

RESEARCH

Open Access



Genome-wide characterization of the *bHLH* gene family in *Gynostemma pentaphyllum* reveals its potential role in the regulation of gypenoside biosynthesis

Yanhong Qin^{1,2†}, Jinmei Li^{1,2†}, Jianhua Chen^{1,2}, Shaochang Yao^{1,2}, Liangbo Li^{1,2}, Rongshao Huang^{1,2}, Yong Tan^{1,2}, Ruhong Ming^{1,2*}  and Ding Huang^{1,2*} 

Abstract

Background *Gynostemma pentaphyllum*, an ancient Chinese herbal medicine, serves as a natural source of gypenosides with significant medicinal properties. Basic helix-loop-helix (bHLH) transcription factors play pivotal roles in numerous biological processes, especially in the regulation of secondary metabolism in plants. However, the characteristics and functions of the *bHLH* genes in *G. pentaphyllum* remain unexplored, and their regulatory role in gypenoside biosynthesis remains poorly elucidated.

Results This study identified a total of 111 *bHLH* members in *G. pentaphyllum* (GpbHLHs), categorizing them into 26 subgroups based on shared conserved motif compositions and gene structures. Collinearity analysis illustrated that segmental duplications predominately lead to the evolution of GpbHLHs, with most duplicated *GpbHLH* gene pairs undergoing purifying selection. Among the nine gypenoside-related *GpbHLH* genes, two GpbHLHs (*GpbHLH15* and *GpbHLH58*) were selected for further investigation based on co-expression analysis and functional prediction. The expression of these two selected GpbHLHs was dramatically induced by methyl jasmonate, and their nuclear localization was confirmed. Furthermore, yeast one-hybrid and dual-luciferase assays demonstrated that GpbHLH15 and GpbHLH58 could bind to the promoters of the gypenoside biosynthesis pathway genes, such as *GpFPS1*, *GpSS1*, and *GpOSC1*, and activate their promoter activity to varying degrees.

Conclusions In conclusion, our findings provide a detailed analysis of the *bHLH* family and valuable insights into the potential use of GpbHLHs to enhance the accumulation of gypenosides in *G. pentaphyllum*.

Keywords *Gynostemma pentaphyllum*, Basic helix-loop-helix, Genome-wide characterization, Transcription factor, Gypenoside biosynthesis

[†]Yanhong Qin and Jinmei Li contributed equally to this work.

*Correspondence:

Ruhong Ming
mingrh@gxctcmu.edu.cn
Ding Huang
hdh016@126.com

Full list of author information is available at the end of the article



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Introduction

The basic helix-loop-helix (bHLH) transcription factors (TFs) constitute one of the largest and most important families in eukaryotes, distinguished by their highly conserved bHLH domain [1]. The bHLH domain encompasses approximately 60 amino acids, which are constituted by two functionally distinct regions [2]. The N-terminal basic amino acid region, which varies from 10 to 15 amino acids in length and is essential for binding to the E-box (CANNTG) motif. Conversely, the C-terminal helix-loop-helix (HLH) structural domain, composed of two amphipathic alpha helices linked by a variable-length loop of approximately 40–50 amino acid residues, plays a pivotal role in facilitating homodimeric or heterodimeric complex formations and promoting protein–protein interactions [3, 4]. Additionally, plant MYC-like bHLH proteins contain an additional N-terminal MYC domain and a MYB interacting region, facilitating interactions with bHLH and R2R3-MYB domain proteins, respectively [5, 6].

Extensive research on bHLH TFs in plants has revealed their critical roles in numerous essential physiological processes. These include the transcriptional regulation of key developmental processes, such as stomata differentiation [7], seed coat differentiation [8], trichome/root hair formation [9], fruit dehiscence, and carpel margin development [10]. Additionally, some bHLH TFs are involved in seedling photomorphogenesis [11]. Furthermore, accumulating evidence suggests that bHLH proteins participate in the response to various abiotic stresses, enhancing plant stress tolerance, including tolerance to drought, salinity, low temperature, and mechanical damage [12–15]. Moreover, the regulatory roles of bHLH proteins in the secondary metabolites biosynthesis, such as phenolic acids [16], flavonoids [17, 18], alkaloids [19], and terpenoids [20, 21], have been well characterized in plants.

Gynostemma pentaphyllum (Thunb.) Makino, an herbaceous perennial plant of the Cucurbitaceae family, has a rich history in traditional Chinese medicine and is widely consumed as an herbal tea in Asian countries for its potential to prevent obesity [22]. The leaves of *G. pentaphyllum* plants, particularly the tender leaves, are the primary site for the synthesis of gypenosides, a class of dammarane-type triterpenoid saponins [23]. Gypenosides derived from *G. pentaphyllum* offer a wide range of health benefits, such as anti-cancer, anti-obesity, anti-inflammatory, and sedative-hypnotic effects [24–26]. The gypenoside content in *G. pentaphyllum* plants is a primary quality trait, making it imperative to understand the mechanism underlying gypenoside formation for developing high-yield gypenoside-producing plants. The gypenoside biosynthesis pathway has been extensively

characterized, with four key enzyme-encoding genes, namely, *farnesyl diphosphate synthase* (*GpFPS1*), *squalene synthase* (*GpSS1*), *squalene epoxidase* (*GpSE2*), and *2,3-oxidosqualene cyclase* (*GpOSC1*), playing pivotal roles [23]. Moreover, the overexpression or interference with any of these genes (*FPS*, *SS*, *SE*, and *OSC*) has been demonstrated to significantly impact the production of dammarane-type triterpenoid saponins [27–30].

Given that the overexpression of bHLH TFs has been demonstrated to effectively regulate the selective biosynthesis of specific secondary metabolism in plants [31], investigating the regulatory roles of *bHLH* genes in gypenoside biosynthesis offers promising avenues for increasing gypenosides content in *G. pentaphyllum*. In this study, we identified 111 *GpbHLH* genes using a publicly available genomic database and carried out an exhaustive analysis of the bHLH family in *G. pentaphyllum*. Our analysis encompassed protein properties, phylogenetic relationships, gene structure, conserved domains, motifs composition, *cis*-acting elements, and *bHLH* expression patterns. The comprehensive analysis identified two bHLH proteins (*GpbHLH15/58*) as potential regulators of gypenoside biosynthesis. Their expression levels increased significantly under methyl jasmonate (MeJA) treatment, localizing primarily within the nucleus. Yeast one-hybrid (Y1H) and dual-luciferase assays demonstrated that *GpbHLH15/58* could bind to the promoters of key enzyme-encoding genes (*GpFPS1*, *GpSS1*, and *GpOSC1*) in gypenosides biosynthetic pathway, and activate their expression to varying degrees. In summary, our findings expand on the understanding of the *bHLH* gene family in *G. pentaphyllum* and shed light on the underlying molecular mechanisms of gypenoside regulation, offering promising strategies for the improvement of gypenosides content.

Results

Identification and characterization of bHLH TFs in *G. pentaphyllum* and their evolutionary relationship

To comprehensively identify bHLH members in *G. pentaphyllum* at the genome-wide level, we employed *Arabidopsis* bHLH proteins as queries to search the *G. pentaphyllum* protein database. Additionally, we conducted a Hidden Markov Model search using the bHLH domain to globally search the *G. pentaphyllum* genome. Subsequently, the presence of the bHLH conserved domain was verified, and protein sequences lacking the bHLH domain were excluded. After removing the redundant sequences, we identified a total of 111 putative *GpbHLH* genes. For consistency, these *bHLH* genes were renamed *GpbHLH1–GpbHLH111* based on their chromosomal locations. Detailed information about these 111 *GpbHLH* genes, including gene name, gene

ID, protein length, theoretical isoelectric point (pI), molecular weight (MW), and subcellular localization, is presented in Table S1. The proteins encoded by the 111 *GpbHLH* genes ranged in length from 91 to 1077 amino acids, with PI values ranging from 4.76 (*GpbHLH83*) to 10.57 (*GpbHLH74*) and MW values spanning from 10.18 (*GpbHLH18*) to 119.76 (*GpbHLH38*) kDa. Predictions using the Cell-PLoc indicated that the majority of *GpbHLH* proteins were localized in the nucleus, although a few were predicted to be extracellular, suggesting potential diverse tissue-specific expressions and functional roles among these *GpbHLH* genes.

To elucidate the evolutionary relationships, we constructed an unrooted maximum likelihood (ML) phylogenetic tree based on bHLH protein sequences from *A. thaliana* and *G. pentaphyllum* (Fig. 1). The phylogenetic analysis clustered the *bHLH* genes into 26 well-conserved subgroups following the taxonomy proposed for *A. thaliana*. Among the 111 *GpbHLH* proteins, subgroup XII was the most abundant with 15 bHLH members, followed by subgroup Ia with nine members. Contrastingly,

subgroups II and XIII contained the fewest *GpbHLH* proteins, each with only one member. Additionally, the number of genes in the specific subfamilies differed between *A. thaliana* and *G. pentaphyllum*. For example, subgroup XIII contained 17 *AtbHLHs* in *A. thaliana* but only one *GpbHLH* in the same subgroup, highlighting significant interspecific divergences with the *bHLH* gene family between the two species.

Conserved domains, motif and gene structure analysis of the *GpbHLH* gene family

To validate the findings from the phylogenetic analysis (Fig. 2A), we analyzed the conserved domains within the *GpbHLH* proteins. As shown in Fig. 2B, all *GpbHLH* proteins contained the bHLH domain. Notably, members of subgroups IIIa, IIIc, IIIe, and IIIf possessed the MYC domain, in contrast to subgroups IIIb and IIIc. This observation is consistent with the notion that MYC-like bHLH TFs are distributed in the bHLH subfamily III in *A. thaliana* [32].

