Article

Identification and Characterization of the Basic Helix-Loop-Helix Transcription Factor Family in *Pinus massoniana*

Yu Chen, Peihuang Zhu, Fan Wu, Xiaofeng Wang, Jinfeng Zhang and Kongshu Ji *

Key Laboratory of Forestry Genetics & Biotechnology of Ministry of Education, Co-Innovation Center for Sustainable Forestry in Southern China, Nanjing Forestry University, Nanjing 210037, China; luxiansenaichiyu@163.com (Y.C.); zphzhupeihuang@163.com (P.Z.); frankiew2020@163.com (F.W.); wxf860724@163.com (X.W.); 05zhangjinfeng@163.com (J.Z.)

***** Correspondence: ksji@njfu.edu.cn; Tel.: +86-025-8542-7308

Abstract: The basic helix-loop-helix (bHLH) protein transcription factor family is the most widely distributed transcription factor family in eukaryotes. Members of this family play important roles in secondary metabolic biosynthesis, signal transduction, and plant resistance. Research on the bHLH family in animals is more extensive than that in plants, and members of the family in plants are classified according to the classification criteria for those in animals. To date, no research on the *bHLH* gene family in *Pinus massoniana* (Masson pine) has been reported. In this study, we identified 88 *bHLH* genes from four transcriptomes of Masson pine and performed bioinformatics analysis. These genes were divided into 10 groups in total. RT-PCR analysis revealed that the expression levels of the six genes increased under abiotic stress and hormone treatments. These findings will facilitate further studies on the functions of bHLH transcription factors.

Keywords: *Pinus massoniana*; bHLH; abiotic stress; hormone treatment; gene expression

1. Introduction

During the growth and development of plants in nature, to adapt to and resist various biotic and abiotic stresses, a set of adaptive mechanisms involving various physiological and biochemical reactions are employed [1–3]. These mechanisms are usually involved in the transcriptional regulation of gene-specific expression, which plays a very important role in plant responses to environmental stimuli [4]. Transcription factors (TFs) are the most important class of regulatory factors in plants and can maintain normal life activities in these organisms [5]. In recent years, a large number of TFs related to biotic and abiotic stresses have been isolated from higher plants. These TFs are widely involved in regulating the expression of genes related to drought, high salt, high temperature, low temperature, hormone and pathogen responses. In plants, the basic helix-loop-helix (bHLH) family is the second largest TF family after the MYB family [4]. Members of this family not only are generally involved in aspects of plant growth and metabolism, including light morphogenesis, light signal transduction, and secondary metabolism, but also play an important role in plant responses to adversity [6].

bHLH family members contain two highly conserved and different functional domains, namely, the basic region and the helix-loop-helix (HLH) region, with a total of approximately 50–60 amino acids [7]. The basic region consists of approximately 15 amino acids, is located at the N-terminus of the domain, and can recognize and bind the E-box and G-box. The HLH region is composed of 40–50 amino acids and is located at the C-terminus of the domain. It can form dimers of two HLH proteins and regulate the expression of downstream target genes [8–10].

The bHLH family is the most widely distributed TF family in eukaryotes, existing in all animals, plants and fungi [4]. It was first discovered in mice that the TFs E12 and E47 contain a conserved hydrophobic HLH region [11]. Since then, the bHLH family has been identified and analyzed in many eukaryotes, but compared with those on animal bHLH TFs, studies on plant bHLH TFs are scarce. To date, the bHLH family has been identified in various plants, such as *Arabidopsis thaliana* [12], *Oryza sativa* [13], *Zea mays* [14], *Panax ginseng* [15], *Arachis hypogaea* [16], *Vitis vinifera* [17] and *Salvia miltiorrhiza* [18]. In a study on bHLH TF functioning in plant resistance, LIU et al. [19] found that transgenic *A. thaliana* overexpressing bHLH122 showed stronger resistance to drought, NaCl and osmotic stresses than wild-type (WT) *A. thaliana*. ICE1 encodes a bHLH TF similar to MYC. At low temperatures, *AtICE1* binds to the MYC recognition site in the promoter sequence of the *CBF3* gene and induces the expression of *CBF3*. The function of this TF is to increase the resistance of *A. thaliana* to low-temperature stress [20]. In a functional study of bHLH involvement in signal transmission and secondary metabolite biosynthesis, the bHLH04, bHLH05 and bHLH06 proteins were shown to be key regulators of indolic glucosinolate (GSL) biosynthesis in *A. thaliana* [21]. In *Panax notoginseng*, *PnbHLH1* improves triterpenoid biosynthesis by interacting with E-box core sequences in the promoter region of target genes [22].

Pinus massoniana (Masson pine) is an important oleoresin-producing species in China and a pioneer species for restoring forests on barren hills. It has important economic and ecological value. More than 70% of the annual oleoresin production comes from Masson pine, whose oleoresin yields turpentine and rosin, both of which are important raw materials for the chemical industry. Oleoresin is a resin secreted by conifers that includes terpenoids, such as monoterpenoids, sesquiterpenoids and diterpenoids, as the main component and has a specific smell. It is a secondary product of pine physiological metabolism [23,24]. The components of oleoresin play an important role in the defense system of conifers. After being exposed to biological or nonbiological stimuli, conifers release oleoresin from resin ducts and at the same time induce the synthesis of oleoresin [25,26]. Therefore, the biosynthetic regulators of secondary metabolism play a very important role in plant adaptation to environmental stress. At present, because genomic data for conifers are lacking, reports on the *bHLH* gene family are very rare. In this study, we screened 88 bHLH TFs in Masson pine and analyzed them with a bioinformatics approach based on transcriptome data. In addition, we analyzed the expression of some representative genes under different treatments. The results provide a new reference for functional studies of the *bHLH* gene family.

2. Materials and Methods

2.1. Identification of PmbHLH Sequences

bHLH sequences of *A. thaliana* were downloaded from the Plant Transcription Factor database (http://planttfdb.cbi.pku.edu.cn/). Transcriptome data for Masson pine were derived from the previously determined CO_2 stress transcriptome [27], drought stress transcriptome (PRJNA595650), pine wood nematode inoculation transcriptome (PRJNA660087) and young shoots transcriptome (PRJNA655997). A Markov model (HMM) was used to screen out the bHLH gene files from the four transcriptomes by the hmmsearch program (HMMER 3.0). Pfam (http://pfam.xfam.org/) and CD-search (https://www.ncbi.nlm.nih.gov/cdd/) were used to screen out protein sequences of Masson pine with the bHLH domain (PF00010). Finally, sequences with complete bHLH domains were selected and sequences with more than 97% similarity between different databases were deleted. The online ExPASy ProtParam tool (https://web.expasy.org/protparam/) was used to analyze the physical and chemical properties of the bHLH protein sequences of Masson pine.

2.2. Phylogenetic and Protein Domain Analyses of PmbHLH Proteins

The bHLH protein sequences of Masson pine and Arabidopsis were subjected to multiple alignment using the ClustalW algorithm in MEGA7 [28] and used to build a phylogenetic tree based on the Maximum likelihood (ML) method, a bootstrap test performed with 1000 replicates and Jones–Taylor–Thornton (JTT) model in IQ-TREE. Then, the phylogenetic tree was edited for visualization purposes with the online software EvolView (https://www.evolgenius.info/evolview/#login). In order to better determine the distribution of conserved motifs and grouping of bHLH proteins, the online program MEME (http://meme-suite.org/tools/meme) was used to analyze the conserved motifs of bHLH protein sequences of *A. thaliana* and Masson pine, with the number of motifs set to 20. Finally, TBtools [29] was used to draw the phylogenetic tree and conserved domain combination diagram.

