Genome-Wide Identification and Analysis of Class III Peroxidases in *Betula pendula*

Kewei Cai
Northeast Forestry University
Song Chen
Northeast Forestry University
Xiyang Zhao
Northeast Forestry University
Su Chen (chensunefu@163.com)
State Key Laboratory of Tree Genetics and Breeding, Northeast Forestry University, Harbin 150040, China

Research article

Keywords: Betula pendula, class III peroxidases, phylogenetic analysis, chromosomal location, expression pattern

Posted Date: May 26th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-29021/v1

Abstract

Background: Class III peroxidases (POD) proteins are widely present in the plant kingdom that are involved in a broad range of physiological processes including stress responses and lignin polymerization throughout the plant life cycle. However, little is known about the *POD* genes in *Betula pendula*, although it has been characterized in *Arabidopsis*, rice and maize. The *POD* genes remain to be determined in *Betula pendula*.

Results: A total of 90 nonredundant *POD* genes were identified in *Betula pendula*. (designated *BpPODs*). These *POD* genes were divided into twelve groups based on their phylogenetic relationships. The *BpPODs* are unevenly distributed on the 14 chromosomes. In addition, some *BpPOD* genes were located sequentially in tandem on chromosomes, inferred that tandem duplication contributes to the expansion of the *POD* genes family in *Betula pendula*. Analysis of the distribution of conserved domains of BpPOD proteins showed that all these proteins contain highly conserved motifs. We also investigated their expression patterns in different tissues, the results show that some *BpPOD* genes might play significant roles in root, xylem, leaf and flower. Furthermore, under low temperature conditions, some *BpPOD* genes showed different expression patterns at different times.

Conclusions: Comprehensive study of the *POD* genes suggests that their functional diversity during *Betula pendula* growth and development. Our findings provide a basis for further functional analysis on *POD* genes family in *Betula pendula.*

Background

Peroxidases or peroxide reductases (POD, EC number 1.11.1.x) are a large group of oxidases existing in animals, plants and microorganisms, which catalyzes the oxidation of a particular substrate by hydrogen peroxide [1]. Among them, class III peroxidases, belonging to the haem peroxidase subfamily, exist only in plants and have an extremely widespread presence in the plant kingdom [2]. The Class III peroxidase in plants are also reported as POX [3, 4], GPX [5], Prx [6], ClassIII PRX [7], and POD [8, 9]. Most plant species contain dozens of Class III peroxidases, for example, switchgrass [7] genome contains more than 200 *POD* coding genes, and *Populus* [10], rice and *Arabidopsis* contain 93, 138 and 73 members of POD family, respectively [6, 11] .

POD are involved in a broad range of physiological processes throughout the plant life cycle, probably due to the high number of enzymatic isoforms (isoenzymes) and to the versatility of their enzyme-catalysed reactions [12]. Recent studies indicate that POD has two most important functions in plants: on the one hand, it is related to the normal morphogenesis of plants and plays a role in the growth and development of plants. On the other hand, it is related to the resistance of plants, including disease resistance, cold resistance, drought resistance, *etc.*, and it is one of the important protective enzymes of plant protective enzymes [13, 14]. Although it is known that POD play a key role in cell growth and

response to abiotic stress, the specific function of each member of the family is still elusive. The comprehensive researches are necessary to explore the role of POD in plant growth and defense.

The gene family always arose from multiple ways including tandem duplication, duplicative transposition, and whole genome duplication (WGD), which was followed by mutation and divergence [15]. During the last decade, several molecular biology approaches have been developed to isolate, characterize and study the expression of *POD* genes family in plants [6]. *Betula pendula* is a pioneer boreal tree that can be induced to flower within one year [16], it plays an important role in people life. Up to now, however, no genome-wide characterization of the POD family in *B. pendula*. It has been shown that POD is related to the synthesis of lignin [17] and cork [18, 19], and lignin is considered as an important defense means against invasion and expansion of pathogens [20, 21]. At the same time, a large number of experimental evidences of stress treatment showed that under the stress of drought and low temperature, the expression of POD increased significantly [22, 23].

Since *Betula pendula* is a widespread species and has many applications in the pulp and paper industry, it is necessary to study its development and physiology [24]. To understand the role of POD family in lignin synthesis and resistance to biotic and abiotic stresses in *B. pendula* will greatly contribute to its application in industrial production. Fortunately, *B. pendula* has attracted much attention, particularly by the availability of its genome sequences [25], which gave us an opportunity to carry out a comprehensive genome-wide analysis for exploring the potential functions of the *POD* genes in *B. pendula.*

In the present study, a genome-wide analysis of *POD* genes family from *B. pendula* was conducted via genomic sequence, including *BpPOD* gene models, phylogenetic relationship, conserved motif, chromosome location and other structural features [26]. we performed for the first time the comprehensive analysis to the *POD* genes involved in lignin synthesis and abiotic stress response in *B. pendula.* Our study provides important insights for further study of the potential role of *POD* genes family in *B. pendula* growth and development.

Results

Identification of *POD* genes

To identify members of *POD* family in *B. pendula*, we used the 73 *POD* genes of *Arabidopsis* to obtain the best hits in the *B. pendula* genome by BLASTP. A total of 90 putative *PODs* were identified in the *B. pendula* genome. We further examined the conserve domains of proteins encoded by these genes using Pfam [27] and SMART [28] database. The results revealed that all the genes have classical *POD* domain structures, which demonstrates the reliability of the results. The *B. pendula* genome contains more *PODs* than *Arabidopsis* (73) [6], but fewer than *Populus euphratica* (93) [29], chinese pear (94) [26], and rice (138) [11]. We defined the *BpPODs* as *BpPOD1* to *BpPOD90*. The isoelectric point (PI) varied from 4.28 to 9.6 with a mean of 7.25 and >7.0 of 52.2% POD proteins. Their detailed information, including chromosome location, gene name and molecular weight (MW) gene size of each *BpPOD* gene/protein, was listed in **Table1**.

Phylogenetic analyses of POD proteins in *B. pendula*

To investigate the evolutionary history and phylogenetic relationships among the members of POD family in *B. pendula*, a phylogenetic tree was constructed with the Neighbor-Joining method based on multiple sequence alignment of the 90 BpPODs, with 1000 bootstrap replicates (**Figure 1**). The BpPOD proteins were divided into twelve major subgroups with high bootstrap probabilities, designated group I to group XII. The *POD* genes of each subgroup is unevenly distributed, with the number of members varies from 4 to 15. Subgroup VIII contains the most members (15), subgroup X, XI, XII contains the least number of members, with only 4 members.

Analysis of conserved amino acid motifs

To understand the functional regions of BpPODs, conserved amino acid motifs analyses of BpPOD proteins were performed. A total of eight conserved amino acid motifs were identified in the BpPOD proteins (**Figure 2**). All BpPOD proteins contain at least one conserved amino acid motif. For example, BpPOD55 only contains motif 8, BpPOD83 contains motif 1 and 7, while BpPOD10 proteins contain all the eight conserved amino acid motifs.

Most of the closely related members have the same motif compositions, suggesting that there are functional similarities between POD proteins within the same subgroup [30]. We found that motifs 1, 2, 3, 4, 5 and 7 appeared in nearly all members of BpPOD proteins, these motifs might be important for the functions of BpPOD proteins.

