



OPEN Phylogenetic and functional analysis of MYB genes unraveling its role involved in anthocyanin biosynthesis in *H. macrophylla*

Yingying Liao^{1,2,3,5}, Wenwen Zhao^{1,2,3,5}, Yu Wang^{1,2,3}, Sumeiqi Zhao^{1,2,3}, Shiqin Gu^{1,2,3}, Chao Zhang⁴, Jiqing Peng^{1,2,3}✉, Song Sheng^{1,2,3}✉ & Sen Wang^{1,2,3}✉

Flower color serves as a pivotal ornamental trait in plants, significantly impacting their aesthetic value. *H. macrophylla* is particularly esteemed for its large, vibrant flowers and the remarkable variability in color it exhibits. In this study, the MYB gene family within *H. macrophylla* was comprehensively examined, resulting in the identification of 72 MYB genes through genomic screening utilizing the Hidden Markov Model (HMM) method. A physicochemical analysis revealed a diverse range of properties among these proteins, with sizes varying from 65 to 770 amino acids; notably, six of these proteins were classified as stable. Predominantly, MYB proteins were localized within the nucleus, although some were also detected in other organelles, indicating a multifaceted role in cellular function. Phylogenetic analysis established that 13 MYB genes from *H. macrophylla* can be categorized into five distinct subfamilies based on the classifications observed in *Arabidopsis thaliana*, suggesting their involvement in the biosynthetic pathways of anthocyanins and proanthocyanidins. Expression analysis indicated a significant correlation between specific R2R3-MYB transcription factors, particularly HmMYB54, and the accumulation of anthocyanins during various flowering stages. qPCR results showed that the expression of *HmMYB54* was consistent with the accumulation trend of 'Forever Summer' colors in hydrangea, while *HmMYB61* was just the opposite, which was consistent with RNAseq, suggesting that they regulated the formation of 'Forever Summer' blue flowers positively and negatively, respectively. A proposed regulatory model has been articulated to elucidate the mechanisms by which these MYB genes affect flower color, particularly through their interactions with essential structural genes in the anthocyanin biosynthesis pathway. Thus, this study not only enhances the understanding of the genetic basis of flower coloration in *H. macrophylla* but also highlights the adaptive significance of MYB transcription factors in response to environmental challenges.

Keywords *H. macrophylla*, MYB, Transcription factor, Anthocyanin

Flower color is one of the most prominent traits in plants and is a decisive factor in assessing ornamental value. The expression of flower color in plants is influenced by several factors, including the types of pigments, the intracellular environment of flower cells, metal ions, and external conditions, with the types and content of pigments being the primary determinants^{1,2}. Plant pigments can be classified into four main categories: flavonoids, carotenoids, chlorophylls, and betacyanins. Betacyanins can give flowers yellow, orange, and red hues, existing only in the order Caryophyllales. Flavonoids primarily consist of anthocyanins, flavonol glycosides, and flavones³. Among them, compounds such as flavones, dihydroflavones, and chalcones lead to light yellow to yellow flower colors⁴.

Anthocyanins are compounds formed by the glycosidic bond between anthocyanins and sugars, representing a class of natural water-soluble pigments that are widely present in the cell sap of flower, fruit, stem, leaf, and root organs. The six most common anthocyanins are pelargonidin, cyanidin, delphinidin, peonidin, malvidin,

¹Central South University of Forestry and Technology, Changsha 410004, China. ²Yuelushan Laboratory, Hunan Agricultural University, Qiushi Building, Furong District, No. 1, Nongda Road, Furong District, Changsha, Hunan Province, China. ³The Belt and Road" international union research center for tropical Arid Non-wood forest in Hunan Province, 498 South Shaoshan Road, Changsha 410004, Hunan, China. ⁴Beijing Jiaotong University School of Software Engineering, Beijing 100044, China. ⁵Yingying Liao and Wenwen Zhao contributed equally to this work. ✉email: pengjiqing17@csuft.edu.cn; songsheng@csuft.edu.cn; t20061111@csuft.edu.cn

and petunidin, which exhibit various colors from red, purple-red to blue. These colors are largely related to the number of hydroxyl groups on the B ring: the greater the number, the bluer the color. Furthermore, methylation of anthocyanins can lead to a slight reddening effect. The biosynthetic pathway of anthocyanins has been extensively studied, regulated by structural and regulatory genes. Structural genes are responsible for encoding the enzymes involved in the biosynthesis of anthocyanins, while regulatory genes mainly include those encoding three types of transcription factors: MYB, bHLH, and WD40^{5,6}.