2.3. Subcellular Localization Analysis

The subcellular localization of PmbHLH proteins was predicted and analyzed with CELLO (http://cello.life.nctu.edu.tw/) and PSORT (https://psort.hgc.jp/). To verify the predicted results, we designed primers for the open reading frame (ORF) of *PmbHLH44* (MT989428) in group III and obtained coding sequences from cDNA, which was reverse transcribed from RNA extracted from 15 years-old Masson pine twigs. The primer sequences can be found in Supplementary File S1. The obtained sequences were cloned into pBI121 green fluorescent protein (GFP) C-terminal fusion expression vector, transferred into *Agrobacterium tumefaciens* GV3101 and injected into *Nicotiana benthamiana* leaves for transient expression. After dark cultivation for 24 h, images were taken with a LSM710 confocal laser-scanning microscope (Zeiss, Jena, Germany).

2.4. Transcriptional Profile Analysis

The expression levels of *bHLH* genes of Masson pine were quantified by transcriptome sequencing data (RNA-seq) collected after pine wood nematode inoculation. The pine wood nematode inoculation transcriptome includes five sets of data: 0, 3, 10, 20, and 35 days after treatment. Through ID conversion and data extraction, the gene expression level was calculated by determining the expected fragments per transcript per million mapped reads (FPKM). TBtools software was used to create *bHLH* gene expression heat maps of Masson pine.

2.5. Abiotic Stress and Hormone Treatment of Plant Materials

Two-year-old Masson pine seedlings obtained from the seed orchard of the Baisha state-owned forest farm, Shanghang, Fujian Province, China, were used in this study. To study the expression level of *PmbHLH*, we subjected the Masson pine seedlings to six treatments. The six treatments included abiotic stress and hormone treatments, namely, 15% PEG 6000-osmotic stress, mechanical injury, 10 mM H_2O_2, 100 μM methyl jasmonate (MeJA), 500 μM ethephon (ETH) and 1 mM salicylic acid (SA). The osmotic stress treatment was an immersion treatment, the mechanical damage treatment involved cutting the upper half of the needle, and the rest of the treatments were spray treatments. Three uniformly growing seedlings were selected for each treatment as three biological replicates. The needles were sampled at 0, 3, 6, 12, 24 and 48 h, immediately frozen in liquid nitrogen and stored at −80 °C. Samples collected at 0 h without any treatment were used as controls.

2.6. RNA Extraction and qRT-PCR Analysis

The tissue samples were ground into a fine powder in liquid nitrogen. Total RNA was isolated from all samples using an RNAprep Pure Plant Plus Kit (DP441, Tiangen Biotech, Beijing, China) following the manufacturer's instructions and stored at −80 °C. The 1st strand cDNA synthesis kit (11141, Yeasen Biotech, Shanghai, China) was used to reverse transcribe RNA into cDNA, and the cDNA was stored at −80 °C. Primer 5.0 was used to design primers for qRT-PCR, and the amplicon size was 120–200 bp. The sequences of the internal control gene alpha-tubulin (*TUA*) (KM496535.1)

and gene-specific primers for qPCR can be found in Supplementary File S1. PCR amplifications were performed in a StepOne Plus real-time thermal cycler (Applied Biosystems, Foster City, CA, USA) in a final volume of 20 μL containing 10 μL (2×) of SYBR Green qRT-PCR Master Mix (QPK-201, Toyobo Bio-Technology, Shanghai, China), 2.0 μL of cDNA, 7.2 μL of sterile water, and 0.4 μL of each primer (200 nM). The data represent three biological replicates and three technical replicates. The conditions for amplification were 10 min of denaturation at 95 °C, followed by 40 cycles of 95 °C for 15 s, 58 °C for 30 s, and 72 °C for 30 s, after which a melting curve was produced at 60–95 °C. Finally, the relative expression levels of genes were calculated using the $2^{-\triangle\triangle Ct}$ method, and t-test was used to analyze the significance of the difference between the values of each time period and 0 h.

3. Results

3.1. Identification of PmbHLH Protein Sequences

The CO_2 stress transcriptome was a transcriptome sampled and sequenced from seedling needles after treatment with a high concentration of CO_2 for 6, 12 and 24 h. The drought stress transcriptome used a weighing method to control the four gradients of normal, mild, moderate and severe soil water content. The needles were collected 60 days later for sequencing. The pine wood nematode inoculation transcriptome was sequenced from needles on the 0th, 3rd, 10th, 20th and 35th days after the inoculation of pine wood nematode, and the young shoots transcriptome was sequenced with tender shoots without any processing. From the four transcriptomes, 198 bHLH protein sequences of Masson pine were screened out. After deleting the repeated sequences and sequences with incomplete domains, 88 PmbHLH protein sequences were retained (Supplementary File S2). The PmbHLH proteins encoded 118–990 amino acids, with the predicted molecular weight ranging from 13.63 to 107.24 kDa. The pI values ranged from 4.93 to 11.77. Detailed parameter information for physical and chemical property analysis of PmbHLH proteins is provided in Supplementary File S3.