Analysis of chromosomal location

To investigate the genome organization and distribution of *BpPODs* on different chromosomes of *B. pendula*, a chromosome map was constructed. The results show that the 90 *BpPOD* genes were distributed among 14 chromosomes, as shown in **Figure 3**, the physical locations of these *BpPODs* on chromosomes were scattered and uneven. Chromosome 1 and 8 contains the most *BpPOD* genes (14), followed by chromosome 13 (10). Eight *BpPOD* genes were simultaneously distributed on chromosomes 5 and 7, whereas chromosomes 14 had only one and chromosomes 11 does not include the *POD* genes. In addition, some chromosomes exhibit a relatively high density of *BpPOD* genes, such as the bottoms of chromosomes 13 and the top of chromosome 8.

Gene duplication, including segmental and tandem duplication, is considered to be one of the primary driving forces in the evolution of genomes [31, 32]. In this study, among the 90 *BpPOD* genes identified, a large number of *BpPOD* genes have the same duplicated regions (**Figure 4**). Generally, a gene cluster is the result of gene tandem duplication [33]. In this study, we found that some *BpPOD* genes were adjacent to each other (**Figure 3**). For instance, *BpPOD17-20*, *BpPOD22-29* and *BpPOD11-15* were located sequentially in tandem on chromosomes 5, 8, and 13, respectively, implying that these genes might arise from recent tandem duplication events [34]. The result indicated that tandem duplications play main contributors in the expansion of the *BpPOD* genes family. However, in previous studies, segmental

duplication and tandem duplication were identified in maize POD family [30]. This indicates that there are significant differences in the *POD* genes expansion pattern in *B. pendula* and maize, which strongly implied that POD family members have different expansion patterns among different species. It may be the reason why the POD family members (90) in *B. pendula* were less than those in the maize (119) [26].

Tissue-specifc expression of *BpPOD* genes in *B. pendula*

To better understand the functions of *POD* genes in the growth and development of *B. pendula*, their expression profiles in different tissues (including root, xylem, young leaf and flower) were analyzed with publicly available transcriptome data. Of the 90 *BpPOD* genes, 69 genes were expressed in one or more birch tissues, while 21 *BpPOD* genes exhibited no expression in various individual tissues. The heat map (**Figure 5**) demonstrated that most *BpPOD* genes had tissue-specific or preferential expression patterns. *BpPOD6*, *BpPOD21* and *BpPOD37* were highly expressed in xylem, suggesting that they may play specific roles in xylem development. Several *BpPOD* genes were expressed in root during development, revealing the significant roles of these genes in root growth, such as *BpPOD62, BpPOD63* and *BpPOD65*. *BpPOD78* and *BpPOD19* showed higher expression levels in young leaf and flower, respectively, implying their specific roles in leaf and flower development. The expression level of *BpPOD6* was high in xylem and low in root, leaf and flower. In contrast, *BpPOD67, BpPOD68, BpPOD80* and *BpPOD81* had no expression in any of the investigated tissues, suggesting that these genes are not involved in the development of these tissues. *BpPOD21, BpPOD59* and *BpPOD62* were highly expressed in developing xylem, root, leaf and flower. In conclusion, the variations in the expression of *BpPOD* genes in different tissues revealed that *POD* genes may be involved in several processes during *B. pendula* growth and development.

Responses of *BpPOD* genes expression to cold treatment

Several roles have been attributed to plant peroxidases in response to biotic and abiotic stresses [35]. In recent years, the number of studies on *POD* genes response to abiotic stress have been reported [30]. For example, *Arabidopsis* overexpressing *AtPOD3* showed an increase in dehydration and salt tolerance, whereas the antisense suppression of *AtPOD3* exhibited dehydration and salt sensitive phenotypes [36]. The expression of *POD* genes is induced by various environmental stresses, such as metal, pathogens, humidity, temperature, anoxia and potassium deficiency [35], suggesting that *POD* genes are involved in plant defense. In this study, we examined the expression levels of the *BpPOD* genes in response to low temperature stress. As shown in **Figure 6**, the result indicated that the expression of most *BpPOD* genes was altered under cold treatment. After cold treatment, the expression levels of *BpPOD4, BpPOD13, BpPOD15, BpPOD17* and *BpPOD21* were significantly induced at a relatively early stage (0.5 h after treatment), this suggests that these genes play a more important role in *B. pendula* under cold treatment than other genes. The expression levels of *BpPOD19*, *BpPOD21, BpPOD39* and *BpPOD47* were increased after 1.5 h treatment of low temperature. *BpPOD50* and *BpPOD58* did not response to cold treatment at the beginning (0.5 h), and were slightly increased after 2 h exposure to low temperature. The low temperature responsive *BpPODs* may play important roles in birch under cold stress.

Discussion

It is reported that members of Class III Peroxidases gene family are involved in the regulation of a variety of processes [6, 29], and play a key role in biological and abiotic stress responses during plant growth and development [30]. Systematic and comprehensive analyses of *POD* genes families have been published for *Arabidopsis thaliana* [6], *Populus trichocarpa* [29], *Zea mays* [30] and *Oryza sativa* [11], but a genome-wide study of the POD family has not previously been reported in *B. pendula*. The published genome data of *B. pendula* [25] provides a useful tool for analysis of the *BpPOD* genes family in birch.

In the present study, 90 non-redundant *POD* genes were identified in *B. pendula*, the number is higher than that in *Arabidopsis* (73), but lower than that in maize (119) and rice (138), which indicates that the *POD* genes in *B. pendula* have expanded compared to those in *Arabidopsis*. Subsequently, we performed analyses of the phylogenic relationships, chromosomal locations, conserved motifs and expression profiles [37].

In the process of genome evolution, tandem duplication and segmental duplication were the main factors that led to the expansion of gene family [32]. Certain studies have shown that tandem duplication was largely responsible for the expansion of brich gene families [30], such as tandem duplication are the main reason for the expansion of *B. pendula* NAC gene family [15]. By contrast, segmental duplication has contributed significantly to the expansion of this gene family in *populus* [34] and chinese pear [26]. Interestingly, in this study, we found that some *BpPOD* genes were adjacent to each other, suggesting tandem duplications play main contributors to the expansion of the *BpPOD* genes family. However, in maize, the segmental duplication and tandem duplication almost identically contributed to the *POD* genes family expansion [30]. These results also explain why the number of *POD* genes in *B. pendula* (90) , pear and *populus* were less than those in maize (119). According to the above analysis, we speculated that the expansion of the *POD* genes family differed between different plants.

The 90 BpPOD proteins possess ten highly conserved motifs. MEME analysis revealed that different conserved motifs are present in each of the BpPOD proteins. Notably, most BpPOD proteins contain all the conserved motifs, while only a few BpPOD proteins contain one or two motifs, which means that these motifs may be involved in the important basic function of the POD protein. However, a few motifs with unknown functions are present in nearly every subgroup, these motifs might play important roles in the BpPOD family.

