Article

Genome-Wide Identification and Functional Analysis of the Basic Helix-Loop-Helix (bHLH) Transcription Family Reveals Candidate *PtFBH* Genes Involved in the Flowering Process of *Populus trichocarpa*

Yang Ye ¹, Haodong Xin ¹, Xiting Gu ¹, Jianwen Ma ¹ and Lingli Li ^{1,2,3,*}

¹ College of Forestry, Northwest A&F University, Yangling 712100, China; dayeyang2020@163.com (Y.Y.); haodongxinn@163.com (H.X.); xitingnwf@163.com (X.G.); Jianwenmaxn@126.com (J.M.)

² Shaanxi Province Key Laboratory of Economic Plant Resources Development and Utilization, Yangling 712100, China

³ Research and Industrialization Development Research Center of the Shantung Maple, Yangling 712100, China

* Correspondence: lill@nwf.edu.cn; Tel./Fax: +86-029-8708-2230

Citation: Ye, Y.; Xin, H.; Gu, X.; Ma, J.; Li, L. Genome-Wide Identification and Functional Analysis of the Basic Helix-Loop-Helix (bHLH)

Transcription Family Reveals Candidate *PtFBH* Genes Involved in the Flowering Process of *Populus trichocarpa*. *Forests* **2021**, *12*, 1439.

<https://doi.org/10.3390/f12111439>

Academic Editor: Tadeusz Malewski

Received: 11 August 2021

Accepted: 18 October 2021

Published: 22 October 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: As one of the largest TF families in plants, the basic helix-loop-helix (bHLH) family plays an important part in the growth and development of many plants. FLOWERING BHLH (FBH) encodes a bHLH-type transcriptional factor related to the flowering process. Poplar is a model woody plant as well as an important economic tree species with a small genome. However, the characteristics of the *bHLHs* and *FBHs* gene family in the newest version of *Populus trichocarpa* genome have not been analyzed yet. We identified 233 *PtbHLHs* and 10 *PtFBHs* in the newest version genome, and *PtbHLHs* were classified into 21 groups with FBH subfamily occupying one, supported by phylogenetic analysis, exon–intron patterns, and conserved protein motifs. These *PtHLHs* were distributed on 19 chromosomes unevenly and expressed in nucleus mainly. Gene duplication and synteny analysis have indicated that the *PtbHLHs* gene family has undergone strong purification selection during the evolution process. The *cis*-elements analysis has suggested that *PtbHLHs* may be related to the growth and development. Conserved residues of FBHs among *Arabidopsis* and poplar were also identified. Expression of 227 *PtHLH* genes (6 unmatched, 13 no expressed) showed diverse patterns in different tissues, implying their multiple functions. Protein–protein interaction network prediction and expression patterns in three states of the flowering process (Flowers-Dormant, Flowers-Expanding and Flowers-Expanded) suggested that some members of *PtbHLH* and *PtFBH* family may be involved in the flowering process. Our comprehensive and systematic analysis can provide some valuable clues and basic reference toward further investigations on physiological and molecular functions of *PtbHLHs*.

Keywords: *Populus trichocarpa*; genome-wide analysis; *PtHLHs*; *PtFBHs*; expression analysis; flowering process

1. Introduction

There are four types of transcription factors (TFs) in higher plants categorized by the sequence of arginine and lysine residues in the DNA binding region: Zinc finger, HTH, bZIP, and HLH [1–4]. As one of the largest TF families in plants [5], the basic helix-loop-helix (bHLH) family plays an important part in the growth and development of many plants. Each member of bHLH family contains a bHLH domain, which consists of around 50–60 conserved amino acid sequences. The bHLH domain contains two functional regions: a basic amino acid region and an HLH region. The basic region is composed of approximately 15 amino acids at the N-terminal and responsible for binding to the *cis*-element E-box (CANNTG) or G-box (CACGTG) while the HLH region of a length of approximately 50 amino acids in a structure of two alpha helices separated by a loop

of variable length is located at the C-terminal [6–8]. It has been reported that two HLH proteins form homodimeric or heterodimeric complexes by the interaction of the HLH region's hydrophobic amino acids to regulate the expression of downstream target genes. CIBs have been proved to form heterodimers to engage in regulating CRY2-dependent flowering [9]. Thus, the bHLH transcription factors generally function as a dimer [10,11].

FLOWERING BHLH (FBH) encodes a bHLH-type transcriptional factor related to the flowering process. In *Arabidopsis*, there are four FBHs identified and proved to target the CO (CONSTANS) gene preferentially binding to the E-box *cis*-elements to upregulate the expression of CO levels [12]. FBH1 acts as a transcriptional modulator of warm temperature signals and clock responses by binding to the promoter of the key clock gene CIRCADIAN CLOCK-ASSOCIATED 1 (CCA1) and regulating its expression [13]. A study has indicated that a member of the bHLH transcription factor family in petunia PhFBH4 targeted PhACS1 through the G-box *cis*-element directly and then participated in modulating the ethylene biosynthesis pathway at flower senescence [14]. While, ScFBH1, ScFBH2, and ScFBH3 can control ACC synthase expression via binding to the promoter of a sugarcane type 3'ACS isozyme gene (ScACS2) and form homodimeric and heterodimers in the nucleus, transcriptionally regulating ethylene biosynthesis in sugarcane [15]. Therefore, reasonably, a branch of FBH from bHLH family should be taken out to explore their relationship with the plant flowering pathway.

Poplar, a model woody plant as well as an important economic tree species, has a small genome and is relatively easy to manipulate in terms of genetics and molecular biology [16]. Therefore, it is meaningful to promote early flowering of poplar and shorten its juvenile period, which plays a positive role in accelerating the life cycle of poplar in order to solve the problem of genetic breeding and related research progress limited by poplar's long juvenile stage years. At the same time, understanding the mechanism of poplar flowering can provide a theoretical basis for solving the floating catkins problem.