3.2. Phylogenetic Analysis

The Maximum likelihood (ML) method was used to construct a phylogenetic tree of 88 PmbHLH protein sequences from Masson pine. As shown in Figure 1, the 88 bHLH protein sequences were divided into ten large groups, including I, III, IV, V, VII, VIII, IX, XI, XII and XIV. The group with the largest number of PmbHLH proteins was I, with a total of 24 members, followed by III and XII. The group with the smallest number of PmbHLH proteins was VIII, with only one member. The long branch lengths of PmbHLH41, PmbHLH43, PmbHLH39, PmbHLH60 and PmbHLH62 indicate that they are quite different from other proteins. The high branch support values on the node indicate that the reliability of the phylogenetic tree is high. The bHLH transcription factor family has many members and the evolutionary model is complicated, which may result in low branch support values of the phylogenetic tree root. In addition, a phylogenetic tree constructed by 88 Masson pine bHLH protein sequences and 134 Arabidopsis bHLH protein sequences is provided in Supplementary File S4.
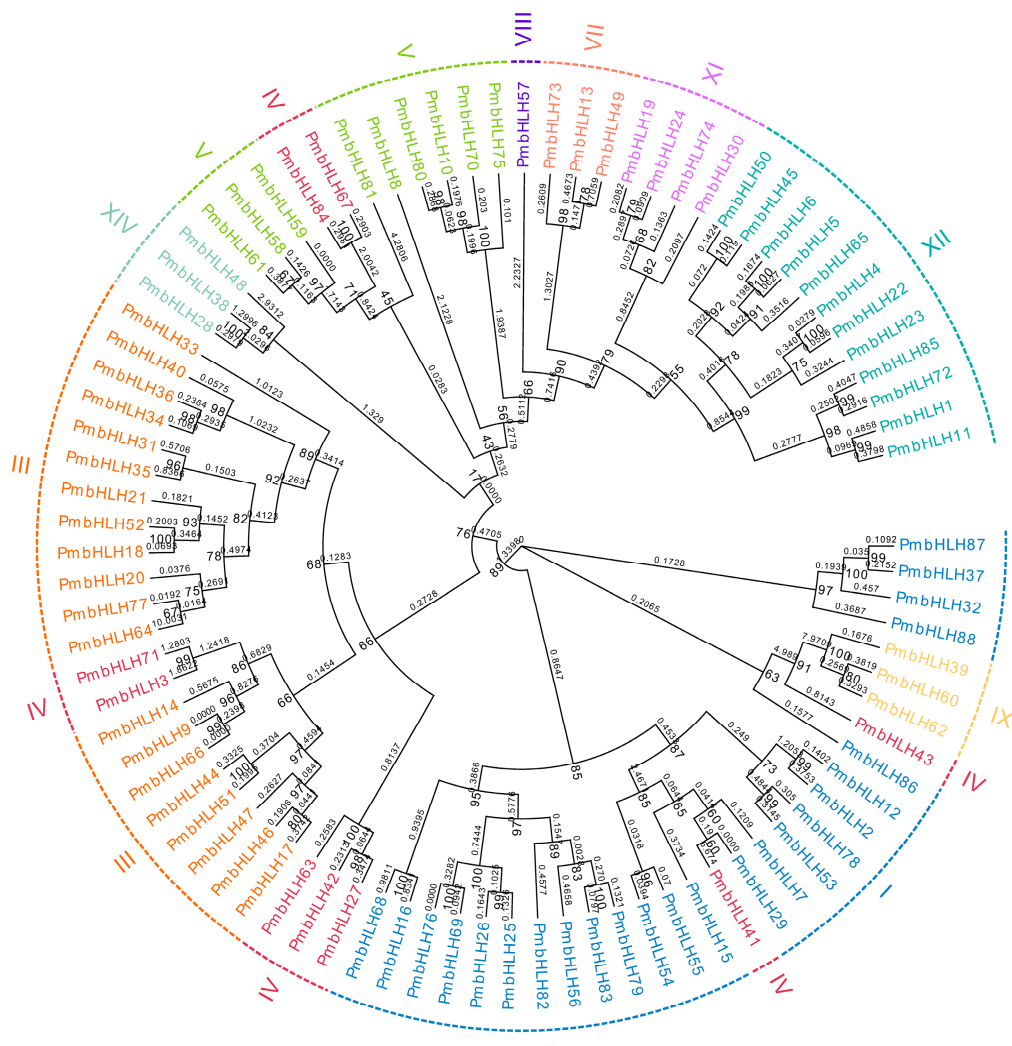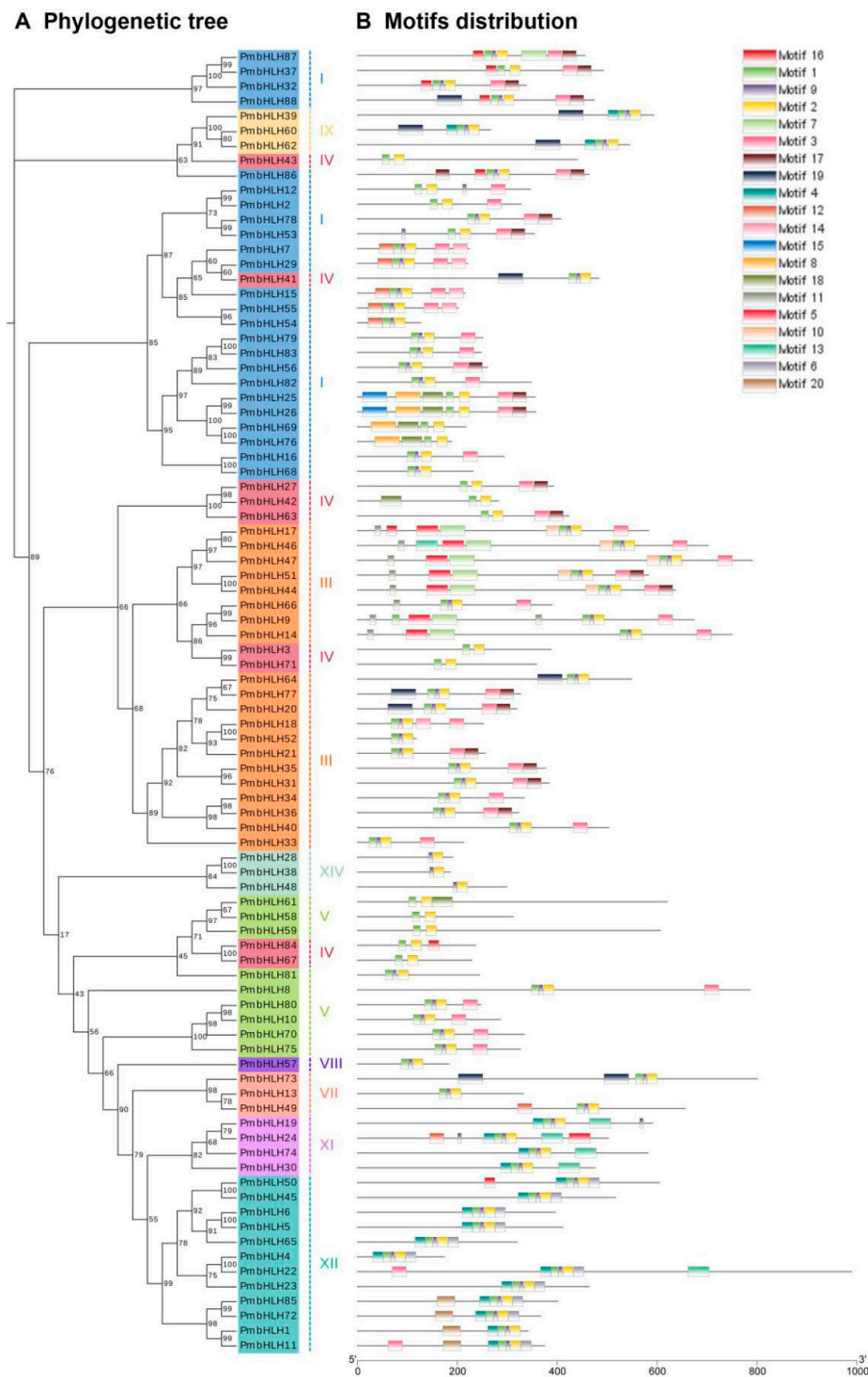
Figure 1. Phylogenetic analysis of bHLH proteins in *P. massoniana*. Gene names in different colors correspond to groups of the same color in the outermost circle, and the numbers on the node in the figure indicate branch support values. The decimals on the branches represent the different lengths of the branches, indicating the differences between the proteins.

3.3. Analysis of PmbHLH Protein Domains

According to MEME software identification of the conserved motifs of the 88 PmbHLH protein sequences, all of these proteins have highly conserved bHLH domains. The conserved motifs are shown in Supplementary File S5. The bHLH domain contains two different functional regions, the N-terminal basic amino acid region and the C-terminal HLH region, which are the parts composed of motifs 1, 2 and 9 in Figure 2. Motif 1 and 2 are the two main parts of the domain, which are longer than motif 9, and motif 9 is the very short part of the domain between motif 1 and 2. As shown in Figure 2, almost all 88 PmbHLH proteins contained motifs 1 and 2, except for the XIV group. Most of these proteins contain motif 1, 2 and 9 at the same time, indicating that their domains are more complete. Among them, each group generally contained partly the same motifs. For example, part of the proteins in group III contained not only motifs 1, 2, 9 but also motifs 5, 7, 10, and 14; part of proteins in group V and III both contained motif 14. Meanwhile, the proteins in group XI contained motifs 4, and 13, and proteins in group XIV contained only motif 2. Therefore, there were certain differences between these groups, suggesting different functions. Additionally, there were certain similarities between groups. For example, proteins in groups XI and XII had similar motifs. The proteins of group IX and part of group III also had similar conserved motifs. Therefore, the distribution of conserved motifs

may be related to protein function, and the differences between groups may be related to the diversity of gene functions during evolution. At the same time, we also used MEME to analyze the conservative motifs of 88 Masson pine bHLH proteins and 134 Arabidopsis bHLH proteins (Supplementary File S6). With reference to the distribution of conserved motifs and the grouping of Arabidopsis *bHLH* genes, we can better determine the group to which each *PmbHLH* gene belongs.