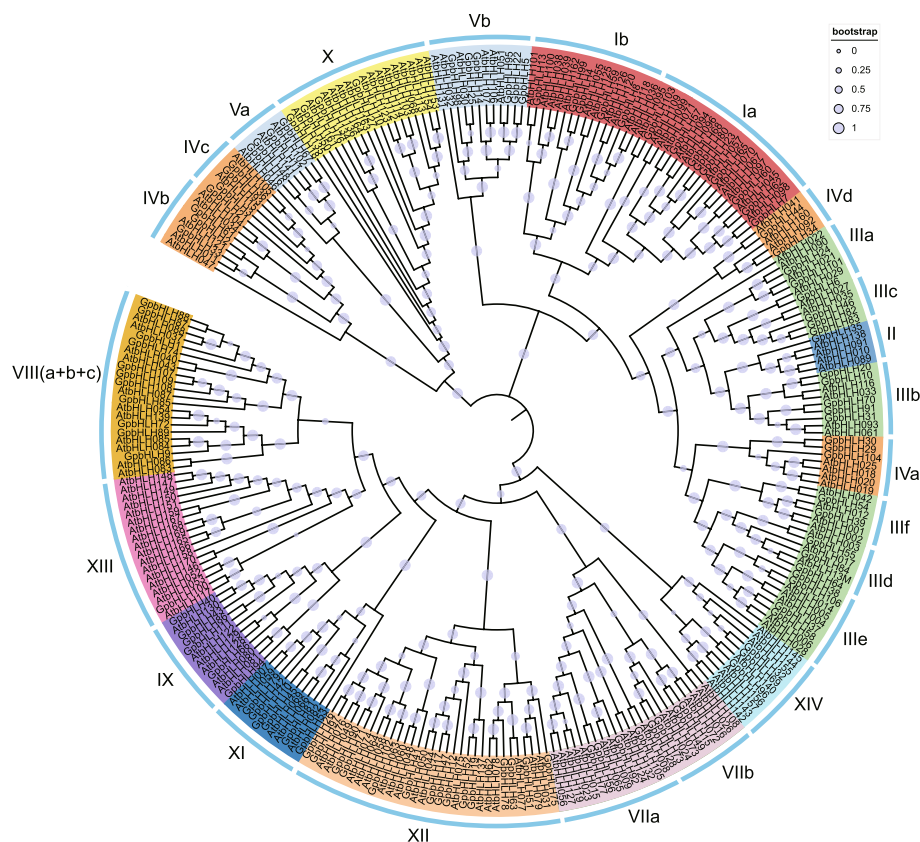


Fig. 1 Phylogenetic relationship of bHLH proteins between *Gynostemma pentaphyllum* and *Arabidopsis thaliana*. The unrooted evolutionary tree was built by using the FastTree software with the maximum likelihood (ML) method. The reliability of each node was determined using the Shimodaira–Hasegawa test, which is highly correlated with commonly used bootstrap values. The 111 *GpbHLH* proteins are classified into 26 subfamilies, in line with the classification method of the 167 *AtbHLH* proteins, represented by different colored subgroups within the tree. Bootstrap values are visualized by a circle at each branch, with those below 50% not displayed in the figure

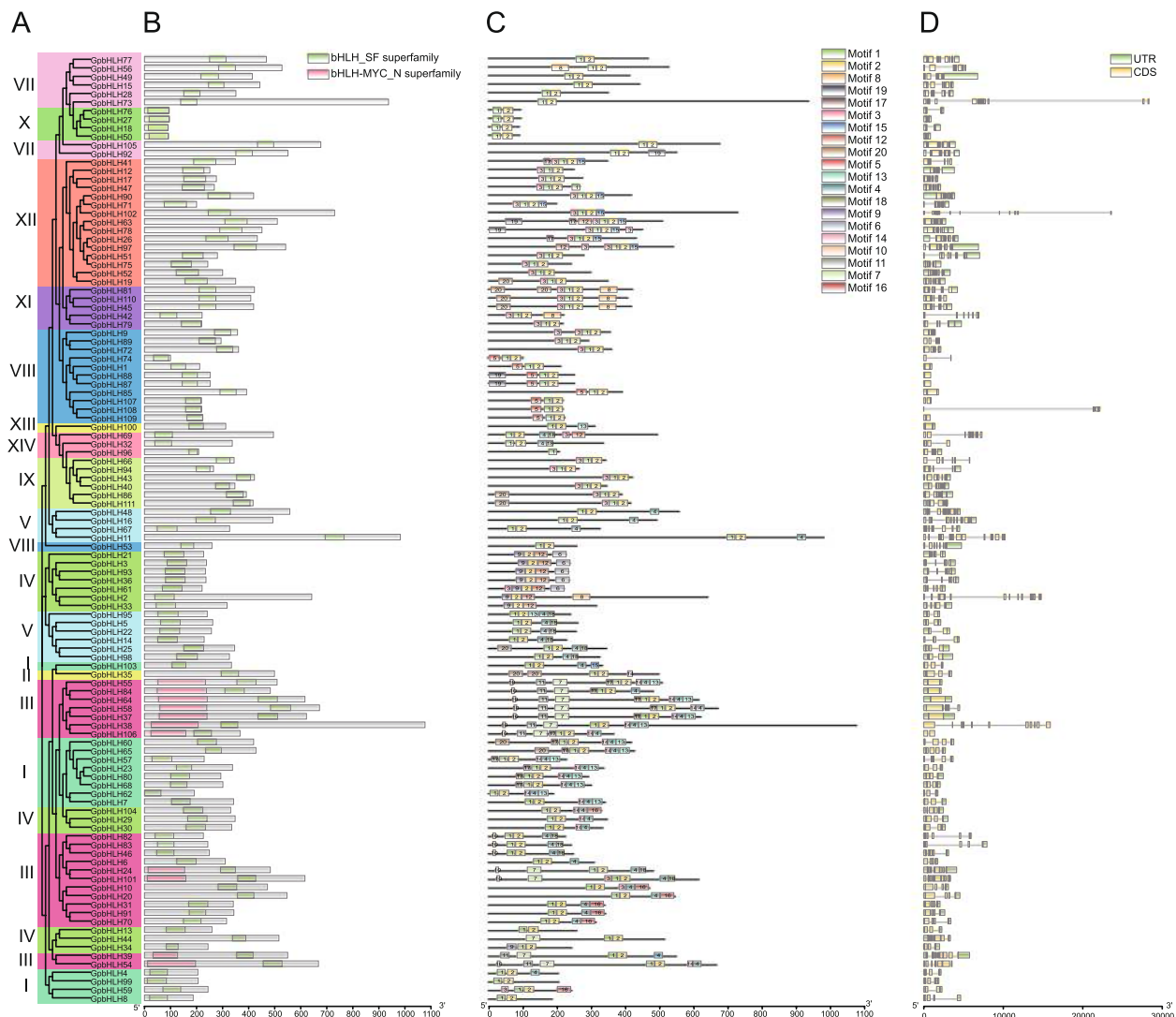


Fig. 2 Analysis of conserved motifs, protein domains, and GpbHLH gene structures based on the phylogenetic relationships. **A** The phylogenetic tree of 111 GpbHLHs. **B** Conserved domains were identified in GpbHLH proteins, and their distributions were visualized using the TB tools software. **C** The conserved motifs distribution of GpbHLH proteins. The best matched sequences of each conserved motifs were provided in Supplementary Table S3. **D** Gene structure of *GpbHLH* genes. Green rectangles and yellow rectangles indicate the untranslated regions (UTRs) and exons, respectively. Introns are presented by black lines. The scale bar is indicated at the bottom of the diagram

To gain insight into the sequence characteristics of GpbHLH proteins in *G. pentaphyllum*, we performed motif analysis using MEME, which identified a total of 20 conserved motifs (Table S2). GpbHLH proteins within the same phylogenetic group shared similar motif compositions and arrangements (Fig. 2C). For instance, subgroup IV was characterized by the presence of motifs 6, 9, and 12, while subgroup II contained two copies of motif 7, and subgroups IIId/e exhibited motifs 7, 10, 11, and 17. The presence of these specific motifs in particular subgroups underscores the relative conservation of motif compositions and arrangements within the same

subgroup. Notably, motifs 1 and 2 were associated with the bHLH domain whereas motifs 7, 10, and 11 could form MYC structural domains individually (Fig. 2C).

To explore the exon/intron structure of the 111 *GpbHLH* genes, coding sequence (CDS)/intron patterns were visualized using feature structure of a generic file format (GFF) file (Fig. 2D). *GpbHLH* genes exhibited variable exon numbers, ranging from 1 to 15. Additionally, *GpbHLH* genes in the same subgroup presented highly similar gene structures but different among different subfamilies. For example, subgroup I contained *GpbHLH* genes with two or three exons, whereas subgroup X

harbored genes with two exons. Consequently, a majority of the *GpbHLH* genes in the same subgroup exhibited similar motif compositions, gene structures, and conserved domain distribution, which further validated the evolutionary relationship analysis results.

Chromosomal distribution, gene duplication, and synteny analysis of *GpbHLH* genes

An analysis of the genomic locations revealed that the 111 *GpbHLH* genes were widely distributed across 11 chromosomes of *G. pentaphyllum* (Fig. 3A). Interestingly, the number of *GpbHLH* genes on each chromosome showed no correlation with chromosome length. Specifically, chromosome 7 hosted the largest number of genes,

with 20 genes (*GpbHLH60–GpbHLH79*), followed by chromosome 6 (18 *GpbHLH* genes). In contrast, chromosome 10 had only two *GpbHLH* genes.