MYB transcription factors (TFs) are present in all eukaryotes studied to date and are among the largest members of the plant transcription factor family. R2R3-MYB TFs are extensively involved in the regulation of various biological processes⁷. R2R3-MYB proteins form MBW complexes with basic helix-loop-helix (bHLH) proteins and WD-repeat proteins to co-regulate anthocyanins^{8–10}. Christian Dubos classified Arabidopsis R2R3-MYB into 23 subfamilies, with genes in subfamilies S4–S7 mainly controlling the synthesis of flavonols, anthocyanins, and proanthocyanidins (PAs)¹¹. The S6 subfamily gene AtMYB75 in Arabidopsis enhances secondary metabolites (anthocyanins and flavonols); the S6 subfamily gene Ze MYB9 in zinnia positively regulates anthocyanin synthesis; the S5 subfamily gene LcsMYB123 in *Schima superba* has a potential role in flavonol biosynthesis; and the S4 subfamily gene CsMYB4a in tea plants inhibits the biosynthesis of phenylpropanoids and flavonoids^{12–15}. In addition, in Purple pepper, overexpression of CaMYB5 not only induced the accumulation of delphinidin and anthocyanin, but also improved the expression level of anthocyanin biosynthesis structural gene in fruit¹⁶. In roses, RcMYB114, RcbHLH and RcWD40 promote anthocyanin accumulation and determine the color of roses by forming MBW (MYB-bHLH-WD40) complex¹⁷. In *bletilla striata*, the over-expression of BsMYB36 increased the content of flavonoids, and the over-expression of BsMYB51 promoted the accumulation of flavonoids, anthocyanins and proanthocyanidins¹⁸.

The investigation into the molecular mechanisms governing flower color synthesis in *H. macrophylla* is essential for understanding the complex biological processes that influence its ornamental traits. This species, a woody plant belonging to the family Hydrangeaceae, is primarily characterized as an evergreen or deciduous shrub, valued for its vibrant flower colors and large flower clusters. Its applications span potted plants, cut flowers, and landscape greening, highlighting its popularity in horticulture^{19,20}. In addition to its ornamental appeal, *H. macrophylla* is rich in flavonoids, isoflavones, and cyclic ether terpenes, which are noted for their health benefits, including heat-clearing and detoxifying properties, as well as diuretic and anti-swelling effects²¹. Within Chinese traditional culture, this plant is emblematic of hope, reunion, and love, thereby solidifying its role in various festivals and celebrations. The convergence of its aesthetic, medicinal, and cultural values positions *H. macrophylla* as a significant subject of research, particularly in understory floral studies. The synthesis of flower color in *H. macrophylla* is intricately influenced by a myriad of molecular mechanisms, which merit thorough investigation, especially through genetic engineering approaches. Research has elucidated key genes involved in anthocyanin biosynthesis, which are critical for the development of blue flower coloration. This has been demonstrated by the finding that the anthocyanin biosynthesis pathway is pivotal, whereas the carotenoid pathways play a comparatively minimal role in color formation²². Moreover, it has been shown that aluminum ions interact with anthocyanins to regulate color changes; under aluminum stress, the downregulation of various anthocyanin biosynthesis-related genes adversely affects flower pigmentation²³. Transcriptomic analyses have identified several hub genes, including F3'5'H and DFR, which are integral to the color variations observed across different pH levels, thereby underscoring the significance of environmental factors in the regulation of flower color²⁴. Additionally, epigenetic modifications, particularly DNA methylation, have been implicated in pH-dependent color changes, suggesting that such mechanisms contribute further to the observed variations in flower color²⁵. The transport of aluminum ions into flower tissues has also been investigated to elucidate how these ions enhance color intensity, presenting potential avenues for breeding practices aimed at improving ornamental traits²⁶. Notably, while traditional breeding methods face challenges in producing novel colors, molecular techniques offer a means to enhance breeding efficiency and enable more targeted modifications²⁷. Therefore, comprehensive insights into these molecular mechanisms not only enrich the understanding of flower color synthesis in *H. macrophylla* but also facilitate innovative approaches in the breeding of new ornamental varieties.