Figure 2. Phylogenetic analysis and motif distribution of bHLH proteins in *P. massoniana*. (**A**) The phylogenetic tree constructed for 88 PmbHLH proteins. (**B**) The conserved motifs of each protein as identified by MEME. The 20 different colors represent 20 different conserved motifs.
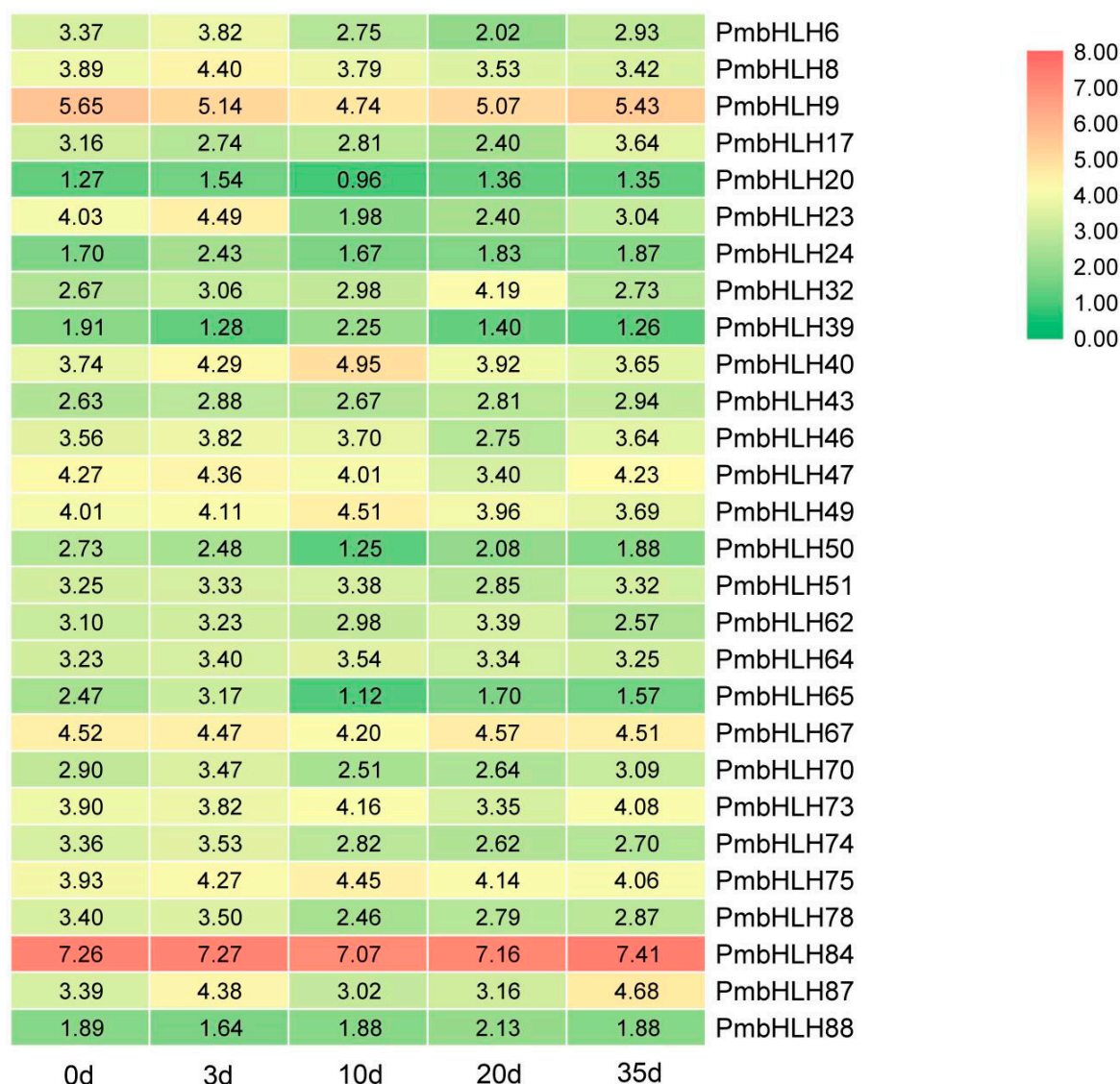
3.4. Subcellular Localization Analysis

According to the results of CELLO and PSORT prediction of the subcellular localization of PmbHLH1-PmbHLH88 proteins, the proteins were almost always localized in the nucleus. The prediction results for these proteins are shown in Supplementary File S7. To verify the prediction results, we selected the PmbHLH44 protein in group III related to the synthesis of secondary metabolites, fused GFP to the ORF of *PmbHLH44* and constructed a 121-PmbHLH44-GFP vector. After transient expression in *N. benthamiana*, fluorescence signals were observed in the nucleus by confocal laser-scanning microscopy (Figure 3), which was consistent with the predicted results. The 121-GFP no-load produced green fluorescence in the entire cell, which is obviously different from the expression pattern of PmbHLH44. PmbHLH44 only emits green fluorescence in the nucleus, proving that this protein is a nucleus-localized protein.
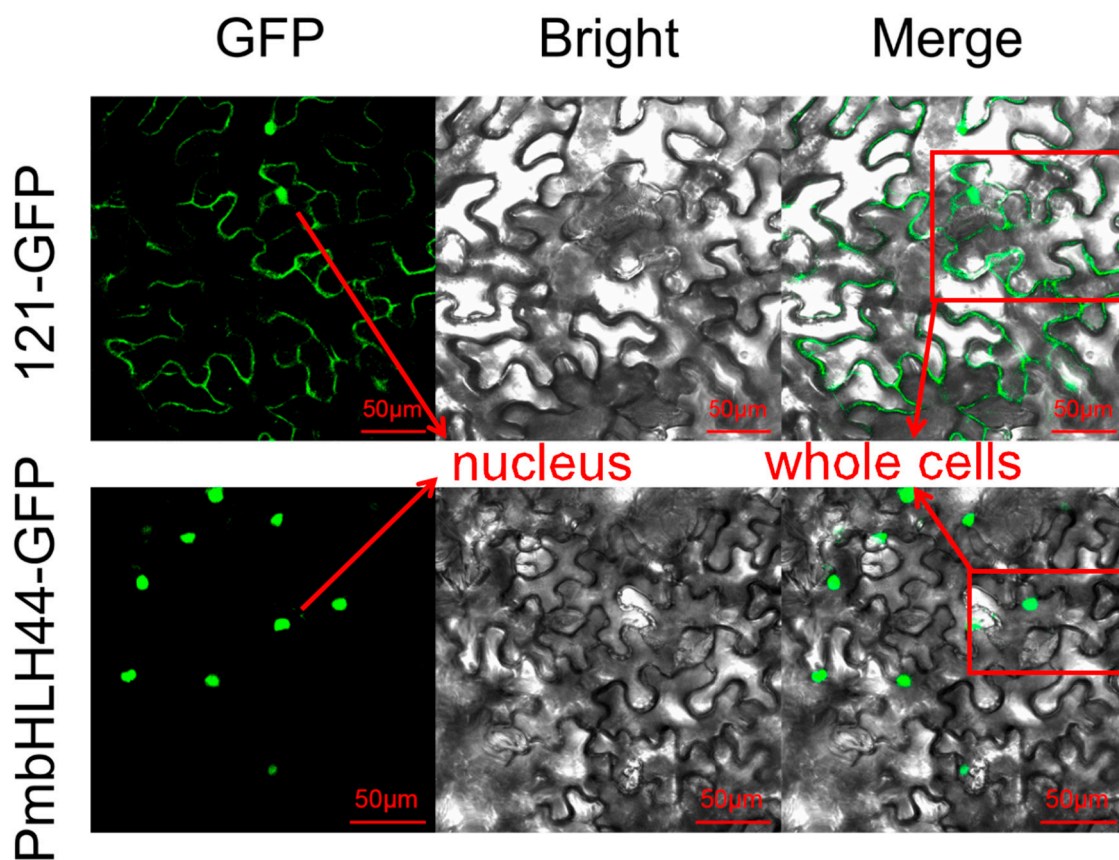
Figure 3. Subcellular localization assays of PmbHLH44 protein in tobacco leaf cells. Empty 121-GFP is localized in the whole cell, and 121-PmbHLH44-GEP is localized in the nucleus.