Gene duplication events are recognized as major drivers of genetic diversity in higher plants, contributing to species evolution. To explore whether gene duplication events contributed to the evolution of the *GpbHLH* gene family, a synteny analysis was performed (Fig. 3A). Notably, only six tandem duplication clusters were identified on the *G. pentaphyllum* chromosome, whereas 31 pairs of duplicated segments were identified within the *GpbHLH* gene family (Table S3), suggesting that segmental duplication events likely played a significant role in expanding the *GpbHLH* gene family. Additionally, the

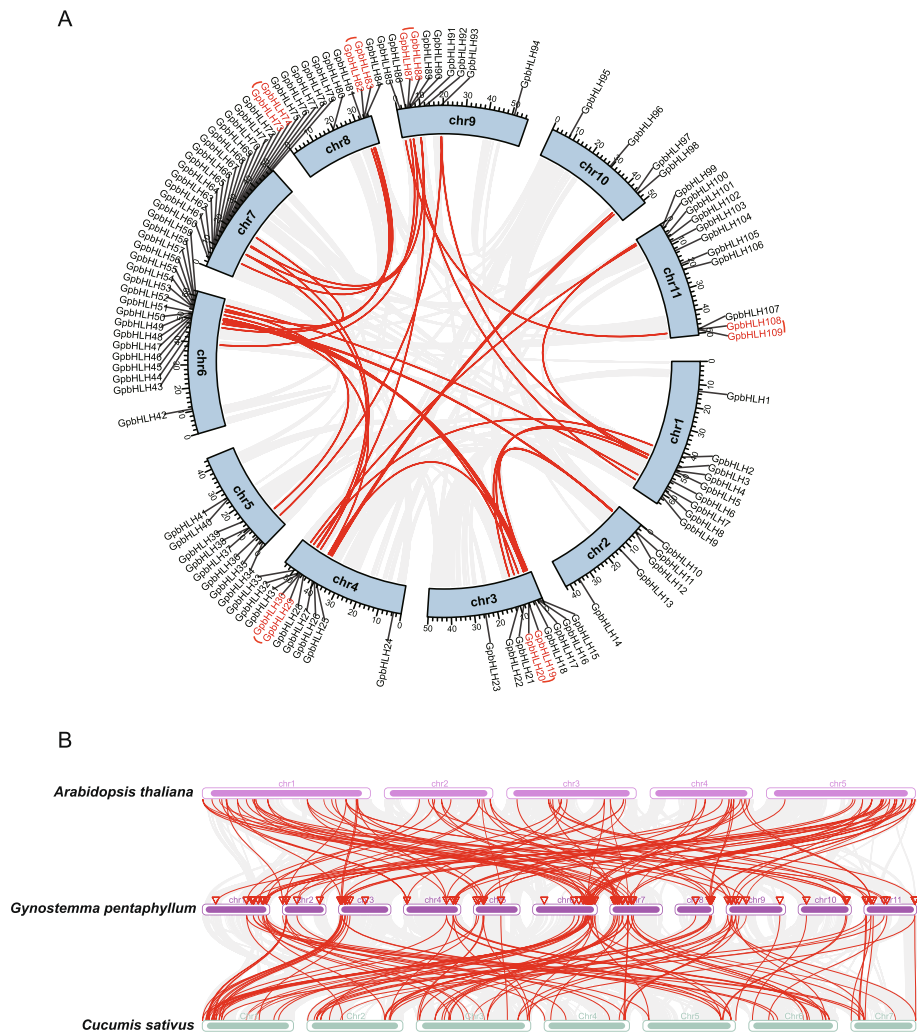


Fig. 3 Distribution and collinearity analysis of *GpbHLH* genes. **A** Chromosomal locations and the intra- and inter-species collinear analysis of *bHLH* genes. The red lines inside the circle indicate segmentally duplicated *GpbHLH* gene pairs, while the tandem duplication gene pairs are labeled in red. **B** Inter-species collinear analysis of *bHLH* genes among *G. pentaphyllum*, *A. thaliana*, and *C. sativus*. Gray lines represent all collinear blocks, whereas syntenic *bHLH* gene pairs are highlighted as red lines

non-synonymous (Ka) and synonymous (Ks) substitution ratios for the duplicated gene pairs were calculated to assess the selective pressure during evolution (Table S3). All the segmentally and tandemly duplicated *GpbHLH* gene pairs exhibited Ka/Ks values < 1, indicating that these *GpbHLH* genes primarily underwent purifying selection.

Furthermore, we conducted interspecific collinearity analysis to identify orthologous *bHLH* genes among different species (Fig. 3B). It was observed that 99 *GpbHLH* genes exhibited collinearity relationships with 87 *bHLH* genes in *A. thaliana* and 96 *bHLH* genes in *Cucumis sativus*. The numbers of orthologous pairs between *G. pentaphyllum* and *A. thaliana*, and *G. pentaphyllum* and *C. sativus* were 128 and 151, respectively (Table S4). These results suggest that *GpbHLH* genes likely share functional similarities with *bHLH* genes in *C. sativus*.

Gypenoside-related *GpbHLH* genes and their *cis*-element analysis

Gypenosides are synthesized in the leaves of *G. pentaphyllum*, especially in tender leaves [33, 34]. In this study, based on gene expression profiling, the expression patterns of gypenoside biosynthesis pathway genes (*GpFPS1*, *GpSS1*, *GpSE2*, and *GpOSC1*) were highly expressed in the tender leaf (Table S5), suggesting that tender leaf was the main tissue for biosynthesis of gypenosides skeletons in *G. pentaphyllum*. To identify *bHLH* TFs that play roles in regulating tender leaf-specific gypenoside

biosynthesis, weighted gene co-expression network analysis (WGCNA) was conducted on seven different tissues (Fig. 4A). A total of 16 modules with similar expression patterns were identified (Fig. 4B). Furthermore, the MEred module shown positive correlation with tender leaf ($r > 0.75$), followed by mature leaf ($r > 0.45$), was regarded as a gypenoside-related module. In the gypenoside-related module, a total of 1774 genes including four gypenoside biosynthesis pathway genes (*GpFPS1*, *GpSS1*, *GpSE2*, and *GpOSC1*). In addition, nine *GpbHLH* genes were also identified in the gypenoside-related module. The FPKM value and WGCNA module of each gene is provided in Table S5. These findings suggest that the nine *GpbHLH* genes may be involved in the regulation of gypenoside biosynthesis.

To analyze the potential *cis*-elements of the nine gypenoside-related *GpbHLH* genes, we extracted 2000 bp upstream promoter region sequences and employed the PlantCARE online tool for *cis*-acting element prediction (Fig. 5A). In total, 152 *cis*-acting elements were identified and categorized into three groups based on their functions: light signal responses (49 elements), phytohormone responses (70 elements), and stress responses (33 elements) (Table S6). In the light signal responses group, the most prevalent *cis*-element was box4, accounting for 25.92%, followed by G-box at 12.96% (Fig. 5B). The *cis*-elements were associated with responses to various plant hormones, such as abscisic acid (ABA), jasmonic acid, auxin,

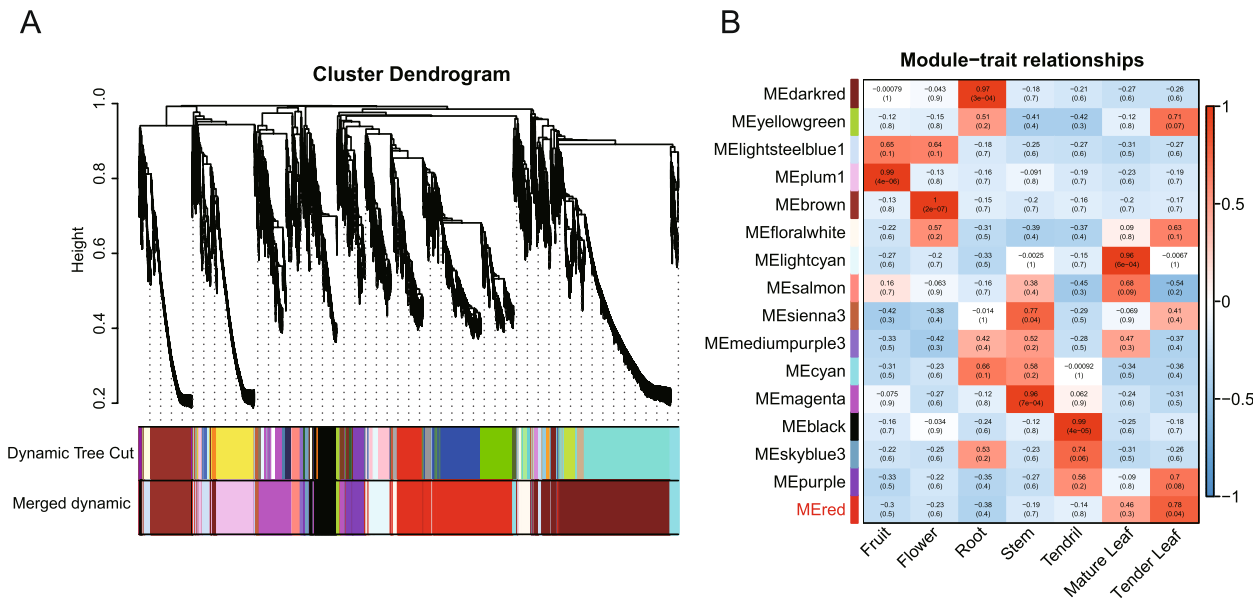


Fig. 4 Co-expression network showing gypenoside-related module identified by WGCNA. **A** Hierarchical clustering tree (cluster dendrogram) showing 16 modules of co-expressed genes with color annotation. **B** Correlation between each module and genes expressed in different tissues. Pearson correlation coefficients and *p* values (in brackets) are indicated in the grid. A low to high degree of Pearson correlation coefficient between a specific module and the tissue is indicated by a change in color from blue to red