Genome-wide analysis of important functional genes is an important and well-established approach [17–22]. Since the *P. trichocarpa* genome has been completely sequenced, it is feasible and can make a lot sense to perform a genome-wide analysis of *PtbHLHs*, providing a biological reference for the further research of the function of the *PtbHLHs* gene family in poplar. A previous study has focused on 202 *bHLH* genes in poplar from version v3.0 of genomic sequences in PlantTFDB website [23]. In this study, we identified 233 members of the bHLH gene family from the newest version v4.1 of genomic sequences in *P. trichocarpa* and performed detailed bioinformatics analysis. We also studied the expression pattern of the *PtbHLH* genes of poplar in different tissues and flowering states. Our comprehensive and systematic analysis can provide some valuable clues and basic reference toward further investigations on physiological and molecular functions of *PtbHLHs*.

2. Materials and Methods

2.1. Identification and Phylogenetic Analyses of the bHLH in Poplar

We downloaded the newest version v4.1 of genomic sequences in *P. trichocarpa* from the Phytozome v12.1 database (<https://phytozome.jgi.doe.gov/>, accessed on 1 September 2021). The HMM files of the typical bHLH protein domain (PF0010) were downloaded from Pfam (E-value < 0.001). Potential poplar bHLH proteins were scanned against the poplar genome by use of the *hmmsearch* in HMMER 3.0 program (<http://hmm.wustl.edu/>, accessed on 1 September 2021) [24]. Then, sequences containing the bHLH or bHLH-MYC_N domains were retained for the *hmm* build construction of HMM which we used for the next round of search. In this way, we went eight rounds with no new sequence appearing in the last round. The intersection of all the searched sequences containing the bHLH or bHLH-MYC_N domains belongs to *PtbHLH* family. All the possible *PtbHLHs* were confirmed using the CDD, SMART, and PFAM databases. The bHLH proteins of *Arabidopsis* were downloaded from the *Arabidopsis* Information Resource (TAIR) (www.arabidopsis.org,

accessed on 1 September 2021) using the 153 *bHLH* gene locus gained from PlantTFDB database (<http://planttfdb.cbi.pku.edu.cn/>, accessed on 1 September 2021).

MEGA-X was used for the evolutionary and phylogenetic analyses (<https://www.megasoftware.net/>, accessed on 1 September 2021) [25]. A total of 233 *PtbHLH* amino acid sequences, were used to generate an unrooted phylogenetic tree using the neighbor-joining method with 1000 bootstrap replications, JTT substitution model, and 50% partial detection. Multiple sequence alignments were performed by MUSCLE [26] and then those sequences were clipped to 71 bp each using TBtools by Multiple Alignment Trimming before building trees [27].

2.2. Chromosome Location and Gene Duplication Analysis

Each *PtbHLH* gene was matched with the chromosomes of poplar using TBtools program based on the genome annotations of poplar [28]. The MCScanX software was used to identify duplicated genes with *E* value set to e^{-5} [29]. We used BioEdit Sequence Alignment Editor [30] to align the CDS sequence of the replicated gene pair, then DnaSP5 (<https://dnasp.software.informer.com/>, accessed on 1 September 2021) to calculate the *Ks* and *Ka* of the replicated gene pair, and the evolution time (*T*) calculation was based on the *Ks* value: $T = Ks/2\lambda$, $\lambda = 9.1 \times 10^{-9}$ [31]. Members of *bHLHs* in *Arabidopsis* and rice were obtained from PlantTFDB database (<http://planttfdb.gao-lab.org/family.php?fam=bHLH>, accessed on 1 September 2021).

2.3. Protein Properties and Structure Analysis

Physical and chemical features of each *PtbHLH* protein were predicted by use of the ExPASy website (<http://web.expasy.org/protparam/>, accessed on 1 September 2021) [32]. The MEME tool (<http://meme-suite.org/tools/meme>, accessed on 1 September 2021) was used to identify 15 conserved motifs of *bHLHs* in *P. trichocarpa* with default parameters [33]. The gene structure was visualized by using TBtools program based on the genome annotations of poplar.

2.4. Cis-Acting Element Analysis and Multiple Sequence Alignment of the *PtbHLHs*

For each *PtbHLH* gene, the gene upstream 2000 bp DNA sequence of the start codon was obtained from the Phytozome v12.1 database. We then used the PlantCRAE to extract *cis*-acting elements that will be visualized by use of the TBtools software [34]. The *cis*-elements of 233 identified *PtbHLH* genes were classified as Light responsive, Hormone responsive, Stress responsive, Hormone responsive, and Light responsive.

2.5. Subcellular Location and Protein–Protein Interaction Network Prediction

TargetP Server online software (<http://www.cbs.dtu.dk/services/TargetP/>, accessed on 1 September 2021) was used to predict the subcellular location of *bHLH* proteins. The 16 full-length FBH protein sequences (six *Arabidopsis* FBHs and 10 poplar FBHs) were selected for multiple sequence alignment using the MAFFT (<https://www.ebi.ac.uk/Tools/msa/mafft/>, accessed on 1 September 2021) online tool [35]. The alignment results were displayed by JalView software (<http://www.jalview.org/download>, accessed on 1 September 2021) [36]. Proteins of 10 *PtFBHs* and 18 *PtbHLHs* possibly involved in flowering development were used for predicting protein–protein interactions by the STRING website (<https://string-db.org/>, accessed on 1 September 2021). We chose the informative figures gained to display. From these figures obtained from STRING database we can predict several relatively reliable interactions and functions of *PtbHLHs*.