Gene expression patterns are an important aspect of the study of gene function [38]. High-throughput microarray technology provides a good platform for the study of genome-wide gene expression patterns [39]. We used publicly available genome-wide transcript profiling data from *B. pendula* tissues as a resource to investigate the expression patterns of *BpPODs*. Most *BpPOD* genes exhibited variable expression patterns, suggesting functional diversification of *BpPOD* genes. Twenty-one *BpPOD* genes exhibited no expression in four individual tissues, indicating that *BpPOD* genes are expressed under specific conditions or at specific developmental stages. In this study, we found that of the 90 *BpPOD* genes, the most abundant expression was in the root, followed by the xylem. The result showed that most

highly expressed *POD* genes might play significant roles in root. The highest expression levels of *BpPOD6, BpPOD21* and *BpPOD37* genes were found in xylem. It was suggested that these three genes were participated in regulation of the xylem synthesis in *B. pendula*. *BpPOD59* is most expressed in flowers and leaves, suggesting that it may be related to leaf spreading and flowering formation in *B. pendula.* In addition, several *BpPOD* genes were expressed in all tissues, suggesting that they might play basic roles in *B. pendula*. In conclusion, the expression profiling in this study provides an important basis for further studying of expression and biological functions of the *POD* genes family in *B. pendula.*

The growth and development of plants are usually affected by abiotic stress, such as drought, low temperature and high salinity [39]. A lot of stress-related genes were induced to adapt to these abiotic stresses [40, 41]. A large number of experimental studies [30] on stress treatment showed that under the stress of low temperature and other conditions, *POD* genes expression increased significantly [22, 23]. However, no *POD* genes responding to cold treatment have been reported in *B. pendula*. Thus, we performed a survey of the expression patterns of the *POD* genes in *B. pendula* under cold treatment. The results suggested that fifty *BpPOD* genes were responsive to cold treatment. Most *BpPOD* genes were highly expressed at a relatively early stage (0.5 h after treatment), and with the extension of time, the expression reached the highest level. This indicated that these genes might play an important role in *B. pendula*. By contrast, the expression level of *BpPOD30* and *BpPOD8* genes gradually increased at 2 h after treatment, indicating that these genes are involved in the late reaction of cold treatment. In addition, the expression of a few *BpPOD* genes decreased under cold treatment, we speculate that these genes may also have defense and other specific functions in *B. pendula.* These results indicated that most *POD* genes were induced by low temperature and might contribute to the defense against abiotic stresses in *B. pendula*.

Conclusion

In short, a total of 90 *POD* genes were identified in *B. pendula* and divided into twelve major subgroups. A total of eight conserved amino acid motifs were identified in the BpPOD proteins. Chromosomal location and microsynteny analysis suggested that these *BpPOD* genes were unevenly distributed in fourteen chromosomes. Tandem duplication were identified as the main patterns contributors to the expansion of *POD* genes expansion in *B. pendula.* Finally, expression patterns analysis revealed that some *BpPOD* genes might play significant roles in root, xylem, leaf and flower. Furthermore, under low temperature conditions, some *BpPOD* genes showed different expression patterns at different times. This present study increases our understanding of POD genes in *B. pendula* and lays the foundation for further clarify of the biological functions of these POD proteins in other plants.

Methods

Identification of *B. pendula* peroxidase genes

To identify *B. pendula* peroxidase genes, the *B. pendula* genome sequences were downloaded [25]. The protein sequences of POD family members in the genome of *Arabidopsis* were retrieved from the TAIR database. The 73 *Arabidopsis POD* members were used as queries to identify the candidate sequences of *B. pendula POD* genes using BLASTP. To verify the reliability of the results, all the acquired candidate sequences were examined for the presence of the POD domain using Pfam [42] and SMART [43]. Finally, all candidate *POD* sequences were aligned using ClustalW [44] and remove potentially redundant genes. The theoretical molecular weights (MWs) and isoelectric points (pIs) of the proteins were collected using the online ExPASy program.

Phylogenetic analyses of *B. pendula* peroxidase genes

To investigate the phylogenetic relationships of the peroxidase genes of *B. pendula*, a phylogenetic tree was constructed. Prior to phylogenetic analysis, multiple sequence alignments were generated using MUSCLE [45]. Subsequently, the RAxML [46] was employed to construct an unrooted phylogenetic tree based on alignments using the Neighbor-Joining (NJ) method with the following parameters: model (p-distance), bootstrap (1000 replicates), and gap/missing data (pairwise deletion).

Analysis of conserved motif

To determine conserved motifs structures of the BpPOD proteins , the conserved motifs were detected using the online MEME Tool [47]. The conserved motifs were analyzed with the SMART and PFAM programs.

Analysis of chromosomal location

Analysis of RNA-seq expression

To determine the expression patterns of *BpPOD* genes in different tissues, we downloaded transcriptome data (PRJNA535361) [15] from the NCBI SRA database. The clean reads of each sample were obtained by filtering out reads of low quality. All the clean reads were aligned to the *B. pendula* reference genome using bowtie2 [48]. The RNA-seq (RNA-sequencing) data were analyzed using the RSEM (RNA-seq by Expectation-Maximization) pipeline [49] and the data were processed using a paired-end sequencing mode. RSEM could compute transcript abundance, estimating the number of RNA-seq fragments corresponding to each gene, and normalized expression values as TMM (trimmed mean of M-values).

Abbreviations

POD: class III peroxidases *B. pendula*: *Betula pendula* *BpPODs*: *POD* genes in *Betula pendula* RNA-seq: RNA sequencing

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

All data generated or analysed during this study are included in this published article.

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

This research was supported by the National Natural Science Foundation of China (31870659), the Fundamental Research Funds for the Central Universities (2572019CG08) and Heilongjiang Touyan Innovation Team Program (Tree Genetics and Breeding Innovation Team).

Authors' contributions

CKW was a major contributor in writing the manuscript. CS analyzed the data and make figures. All authors read and approved the final manuscript.

Acknowledgements

Not applicable

References

1. Dunford HB, Stillman JS: On the Function and Mechanism of Action of Peroxidases. *Coordination Chemistry Reviews* 1976, 19(3):187–251.
2. Passardi F, Theiler G, Zamocky M, Cosio C, Rouhier N, Teixera F, Margis-Pinheiro M, Ioannidis V, Penel C, Falquet L: PeroxiBase: The peroxidase database. *Phytochemistry* 2007, 68(12):1605-1611.
3. Mei W, Qin Y, Song W, Li J, Zhu Y: Cotton GhPOX1 encoding plant class III peroxidase may be responsible for the high level of reactive oxygen species production that is related to cotton fiber elongation. *Journal of Genetics and Genomics* 2009, 36(3):141-150.
4. Martínez-Rubio R, Acebes JL, Encina A, Krknen A: Class III peroxidases in cellulose deficient cultured maize cells during cell wall remodeling. *Physiol Plant* 2018, 164(1).

5. Zhang Z, Xin W, Wang S, Zhang X, Wang Q: Xylem sap in cotton contains proteins that contribute to environmental stress response and cell wall development. *Funct Integr Genomics* 2014, 15(1):17-26.
6. Tognolli M, Penel C, Greppin H, Simon P: Analysis and expression of the class III peroxidase large gene family in Arabidopsis thaliana. *Gene* 2002, 288(1-2):0-138.
7. Moural TW, Lewis KM, Barnaba C, Zhu F, Kang CH: Characterization of Class III Peroxidases from Switchgrass. *Plant Physiology* 2016, 173(1).
8. Delannoy E, Jalloul LA, Assigbetsé K, Marmey P, Nicole M: Activity of Class III Peroxidases in the Defense of Cotton to Bacterial Blight. *Molecular Plant Microbe Interactions* 2003, 16(11):1030-1038.
9. Rácz A, Hideg É, Czégény G: Selective responses of class III plant peroxidase isoforms to environmentally relevant UV-B doses. *Journal of Plant Physiology* 2018, 221:101-106.
10. Comparative genomic analysis of the WRKY III gene family inpopulus, grape,arabidopsisand rice. *Biology Direct* 2015, 10(1):48.
11. Passardi F, Longet D, Penel C, Dunand C: The class III peroxidase multigenic family in rice and its evolution in land plants. *Phytochemistry* 2004, 65(13):1879-1893.
12. Passardi F, Cosio C, Penel C, Dunand C: Peroxidases have more functions than a Swiss army knife. 2005.
13. Welinder KG: Superfamily of plant, fungal and bacterial peroxidases. *Current Opinion in Structural Biology* 1992, 2(3):388-393.
14. Hiraga S, Sasaki K, Ito H, Ohashi Y, Matsui H: A large family of class III plant peroxidases. *Plant and Cell Physiology* 2001, 42(5):462-468.
15. Chen S, Lin X, Zhang D, Li Q, Chen S: Genome-Wide Analysis of NAC Gene Family in Betula pendula. *Forests* 2019, 10(9):741.
16. Alonso-Serra J, Safronov O, Lim KJ, Fraser-Miller SJ, Blokhina OB, Campilho A, Chong SL, Fagerstedt K, Haavikko R, Helariutta Y: Tissue-specific study across the stem reveals the chemistry and transcriptome dynamics of birch bark. *New Phytologist* 2019, 222.
17. Huystee RBV: Some Molecular Aspects of Plant Peroxidase Biosynthetic Studies. *Annual Review of Plant Biology* 2003, 38(1):205-219.
18. Mittler R, Zilinskas BA: Molecular-Cloning And Characterization Of A Gene Encoding Pea Cytosolic Ascorbate Peroxidase. *Journal of Biological Chemistry* 1992, 267(30):21802-21807.
19. Christensen JrH, Bauw G, Welinder KG, Montagu MV, Boerjan W: Purification and Characterization of Peroxidases Correlated with Lignification in Poplar Xylem. *Plant Physiology* 1998, 118(1):125-135.
20. Lewis NG: Lignin: Occurrence, Biogenesis And Biodegradation. *Annu Rev Plant Physiol Plant Mol Biol* 1990, 41(1):455-496.
21. Agostini E, Forchetti SMD, Tigier HA: Production of peroxidases by hairy roots of Brassica napus. *Plant Cell Tissue & Organ Culture* 1997, 47:177-182.
22. Botella MA, Quesada MA, Kononowicz AK, Bressan RA, Valpuesta V: Characterization and in situ localization of a salt-induced tomato peroxidase mRNA. *Plant Molecular Biology* 1994, 25(1):105-