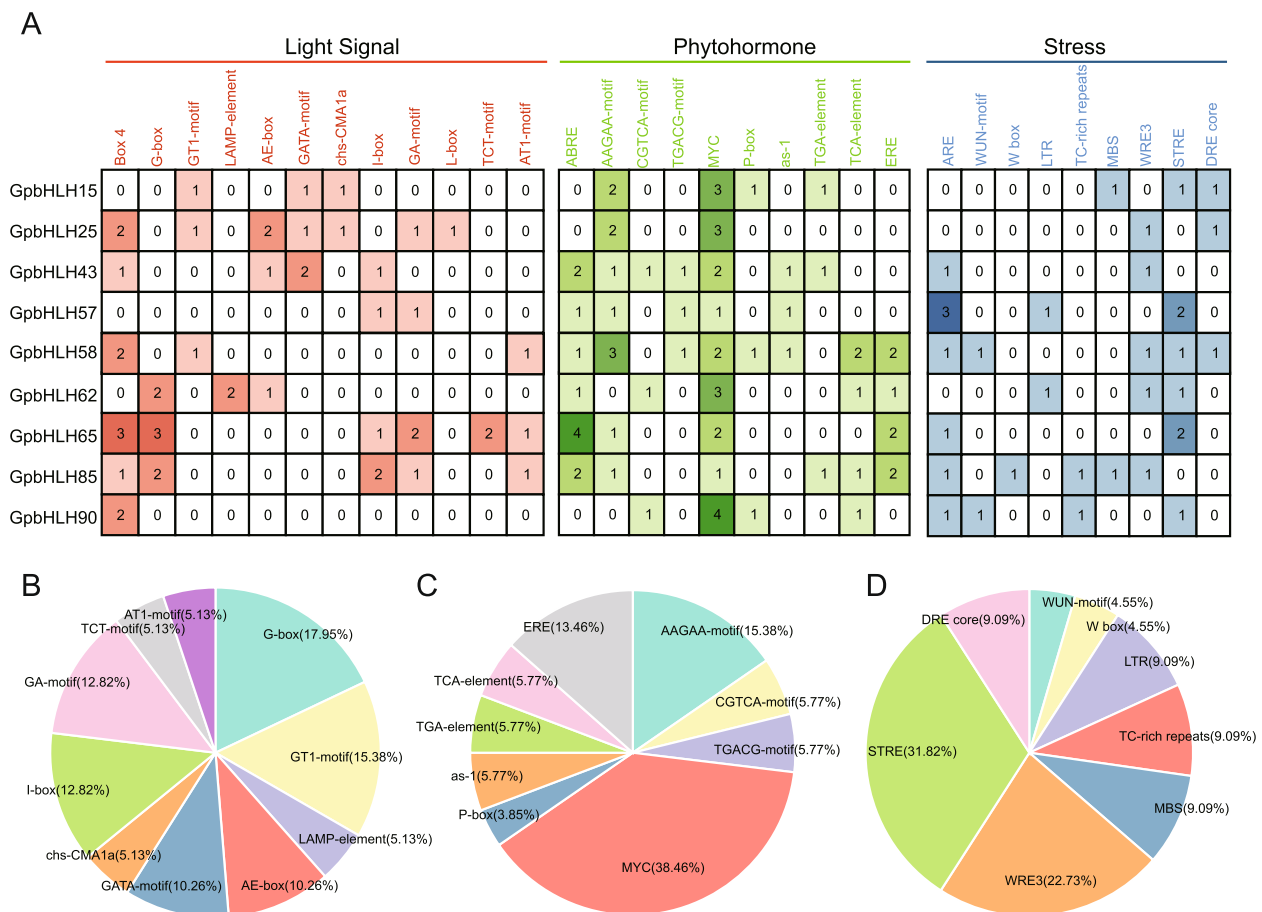


Fig. 5 Promoter *cis*-elements analysis of the nine gypenoside-related *GpbHLH* genes. **A** The identified *cis*-acting elements were divided into three functional groups and represented by red, green and blue shades of colors. The numbers indicate the quantity of corresponding *cis*-acting elements in each *GpbHLH* gene promoter region. Distribution of *cis*-acting elements with different categories within each group of *cis*-acting elements with different specific functions, including light signal (**B**), hormone response (**C**), and stress response (**D**)

gibberellin, salicylic acid, and ethylene. Among these, CGTCA-motif, TGACG-motif, and MYC-motif were the most widely distributed MeJA-responsive elements (40.62%), while ABRE (ABA-responsive element) and AAGAA-motif represented ABA-responsive elements at 31.25% (Fig. 5C). Among the *cis*-elements involved in stress responses, some functioned in response to specific stress conditions. For instance, anaerobic induced response element (ARE) was essential for anaerobic induction, accounting for 26.67%. Additionally, wound responsiveness element (WRE3), low-temperature responsiveness (LTR), and dehydration responsive element (DRE) were also identified as stress responses (Fig. 5D). In summary, our findings suggest that the nine gypenoside-related *GpbHLH* genes play important roles in various biological processes, especially light signal response, hormone response, and stress response.

Prediction of *GpbHLH*s involved in the regulation of gypenoside biosynthesis

Gypenoside biosynthesis is speculated to be strongly induced by MeJA [35]. Moreover, the analysis of the nine gypenoside-related *GpbHLH* gene promoters revealed the presence of MeJA-responsive *cis*-elements, suggesting that these *GpbHLH* genes can be induced by MeJA. To verify this hypothesis, we measured the transcript levels of the four gypenoside biosynthesis pathway genes (*GpFPS1*, *GpSS1*, *GpSE2*, and *GpOSC1*) and the nine candidate *GpbHLH* genes in MeJA-treated leaves of *G. pentaphyllum*. As shown in Fig. 6A and B, *GpFPS1*, *GpSS1*, *GpSE2*, *GpOSC1*, *GpbHLH15*, *GpbHLH58*, *GpbHLH65*, and *GpbHLH85* exhibited strong induction by MeJA compared to the control. In addition, Pearson correlation analysis revealed strong associations between the four gypenoside biosynthesis pathway genes and two *GpbHLH*s (*GpbHLH15/58*) in response to MeJA (Fig. S1). Based on previous WGCNA results and the response of

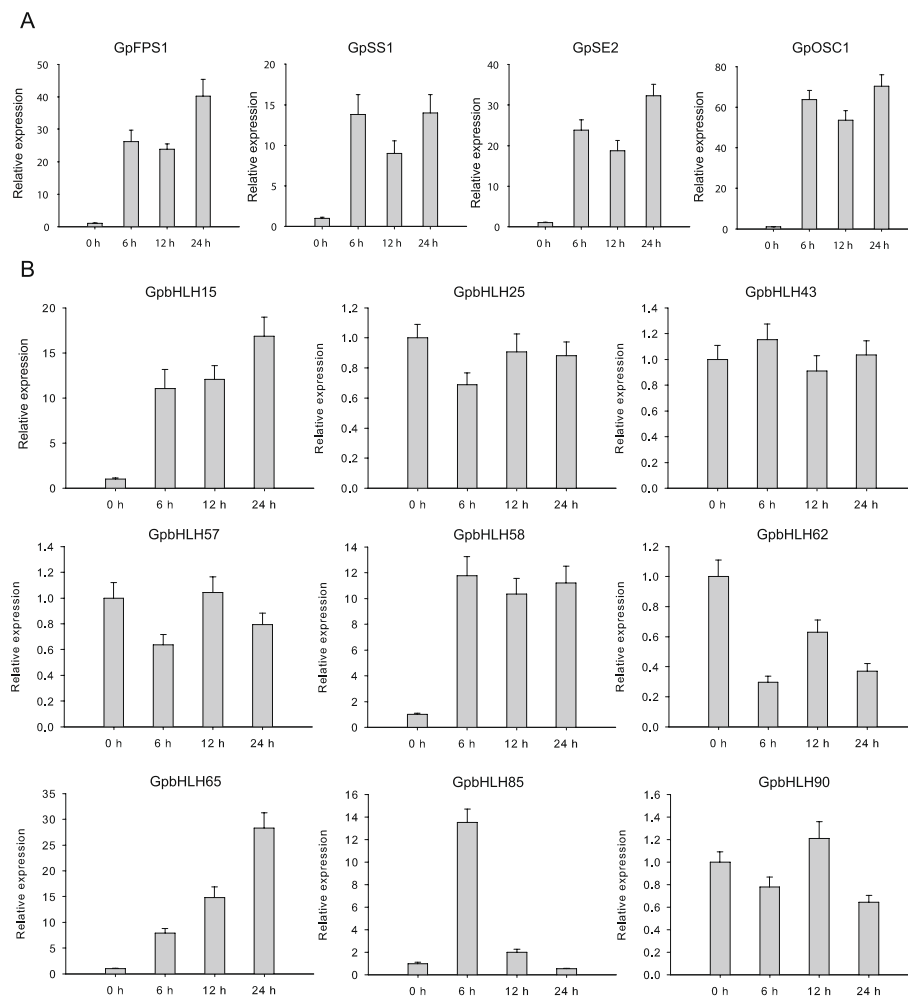


Fig. 6 Expression patterns of gypenoside biosynthesis pathway genes and nine gypenoside-related *GpbHLH* genes under MeJA treatment. **A** The relative expression abundances of the four gypenoside biosynthesis pathway genes (*GpFPS1*, *GpSS1*, *GpSE2*, and *GpOSC1*) were determined using qRT-PCR in response to MeJA treatments. **B** The relative expression abundances of the nine gypenoside-related *GpbHLH* genes were determined using qRT-PCR in response to MeJA treatments. Values represent the mean $\pm$ SE ($n = 3$ biological replicates)

GpbHLH genes to MeJA, we speculate that *GpbHLH15*, *GpbHLH58* are likely involved in the regulation of MeJA-mediated gypenoside biosynthesis.