2.6. Expression Patterns of *PtbHLH* Genes in Different Tissues

We gained the gene expression database available on PopGenIE website (<ftp://plantgenie.org/Data/PopGenIE/>, accessed on 1 September 2021) to explore the *PtbHLH* gene expression patterns in various tissues and different growth stages [37]. In the database, six *PtbHLH* genes could not be matched and were omitted. Flowers-Expanded,

Flowers-Dormant, Flowers-Expanding, Roots-Control, Seeds-Mature, Leaves-Control, and Cambium-Phloem-Dormant were sampled to investigate the expression patterns of *PtbHLH* genes. Subsequently, log2FC values were used for evaluating the gene expression and the heatmap was generated using TBtools. The RT-qPCRs were performed to confirm the DEGs analysis. We used the Plant RNA kit (Aidlab-Biotech, Beijing, China) to extract the total mRNA. The RT-qPCR reactions were performed in a real-time PCR system using SYBR Green (Applied Biosystems, Waltham, MA, USA) according to the manufacturer's instructions. The *Actin2/7* gene was used as a reference, the RNA relative expression of each gene was calculated using the $2^{-\Delta\Delta CT}$ method. The RT-qPCR reactions were repeated three times. The specific primers are shown in Table S8. The experimental data were subjected to analysis of variance by Statistical Product and Service Solutions 17.0 (SPSS). For statistical analyses, Student's *t*-test was used to generate every *p* value.

3. Results

3.1. Identification, Characterization and Subcellular Localization of the bHLHs in Poplar

A total of 233 *bHLH* genes, named *PtbHLH1*–*PtbHLH233* according to their accession number, were identified in *P. trichocarpa* (Table S1). Compared with *Arabidopsis* (162), maize (208), peach (95), apple (188), and grape (94), poplar boasts the largest number of *bHLH* family numbers [38–42].

We then analyzed physicochemical properties and predicted subcellular location of proteins encoded by these 233 *PtbHLH* genes (Tables S2 and S3). *PtbHLH* proteins are predicted to locate in various organelles such as nucleus, inner membrane system, mitochondria, and chloroplasts. According to statistics, 208 *PtbHLH* proteins are mainly expressed in plant cell nuclei, accounting for 89% of the total number. Therefore, it is speculated that *PtbHLHs* mainly play a role in the nucleus. Their number of amino acids varies from 84 (*PtbHLH107*) to 970 (*PtbHLH11*). Molecular weights fall into the range of 9359.95 (*PtbHLH107*) to 104,899.77 kDa (*PtbHLH11*). The protein is basic when its value of theoretical isoelectric point (pI) is over 7, otherwise it is acidic. We predict 86 basic proteins and 147 acidic proteins from 233 *PtbHLHs*, whose pI values range from 4.44 to 11.74. Interestingly, we found that only one of the proteins have the grand average of hydropathicity index greater than 0 (*PtbHLH107*), which means it is a hydrophobic protein, and the other indices are less than 0, which means they are hydrophilic proteins. The values of aliphatic index run the gamut from 50.64 to 130.00. The values of protein instability index are between 23.05 and 89.92. Nine *PtbHLH* are stable proteins (*PtbHLH101*, *PtbHLH210*, *PtbHLH107*, *PtbHLH198*, *PtbHLH128*, *PtbHLH11*, *PtbHLH232*, *PtbHLH185*, and *PtbHLH189*), whose values of protein instability index are less than 40, and 224 proteins unstable. Regarding the formulas of *PtbHLH* proteins, they are all listed in the form (Table S2).

3.2. Chromosomal Distribution and Synteny Analysis of *PtbHLH* Genes

According to the annotation of the poplar genome database, we found that 233 *PtbHLH* genes are widely distributed on 19 chromosomes of poplar and gained the gene density on every chromosome. Then a chromosomal distribution of the *PtbHLH* genes was mapped to locate the position of each member (Figure 1). The map shows that chromosome 2 harbors the largest number of *PtbHLH* genes (27) while chromosome 16 and chromosome 17 both contain the least (7). In summary, *PtbHLHs* were found on all 19 chromosomes and on most chromosomes the number of the genes is between 7 and 12.

Figure 1. Chromosomal locations of *bHLH* gene family in poplar. Chromosome numbers are shown at the left of the chromosomes. *bHLH* genes are labeled at the right of the chromosomes. Scale bar denotes the chromosome lengths (Mb).

Our study also analyzed the gene duplication events in *PtbHLH* genes. The MCScanX software was used to perform a collinearity analysis. The results showed that 223 *PtbHLH* genes came from whole genome duplication (WGD) or segmental duplication and the other 10 *PtbHLH* genes came from dispersed duplication (Table S4). There were six groups of tandem duplicated *PtbHLH* genes (*PtbHLH91/PtbHLH92*, *PtbHLH97/PtbHLH96*, *PtbHLH222/PtbHLH223*, *PtbHLH229/PtbHLH231*, *PtbHLH229/PtbHLH230*, *PtbHLH230/PtbHLH231*) and 110 pairs of segmentally duplicated *PtbHLH* genes (Figure 2, Table S4). Gene duplication greatly promoted the expansion of the *bHLH* gene family in poplar genome, and WGD/segmental duplication events played a major driving role. In addition, we performed homology analysis to identify the conserved chromosomal segments within the poplar genome in order to explore the evolutionary process of the *PtbHLHs*. The distribution of homologous *PtbHLH* gene pairs on chromosomal distribution is random. The evolution date of WGD/fragment replication events can be estimated by the Ks value (frequency of synonymous mutations) [43]. As previously studied, the *P. trichocarpa* genome has undergone at least three genome-wide replication events [16]. Therefore, we used the Ks value to estimate the evolution date of gene duplication events in the *PtbHLH* gene family (Table S5). The Ks value means that most of the *PtbHLH* genes were duplicated around the recent WGD event while some other genes originated from ancient WGD. The intensity and direction of selection can be represented by the ratio of Ka/Ks [44]. The Ka/Ks ratio of all homologous *PtbHLH* genes is less than 1, indicating that the *PtbHLH* gene family has undergone strong purification selection during the evolution process. According to the *PtbHLH* gene collinearity relationship between poplar and *Arabidopsis* or rice, we can easily know that there are a large number of homologous *bHLH* genes among them (Figure S1).

