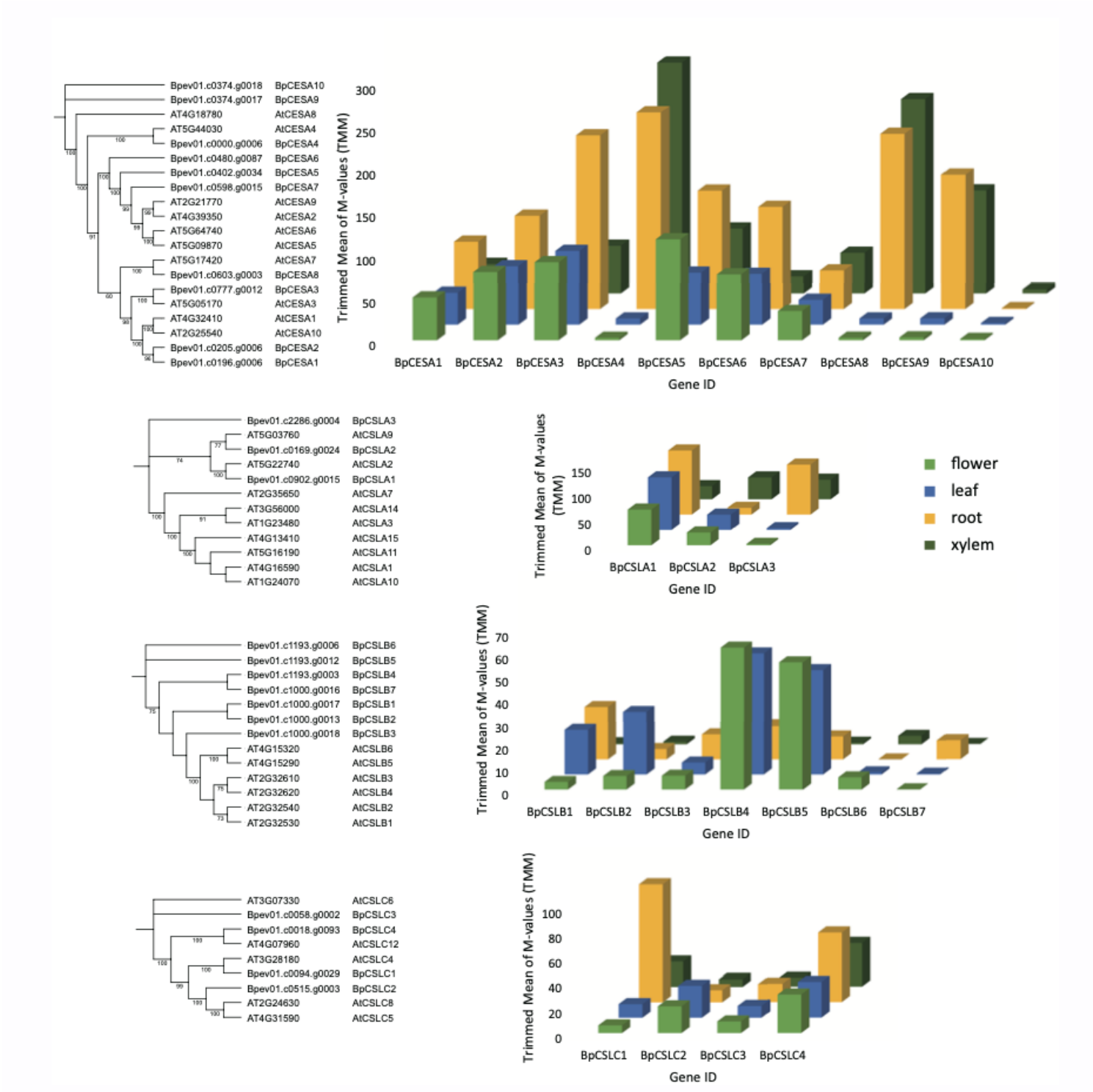

Figure 2

Tissue-specific expression profiles and phylogenetic analysis of BpCESA, BpCSLA, BpCSLB and BpCSLC families in B. pendula. The expression was analyzed in three independent biological replicates of each tissue, and the phylogenetic tree (1,000 bootstraps) was constructed by RAxML using the maximum likelihood algorithm.