To evaluate the potential functions of these candidate *GpbHLHs*, we used the functional characteristics of homologous *AtbHLHs* in the same subgroup for prediction. In this study, *GpbHLH15* showed a high degree of homology to *Arabidopsis* PIF7 (*AtbHLH072*), which is known to serve as a regulatory hub that relays environmental signals to the plant transcriptional network [36]. Moreover, *GpbHLH58* is speculated to be a homolog of *AtbHLHs* involved in JA signal transduction, especially MYC2. *GpbHLH65* was correspondingly clustered with the *AtbHLHs* in subgroup 1a and predicted to be a homolog of *AtbHLH067*, which negatively regulates cold tolerance and interacts physically with ICE1 [37]. Lastly, *GpbHLH85* was grouped with *AtbHLH037*

(*HEC2*), *AtbHLH043* (*HEC3*), and *AtbHLH088* (*HEC1*), all of which are involved in stomata differentiation, transmitting tract formation, and stigma development [38] (Fig. 1). In addition to their roles in growth and development regulation, PIF and MYC2 homologs are known to regulate various secondary metabolites, including flavonoids and terpenoids [31, 39], suggesting that *GpbHLH15* and *GpbHLH58* are the most likely regulators of MeJA-mediated gypenoside biosynthesis. However, further evidence is required to confirm the hypothesis.

Gypenoside biosynthesis genes were transcriptionally activated by *GpbHLH15* and *GpbHLH58*

TFs typically function as transcriptional activators or repressors in the nucleus. To investigate the localization of *GpbHLH15* and *GpbHLH58* proteins, we conducted subcellular localization experiments in *N. benthamiana*

leaves. As shown in Fig. 7, the yellow fluorescent protein (YFP) signals from the GpbHLH15-YFP and GpbHLH58-YFP fusion constructs were exclusively observed in the nucleus, indicating that GpbHLH15 and GpbHLH58 are both nuclear-localized TFs.

In a previous study, bHLH-recognizing elements (BREs) (CANNTG) were identified in the promoters of four gypenoside biosynthesis enzyme genes (*GpFPS1*, *GpSS1*, *GpSE2*, and *GpOSC1*), implying that these genes could be potential targets of GpbHLHs [23]. To assess the ability of GpbHLH15 and GpbHLH58 to bind to the promoters of these four gypenoside biosynthesis enzyme genes, we utilized Y1H assays (Fig. 8A). Bit yeast cells co-transformed with the fusion vectors (AD-GpbHLH15 and AD-GpbHLH58) grew on the selective medium supplemented with *Aureobasidin A* (AbA). The results demonstrated that GpbHLH15 could bind to the promoter of *GpSS1* in vitro, and the interactions between GpbHLH58 and the structural genes (*GpFPS1*, *GpSE2*, and *GpOSC1*) were also detected.

To further investigate the regulatory roles of GpbHLH15 and GpbHLH58 on *GpFPS1*, *GpSS1*, *GpSE2*, and *GpOSC1*, dual-luciferase assays were conducted (Fig. 8B). The results revealed that GpbHLH15 activated the expression of *GpSS1*, while GpbHLH58 and each of the firefly luciferase (LUC) reporter constructs driven by *GpFPS1*, *GpSS1*, and *GpOSC1* were co-infiltrated, higher levels of activation luciferase activities were observed. These findings demonstrate that GpbHLH15 activates

the expression of *GpSS1*, whereas the GpbHLH58 protein activates the promoter activity of *GpFPS1*, *GpSS1*, and *GpOSC1*.

Discussion

Recently, owing to its significant health benefits, the cultivation of *G. pentaphyllum* has gained significance in the agricultural industry. Moreover, the content of gypenosides is directly related to the plant's quality and commercial value. However, there is considerable variation in the quality of *G. pentaphyllum*, which poses a challenge to the sustainable development of the industry. To address this issue, understanding the molecular mechanisms underlying gypenoside biosynthesis has become a current research highlight. The *bHLH* gene family, known for its vital role in various physiological processes [12, 40, 41], has been comprehensively studied in various plants. However, research on *bHLH* genes in *G. pentaphyllum* has been limited.

Here, we identified and characterized 111 *bHLH* genes in the *G. pentaphyllum* genome. While this number is similar to that reported in *C. mleo* (118), it is less than that in *C. sativus* (142), even though both plants belong to the Cucurbitaceae family [42, 43]. Notably, *G. pentaphyllum* has a larger genome size (582 Mb) than *C. sativus* (367 Mb) but possesses fewer *bHLH* genes, implying that the number of gene family members is not solely dependent on genome size but also on the species' evolutionary history. The 111 GpbHLH proteins were classified into

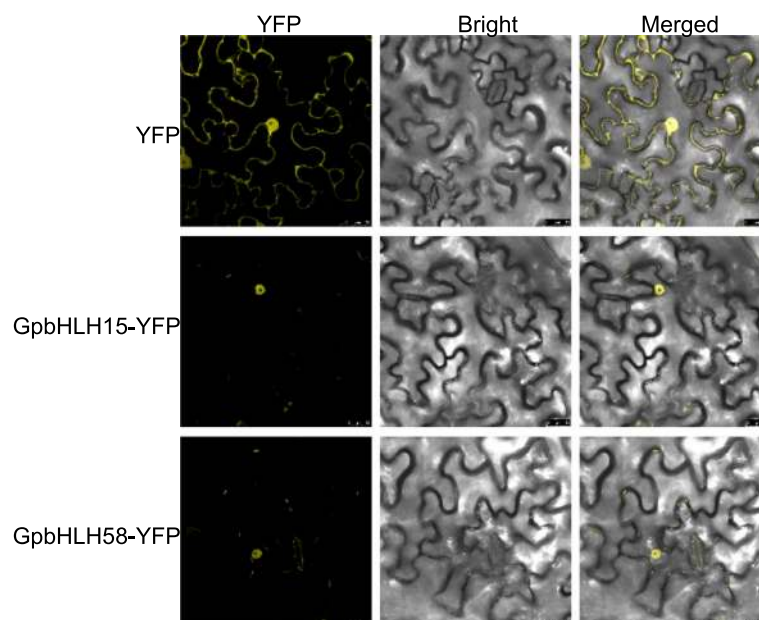


Fig. 7 Subcellular localization of GpbHLH15 and GpbHLH58 proteins. The GpbHLH15/58-YFP fusion constructs were transfected into *N. benthamiana* leaves. YFP indicates the empty vector, which was used as a control. Confocal laser microscopy images were captured three days post-transfection. Scale bar = 25 μ m

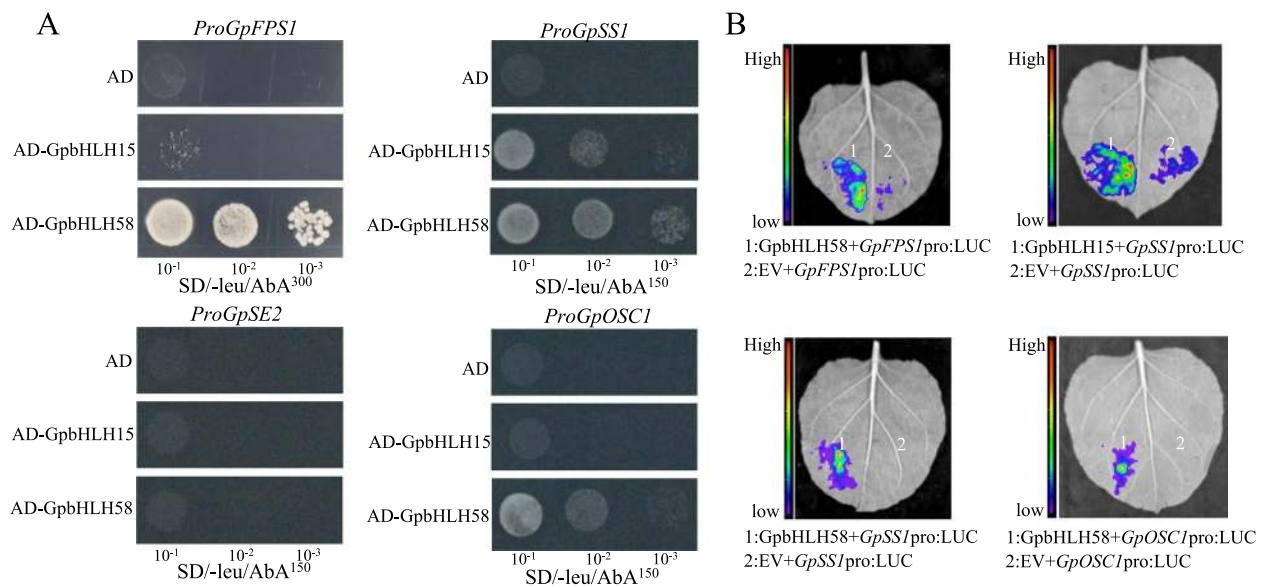


Fig. 8 Interaction of GpbHLH15 and GpbHLH58 proteins with gypenoside biosynthesis gene promoters. **A** Y1H assays demonstrate the binding of GpbHLH15 to the *GpSS1* promoter, whereas GpbHLH58 exhibits binding to the promoters of *GpFPS1*, *GpSS1*, and *GpOSC1*. AD indicates a negative control using the pGADT7 empty vector. The transformed yeast cells were diluted thrice, with 10⁻¹ dilutions each time and then spotted on the selective medium. **B** Dual-luciferase assays were performed using *GpFPS1*, *GpSS1*, and *GpOSC1* promoters driving the *Firefly luciferase* (*LUC*) gene as reporters, along with effectors (GpbHLH15, GpbHLH58). These results reveal that GpbHLH15 activates the promoter of *GpFPS1*, while GpbHLH58 promotes the expression of *GpFPS1*, *GpSS1*, and *GpOSC1*

26 subgroups, consistent with the classification system established for *A. thaliana* [2]. Furthermore, the highly conserved structural features of GpbHLH proteins within each subgroup corroborated the phylogenetic analysis. Thus, collectively, these findings strengthen the reliability of the GpbHLH classification and function prediction based on phylogenetic analysis.