Figure 2. The distribution pattern syntenic analysis of the *bHLH* gene family. The red lines indicated the syntenic gene pairs of the *bHLH* gene family. Scale bar marked on the chromosome indicates chromosome lengths (Mb). Chr: Chromosome.

3.3. Phylogenetic Analyses of the *PtbHLH* in Poplar

A reasonable phylogenetic tree can reveal and help us better understand the evolutionary relationship of gene families. We firstly used 233 *bHLH* protein sequences of poplar to construct an original phylogenetic tree, whose branches were constructed according to its gene structure and motif distribution (Figures 3 and 4).

Figure 3. Phylogenetic tree based on protein sequences of bHLH gene family in poplar. The phylogenetic tree was constructed using the neighbor-joining (NJ) method, with 1000 bootstrap replicates, using Mega-X.

Figure 4. The conserved motifs and gene structure analysis of bHLH gene family in poplar. **(A)** The neighbor-joining (NJ) phylogenetic tree of poplar *bHLH* genes. The phylogenetic tree was constructed by Mega-X with 1000 bootstraps. **(B)** The conserved motifs analysis of poplar *bHLH* genes. A total of 15 motifs were predicted by MEME tool, named Motifs 1–15. **(C)** The gene structure analysis of poplar *bHLH* genes, including UTR, CDS, and intron. The green rectangles represented UTR; the yellow rectangles represented CDS; the grey lines presented introns.

From the phylogenetic tree (Figure 3), *PtbHLHs* can be well divided into 21 groups, named from group I to group XXI. It appears that group I has the maximal number of genes (26) followed by group XII (19). Group XIX, group XX, and group XXI have the minimal number of genes (4). As for groups III–XVIII, there are 17, 17, 16, 15, 15, 12, 11, 11, 10, 10, 8, 7, 7, 7, 7, and 6 *PtbHLH* members, respectively. Fortunately, one subgroup FBH subfamily was classified together all the time, which is exactly what we focused on. As a consequence, we drew the conclusion that there are 10 members in *PtFBH* subfamily.

Finally, we constructed a phylogenetic tree using 16 FBH proteins (Figure 5), of which six were the previously reported in *Arabidopsis* [12]. As is shown in Figure 5, *PtbHLH94* and *PtbHLH206* are closest to FBH4 (AT2G42280) followed by *PtbHLH12* and *PtbHLH131*, *PtbHLH1*, and *PtbHLH57* are closest to FBH3 (AT1G51140), and *PtbHLH173* and *PtbHLH227* are closest to FBH2 (AT4G09180) and FBH1 (AT1G35460).

Figure 5. Multi-sequence alignment and domain analysis of the FBHs in poplar. (A) The phylogenetic analysis of FBHs gene family in poplar and *Arabidopsis*. (B) Conserved domains in FBHs proteins are marked with a black box and indicated. Sequence alignment was generated by DNAMAN 6.0 software. (C) Sequence logo was built by online software WebLogo.

3.4. Conserved Motif and Gene Structure Analysis of *PtbHLHs*

In general, genes in the same group share similar gene and protein structures. We found that the number of exons per *PtbHLH* gene varies from 1 to 11 (Figure 4). *PtbHLH* genes in the same subgroup usually have similar exons and introns: genes in group VII and group XI have 1 exon with no intron, group XXI 2 exons, group VIII 3 exons, group XX 4 exons, group I and group XVII 6–8 exons, group XVI 10–11 exons, and so on. Then we analyzed the conserved motifs of the *PtbHLH* proteins using the MEME website. In total, 15 conserved motifs were identified and designated as motifs 1–15 (Figure 4, Table S6). The motif 1 and motif 2 should be the specific motifs of the *PtbHLH* family because the majority of the protein subgroups contain both while group XIX, group XV, group IX, and group XX do not. Similar to gene structure, *PtbHLH* proteins in the same subgroup share group-specific conserved motifs. As is in Figure 3, motif 3 is only present in all members of group I, group XVII, group XII, and group XVIII and one of the members in group IV; motif 5 is only present in all members of group XVII, group II, and one of the members in group XIII; motif 6 is only observed in group XVI, group X, group XIX, and motif 8 in

group XVI, group X, group XI; motif 11 is unique to group I, motif 12 in group IX, motif 13 in group XV, and motif 14 in group IX.

3.5. *Cis-Elements Analysis in Promoters of the PtbHLH Genes*

From the visualized *cis*-elements in Table S7 and Figure S2, all members of the *PtbHLH* gene family contain a large number of light responsive elements, almost every member contains hormone responsive elements, and most members also contain stress responsive elements. The *cis*-elements in the promoter region may reflect the potential regulation and function of genes. These results suggest that the *PtbHLH* gene family may play an important role in regulation of growth and development.

3.6. *Multiple Alignment Analysis of the FBH Subfamily*

Sequence alignment of 16 proteins from the FBH subfamily in *Arabidopsis* and poplar is shown in the Figure 5 and Figure S3, where only amino acid sites with greater than 30% identity are displayed in color. At the same time, we obtained the LOGO of the conserved HLH domain of poplar and *Arabidopsis* FBH (Figure 5), where we found the bHLH domain here containing 54 conserved amino acid sequences. The structure of the bHLH conservative domain is distributed as follows: 17 bp basic amino acid region at the N-terminal, 15 and 19 bp alpha helixes and separated by a 3 bp loop at the C-terminal. Generally, the loop region is not conservative and its length is variable. We proved the loop in FBHs is a conservative region with 3 bp amino acids when not in bHLHs normally. We also can get the consensus of the HLH domain: K-R-G-X-A-T-H-P-R-S-I-A-E-R-X-R-R (b area); T-R-I-S-X-R/K-X-R/K-K-L-Q-X-L-V-P (H area); N-M-D (L area); K-Q-T-X3-D-M-L-X2-A-V-X3-K-X-L (H area). We concluded that HLH domain sequences were highly conserved between poplar and *Arabidopsis* FBHs.