In summary, this study comprehensively analyzes the MYB gene family in *H. macrophylla*, identifying 72 MYB genes that exhibit diverse physicochemical properties and subcellular localizations, including predominance in the nucleus. Phylogenetic analysis reveals that several MYB transcription factors cluster into specific subfamilies associated with anthocyanin biosynthesis, suggesting potential roles in color formation. Expression profiling indicates that genes such as HmMYB15 and HmMYB54 correlate positively with anthocyanin accumulation during flowering, while HmMYB42 and HmMYB57 respond significantly to external stressors, including Pb exposure. Moreover, the expression analysis highlights key correlations between R2R3-MYB transcription factors and structural genes in anthocyanin synthesis, proposing a regulatory model that delineates the interplay between these factors in flower color development. Collectively, these findings underscore the critical roles of MYB transcription factors in regulating color and stress responses in *H. macrophylla*, paving the way for further investigations into their functional mechanisms.

Materials and methods

Experimental materials

The sterile flower bud stage (FSF1), flowering stage (FSF2) and flowering stage (FSF3) of hydrangea variety 'Forever Summer', sterile flower flowering stage (EMF3) of 'emilia' and sterile flower flowering stage (HMF3) of 'Fendai' were used as experimental materials.

Data sources

The genomic sequence and annotation files of *H. macrophylla* were obtained from databases, while MYB transcriptome data (SRP061814, SRP451343, SRP463671) were sourced from the ENA Browser (ENA Browser (ebi.ac.uk)).

Identification of MYB gene family members in hydrangea and analysis of physicochemical properties

The HMM file for the MYB domain (PF00249) was downloaded from the InterPro website (InterPro (ebi.ac.uk)). Using HMMER 3.0, members of the MYB gene family were screened from the *H. macrophylla* protein sequence library, with a retrieval threshold of $E < 1e-5$. This initial screening yielded 114 candidate MYB sequences. These sequences were then submitted to iTAK for secondary screening, resulting in the identification of 72 genes containing the MYB domain, which were subsequently renamed HmMYB1 to HmMYB72. The physicochemical properties of the 72 MYB proteins were analyzed using the Protein Parameter Calc window in TBtools. Subcellular localization predictions for the MYB proteins were conducted utilizing the online tool WoLF PSORT (WoLF PSORT: Protein Subcellular Localization Prediction (hgc.jp)).

Analysis of conserved motifs and gene structure in hydrangea MYB proteins

The protein sequences of the *H. macrophylla* MYB family were submitted to MEME Suite 5.5.6 (MEME - Submission form (meme-suite.org)), with motif parameters set to 7, allowing for the identification of conserved motifs among MYB family members. To obtain domain information, the protein sequences were submitted to CDD (NCBI Conserved Domain Database (nih.gov)), and visual analysis was performed using TBtools. The Visualize Gene Structure function in TBtools was employed to submit the genomic GFF file of *H. macrophylla* in order to acquire the coding sequence (CDS) location information for the MYB genes.