114.

23. Chittoor JM, Leach JE, White FF: Differential Induction of a Peroxidase Gene Family During Infection of Rice by Xanthomonas oryzae pv. oryzae. *Molecular Plant Microbe Interactions* 1997, 10(7):861-871.
24. Zhang R, Yang C, Wang C, Wei Z, Xia D, Wang Y, Liu G, Wang Y: Time-course analyses of abscisic acid level and the expression of genes involved in abscisic acid biosynthesis in the leaves of Betula platyphylla. *Molecular biology reports* 2012, 39(3):2505-2513.
25. Salojrvi J, Smolander OP, Nieminen K, Rajaraman S, Kangasjrvi J: Genome sequencing and population genomic analyses provide insights into the adaptive landscape of silver birch. *Nature Genetics* 2017, 49(6):904.
26. Lü H, Li J, Huang Y, Zhang M, Zhang S, Wu J: Genome-wide identification, expression and functional analysis of the phosphofructokinase gene family in Chinese white pear (Pyrus bretschneideri). *Gene* 2019.
27. Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, Pang N, Forslund K, Ceric G, Clements J: The Pfam protein families database. *Nucleic Acids Research* 2004, 28(1):263-266.
28. Ivica L, Tobias D, Peer B: SMART 7: recent updates to the protein domain annotation resource.
29. Ren LL, Liu YJ, Liu HJ, Qian TT, Qi LW, Wang XR, Zeng QY: Subcellular Relocalization and Positive Selection Play Key Roles in the Retention of Duplicate Genes of Populus Class III Peroxidase Family. *Plant Cell* 2014, 26(6):2404-2419.
30. Wang Y, Wang Q, Zhao Y, Han G, Zhu S: Systematic analysis of maize class III peroxidase gene family reveals a conserved subfamily involved in abiotic stress response. *Gene* 2015, 566(1):95-108.
31. Moore, Richard, C., Purugganan, Michael, D.: The early stages of duplicate gene evolution. *Proceedings of the National Academy of Sciences of the United States of America* 2003.
32. Cannon SB, Mitra A, Baumgarten A, Young ND, May G: The roles of segmental and tandem gene duplication in the evolution of large gene families in Arabidopsis thaliana. *Bmc Plant Biology* 2004, 4:1-21.
33. Riechmann J, Heard J, Martin G, Reuber L, Jiang C, Keddie J, Adam L, Pineda O, Ratcliffe O, Samaha R: Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. *Science* 2000, 290(5499):P.2105-2110.
34. Elsheery NI: Genome-wide characterization of aspartic protease (AP) gene family in Populus trichocarpa and identification of the potential PtAPs involved in wood formation. *BMC Plant Biology* 2019, 19(276):1-17.
35. Claudia C, Christophe D: Specific functions of individual class III peroxidase genes. *Journal of Experimental Botany* 2008(2):2.
36. Llorente F, López-Cobollo RM, Catalá R, Martínez-Zapater JM, Salinas J: A novel cold-inducible gene from Arabidopsis , RCI3 , encodes a peroxidase that constitutes a component for stress tolerance. *Plant Journal* 2002, 32(1):13-24.

37. Hu X, Liu C, Tian J, Zhang Y, Xin Q, Chen A, Li D, Liu X: Identification, molecular characterization, and expression analysis of the SQUAMOSA PROMOTER BINDING PROTEIN-LIKE (SPL) gene family in Betula platyphylla Suk. *Trees* 2020, 34(1):229-241.
38. Oakley T, Ostman B, Wilson A: Repression and loss of gene expression outpaces activation and gain in recently duplicated fly genes. *Proceedings of the National Academy of Sciences of the United States of America* 2006, 103(31):11637-11641.
39. Han Y, Ding T, Su B, Jiang H: Genome-Wide Identification, Characterization and Expression Analysis of the Chalcone Synthase Family in Maize. *International Journal of Molecular Sciences* 2016, 17(2):161-.
40. The calcium sensor CBL1 integrates plant responses to abiotic stresses. *Plant Journal* 2003, 36(4):457-470.
41. Kasuga M, Liu Q, Miura S, Yamaguchi-Shinozaki K, Shinozaki K: Improving plant drought, salt, and freezing tolerance by gene transfer of a single stress-inducible transcription factor. *Nature Biotechnology* 1999, 17(3):287-291.
42. Finn RD, Mistry J, Schuster-Bckler B, Griffiths-Jones S, Bateman A: PFAM: clans, web tools and services. *Nucleic Acids Research* 2006, 34(Database issue):D247-251.
43. Letunic I, Doerks TT, Bork P: SMART 6: Recent updates and new developments. *Nuclc Acids Research* 2008, 37(Database issue):D229-232.
44. Thompson JD, Higgins DG, Gibson TJ: CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic acids research* 1994, 22(22):4673-4680.
45. Publications M: MUSCLE: multiple sequence alignment with high accuracy and high throughput – Edgar 32 (5): 1792 – Nucleic Acids Research. *My Publications* 2010.
46. Alexandros, Stamatakis: RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 2014.
47. Bailey TL, Elkan C: The value of prior knowledge in discovering motifs with MEME. *Proceedings / International Conference on Intelligent Systems for Molecular Biology; ISMB International Conference on Intelligent Systems for Molecular Biology* 1995, 3:21-29.
48. Langmead B, Salzberg SL: Langmead B, Salzberg SL.. Fast gapped-read alignment with Bowtie 2. Nat Methods 9: 357-359. *Nature Methods* 2012, 9(4):357-359.
49. Li B, Dewey CN: RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *Bmc Bioinformatics* 2011, 12(1):1-16.