3.7. *Expression Profile Analysis of PtbHLH Genes in Different Tissues*

To better characterize tissue-specific gene expression patterns of the *PtbHLH* genes, we compared their transcriptional levels in various tissues by using publicly available RNA-Seq data [37]. Values in a total of five sample tissues were selected from Flowers-Expanded, Roots-Control, Seeds-Mature, Leaves-Control, and Cambium-Phloem-Dormant (6 unmatched, 13 no expressed). Based on the Log2fold change (FC) of each *PtbHLH* gene, we generated a hierarchical clustering analysis and heatmap to characterize the *PtbHLH* family expression profiles (Figure S4).

Thirteen genes were not significantly upregulated or downregulated in all analyzed tissues, which may be due to differences in spatial and temporal expression patterns. From Figure S4, we know the number of differentially expressed genes (DEGs) in tissues: 43 genes were found significantly upregulated and 16 downregulated in Flowers-Expanded tissue; 63 genes upregulated and 16 downregulated in Roots-Control; 48 genes upregulated and 18 downregulated in Seeds-Mature; 30 genes upregulated and 18 downregulated in Leaves-Control; 47 genes upregulated and 66 downregulated in Cambium-Phloem-Dormant. A total of 52 genes were significantly upregulated in just one tissue, suggesting their function in a specific tissue. In addition, 14 genes exhibiting a significant upregulation in three or more of all the analyzed tissues may play key roles in this tissue development and growth of poplar. Besides, four of our *PtbHLH*s are in the list of the most tissue-representative genes generated by previous analysis [37]: both *PtbHLH2* and *PtbHLH130* in wood, *PtbHLH134* in buds, and *PtbHLH16* in seeds. Exploring the expression pattern of *PtbHLH* gene in different tissues provides the basis for the identification of functional genes in poplar.

3.8. *The PtbHLH Protein–Protein Interaction Network Prediction and Some May Be Involved in the Flowering Process*

Two heatmaps were then generated for analyzing their expression patterns in three states of the flowering process: Flowers-Dormant, Flowers-Expanding, and Flowers-Expanded (Figure 6). One heatmap is based on the log2FC values of 43 significantly upregulated genes in Flowers-Expanded together with *PtCO1* (Potri.017G107500), *PtCO2*

(Potri.004G108300), and *PtFT* (Potri.008G077700) (Figure 6A). The other heatmap is based on 10 *PtFBHs* together with *PtCO1*, *PtCO2*, and *PtFT* (Figure 6B). The qRT-PCRs were performed to confirm the results showing similar trends in the DEGs analysis (Figure 6C).

to *PtbHLH14*; *PtbHLH61* is related to *PtbHLH160* and *PtbHLH28* when *PtbHLH160* and *PtbHLH28* co-express and interact with each other; *PtbHLH73* interacts with *PtbHLH36* experimentally (Figure S5).

PtbHLH131, *PtbHLH57*, and *PtbHLH173* were significantly downregulated in the three states of flowering. With information obtained from the STRING database considered together, *PtbHLH57* in co-expression with BTB/POZ domain-containing family protein was thought to be involved with flowering development probably (Figure 6C). The *PtCO1* and *PtCO2* did not show any increase but even decreased. Maybe we can attribute this to a certain genetic diurnal fluctuation expression pattern or different spatial and temporal expression. In conditions of low nitrogen, early flowering phenotype happened because of the phosphorylation state of FLOWERING BHLH 4 in *Arabidopsis* [46], then we speculated that some of the *PtFBHs* exert flowering-related effects in certain conditions.

4. Discussion

We obtained 233 *bHLH* genes in the newest version v4.1 of genomic sequences in *P. trichocarpa* based on the analysis of the conserved domain. Widely distributed on 19 chromosomes of poplar (Figure 1), they were named *PtbHLH1*–*PtbHLH233* according to their accession number on the poplar genome (Table S1). The majority of *PtbHLH* proteins (89%) were predicted to function in the nucleus (Table S3). Gene duplication and synteny analysis indicated that *PtbHLH* gene family has undergone strong purification selection during the evolution process (Figure 2 and Figure S4). Based on the gene structure and protein motifs of 233 *PtbHLHs*, we classified them into 21 groups (Figure 3). Genes in the same group share similar gene and protein structure (Figure 4). *Cis*-elements analysis of the *PtbHLH* genes suggest that the *PtbHLH* gene family may play an important role in regulation growth and development (Table S7 and Figure S2).

Based on publicly available RNA-Seq data [37], we obtained differentially expressed genes (DEGs) in tissues selected from Flowers-Expanded, Roots-Control, Seeds-Mature, Leaves-Control, and Cambium-Phloem-Dormant (Figure S4). Heterologous expression of poplar *PtFBH1* (*PtbHLH173*) could lead to early flowering in *Arabidopsis* under both short and long day conditions [12], and histone demethylase JMJ28, a CO activator removes H3K9me2 from the CO locus through interacting with FBH in *Arabidopsis* [15]. Therefore, we identified 10 members of the *PtFBH* subfamily in *PtbHLHs* additionally and the results were supported by both phylogenetic trees (Figure 5). We then explored their expression patterns in three states of the flowering process (Flowers-Dormant, Flowers-Expanding, and Flowers-Expanded) together with *PtCO1*, *PtCO2*, and *PtFT* (Figure 6). Quite a few *bHLH* genes were found to be involved in flower development in Chinese jujube [47]. In total, 18 members (*PtbHLH73*, *PtbHLH187*, *PtbHLH216*, *PtbHLH93*, *PtbHLH142*, *PtbHLH101*, *PtbHLH18*, *PtbHLH104*, *PtbHLH100*, *PtbHLH161*, *PtbHLH137*, *PtbHLH116*, *PtbHLH165*, *PtbHLH181*, *PtbHLH13*, *PtbHLH114*, *PtbHLH164*, and *PtbHLH180*) may play a role in the whole process of poplar flowering due to their upregulation in all three states just as *PtFT*.