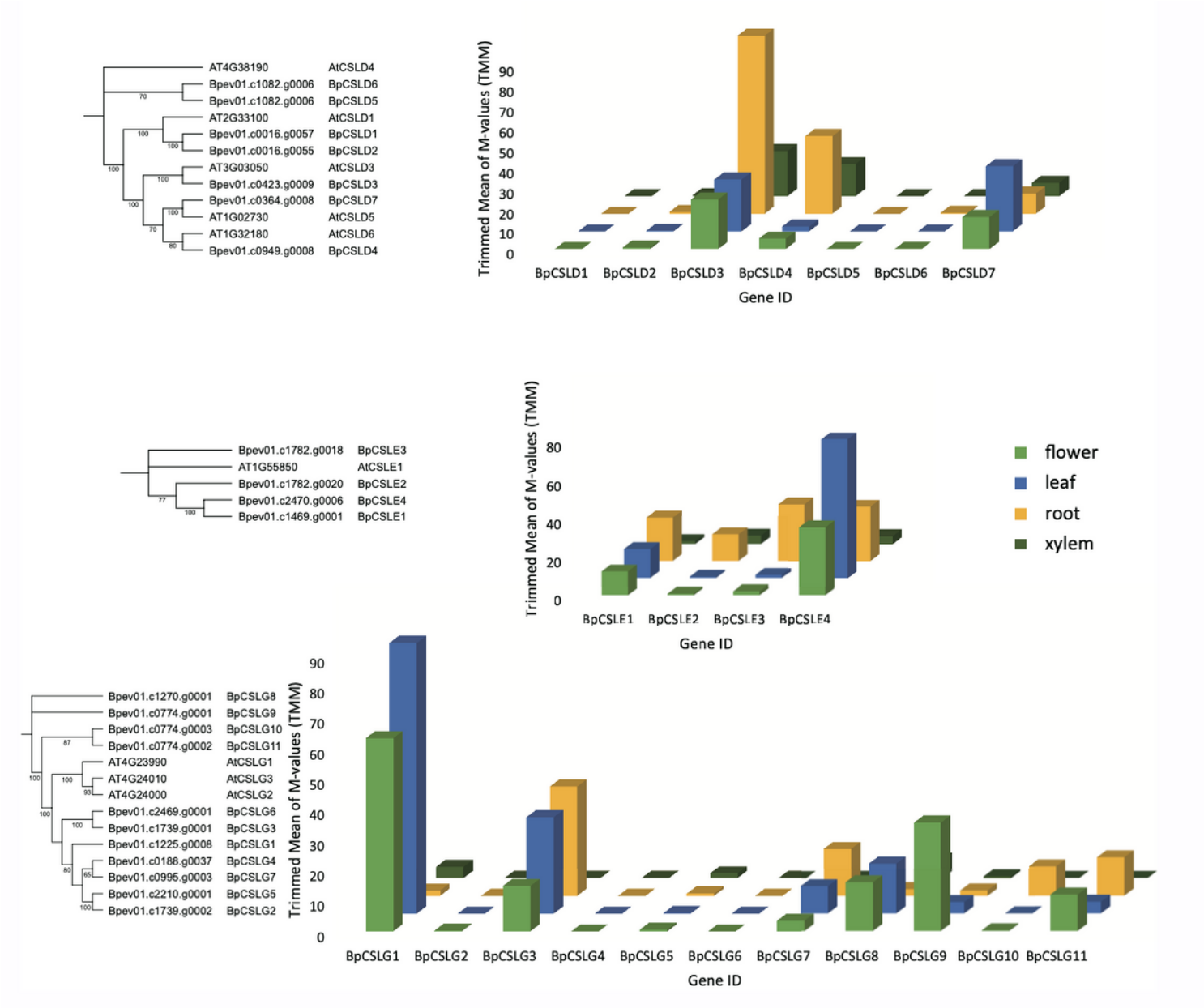

Figure 3

Tissue-specific expression profiles and phylogenetic analysis of BpCSLD, BpCSLE and BpCSLG families in B. pendula. The expression was analyzed in three independent biological replicates of each tissue, and the phylogenetic tree (1,000 bootstraps) was constructed by RAxML using the maximum likelihood algorithm.