Phylogenetic analysis of hydrangea MYB proteins

Using MEGA X version 10.0.3, a phylogenetic tree was constructed for the 72 identified MYB family members from *H. macrophylla* in comparison with 133 MYB genes from Arabidopsis. The neighbor-joining (NJ) method was employed for tree construction, with a bootstrap setting of 1000 replicates. The resulting phylogenetic tree was further refined using the online tool evolview 3.0 (<https://www.evolgenius.info/evolview/#/treeview>), and the tree was classified according to the subfamilies of Arabidopsis MYB family members.

Data analysis of hydrangea MYB transcriptome

The expression HeatMap of hydrangea MYB gene under different conditions can be obtained by putting the transcriptome data into the Heat Map tool of TBtools. Besides, correlation heat map between hydrangea MYB gene and anthocyanin and the correlation heat map with the key structural genes of anthocyanin synthesis pathway can be obtained by uploading data on the online website (<https://www.omicshare.com/tools/>).

Real-time fluorescence quantitative PCR

Using the sterile flower cDNA of hydrangea as a template, according to the specific reaction system and steps of reagent ChamQUniversal SYBR qPCR Master Mix, the cDNA was diluted 10 times with sterile water, and the reaction system and procedure were set up according to the instructions for quantitative experiments. The internal reference gene is *ACTIN* gene with stable expression in hydrangea, which was screened by our research group. After the reaction, the experimental data were analyzed.

Results and analysis

Identification and physicochemical properties of MYB gene family members in *H. macrophylla*

Using the HMM (Hidden Markov Model) method, a total of 114 MYB gene family members were extracted from the *H. macrophylla* genome. After screening for genes containing the MYB domain, 72 MYB genes were identified. The physicochemical properties and subcellular localization of the 72 *H. macrophylla* MYB proteins were analyzed (Table S1). The results indicated that the number of amino acids ranged from 65 to 770, with the largest protein being HmMYB29 and the smallest protein being HmMYB35. The theoretical isoelectric point varied, with a maximum of 10.89 for HmMYB40 and a minimum of 4.12 for HmMYB69. Among the proteins, six (HmMYB3, HmMYB23, HmMYB25, HmMYB35, HmMYB36, and HmMYB61) exhibited instability coefficients less than 40, classifying them as stable proteins, while the remaining 66 proteins had instability coefficients greater than 40, categorizing them as unstable proteins. Furthermore, all 72 *H. macrophylla* MYB proteins had hydrophilicity coefficients less than 0, indicating that they are hydrophilic proteins. According to the subcellular localization results, most of the identified genes were localized in the nucleus. Additionally, HmMYB29 and HmMYB36 were localized to the plasma membrane, HmMYB43 and HmMYB44 were found in the mitochondria, HmMYB61, HmMYB69, and HmMYB70 were localized to the cytoskeleton, while HmMYB24, HmMYB32, HmMYB49, HmMYB65, and HmMYB72 were found in the chloroplasts.

Phylogenetic tree analysis of MYB protein system in hydrangea

To investigate the evolutionary relationships among the MYB gene members, a phylogenetic tree was constructed using 72 MYB genes from *H. macrophylla* and 133 MYB genes from Arabidopsis thaliana with the MEGA software (Fig. 1). Based on the classification criteria of the Arabidopsis MYB gene family¹¹ it was found that 13 MYB transcription factors from *H. macrophylla* are distributed across five subfamilies of the Arabidopsis R2R3-MYB transcription factors. Specifically, HmMYB4 clusters within subfamily 2, while HmMYB3, HmMYB11, HmMYB42, HmMYB61, HmMYB54 and HmMYB57 cluster in subfamily 5, suggesting that these genes may be related to the formation of anthocyanins or proanthocyanidins. HmMYB2, HmMYB7,