We then explored other possible characteristics of *PtFBHs* through protein–protein interaction analysis by STRING database. In tomato, TMF protein interacts with BTB/POZ domain-containing SIBOPs, homologs of BOP transcriptional cofactors in *Arabidopsis*, to promote inflorescence complexity. Elimination of SIBOP function renders tomato inflorescences into single flowers [48]. The predictable co-expression between *PtbHLH57* and BTB/POZ domain-containing family protein has indicated that *PtbHLH57* may be involved in inflorescence complexity regulation. *PtbHLH57* is homologous to *AtFBH3* (AT1G51140) (Figure 5). Overexpression of *FBH3* in *Arabidopsis* has caused early flowering by upregulating CO expression no matter which photoperiod [12]. Since we proved *FBHs* to be conserved between *Arabidopsis* and poplar here, they are likely to function similarly in flower development.

Many *PtbHLHs* work with other members within the family, such as *PtbHLH165* with *PtbHLH127* and *PtbHLH150*, *PtbHLH181* with *PtbHLH14*, and *PtbHLH61* with *PtbHLH160* and *PtbHLH28* when *PtbHLH160* and *PtbHLH28* were in co-expression and interacting

(Figure S5). Maybe *PtbHLH61*, *PtbHLH160*, and *PtbHLH28* function together for a certain biological pathway. The same speculation is on *PtbHLH73* and *PtbHLH36* for they were proved to interact experimentally. Based on our study, further research is needed to explain the detailed function of *PtbHLHs* and *PtFBHs* in *P. trichocarpa*.

5. Conclusions

In our study, 233 bHLH transcription factors of *P. trichocarpa* were subjected to identification, classification, phylogenetic analysis, characterization of conserved motif and gene structure, subcellular localization analysis, gene duplication and synteny analysis, and exploration on *cis*-elements. Then, we also analyzed the expression patterns of 233 *PtbHLH* genes in different tissues and protein interaction network. This analysis of bHLHs in poplar also reveals that some *PtbHLH* and *PtFBH* members may be involved in the flowering process. All in all, our comprehensive and systematic analysis can provide some valuable clues and basic reference toward further investigations on physiological and molecular functions of *PtbHLHs*.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/f12111439/s1>. Table S1. The information of bHLH gene family in poplar; Table S2. Physico-chemical properties of bHLH proteins in poplar; Table S3. Subcellular localization analysis of bHLH proteins in poplar; Table S4. Duplication type of the *PtbHLH* gene; Table S5. *PtbHLH* evolutionary selection pressure prediction; Table S6. Sequence information of 15 identified motifs in MEME analysis; Table S7. *Cis*-elements in members of the *PtbHLH* gene; Table S8. The primers used for qRT-PCR; Figure S1. *PtbHLH* gene collinearity relationship between poplar and Arabidopsis or rice; Figure S2. The *cis*-acting elements analysis of putative promoter of the bHLH gene family in poplar; Figure S3. Multi-sequence alignment and domain analysis of the FBHs in poplar; Figure S4. The expression pattern analysis of *PtbHLHs* genes in poplar; Figure S5. Protein interaction network of *PtFBHs* proteins in poplar.

Author Contributions: Y.Y., H.X., X.G. and J.M. performed the experiments. L.L. and Y.Y. analyzed the data, prepared figures and tables, wrote the paper and reviewed drafts of the paper. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Technology Innovation Project of Shaanxi Province (Project No. 2020QFY10-01), the National Natural Science Foundation of China (No. 31600540), the Agricultural Extension Practice for Young Talents of Northwest A&F University (Project No. TGZX2020-37), and the Key Research and Development Program of Shaanxi Province (2018ZDXM2-01).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is contained within the article or Supplementary Materials. It is also available from the correspondence author (lill@nwaf.edu.cn).

Acknowledgments: We thanks editors and three reviewers very much spending much time to give us helpful comments to improve our work. We thanks Bioeditas Technology Corporation (Shaanxi, China) for technical assistance during the data analysis.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Abdullah, M.; Cheng, X.; Cao, Y.; Su, X.; Manzoor, M.A.; Gao, J.; Cai, Y.; Lin, Y. Zinc Finger-Homeodomain Transcriptional Factors (ZHDs) in Upland Cotton (*Gossypium hirsutum*): Genome-Wide Identification and Expression Analysis in Fiber Development. *Front. Genet.* **2018**, *9*, 357. [CrossRef]
2. Aravind, L.; Anantharaman, V.; Balaji, S.; Babu, M.M.; Iyer, L.M. The many faces of the helix-turn-helix domain: Transcription regulation and beyond. *FEMS Microbiol. Rev.* **2005**, *29*, 231–262. [CrossRef]
3. An, J.-P.; Qu, F.-J.; Yao, J.-F.; Wang, X.-F.; You, C.-X.; Hao, Y.-J. The bZIP transcription factor MdHY5 regulates anthocyanin accumulation and nitrate assimilation in apple. *Hortic. Res.* **2017**, *4*, 1–9. [CrossRef]
4. Toledo-Ortiz, G.; Huq, E.; Quail, P.H. The *Arabidopsis* Basic/Helix-Loop-Helix Transcription Factor Family. *Plant Cell* **2003**, *15*, 1749–1770. [CrossRef] [PubMed]