Table

Table 1. The 90 POD genes identified in B. pendula and their sequence characteristics

Protein Name	Gene ID	Theoretical pI	Molecular weight（Da）
BpPOD1	Bpev01.c0000.g0142	7.16	38520.2
BpPOD2	Bpev01.c0001.g0018	5.76	34481.79
BpPOD3	Bpev01.c0015.g0107	8.52	35715.62
BpPOD4	Bpev01.c0015.g0108	9.06	35517.27
BpPOD5	Bpev01.c0022.g0082	8.05	34206.63
BpPOD6	Bpev01.c0022.g0083	9.11	34146.78
BpPOD7	Bpev01.c0023.g0043	9.32	36436.39
BpPOD8	Bpev01.c0027.g0161	6.43	35985.96
BpPOD9	Bpev01.c0038.g0066	8.86	16650.95
BpPOD10	Bpev01.c0055.g0011	5.7	36849.2
BpPOD11	Bpev01.c0090.g0013	9.15	40130.67
BpPOD12	Bpev01.c0090.g0014	8.72	35448.63
BpPOD13	Bpev01.c0090.g0016	8.9	35155.75
BpPOD14	Bpev01.c0090.g0017	9.03	34839.38
BpPOD15	Bpev01.c0090.g0018	9.21	34931.7
BpPOD16	Bpev01.c0094.g0039	7.57	35824.75
BpPOD17	Bpev01.c0115.g0033	8.28	34790.11
BpPOD18	Bpev01.c0115.g0034	9.21	34709.95
BpPOD19	Bpev01.c0115.g0036	9.57	34410.61
BpPOD20	Bpev01.c0115.g0100	8.13	28980.85
BpPOD21	Bpev01.c0127.g0079	8.51	37428.88
BpPOD22	Bpev01.c0154.g0008	6.98	34749.39
BpPOD23	Bpev01.c0154.g0009	5.97	34913.58
BpPOD24	Bpev01.c0154.g0011	6.17	38375.7
BpPOD25	Bpev01.c0154.g0012	5.71	34090.42
BpPOD26	Bpev01.c0154.g0013	8.56	33988.06
BpPOD27	Bpev01.c0154.g0014	4.92	30751.05
BpPOD28	Bpev01.c0154.g0015	5.79	33695.64
BpPOD29	Bpev01.c0154.g0016	9.09	37699.91
BpPOD30	Bpev01.c0161.g0034	6.95	37831.82

BpPOD31	Bpev01.c0210.g0047	8.01	35734.78
BpPOD32	Bpev01.c0214.g0014	4.7	44989.41
BpPOD33	Bpev01.c0222.g0007	6.09	36320.35
BpPOD34	Bpev01.c0228.g0001	6.29	25755.32
BpPOD35	Bpev01.c0253.g0021	6.31	33855.45
BpPOD36	Bpev01.c0253.g0022	4.75	35040.89
BpPOD37	Bpev01.c0253.g0025	4.28	36363.54
BpPOD38	Bpev01.c0253.g0026	4.8	36734.26
BpPOD39	Bpev01.c0292.g0023	6.75	35088.84
BpPOD40	Bpev01.c0335.g0033	5.16	37438.99
BpPOD41	Bpev01.c0395.g0053	4.8	34822.54
BpPOD42	Bpev01.c0414.g0013	9.23	35888.05
BpPOD43	Bpev01.c0441.g0005	7.52	35297.73
BpPOD44	Bpev01.c0443.g0013	6.51	37358.81
BpPOD45	Bpev01.c0483.g0021	5.6	36858.73
BpPOD46	Bpev01.c0518.g0009	6.34	35401.21
BpPOD47	Bpev01.c0518.g0010	6.22	35012.98
BpPOD48	Bpev01.c0566.g0037	4.74	35256.87
BpPOD49	Bpev01.c0577.g0019	8.86	33926.77
BpPOD50	Bpev01.c0605.g0023	5.58	37438.76
BpPOD51	Bpev01.c0605.g0024	5.92	40256.6
BpPOD52	Bpev01.c0672.g0007	5.31	35421.93
BpPOD53	Bpev01.c0702.g0001	8.28	41401.42
BpPOD54	Bpev01.c0753.g0001	5.97	23067.41
BpPOD55	Bpev01.c0811.g0007	8.7	9122.73
BpPOD56	Bpev01.c0834.g0015	7.95	37636.09
BpPOD57	Bpev01.c0848.g0029	8.46	36912.4
BpPOD58	Bpev01.c0932.g0013	4.69	34485.93
BpPOD59	Bpev01.c0944.g0009	9.6	35965.28
BpPOD60	Bpev01.c0990.g0011	8.86	34411.46
BpPOD61	Bpev01.c0991.g0009	9.37	16644.16
BpPOD62	Bpev01.c1029.g0016	4.71	38697.61
BpPOD63	Bpev01.c1029.g0017	5.2	38867.05

BpPOD64	Bpev01.c1078.g0006	5.67	17097.87
BpPOD65	Bpev01.c1163.g0010	8.1	36508.06
BpPOD66	Bpev01.c1189.g0010	6.93	35457.58
BpPOD67	Bpev01.c1189.g0011	5.94	28953.96
BpPOD68	Bpev01.c1230.g0004	6.41	57999.02
BpPOD69	Bpev01.c1230.g0005	8.95	37658.16
BpPOD70	Bpev01.c1519.g0002	6.99	35815.14
BpPOD71	Bpev01.c1529.g0006	8.89	38531.35
BpPOD72	Bpev01.c1719.g0005	8.42	33743.46
BpPOD73	Bpev01.c1776.g0001	8.38	33425.87
BpPOD74	Bpev01.c1776.g0002	6.44	28814.32
BpPOD75	Bpev01.c1889.g0001	8.46	32601.89
BpPOD76	Bpev01.c1889.g0002	8.75	43372.13
BpPOD77	Bpev01.c1889.g0003	8.05	33592.04
BpPOD78	Bpev01.c1922.g0001	8.42	34940.65
BpPOD79	Bpev01.c1922.g0002	9.41	32305.51
BpPOD80	Bpev01.c2035.g0001	5.3	20474.83
BpPOD81	Bpev01.c2059.g0007	7.56	34908.57
BpPOD82	Bpev01.c2165.g0002	6.38	34883.46
BpPOD83	Bpev01.c2185.g0001	5.01	14822
BpPOD84	Bpev01.c2220.g0001	9.04	29887.06
BpPOD85	Bpev01.c2322.g0001	9.35	35748.32
BpPOD86	Bpev01.c3133.g0001	6.89	9127.53
BpPOD87	Bpev01.c3133.g0002	7.89	61365.35
BpPOD88	Bpev01.c3139.g0001	8.54	34756.42
BpPOD89	Bpev01.c3210.g0001	8.53	33682.38
BpPOD90	Bpev01.c3916.g0001	4.74	38628.55