Homology analysis is a valuable tool for predicting the functions of genes in various species. To date, numerous bHLH proteins in *A. thaliana* have been functionally characterized [32], thus enabling us to anticipate the roles of the bHLH gene family in *G. pentaphyllum*. For instance, bHLHs belonging to the subgroup IIIIf are involved in anthocyanin biosynthesis, trichome and root hair formation [44–46], whereas members from the subgroup IIIb regulate stomata development and enhance freezing tolerance [37, 47]. Furthermore, bHLHs in subgroup XIII, including At159P1R2, At165PAR1, At166PAR2, At167P1R1, and At168P1R3, have been shown to control shade avoidance and auxin responses [48, 49]. Notably, bHLHs belonging to the subgroup XIII appear to have been significantly lost from the *G. pentaphyllum* genome compared to *A. thaliana*. These findings suggest that the shade avoidance-related gene family may have undergone adaptive evolution and the regulation of shade response may have been altered in *G. pentaphyllum*.

The production and modulation of secondary metabolites are important mechanisms that allow plants to adapt to environmental changes [50]. These processes involve complex signal transduction cascades in response to internal hormone signals and external environmental factors, ultimately leading to the activation or inhibition of downstream gene expression and the production of various specialized plant metabolites [51–53]. In our study, the *cis*-acting regulatory element prediction revealed the prevalence of *cis*-elements responsive to plant hormones, such as ABA and MeJA, among the nine gypenoside-related *GpbHLH* genes. Additionally, light-responsive *cis*-elements were widely identified. These findings suggest that gypenoside biosynthesis is influenced by various phytohormones, especially MeJA and ABA, as well as changing environmental conditions, especially light signals. However, further research is needed to comprehensively investigate the interplay among ABA, MeJA, and light signals in gypenoside biosynthesis regulation.

Based on previous studies, we hypothesized that *GpbHLH15* and *GpbHLH58* likely play a role in regulating gypenoside biosynthesis in *G. pentaphyllum*. The subcellular localization experiments revealed that GpbHLH15/58-YFP fusion proteins were localized in the nucleus (Fig. 5), indicating their regulatory function within the nucleus. Moreover, GpbHLH15 and GpbHLH58 could bind to the promoters of gypenoside

biosynthesis genes and activate their promoter activity. In our experiments, the expression of GpbHLH58 (GpMYC2), along with gypenoside biosynthesis enzymes, was induced by extracellular MeJA treatment. A previous study reported that MeJA-induced AtMYC2 acted as a transcriptional activator in the MeJA signaling pathway [54]. Therefore, it is likely that GpbHLH58 functions as a transcriptional activator by binding to the promoters of gypenoside biosynthesis genes in a MeJA-dependent manner.

Conclusions

This study conducted a comprehensive genome-wide characterization and analysis of the bHLH TF family in *G. pentaphyllum* and explored their potential roles in regulating gypenoside biosynthesis. Furthermore, we identified four gypenoside-related *GpbHLH* genes, with a particular focus on GpbHLH15 and GpbHLH58, as potential regulatory factors for gypenoside biosynthesis. Additionally, through Y1H and dual-luciferase assays, we demonstrated that GpbHLH15 and GpbHLH58 could bind to the promoters of enzyme-encoding genes in gypenoside biosynthetic pathway and activate their transcription, highlighting their involvement in the regulation of gypenoside biosynthesis. In conclusion, our findings provide valuable insights into the potential functions of GpbHLHs in the regulation of gypenosides in *G. pentaphyllum*.

Materials and methods

Plant materials and MeJA treatment

The *G. pentaphyllum* plant material used in this study is a variety whose genome was sequenced previously [23]. For MeJA treatment, six weeks old *G. pentaphyllum* plants were used, and then cultured in Hoagland's nutrient solutions containing 100 μ M MeJA. The MeJA-treated leaves were collected from three independent *G. pentaphyllum* plants before treatment (0 h) and 6, 12, 24 h after treatment.

Identification of *GpbHLH* proteins in *G. pentaphyllum*

Firstly, to obtain putative GpbHLH members, bidirectional BLASTP comparisons were performed with an E-value of 10^{-10} by using the protein sequences of 167 AtbHLH as the query sequence. Secondly, the conserved structural domain of HLH (PF00010) was searched in the genome of *G. pentaphyllum* to identify the *GpbHLH* proteins by using HMMER (v3.3.2) with default parameters [55]. After excluding redundant sequences, a set of predicted bHLH candidate sequences were obtained. Finally, to confirm the presence of conserved bHLH domain in the candidate genes, Simple Modular Architecture

Research Tool (SMART) (<http://smart.embl-heidelberg.de/>) was employed to filter with candidate genes.

Sequence analysis, phylogenetic analysis and *GpbHLH* classification

According to the chromosomal position on the *G. pentaphyllum* chromosomes, all the identified GpbHLHs were renamed in order. The physical and chemical properties were determined by the ExPASy ProtParam (<https://web.expasy.org/protparam/>), comprising amino acid length, MW and predicted pI. Subcellular localization predictions were done with the online tool Cell-PLoc (v2.0) (<http://www.csbio.sjtu.edu.cn/bioinf/Cell-PLoc/>).

Multiple sequence alignments were performed using bHLH amino acid sequences of *G. pentaphyllum* and *Arabidopsis thaliana*, with ClustalX (v2.1) with the default parameters. After alignment the evolutionary tree was constructed by the FastTree software (v2.1.1), and the default parameters were adopted with the ML method [56]. The Interactive Tree of Life (iTOL) was used for visualisation and annotation of the phylogenetic tree [57]. For subfamily classification of GpbHLH proteins, the classification method of AtbHLH proteins was adopted, which was based on their evolutionary relationships [2].

Conserved motif and gene structure analysis

For conserved motifs prediction of GpbHLH proteins, a motif discovery and analysis tool MEME (v5.1.1) [58] was used. The maximum motif number was set to 20. The Gene structure (exon/intron) of 111 *GpbHLH* genes were visualized by the TBtools software [59] based on feature structure information in GFF annotation files from the *G. pentaphyllum* genome. Additionally, for building a ML phylogenetic tree containing 111 GpbHLH protein sequences, FastTree software (v2.1.1) was used with default parameters.

Genomic localization and gene duplication of *GpbHLH* genes

The chromosome distribution map of GpbHLH gene family members was drawn as described by Chen [59]. Comparison and annotation of orthologous bHLH genes among *G. pentaphyllum*, *A. thaliana* and *C. sativus* were completed using OrthoVenn2 [60]. Collinearity regions and gene duplication events of the bHLH gene family among different species were assessed as described previously [59]. Ka/Ks substitution of each duplicated bHLH genes was assessed using Tbtools software.

Identification of co-expression modules

The clean reads of seven different tissue of *G. pentaphyllum* including root, stem, tender leaf, mature leaf, flower,

tendril and fruit were downloaded from NCBI database (<https://www.ncbi.nlm.nih.gov>) under the accession number PRJNA720501 and PRJNA631355 in previous study [23, 61]. Then the clean reads were mapped to the reference genome using TBtools. The expression level was evaluated by normalization to fragments per kilobase of exon model per million mapped reads (FPKM) value calculated from the number of aligned reads for each gene. Weighted gene co-expression network analysis (WGCNA) package was used for co-expression analysis.

Quantitative real-time PCR (qRT-PCR) validation

Total RNA extraction and reverse transcription were performed as described previously [35]. qRT-PCR was carried out using LightCycler 96 System (Roche). The *GpActin* gene was set as an internal control [61]. To calculate relative gene expression levels, the $2^{-\Delta\Delta C_t}$ analysis method was adopted. The primers used in this study are given in Supplementary Table S7.

Subcellular localization

To perform subcellular localization analysis of GpbHLH proteins, the CDS of *GpbHLH15* and *GpbHLH58* were amplified and cloned into the pHB-YFP vector to form GpbHLH15/58-YFP fusion proteins. The GpbHLH15/58-YFP fusion constructs were transfected into *Nicotiana benthamiana* leaves via *Agrobacterium tumefaciens* strain GV3101 transformation. After 3 d transfection, the YFP signals were captured using a confocal laser scanning microscope (Zeiss LSM 510 Meta).

Y1H assays

Y1H was performed as described previously [62]. To generate prey vectors, the CDS of *GpbHLH15* and *GpbHLH58* were cloned into pGADT7 plasmid. Meanwhile, the promoter fragments of four gypenoside biosynthetic gene, designated as *proGpFPS1*, *proGpSS1*, *proGpSE2* and *proGpOSC1*, containing the E-box *cis*-elements, were cloned into pAbAi plasmid to construct baits. The 10× dilution series transformed yeast cells were dotted on the SD/-Leu plates with or without AbA.