3.5. Analysis of the Transcriptional Profile of PmbHLH Genes

According to the pine wood nematode inoculation transcriptome data, heat maps were drawn (Figure 4). The heat map for the pine wood nematode inoculation transcriptome showed that the expression level of *PmbHLH84* at 0–35 days was very high when compared with other genes longitudinally, while the expression levels of *PmbHLH20*, *PmbHLH24*, *PmbHLH39*, *PmbHLH50*, *PmbHLH65* and *PmbHLH88* continued to be very low at 0–35 days, and the expression levels of the remaining genes were intermediate. In general, the horizontal expression levels of these genes tended to first increase, then decrease and then increase. The expression levels of *PmbHLH* genes varied, indicating that these genes each have specific functions in specific environments.

0d	3d	10d	20d	35d	
3.37	3.82	2.75	2.02	2.93	PmbHLH6
3.89	4.40	3.79	3.53	3.42	PmbHLH8
5.65	5.14	4.74	5.07	5.43	PmbHLH9
3.16	2.74	2.81	2.40	3.64	PmbHLH17
1.27	1.54	0.96	1.36	1.35	PmbHLH20
4.03	4.49	1.98	2.40	3.04	PmbHLH23
1.70	2.43	1.67	1.83	1.87	PmbHLH24
2.67	3.06	2.98	4.19	2.73	PmbHLH32
1.91	1.28	2.25	1.40	1.26	PmbHLH39
3.74	4.29	4.95	3.92	3.65	PmbHLH40
2.63	2.88	2.67	2.81	2.94	PmbHLH43
3.56	3.82	3.70	2.75	3.64	PmbHLH46
4.27	4.36	4.01	3.40	4.23	PmbHLH47
4.01	4.11	4.51	3.96	3.69	PmbHLH49
2.73	2.48	1.25	2.08	1.88	PmbHLH50
3.25	3.33	3.38	2.85	3.32	PmbHLH51
3.10	3.23	2.98	3.39	2.57	PmbHLH62
3.23	3.40	3.54	3.34	3.25	PmbHLH64
2.47	3.17	1.12	1.70	1.57	PmbHLH65
4.52	4.47	4.20	4.57	4.51	PmbHLH67
2.90	3.47	2.51	2.64	3.09	PmbHLH70
3.90	3.82	4.16	3.35	4.08	PmbHLH73
3.36	3.53	2.82	2.62	2.70	PmbHLH74
3.93	4.27	4.45	4.14	4.06	PmbHLH75
3.40	3.50	2.46	2.79	2.87	PmbHLH78
7.26	7.27	7.07	7.16	7.41	PmbHLH84
3.39	4.38	3.02	3.16	4.68	PmbHLH87
1.89	1.64	1.88	2.13	1.88	PmbHLH88

Scale:
8.00
7.00
6.00
5.00
4.00
3.00
2.00
1.00
0.00

Figure 4. Transcriptional profile of *bHLH* genes under pine wood nematode inoculation in *P. massoniana*.

3.6. Gene Expression Analysis under Different Treatments

Based on the classification of *A. thaliana*, we selected group III related to the synthesis of secondary metabolites and randomly selected six *PmbHLH* genes for RT-PCR expression analysis: *PmbHLH9* (MT989425), *PmbHL14* (MT989426), *PmbHLH17* (MT989427), *PmbHLH46* (MT989429), *PmbHLH51* (MT989430), and *PmbHLH66* (MT989431). The results in Figure 5 show that under MeJA treatment, the expression levels of *PmbHLH9*, *PmbHLH14*, *PmbHLH46* and *PmbHLH66* decreased slightly first and then increased at 24 h, reaching the highest level. The expression levels of *PmbHLH17* and *PmbHLH51* first increased to their highest levels at 3 h, decreased sharply at 6 h, and then increased steadily. Under ETH treatment, the expression levels of *PmbHLH9*, *PmbHLH51* and *PmbHLH66* rose steadily to their highest levels. The expression levels of *PmbHLH17* and *PmbHLH46* first increased and then decreased at 24 h, while the expression level of *PmbHLH14* increased to its maximum at 3 h and then decreased. Under H_2O_2 treatment, the expression levels of the six genes reached their maxima at 6 h and then decreased, but the expression levels of *PmbHLH9*, *PmbHLH17*, *PmbHLH51* and *PmbHLH66* increased again at 24 h. Under osmotic stress, the expression levels of *PmbHLH17* and *PmbHLH51* showed an overall increasing trend, and the expression levels of the remaining four genes increased first and then decreased, showing the same patterns as the six genes after SA treatment. Under mechanical