Figures

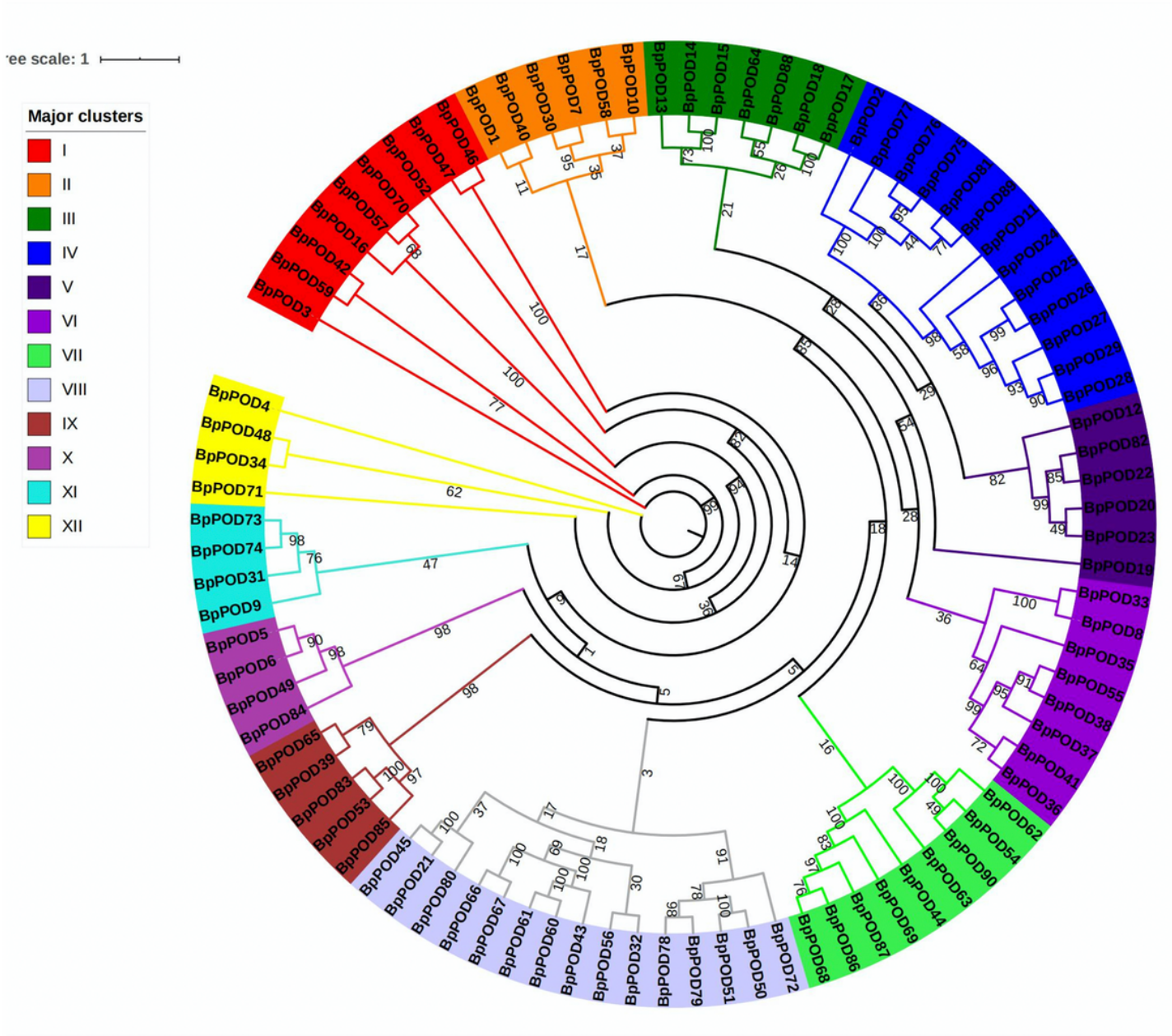

Figure 1

Phylogenetic relationship of the 90 identified BpPOD genes. Unrooted tree constructed using RAxML by the Neighbor-Joining (NJ) method. The tree shows 12 major phylogenetic subgroups (subgroups I to XII).

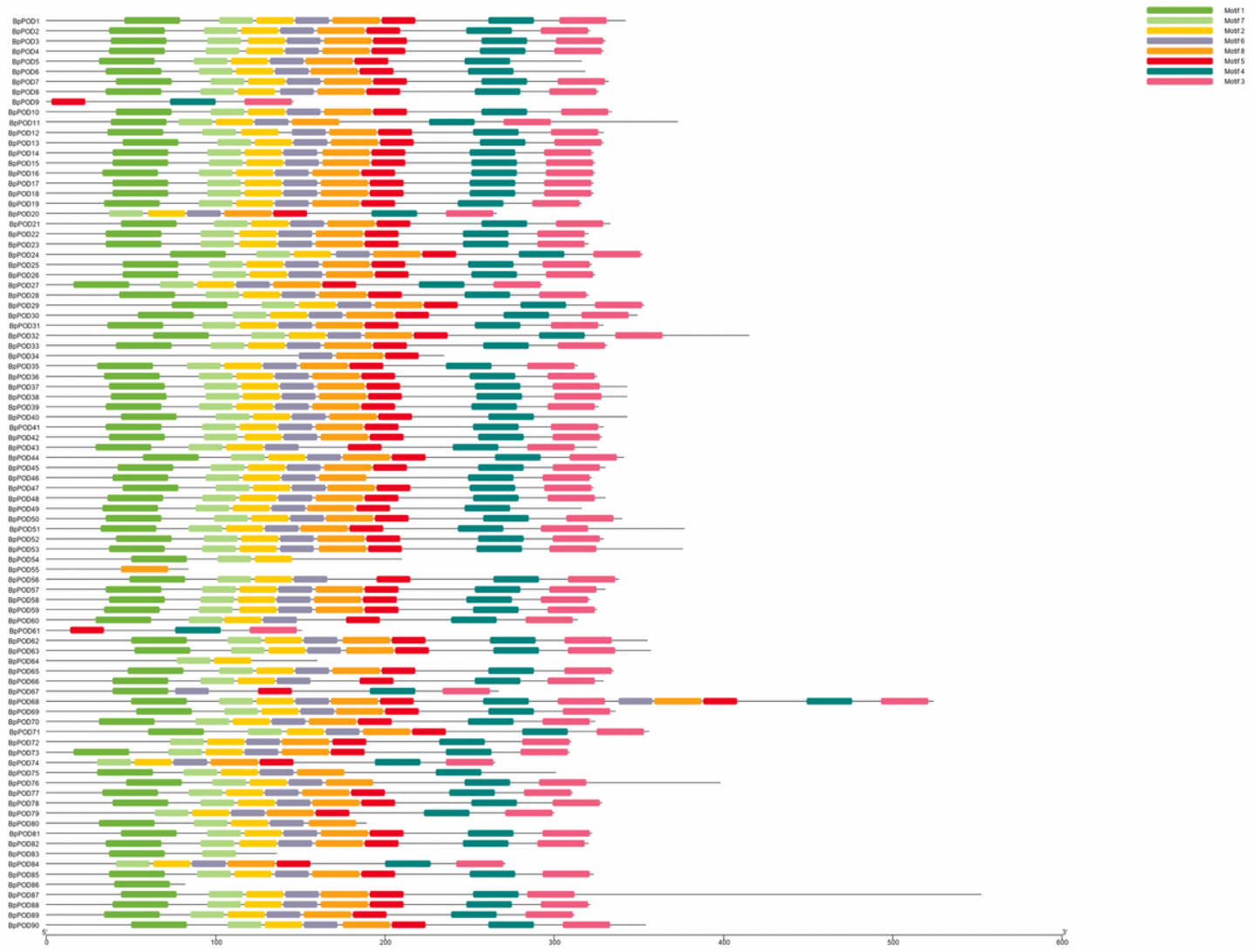

Figure 2

Distribution of eight putative conserved motifs in BpPOD proteins. Conserved motifs are represented by different colored boxes while nonconserved sequences are shown by gray lines. Note that the length of each box does not represent the actual motif size, and the colored boxes were ordered manually according to the results of MEME analysis.