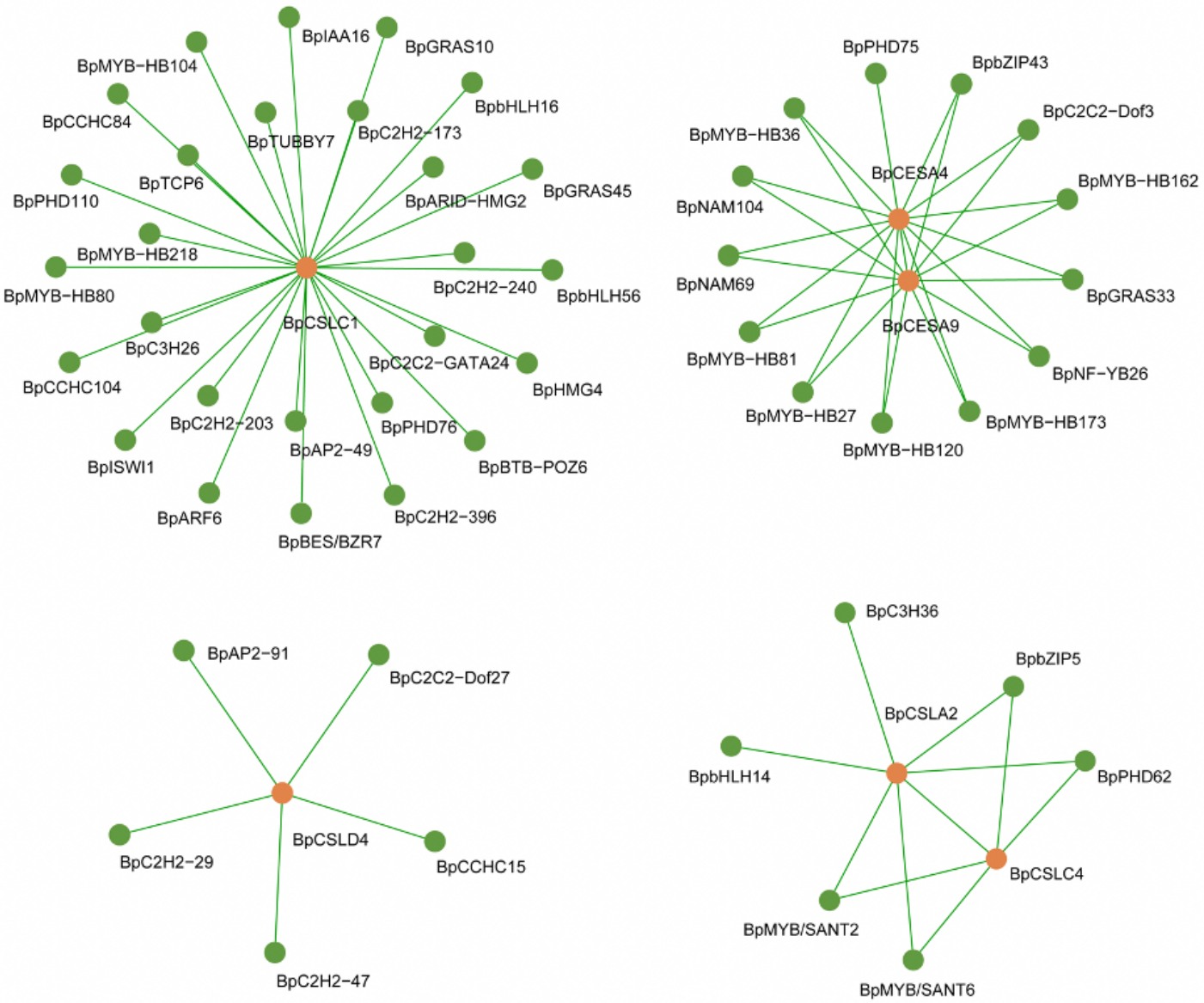

Figure 4

The transcription factor regulatory network calculated by WGCNA. The green dots are transcription factors, and the orange dots are cell wall synthesis genes.