injury stress, only the expression level of *PmbHLH17* increased first and then decreased, and the levels of the remaining genes decreased first, increased and then decreased. The expression levels were highest at 12 h.
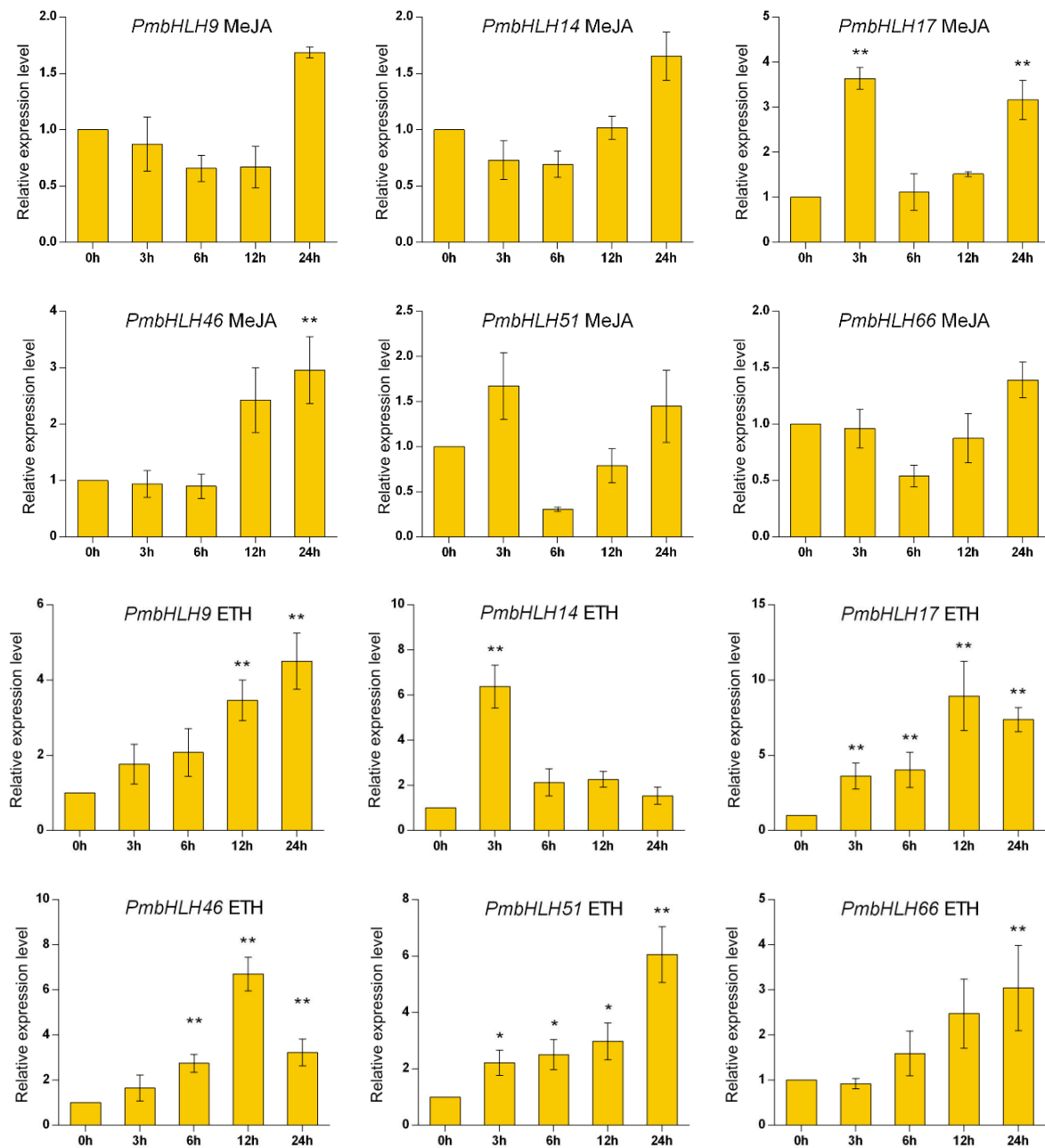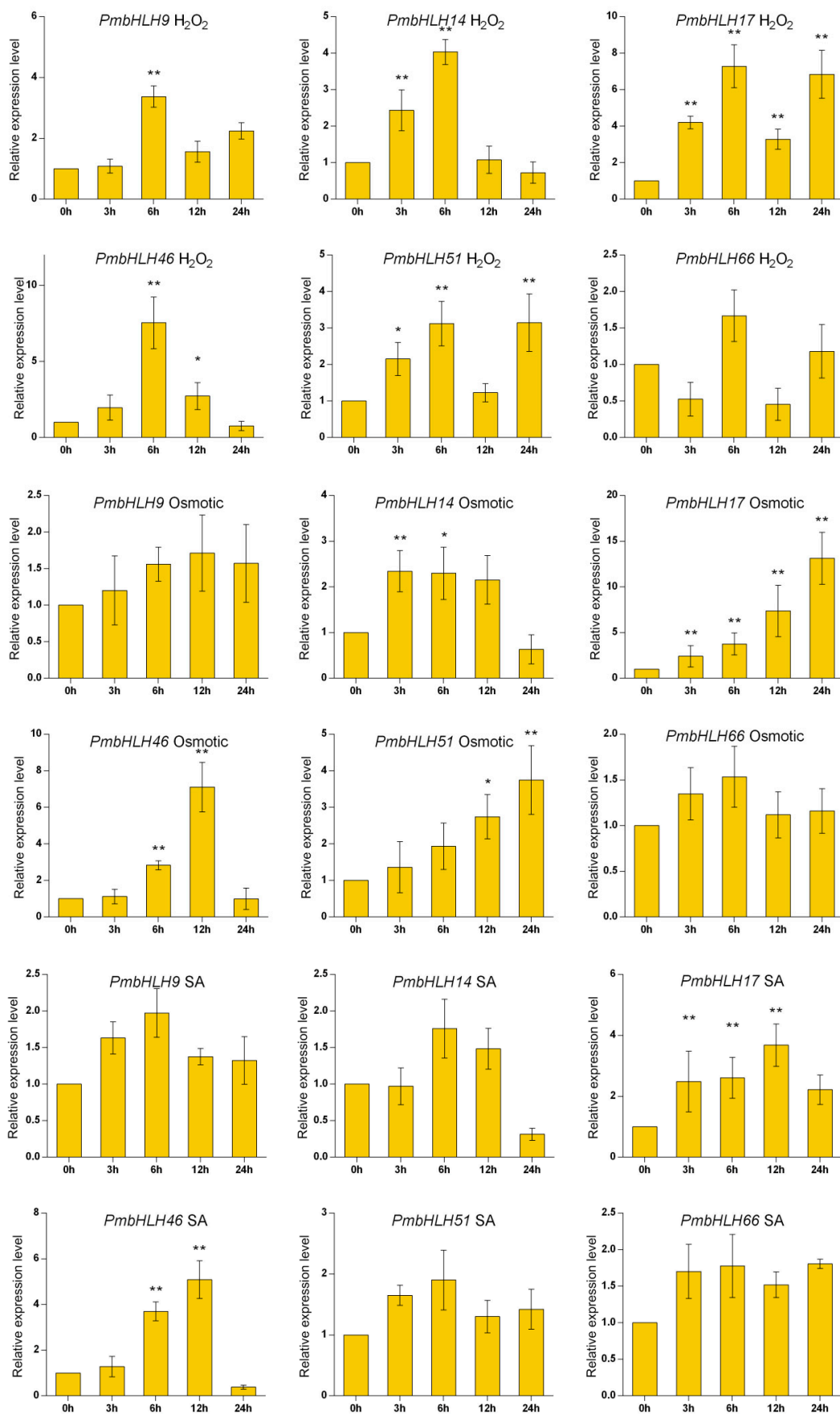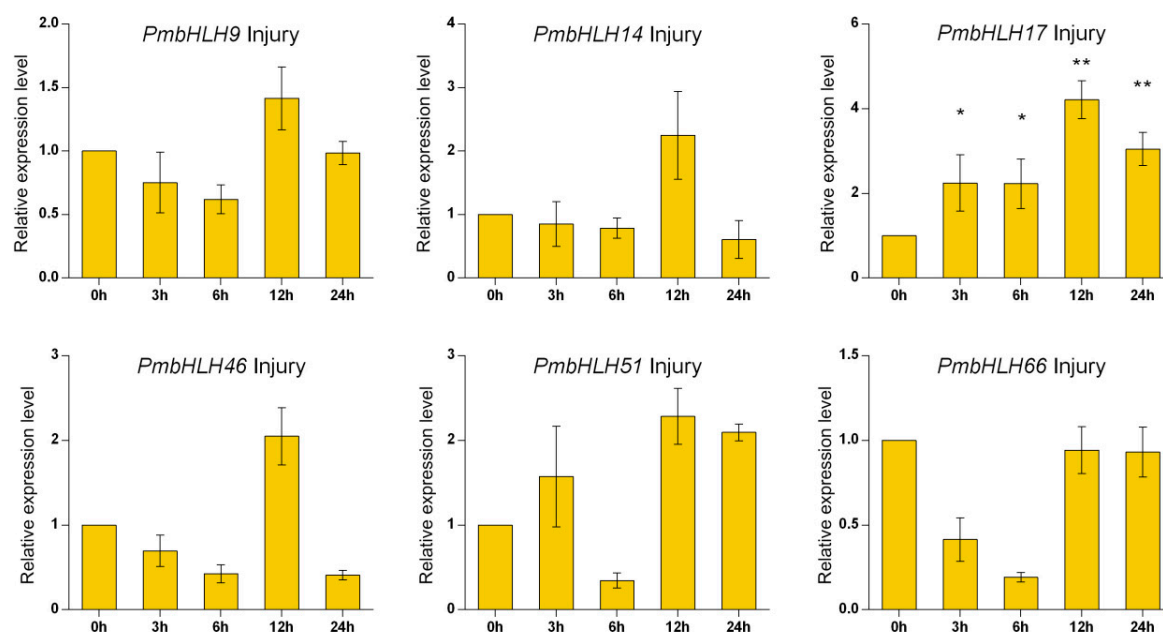

Figure 5. *Cont.*

Figure 5. *Cont.*

Figure 5. Expression of six *PmbHLH* genes under six treatments. Asterisks represent significant differences between each time point and 0 h (* $p < 0.05$, ** $p < 0.01$).