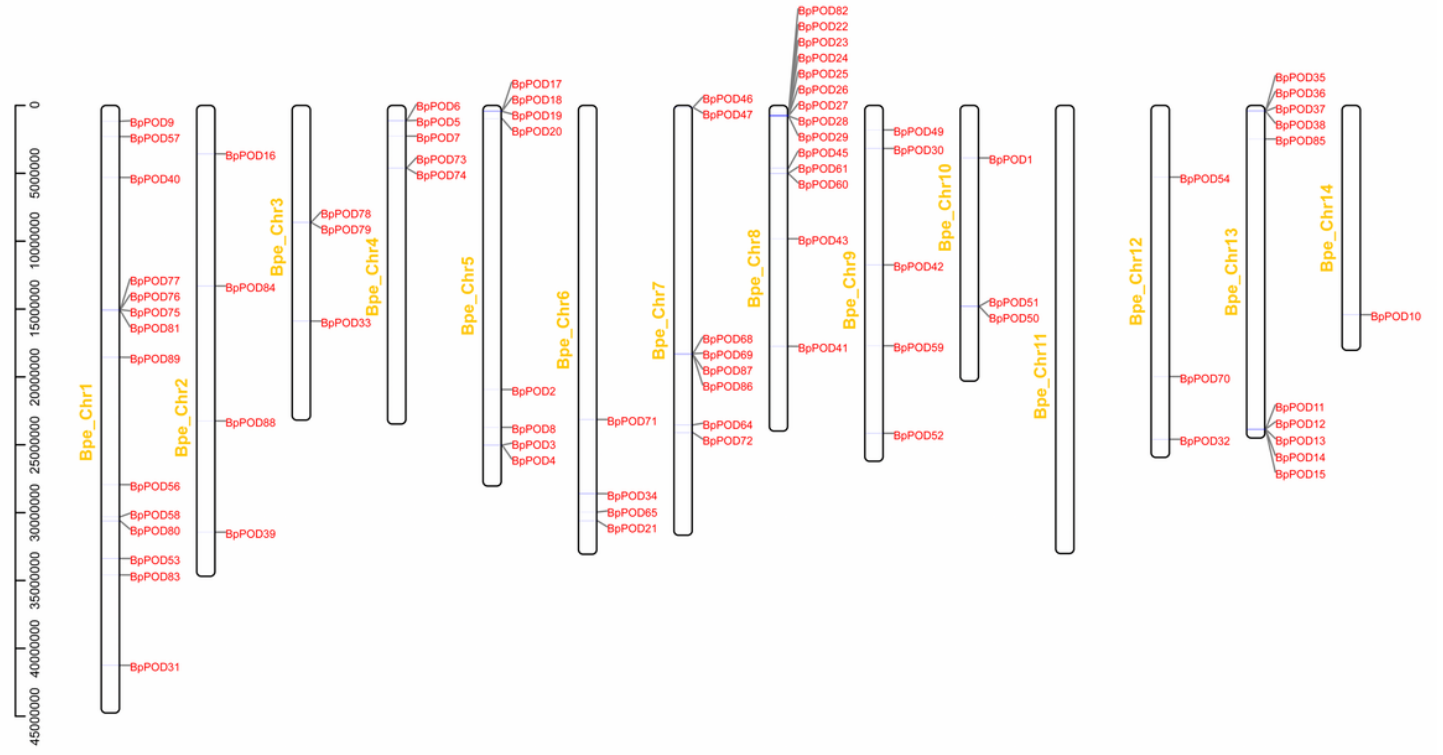

Figure 3

Chromosomal locations of 90 POD genes on 14 B. pendula chromosomes. Each was mapped to the chromosome based on its physical location. The number of chromosomes (chr01-chr14) is marked in yellow. The gene names on the right side of each chromosome correspond to the approximate locations of each POD gene.

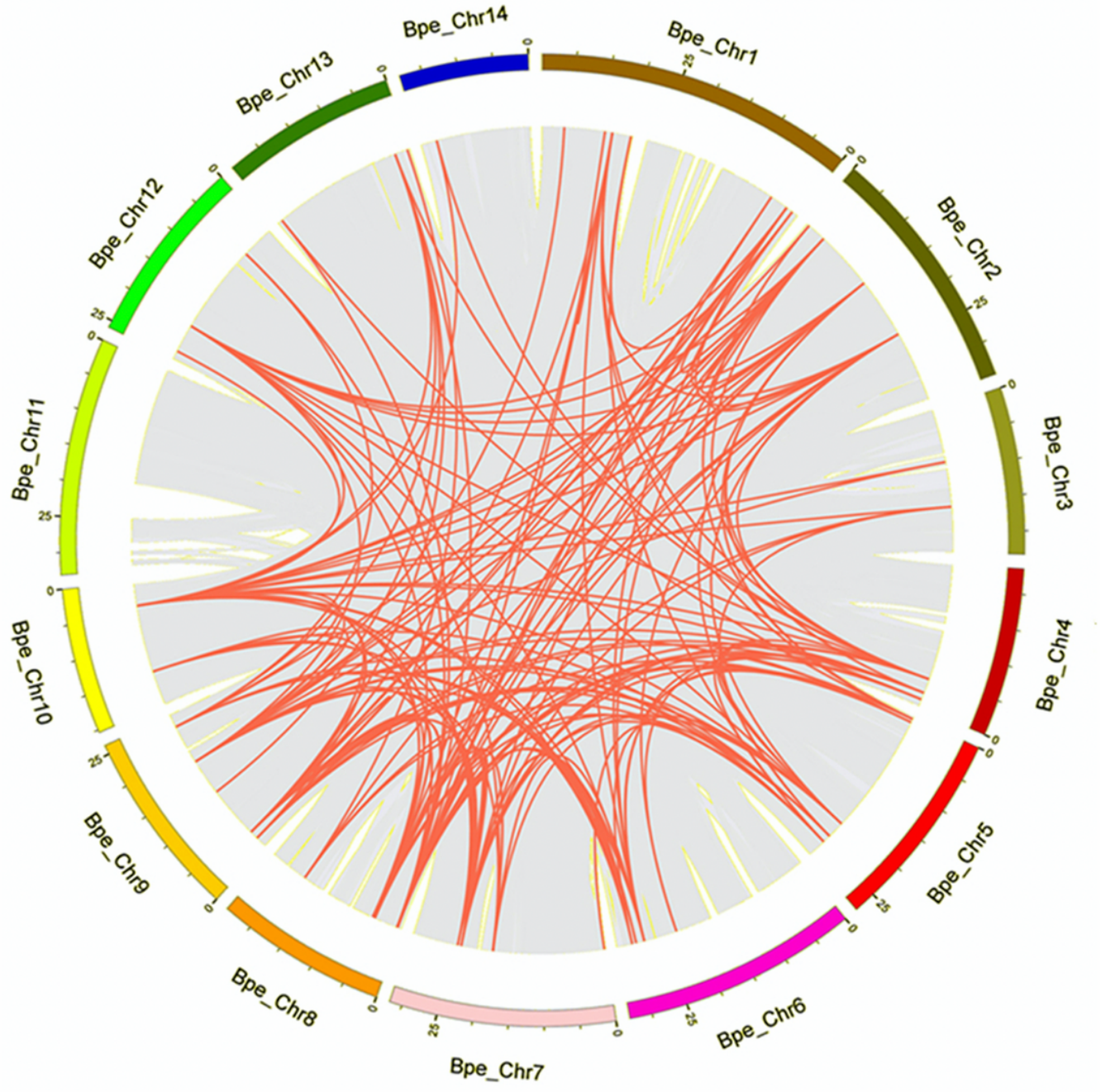

Figure 4

Collinearity analysis of POD genes in B. pendula genome.