Dual luciferase assays

A transcriptional activity assay was conducted as described previously [62]. To generate effector vector, the CDS of *GpbHLH15* and *GpbHLH58* were cloned into pK2GW7 plasmid using the Gateway cloning system, while the promoter fragments of four gypenoside biosynthetic genes (*GpFPS1*, *GpSS1*, *GpSE2* and *GpOSC1*) were inserted into pGreenII 0800-LUC plasmid. Fluorescence was observed and captured using a fluorescence imaging system (IVIS Lumina LT, USA).

Abbreviations

bHLH	Basic helix-loop-helix
GpbHLHs	bHLH members in <i>G. pentaphyllum</i>
TF	Transcription factor
HLH	Helix-loop-helix
MeJA	Methyl jasmonate
Y1H	Yeast one-hybrid
pI	Isoelectric point
MW	Molecular weight
ML	Maximum likelihood
CDS	Coding sequence
GFF	Generic file format
Ka	Non-synonymous
Ks	Synonymous
WGCNA	Weighted gene co-expression network analysis
FPKM	Fragments per kilobase of exon model per million mapped reads
ABA	Absciscic acid
JA	Jasmonic acid
ABRE	ABA-responsive element
ARE	Anaerobic induced response element
WRE3	Wound responsiveness element
LTR	Low-temperature responsiveness
DRE	Dehydration responsive element
YFP	Yellow fluorescent protein
BRE	bHLH-recognizing element
AbA	Aureobasidin A
qRT-PCR	Quantitative real-time PCR
UTR	Untranslated region
LUC	Firefly luciferase

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12870-024-04879-y>.

Additional file 1: Table S1. Characteristics of the 111 bHLH genes in *Gynostemma pentaphyllum* of this study.

Additional file 2: Table S2. Analysis and distribution of conserved motifs in GpbHLH proteins.

Additional file 3: Table S3. Segmentally and tandemly duplicated GpbHLH gene pairs.

Additional file 4: Table S4. One-to-one orthologous pairs of bHLH genes between two species.

Additional file 5: Table S5. The FPKM value of each structural gene from the pathway of gypenoside biosynthesis and bHLH genes identified in the gypenoside-related module.

Additional file 6: Table S6. Promoter analysis of the GpbHLH gene family in *Gynostemma pentaphyllum*.

Additional file 7: Table S7. List of oligonucleotide sequences used in this study.

Additional file 8: Figure S1. Correlations between the four gypenoside biosynthesis pathway genes and two GpbHLHs (GpbHLH15/58) in response to MeJA.

Acknowledgements

We appreciate the linguistic assistance provided by TopEdit (www.topeditsci.com) during the preparation of this manuscript.

Authors' contributions

HD, and MRH conceived the research. QYH, and LJM carried out the experiments. YSC, LLB, HRS, and TY contributed to sample collection and preparation. MRH, and CJH performed bioinformatic analysis and analyzed the data. QYH wrote the original draft of the manuscript, which was revised by LJM, and HD. HD, and MRH supervised the project. All authors agreed with the submission of the finalized manuscript.

Funding

This project was supported by the Natural Science Foundation of Guangxi Zhuang Autonomous Region (2020GXNSFBA297025), Postgraduate Joint Training base of Guangxi (Guixuwei [2021] No. 6), Inheritance and innovation team of Guangxi Traditional Chinese Medicine (2022B005) and the 'Xinglin' Youth Talent of Guangxi University of Chinese Medicine Project Fund (2022C031).

Availability of data and materials

The plant materials were collected from the orchard of Guangxi University of Chinese Medicine with the permission of the owner. The raw transcriptome data have been deposited in the NCBI database (<https://www.ncbi.nlm.nih.gov>) under the accession number PRJNA720501 and PRJNA631355 in previous study [23, 61].

Declarations

Ethics approval and consent to participate

Authors confirmed that the whole study on plants, including the collection of plant material, complied with institutional, national, and international guidelines and legislation. This study did not require ethical approval or consent, plant samples used in the study were not included in the list of national key protected plants and not collected from national park or natural reserve.

Consent for publication

Not applicable.

Competing interests

The authors have declared that no competing interests exist.

Author details

¹College of Pharmacy, Guangxi University of Chinese Medicine, Nanning 530200, China. ²Guangxi Key Laboratory of Zhuang and Yao Ethnic Medicine, Guangxi University of Chinese Medicine, Nanning 530200, China.

Received: 22 December 2023 Accepted: 4 March 2024

Published online: 20 March 2024

References

- Riechmann JL, Heard J, Martin G, Reuber L, Jiang C, Keddie J, et al. *Arabidopsis* transcription factors: genome-wide comparative analysis among eukaryotes. *Science*. 2000;290(5499):2105–10.
- Pires N, Dolan L. Origin and diversification of basic-helix-loop-helix proteins in plants. *Mol Biol Evol*. 2010;27(4):862–74.
- Massari ME, Murre C. Helix-loop-helix proteins: regulators of transcription in eucaryotic organisms. *Mol Cell Biol*. 2000;20(2):429–40.
- Nair SK, Burley SK. Recognizing DNA in the library. *Nature*. 2000;404(6779):715, 717–715718.
- Alex R, Sözeri O, Meyer S, Dildrop R. Determination of the DNA sequence recognized by the bHLH-zip domain of the N-Myc protein. *Nucleic Acids Res*. 1992;20(9):2257–63.
- Grotewold E, Sainz MB, Tagliani L, Hernandez JM, Bowen B, Chandler VL. Identification of the residues in the Myb domain of maize C1 that specify the interaction with the bHLH cofactor R. *Proc Natl Acad Sci*. 2000;97(25):13579–84.
- Chang G, Ma J, Wang S, Tang M, Zhang B, Ma Y, et al. Liverwort bHLH transcription factors and the origin of stomata in plants. *Curr Biol*. 2023;33:2806–13.
- Gonzalez A, Mendenhall J, Huo Y, Lloyd A. TTG1 complex MYBs, MYB5 and TT2, control outer seed coat differentiation. *Dev Biol*. 2009;325(2):412–21.
- Qiu Y, Tao R, Feng Y, Xiao Z, Zhang D, Peng Y, et al. EIN3 and RSL4 interfere with an MYB-bHLH-WD40 complex to mediate ethylene-induced ectopic root hair formation in *Arabidopsis*. *Proc Natl Acad Sci*. 2021;118(51):e2110004118.
- Groszmann M, Paicu T, Alvarez JP, Swain SM, Smyth DR. SPATULA and ALCATRAZ, are partially redundant, functionally diverging bHLH genes required for *Arabidopsis* gynoecium and fruit development. *Plant J*. 2011;68(5):816–29.
- Leivar P, Monte E, Oka Y, Liu T, Carle C, Castillon A, et al. Multiple phytochrome-interacting bHLH transcription factors repress premature seedling photomorphogenesis in darkness. *Curr Biol*. 2008;18(23):1815–23.
- Liang J, Fang Y, An C, Yao Y, Wang X, Zhang W, et al. Genome-wide identification and expression analysis of the bHLH gene family in passion fruit (*Passiflora edulis*) and its response to abiotic stress. *Int J Biol Macromol*. 2023;225:389–403.
- Su L, Zhang Y, Yu S, Geng L, Lin S, Ouyang L, et al. RcbHLH59-RcPRs module enhances salinity stress tolerance by balancing Na(+)/K(+) through callose deposition in rose (*Rosa chinensis*). *Hortic Res*. 2023;10(3):uhac291.
- Xue L, Wei Z, Zhai H, Xing S, Wang Y, He S, et al. The IbPYL8-IbbHLH66-IbbHLH118 complex mediates the abscisic acid-dependent drought response in sweet potato. *New Phytol*. 2022;236(6):2151–71.
- Zhang L, Xiang Z, Li J, Wang S, Chen Y, Liu Y, et al. bHLH57 confers chilling tolerance and grain yield improvement in rice. *Plant Cell Environ*. 2023;46(4):1402–18.
- Liu S, Wang Y, Shi M, Maoz I, Gao X, Sun M, et al. SmbHLH60 and SmMYC2 antagonistically regulate phenolic acids and anthocyanins biosynthesis in *Salvia miltiorrhiza*. *J Adv Res*. 2022;42:205–19.
- Jia N, Wang JJ, Liu J, Jiang J, Sun J, Yan P, et al. DcTT8, a bHLH transcription factor, regulates anthocyanin biosynthesis in *Dendrobium candidum*. *Plant Physiol Biochem*. 2021;162:603–12.
- Li X, Cao L, Jiao B, Yang H, Ma C, Liang Y. The bHLH transcription factor AcB2 regulates anthocyanin biosynthesis in onion (*Allium cepa* L.). *Hortic Res*. 2022;9:uhac128.
- Singh SK, Patra B, Paul P, Liu Y, Pattanaik S, Yuan L. BHLH IRIDOID SYNTHESIS 3 is a member of a bHLH gene cluster regulating terpenoid indole alkaloid biosynthesis in *Catharanthus roseus*. *Plant direct*. 2021;5(1):e00305.
- Mertens J, Pollier J, Vanden Bossche R, Lopez-Vidriero I, Franco-Zorrilla JM, Goossens A. The bHLH transcription factors TSAR1 and TSAR2 regulate triterpene saponin biosynthesis in *Medicago truncatula*. *Plant Physiol*. 2016;170(1):194–210.
- Mohammad, Hurrah IM, Kumar A, Abbas N. Synergistic interaction of two bHLH transcription factors positively regulates artemisinin biosynthetic pathway in *Artemisia annua* L. *Physiol Plant*. 2023;175(1):e13849.
- Xie P, Xie JB, Xiao MY, Guo M, Qi YS, Li FF, et al. Liver lipidomics analysis reveals the anti-obesity and lipid-lowering effects of gynosides from heat-processed *Gynostemma pentaphyllum* in high-fat diet fed mice. *Phytomedicine*. 2023;115:154834.
- Huang D, Ming R, Xu S, Wang J, Yao S, Li L, et al. Chromosome-level genome assembly of *Gynostemma pentaphyllum* provides insights into gynoside biosynthesis. *DNA Res*. 2021;28(5):dsab018.
- Park SH, Huh TL, Kim SY, Oh MR, Tirupathi Pichiah PB, Chae SW, et al. Antiobesity effect of *Gynostemma pentaphyllum* extract (actiponin): a randomized, double-blind, placebo-controlled trial. *Obesity*. 2014;22(1):63–71.
- Shen CY, Ma PY, Zhu JJ, Jiang JG, Liu L, Yi YK, et al. Saponin extracts from *Gynostemma pentaphyllum* (Thunb.) Makino display sedative-hypnotic and anxiolytic effects. *Ind Crops Prod*. 2020;157:112893.
- Su C, Li N, Ren R, Wang Y, Su X, Lu F, et al. Progress in the medicinal value, bioactive compounds, and pharmacological activities of *Gynostemma pentaphyllum*. *Molecules*. 2021;26(20):6249.
- Han JY, Kim HJ, Kwon YS, Choi YE. The Cyt P450 enzyme CYP716A47 catalyzes the formation of protopanaxadiol from dammarenediol-II during ginsenoside biosynthesis in *Panax ginseng*. *Plant Cell Physiol*. 2011;52(12):2062–73.
- Han JY, Kwon YS, Yang DC, Jung YR, Choi YE. Expression and RNA interference-induced silencing of the dammarenediol synthase gene in *Panax ginseng*. *Plant Cell Physiol*. 2006;47(12):1653–62.
- Kim YK, Kim YB, Uddin MR, Lee S, Kim SU, Park SU. Enhanced triterpene accumulation in *Panax ginseng* hairy roots overexpressing mevalonate-5-pyrophosphate decarboxylase and farnesyl pyrophosphate synthase. *ACS Synth Biol*. 2014;3(10):773–9.
- Lee MH, Jeong JH, Seo JW, Shin CG, Kim YS, In JG, et al. Enhanced triterpene and phytosterol biosynthesis in *Panax ginseng* overexpressing squalene synthase gene. *Plant Cell Physiol*. 2004;45(8):976–84.