5. Wang, R.; Zhao, P.; Kong, N.; Lu, R.; Pei, Y.; Huang, C.; Ma, H.; Chen, Q. Genome-Wide identification and characterization of the potato bhlh transcription factor family. *Genes* **2018**, *9*, 54. [\[CrossRef\]](#)
6. Wang, J.; Hu, Z.; Zhao, T.; Yang, Y.; Chen, T.; Yang, M.; Yu, W.; Zhang, B. Genome-wide analysis of bHLH transcription factor and involvement in the infection by yellow leaf curl virus in tomato (*Solanum lycopersicum*). *BMC Genom.* **2015**, *16*, 1–14. [\[CrossRef\]](#) [\[PubMed\]](#)
7. 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. [\[CrossRef\]](#) [\[PubMed\]](#)
8. Jones, S. An overview of the basic helix-loop-helix proteins. *Genome Biol.* **2004**, *5*, 226. [\[CrossRef\]](#) [\[PubMed\]](#)
9. Liu, Y.; Li, X.; Li, K.; Liu, H.; Lin, C. Multiple bHLH Proteins form Heterodimers to Mediate CRY2-Dependent Regulation of Flowering-Time in Arabidopsis. *PLoS Genet.* **2013**, *9*, e1003861. [\[CrossRef\]](#) [\[PubMed\]](#)
10. Martínez-García, J.F.; Huq, E.; Quail, P.H. direct targeting of light signals to a promoter element-bound transcription factor. *Science* **2000**, *288*, 859–863. [\[CrossRef\]](#)
11. 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 silicles*. *Plant Cell* **2000**, *12*, 1863–1878. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Ito, S.; Song, Y.H.; Josephson-Day, A.R.; Miller, R.J.; Breton, G.; Olmstead, R.G.; Imaizumi, T. Flowering bhlh transcriptional activators control expression of the photoperiodic flowering regulator CONSTANS in *Arabidopsis*. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 3582–3587. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Nagel, D.H.; Pruneda-Paz, J.L.; Kay, S.A. FBH1 affects warm temperature responses in the *Arabidopsis* circadian clock. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 14595–14600. [\[CrossRef\]](#)
14. Yin, J.; Chang, X.; Kasuga, T.; Bui, M.; Reid, M.S.; Jiang, C.-Z. A basic helix-loop-helix transcription factor, PhFBH4, regulates flower senescence by modulating ethylene biosynthesis pathway in petunia. *Hortic. Res.* **2015**, *2*, 15059. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Hung, F.-Y.; Lai, Y.-C.; Wang, J.; Feng, Y.-R.; Shih, Y.-H.; Chen, J.-H.; Sun, H.-C.; Yang, S.; Li, C.; Wu, K. The *Arabidopsis* histone demethylase JMJ28 regulates CONSTANS by interacting with FBH transcription factors. *Plant Cell* **2021**, *33*, 1196–1211. [\[CrossRef\]](#)
16. Tuskan, G.A.; DiFazio, S.; Jansson, S.; Bohlmann, J.; Grigoriev, I.; Hellsten, U.; Putnam, N.; Ralph, S.; Rombauts, S.; Salamov, A.; et al. The Genome of Black Cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* **2006**, *313*, 1596–1604. [\[CrossRef\]](#)
17. Kumar, M.; Kherawat, B.; Dey, P.; Saha, D.; Singh, A.; Bhatia, S.; Ghodake, G.; Kadam, A.; Kim, H.-U.; Chung, S.M.; et al. Genome-wide identification and characterization of Pin-Formed (PIN) gene family reveals role in developmental and various stress conditions in *Triticum aestivum* L. *Int. J. Mol. Sci.* **2021**, *22*, 7396. [\[CrossRef\]](#)
18. Kesawat, M.S.; Kherawat, B.S.; Singh, A.; Dey, P.; Kabi, M.; Debnath, D.; Saha, D.; Khandual, A.; Rout, S.; Chung, S.M.; et al. Genome-Wide Identification and Characterization of the Brassinazole-resistant (BZR) gene family and its expression in the various developmental stage and stress conditions in wheat (*Triticum aestivum* L.). *Int. J. Mol. Sci.* **2021**, *22*, 8743. [\[CrossRef\]](#)
19. Tong, T.; Fang, Y.X.; Zhang, Z.; Zheng, J.; Zhang, X.; Li, J.; Niu, C.; Xue, D.; Zhang, X. Genome-wide identification and ex-pression pattern analysis of the KCS gene family in barley. *Plant Growth Regul.* **2021**, *93*, 1–15. [\[CrossRef\]](#)
20. Jiang, M.; Chen, H.; Liu, J.; Du, Q.; Lu, S.; Liu, C. Genome-wide identification and functional characterization of natural an-tisense transcripts in *Salvia miltiorrhiza*. *Sci. Rep.* **2021**, *11*, 1–14.
21. Baruah, P.M.; Krishnatreya, D.B.; Bordoloi, K.S.; Gill, S.S.; Agarwala, N. Genome wide identification and characterization of abiotic stress responsive lncRNAs in *Capsicum annuum*. *Plant Physiol. Biochem.* **2021**, *162*, 221–236. [\[CrossRef\]](#)
22. Wang, M.; Chen, B.; Zhou, W.; Xie, L.; Wang, L.; Zhang, Y.; Zhang, Q. Genome-wide identification and expression analysis of the AT-hook Motif Nuclear Localized gene family in soybean. *BMC Genom.* **2021**, *22*, 1–26. [\[CrossRef\]](#)
23. Zhao, K.; Li, S.; Yao, W.; Zhou, B.; Li, R.; Jiang, T. Characterization of the basic helix-loop-helix gene family and its tis-sue-differential expression in response to salt stress in poplar. *PeerJ* **2018**, *6*, e4502. [\[CrossRef\]](#) [\[PubMed\]](#)
24. Potter, S.C.; Luciani, A.; Eddy, S.R.; Park, Y.; López, R.; Finn, R.D. HMMER web server: 2018 update. *Nucleic Acids Res.* **2018**, *46*, W200–W204. [\[CrossRef\]](#) [\[PubMed\]](#)
25. Kumar, S.; Stecher, G.; Tamura, K. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* **2016**, *33*, 1870–1874. [\[CrossRef\]](#)
26. Edgar, R.C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **2004**, *32*, 1792–1797. [\[CrossRef\]](#)
27. Steenwyk, J.L.; Buida, T.J., 3rd; Li, Y.; Shen, X.X.; Rokas, A. ClipKIT: A multiple sequence alignment trimming software for accurate phylogenomic inference. *PLoS Biol.* **2020**, *18*, e3001007. [\[CrossRef\]](#)
28. Chen, C.; Chen, H.; Zhang, Y.; Thomas, H.R.; Frank, M.H.; He, Y.; Xia, R. TBtools—An integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant* **2020**, *13*, 1194–1202. [\[CrossRef\]](#) [\[PubMed\]](#)
29. Wang, Y.; Tang, H.; DeBarry, J.; Tan, X.; Li, J.; Wang, X.; Lee, T.-H.; Jin, H.; Marler, B.; Guo, H.; et al. MCScanX: A toolkit for detection and evolutionary analysis of gene syteny and collinearity. *Nucleic Acids Res.* **2012**, *40*, e49. [\[CrossRef\]](#) [\[PubMed\]](#)
30. Hall, T.A. BioEdit: A User-Friendly Biological Sequence Alignment Editor and Analysis Program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* **1999**, *41*, 95–98.
31. Lynch, M.; Conery, J.S. The evolutionary fate and consequences of duplicate genes. *Science* **2000**, *290*, 1151–1155. [\[CrossRef\]](#)
32. Artimo, P.; Jonnalagedda, M.; Arnold, K.; Baratin, D.; Csardi, G.; de Castro, E.; Duvaud, S.; Flegel, V.; Fortier, A.; Gasteiger, E.; et al. ExPASy: SIB bioinformatics resource portal. *Nucleic Acids Res.* **2012**, *40*, W597–W603. [\[CrossRef\]](#) [\[PubMed\]](#)