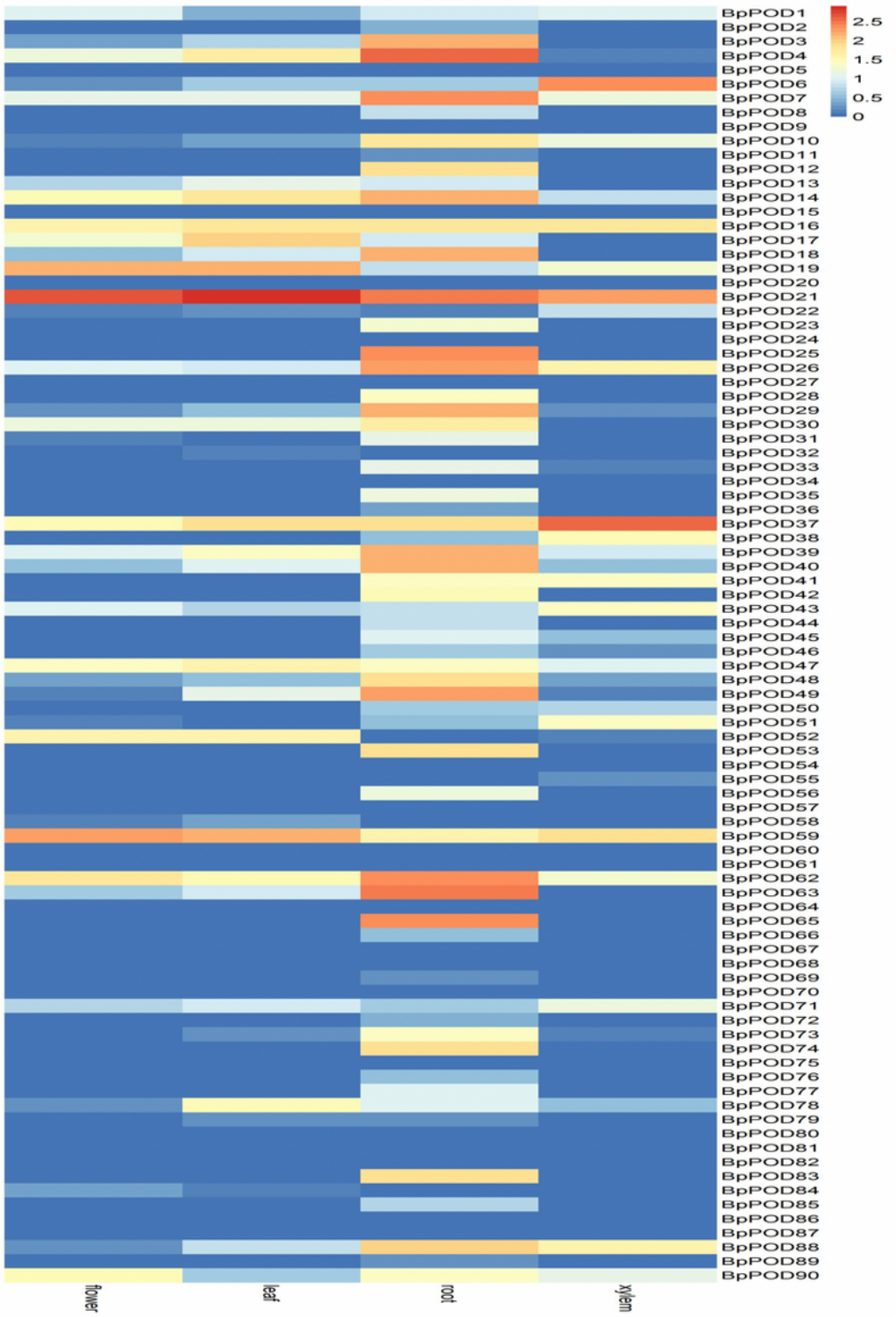

Figure 5

Expression profiles of BpPOD genes across different tissues. Different organs/tissues are exhibited below each column. The BpPOD genes were listed at the right of the expression array, and the expression values mapped to a color gradient from low (blue) to high expression (red) are shown at the right of the figure.

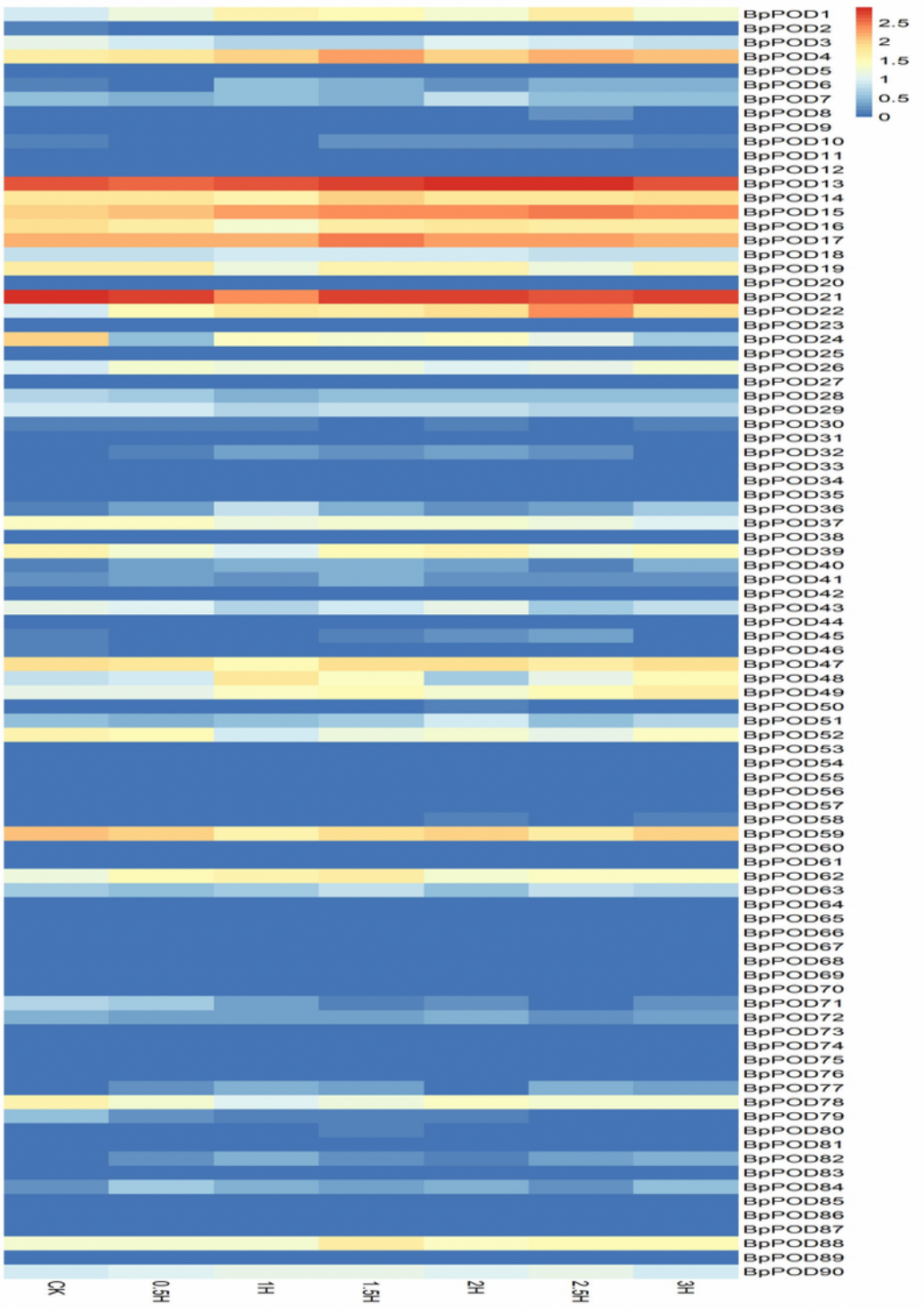

Figure 6

Responses of BpPOD genes expression to cold treatment. Different time are exhibited below each column. The BpPOD genes were listed at the right of the expression array, and the expression values mapped to a color gradient from low (blue) to high expression (red) are shown at the right of the figure.