31. Luo L, Wang Y, Qiu L, Han X, Zhu Y, Liu L, et al. MYC2: A master switch for plant physiological processes and specialized metabolite synthesis. *Int J Mol Sci.* 2023;24(4):3511.
32. Feller A, Machemer K, Braun EL, Grotewold E. Evolutionary and comparative analysis of MYB and bHLH plant transcription factors. *Plant J.* 2011;66(1):94–116.
33. Guo H, Li R, Liu S, Zhao N, Han S, Lu M, et al. Molecular characterization, expression, and regulation of *Gynostemma pentaphyllum* squalene epoxidase gene 1. *Plant Physiol Biochem.* 2016;109:230–9.
34. Chen Q, Ma C, Qian J, Lan X, Chao N, Sun J, et al. Transcriptome sequencing of *Gynostemma pentaphyllum* to identify genes and enzymes involved in triterpenoid biosynthesis. *Int J Genomics.* 2016;2016:7840914.
35. Huang D, Ming R, Xu S, Yao S, Li L, Huang R, et al. Genome-Wide Identification of R2R3-MYB transcription factors: discovery of a “dual-function” regulator of gypenoside and flavonol biosynthesis in *Gynostemma pentaphyllum*. *Front Plant Sci.* 2022;12:3094.
36. Burko Y, Willige BC, Seluzicki A, Novák O, Ljung K, Chory J. PIF7 is a master regulator of thermomorphogenesis in shade. *Nat Commun.* 2022;13(1):4942.
37. Li H, Ding Y, Shi Y, Zhang X, Zhang S, Gong Z, et al. MPK3- and MPK6-mediated ICE1 phosphorylation negatively regulates ICE1 stability and freezing tolerance in *Arabidopsis*. *Dev Cell.* 2017;43(5):630–642 e634.
38. Zhu L, Xin R, Bu Q, Shen H, Dang J, Huq E. A negative feedback loop between PHYTOCHROME INTERACTING FACTORS and HECATE proteins fine-tunes photomorphogenesis in *Arabidopsis*. *Plant Cell.* 2016;28(4):855–74.
39. Zhang Q, Wu N, Jian D, Jiang R, Yang C, Lan X, et al. Overexpression of AaPIF3 promotes artemisinin production in *Artemisia annua*. *Ind Crops Prod.* 2019;138:111476.
40. Carretero-Paulet L, Galstyan A, Roig-Villanova I, Martínez-García JF, Bilbao-Castro JR, Robertson DL. Genome-wide classification and evolutionary analysis of the bHLH family of transcription factors in *Arabidopsis*, poplar, rice, moss, and algae. *Plant Physiol.* 2010;153(3):1398–412.
41. Xu J, Xu H, Zhao H, Liu H, Xu L, Liang Z. Genome-wide investigation of bHLH genes and expression analysis under salt and hormonal treatments in *Andropogon paniculata*. *Ind Crops Prod.* 2022;183:114928.
42. Li J, Wang T, Han J, Ren S. Genome-wide identification and characterization of cucumber bHLH family genes and the functional characterization of CsbHLH041 in NaCl and ABA tolerance in *Arabidopsis* and cucumber. *BMC Plant Biol.* 2020;20(1):272.
43. Tan C, Qiao H, Ma M, Wang X, Tian Y, Bai S, et al. Genome-wide identification and characterization of melon bHLH transcription factors in regulation of fruit development. *Plants.* 2021;10(12):2721.
44. Baudry A, Heim MA, Dubreucq B, Caboche M, Weisshaar B, Lepiniec L. TT2, TT8, and TTG1 synergistically specify the expression of BANYULS and proanthocyanidin biosynthesis in *Arabidopsis thaliana*. *Plant J.* 2004;39(3):366–80.
45. 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(10):1863–78.
46. Zhang F, Gonzalez A, Zhao M, Payne CT, Lloyd A. A network of redundant bHLH proteins functions in all TTG1-dependent pathways of *Arabidopsis*. *Development.* 2003;130(20):4859–69.
47. Kanaoka MM, Pillitteri LJ, Fujii H, Yoshida Y, Bogenschütz NL, et al. SCREAM/ICE1 and SCREAM2 specify three cell-state transitional steps leading to *Arabidopsis* stomatal differentiation. *Plant Cell.* 2008;20(7):1775–85.
48. Roig-Villanova I, Bou-Torrent J, Galstyan A, Carretero-Paulet L, Portolés S, Rodríguez-Concepción M, et al. Interaction of shade avoidance and auxin responses: a role for two novel atypical bHLH proteins. *EMBO J.* 2007;26(22):4756–67.
49. Zhou P, Song M, Yang Q, Su L, Hou P, Guo L, et al. Both PHYTOCHROME RAPIDLY REGULATED1 (PAR1) and PAR2 promote seedling photomorphogenesis in multiple light signaling pathways. *Plant Physiol.* 2014;164(2):841–52.
50. Zhang J, Yang X, Zhang X, Zhang L, Zhang Z, Zhang Y, et al. Linking environmental signals to plant metabolism: The combination of field trials and environment simulators. *Mol Plant.* 2022;15(2):213–5.
51. Back K. Melatonin metabolism, signaling and possible roles in plants. *Plant J.* 2021;105(2):376–91.
52. Yue P, Jiang Z, Sun Q, Wei R, Yin Y, Xie Z, et al. Jasmonate activates a CsMPK6-CsMYC2 module that regulates the expression of β -citralin biosynthetic genes and fruit coloration in orange (*Citrus sinensis*). *Plant Cell.* 2023;35(4):1167–85.
53. Zhu C, Zhang S, Zhou C, Chen L, Zaripov T, Zhan D, et al. Integrated transcriptome, microRNA, and phytochemical analyses reveal roles of phytohormone signal transduction and ABC transporters in flavor formation of oolong tea (*Camellia sinensis*) during solar withering. *J Agric Food Chem.* 2020;68(45):12749–67.
54. Abe H, Urao T, Ito T, Seki M, Shinozaki K, Yamaguchi-Shinozaki K. *Arabidopsis* AtMYC2 (bHLH) and AtMYB2 (MYB) function as transcriptional activators in abscisic acid signaling. *Plant Cell.* 2003;15(1):63–78.
55. Finn RD, Clements J, Eddy SR. HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res.* 2011;39:W29–37.
56. Price MN, Dehal PS, Arkin AP. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol Biol Evol.* 2009;26(7):1641–50.
57. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* 2021;49(W1):W293–6.
58. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, et al. MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.* 2009;37:W202–208.
59. Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, et al. TBtools: An integrative toolkit developed for interactive analyses of big biological data. *Mol Plant.* 2020;13(8):1194–202.
60. Xu L, Dong Z, Fang L, Luo Y, Wei Z, Guo H, et al. OrthoVenn2: a web server for whole-genome comparison and annotation of orthologous clusters across multiple species. *Nucleic Acids Res.* 2019;47(W1):W52–8.
61. Xu S, Yao S, Huang R, Tan Y, Huang D. Transcriptome-wide analysis of the AP2/ERF transcription factor gene family involved in the regulation of gypenoside biosynthesis in *Gynostemma pentaphyllum*. *Plant Physiol Biochem.* 2020;154:238–47.
62. Huang D, Wang X, Tang Z, Yuan Y, Xu Y, He J, et al. Subfunctionalization of the Ruby2–Ruby1 gene cluster during the domestication of citrus. *Nat Plants.* 2018;4(11):930–41.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.