4. Discussion

The *bHLH* gene family is widespread in animals and plants and is the second largest gene family in plants. Its members play important roles in the metabolism, physiology and development of higher organisms. bHLH TFs were first classified in animals, and the research on these TFs is more extensive in animals than in plants. In plants, bHLH TFs are classified according to animal classification standards and bioinformatics analysis [30]. The *bHLH* gene family has been extensively studied in many plant species, such as *Z. mays* [31], *O. sativa* [13], *Solanum lycopersicum* [32] and *Malus domestica* [30], but has not been reported in *P. massoniana*. Therefore, we identified 88 *PmbHLH* genes from four transcriptomes of *P. massoniana*. These genes were divided into 10 groups based on evolutionary relationships. The domains of these genes were analyzed by MEME. All the genes contained bHLH domains, and the domains in the same group were similar. In addition, subcellular localization analysis, RT-PCR, and transcriptional profiling were also performed.

In this study, 88 bHLH-encoding proteins were identified from the transcriptomes of *P. massoniana* and named PmbHLH1-PmbHLH88. Evolutionary analysis of 88 *P. massoniana bHLH* genes yielded a phylogenetic tree divided into 10 groups. In a study of peach bHLH TFs, it was found that proteins were lost during evolution, indicating that some nonconserved bHLH subfamilies in certain plants may have evolved specifically to meet the needs of plant growth or stress resistance [33]. In a study of apple *bHLH* genes, it was found that the conserved domains of genes in the same group were almost the same, and there were large differences between different groups [30]. Maize ZmbHLH TFs in the same group have similar gene structures and similar conserved protein motifs, indicating that members of the same group are closely evolutionarily related [34]. These results are consistent with the results of the present study. However, special patterns were observed in some groups, such as the *PmbHLH16* gene in group I, which was different from other members and was more similar to members of group V. This gene may be a transitional form between the two groups in the evolutionary trajectory of a specific regulatory function, consistent with the conclusion of the study on apple *bHLH* genes. This system may provide a model for studying other groups in the *bHLH* gene family [30].

The results of subcellular localization prediction and tobacco transient expression showed that PmbHLHs are almost all nucleus-localized proteins. Only PmbHLH57 is located in the extracellular space. This is consistent with the subcellular localization results for rice bHLH TFs [13].

To further explore the physiological functions of bHLH TFs in plant growth and development, we analyzed the transcription profiles of the PmbHLH family under pine wood nematode inoculation. Within 35 days after inoculation with pine wood nematodes, the expression levels of these genes did not change very much. Through longitudinal comparison, some genes had high expression levels from beginning to end, and the levels of some genes continued to be very low. The overall expression change trends of all the genes were consistent, indicating that these genes are somewhat related to resistance to pine wood nematodes, but they do not play a particularly obvious regulatory role.

We treated *P. massoniana* with abiotic stress and hormones and then performed RT-PCR analysis. The expression levels of the six selected genes under MeJA, ETH, H_2O_2, osmotic, SA and injury stress increased in the later stage. PmbHLH9, PmbHLH51 and PmbHLH66 proteins eventually all increased their expression levels under ETH stress, indicating that these three proteins were significantly expressed under the induction of ETH. PmbHLH17 and PmbHLH51 proteins can be expressed significantly under osmotic stress. The expression levels of bHLH TFs in many plants have been analyzed under these treatments. For example, the *OsbHLH6* gene is significantly expressed in rice leaves under induction by MeJA. Studies have also confirmed that OsbHLH TFs are involved in regulating the SA and JA signaling pathways in rice [35]. The expression patterns under abscisic acid (ABA) and ETH treatment indicated that seven genes in strawberry are involved in the biosynthesis of fruit anthocyanins [36]. The expression level of TabHLH49 protein in wheat under osmotic stress was significantly higher than that of the control, and after deleting the C-terminus (amino acids 323–362), TabHLH49 protein exhibited no transcriptional activity in yeast, indicating that TabHLH49 positively regulates the expression of the dehydrin *WZY2* gene [37]. These results are consistent with the results of the present study, indicating that bHLH TFs play certain roles in drought tolerance, plant immune response and secondary metabolite synthesis.

5. Conclusions

In this study, we identified 88 *PmbHLH* genes from four transcriptomes of *P. massoniana* and performed bioinformatics analysis. These genes were divided into 10 groups in total, each gene contained a bHLH domain, and the domains in the same group were more similar than those in different groups. Subcellular localization prediction and experiments revealed that these genes encode mostly nucleus-localized proteins. RT-PCR showed that the expression levels of the six selected genes under the six treatments all increased in different time periods. Some proteins can be significantly expressed under specific stress. Therefore, members of the bHLH TF family play crucial roles in plant abiotic stress and hormone responses. This is of great significance to the breeding of the biosynthesis of resin and resistance to pine wood nematode disease in Masson pine. At present, there are very few studies on bHLH transcription factors of Pinaceae. This study will provide a theoretical basis for functional studies of the *bHLH* gene family, be very useful for studying the secondary metabolism of Pinaceae and provide a potential strategy for further breeding *P. massoniana*.

Author Contributions: Conceptualization, Y.C. and P.Z.; methodology, Y.C. and P.Z.; formal analysis, Y.C.; resources, P.Z., X.W., F.W. and J.Z.; writing—original draft preparation, Y.C.; writing—review and editing, Y.C., P.Z. and K.J. All authors have read and agreed to the published version of the manuscript.