33. Bailey, T.L.; Williams, N.; Misleh, C.; Li, W.W. MEME: Discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Res.* **2006**, *34*, W369–W373. [[CrossRef](#)]
34. Lescot, M.; Déhais, P.; Thijs, G.; Marchal, K.; Moreau, Y.; Van de Peer, Y.; Rouzé, P.; Rombauts, S. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res.* **2002**, *30*, 325–327. [[CrossRef](#)] [[PubMed](#)]
35. Katoh, K.; Rozewicki, J.; Yamada, K.D. MAFFT online service: Multiple sequence alignment, interactive sequence choice and visualization. *Brief. Bioinform.* **2019**, *20*, 1160–1166. [[CrossRef](#)]
36. Clamp, M.; Cuff, J.; Searle, S.M.; Barton, G. The Jalview Java alignment editor. *Bioinformatics* **2004**, *20*, 426–427. [[CrossRef](#)] [[PubMed](#)]
37. Sundell, D.; Mannapperuma, C.; Netotea, S.; Delhomme, N.; Lin, Y.C.; Sjödin, A.; Van de Peer, Y.; Jansson, S.; Hvidsten, T.R.; Street, N.R. The Plant Genome Integrative Explorer Resource: PlantGenIE.org. *New Phytol.* **2015**, *208*, 1149–1156. [[CrossRef](#)]
38. Bailey, P.C.; Martin, C.; Toledo-Ortiz, G.; Quail, P.H.; Huq, E.; Heim, M.A.; Jakoby, M.; Werber, M.; Weisshaar, B. Update on the Basic Helix-Loop-Helix Transcription Factor Gene Family in *Arabidopsis thaliana*. *Plant Cell* **2003**, *15*, 2497–2502. [[CrossRef](#)] [[PubMed](#)]
39. 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](#)]
40. 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.
41. 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](#)] [[PubMed](#)]
42. Wang, P.F.; Su, L.; Gao, H.H.; Jiang, X.L.; Wu, X.Y.; Li, Y.; Zhang, Q.Q.; Wang, Y.M.; Ren, F.S. 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](#)] [[PubMed](#)]
43. Qiao, X.; Li, M.; Li, L.; Yin, H.; Wu, J.; Zhang, S. Genome-wide identification and comparative analysis of the heat shock transcription factor family in Chinese white pear (*Pyrus bretschneideri*) and five other Rosaceae species. *BMC Plant Biol.* **2015**, *15*, 1–16. [[CrossRef](#)] [[PubMed](#)]
44. Yang, Z. PAML 4: Phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **2007**, *24*, 1586–1591. [[CrossRef](#)]
45. 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](#)] [[PubMed](#)]
46. Sanagi, M.; Aoyama, S.; Kubo, A.; Lu, Y.; Sato, Y.; Ito, S.; Abe, M.; Mitsuda, N.; Ohme-Takagi, M.; Kiba, T.; et al. Low nitrogen conditions accelerate flowering by modulating the phosphorylation state of flowering BHLH 4 in *Arabidopsis*. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2022942118. [[CrossRef](#)]
47. Li, H.; Gao, W.; Xue, C.; Zhang, Y.; Liu, Z.; Zhang, Y.; Meng, X.; Liu, M.; Zhao, J. Genome-wide analysis of the bHLH gene family in Chinese jujube (*Ziziphus jujuba* Mill.) and wild jujube. *BMC Genom.* **2019**, *20*, 1–13. [[CrossRef](#)]
48. Xu, C.; Park, S.J.; Van Eck, J.; Lippman, Z.B. Control of inflorescence architecture in tomato by BTB/POZ transcriptional regulators. *Genes Dev.* **2016**, *30*, 2048–2061. [[CrossRef](#)]