References

1. Hou, L.; Liu, W.; Li, Z.; Huang, C.; Fang, X.; Wang, Q.; Liu, X. Identification and expression analysis of genes responsive to drought stress in peanut. *Russ. J. Plant Physiol.* **2014**, *61*, 842–852. [CrossRef]
2. Shen, X.; Wang, Z.; Song, X.; Xu, J.; Jiang, C.; Zhao, Y.; Ma, C.; Zhang, H. Transcriptomic profiling revealed an important role of cell wall remodeling and ethylene signaling pathway during salt acclimation in Arabidopsis. *Plant Mol. Biol.* **2014**, *86*, 303–317. [CrossRef]
3. Yuan, F.; Chen, M.; Yang, J.; Leng, B.; Wang, B. A system for the transformation and regeneration of the recretohalophyte Limonium bicolor. *Vitr. Cell. Dev. Biol.–Plant* **2014**, *50*, 610–617. [CrossRef]
4. Feller, A.; Machemer, K.; Braun, E.L.; Grotewold, E. Evolutionary and comparative analysis of MYB and bHLH plant transcription factors. *Plant J.* **2011**, *66*, 94–116. [CrossRef]
5. Liu, X.Y.; Wang, W.S. Research progress of plant bHLH transcription factor family. *Curr. Biotechnol.* **2011**, *1*, 391–397.
6. Sun, X.; Wang, Y.; Sui, N. Transcriptional regulation of bHLH during plant response to stress. *Biochem. Biophys. Res. Commun.* **2018**, *503*, 397–401. [CrossRef]
7. Li, F.; Liu, W. Genome-wide identification, classification, and functional analysis of the basic helix-loop-helix transcription factors in the cattle, Bos Taurus. *Mamm. Genome* **2017**, *28*, 176–197. [CrossRef]
8. Jones, S. An overview of the basic helix-loop-helix proteins. *Genome Biol.* **2004**, *5*, 226. [CrossRef]
9. Nesi, N.; Debeaujon, I.; Jond, C.; Pelletier, G.; Caboche, M.; Lepiniec, L. The TT8 gene encodes a basic helix-loop-helix domain protein required for expression of DFR and BAN genes in Arabidopsis Siliques. *Plant Cell* **2000**, *12*, 1863–1878. [CrossRef]
10. Atchley, W.R.; Terhalle, W.; Dress, A.W.M. Positional dependence, cliques, and predictive motifs in the bHLH protein domain. *J. Mol. Evol.* **1999**, *48*, 501–516. [CrossRef] [PubMed]
11. Murre, C.; McCaw, P.S.; Baltimore, D. A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins. *Cell* **1989**, *56*, 777–783. [CrossRef]
12. Heim, M.A.; Jakoby, M.; Werber, M.; Martin, C.; Weisshaar, B.; Bailey, P.C. The basic helix-loop-helix transcription factor family in plants: A genome-wide study of protein structure and functional diversity. *Mol. Biol. Evol.* **2003**, *20*, 735–747. [CrossRef] [PubMed]
13. Li, X.; Duan, X.; Jiang, H.; Sun, Y.; Tang, Y.; Yuan, Z.; Guo, J.; Liang, W.; Chen, L.; Yin, J.; et al. Genome-wide analysis of basic/helix-loop-helix transcription factor family in rice and Arabidopsis. *Plant Physiol.* **2006**, *141*, 1167–1184. [CrossRef] [PubMed]
14. Albert, N.W.; Lewis, D.H.; Zhang, H.; Irving, L.J.; Jameson, P.E.; Davies, K.M. Light-induced vegetative anthocyanin pigmentation in Petunia. *J. Exp. Bot.* **2009**, *60*, 2191–2202. [CrossRef]
15. Chu, Y.; Xiao, S.; Su, H.; Liao, B.; Zhang, J.; Xu, J.; Chen, S. Genome-wide characterization and analysis of bHLH transcription factors in *Panax ginseng*. *Acta Pharm. Sin. B* **2018**, *8*, 666–677. [CrossRef]
16. Gao, C.; Sun, J.; Wang, C.; Dong, Y.; Xiao, S.; Wang, X.; Jiao, Z. Genome-wide analysis of basic/helix-loop-helix gene family in peanut and assessment of its roles in pod development. *PLoS ONE* **2017**, *12*, e0181843. [CrossRef] [PubMed]
17. Wang, P.; Su, L.; Gao, H.; Jiang, X.; Wu, X.; Li, Y.; Zhang, Q.; Wang, Y.; Ren, F. Genome-wide characterization of bHLH genes in grape and analysis of their potential relevance to abiotic stress tolerance and secondary metabolite biosynthesis. *Front. Plant Sci.* **2018**, *9*, 64. [CrossRef]
18. Zhang, X.; Luo, H.; Xu, Z.; Zhu, Y.; Ji, A.; Song, J.; Chen, S. Genome-wide characterisation and analysis of bHLH transcription factors related to tanshinone biosynthesis in *Salvia miltiorrhiza*. *Sci. Rep.* **2015**, *5*, 11244. [CrossRef]
19. Liu, W.; Tai, H.; Li, S.; Gao, W.; Zhao, M.; Xie, C.; Li, W.X. bHLH122 is important for drought and osmotic stress resistance in Arabidopsis and in the repression of ABA catabolism. *New Phytol.* **2014**, *201*, 1192–1204. [CrossRef]
20. Chinnusamy, V.; Ohta, M.; Kanrar, S.; Lee, B.H.; Hong, X.; Agarwal, M.; Zhu, J.K. ICE1: A regulator of cold-induced transcriptome and freezing tolerance in Arabidopsis. *Genes Dev.* **2003**, *17*, 1043–1054. [CrossRef]
21. Frerigmann, H.; Berger, B.; Gigolashvili, T. bHLH05 is an interaction partner of MYB51 and a novel regulator of glucosinolate biosynthesis in Arabidopsis. *Plant Physiol.* **2014**, *166*, 349–369. [CrossRef] [PubMed]
22. Zhang, X.; Ge, F.; Deng, B.; Shah, T.; Huang, Z.; Liu, D.; Chen, C. Molecular cloning and characterization of PnbHLH1 transcription factor in *Panax notoginseng*. *Molecules* **2017**, *22*, 1268. [CrossRef] [PubMed]

23. Strom, B.L.; Goyer, R.A.; Ingram, L.L.; Boyd, G.D.L.; Lott, L.H. Oleoresin characteristics of progeny of loblolly pines that escaped attack by the southern pine beetle. *For. Ecol. Manag.* **2002**, *158*, 169–178. [CrossRef]

24. Martin, D.; Bohlmann, J. Molecular biochemistry and genomics of terpenoid defenses in conifers. *Chem. Ecol. Phytochem. For. Ecosyst.* **2005**, 29–56. [CrossRef]

25. Nagy, N.E.; Franceschi, V.R.; Solheim, H.; Krekling, T.; Christiansen, E. Wound-induced traumatic resin duct development in stems of Norway spruce (Pinaceae): Anatomy and cytochemical traits. *Am. J. Bot.* **2000**, *87*, 302–313. [CrossRef]

26. McKay, S.A.; Hunter, W.L.; Godard, K.A.; Wang, S.X.; Martin, D.M.; Bohlmann, J.; Plant, A.L. Insect attack and wounding induce traumatic resin duct development and gene expression of (-)-pinene synthase in Sitka spruce. *Plant Physiol.* **2003**, *133*, 368–378. [CrossRef] [PubMed]

27. Wu, F.; Sun, X.; Zou, B.; Zhu, P.; Lin, N.; Lin, J.; Ji, K. Transcriptional analysis of Masson Pine (*Pinus massoniana*) under high CO_2 stress. *Genes* **2019**, *10*, 804. [CrossRef]

28. Kumar, S.; Stecher, G.; Tamura, K. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* **2016**, *33*, 1870–1874. [CrossRef]

29. Chen, C.J.; Chen, H.; Zhang, Y.; Thomas, H.R.; Xia, R. TBtools: An integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant* **2020**, *13*, 1194–1202. [CrossRef]

30. Mao, K.; Dong, Q.; Li, C.; Liu, C.; Ma, F. Genome wide identification and characterization of apple bHLH transcription factors and expression analysis in response to drought and salt stress. *Front. Plant Sci.* **2017**, *8*, 480. [CrossRef]

31. Ludwig, S.R.; Habera, L.F.; Dellaporta, S.L.; Wessler, S.R. Lc, a member of the maize R gene family responsible for tissue-specific anthocyanin production, encodes a protein similar to transcriptional activators and contains the myc-homology region. *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 7092–7096. [CrossRef] [PubMed]

32. Sun, H.; Fan, H.J.; Ling, H.Q. Genome-wide identification and characterization of the bHLH gene family in tomato. *BMC Genom.* **2015**, *16*, 9. [CrossRef] [PubMed]

33. Zhang, C.; Feng, R.; Ma, R.; Shen, Z.; Cai, Z.; Song, Z.; Peng, B.; Yu, M. Genome-wide analysis of basic helix-loop-helix superfamily members in peach. *PLoS ONE* **2018**, *13*, e0195974. [CrossRef] [PubMed]

34. Zhang, T.; Lv, W.; Zhang, H.; Ma, L.; Li, P.; Ge, L.; Li, G. Genome-wide analysis of the basic Helix-Loop-Helix (bHLH) transcription factor family in maize. *BMC Plant Biol.* **2018**, *18*, 235. [CrossRef]

35. Meng, F.; Yang, C.; Cao, J.; Chen, H.; Pang, J.; Zhao, Q.; Wang, Z.; Qing Fu, Z.; Liu, J. A bHLH transcription activator regulates defense signaling by nucleo-cytosolic trafficking in rice. *J. Integr. Plant. Biol.* **2020**. [CrossRef]

36. Zhao, F.; Li, G.; Hu, P.; Zhao, X.; Li, L.; Wei, W.; Feng, J.; Zhou, H. Identification of basic/helix-loop-helix transcription factors reveals candidate genes involved in anthocyanin biosynthesis from the strawberry white-flesh mutant. *Sci. Rep.* **2018**, *8*, 2721. [CrossRef]

37. Liu, H.; Yang, Y.; Liu, D.; Wang, X.; Zhang, L. Transcription factor TabHLH49 positively regulates dehydrin WZY2 gene expression and enhances drought stress tolerance in wheat. *BMC Plant Biol.* **2020**, *20*, 259. [CrossRef]