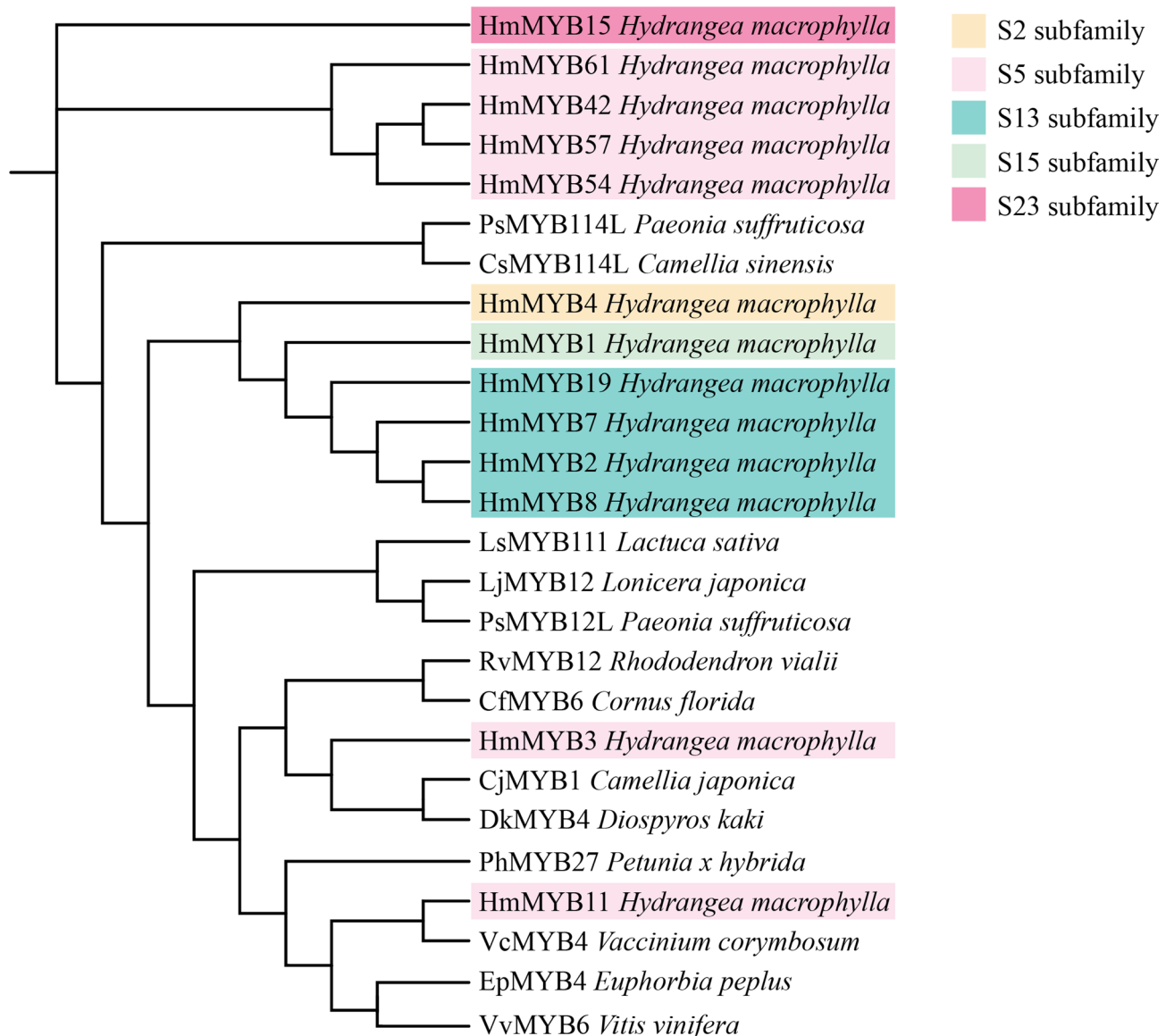


Fig. 2. Phylogenetic tree analysis of MYB protein between hydrangea and other plants Note: GenBank accession number of protein sequences of other species is PsMYB114L (QBK15079.1), CsMYB114L (XP_028068989.1), LsMYB111 (XP_023741633.1), LjMYB12 (QER90717.1), PsMYB12L (QBK15080.1), RvMYB12 (XP_058189634.1), CfMYB6 (XP_059634405.1), CjMYB1 (UZK92551.1), DkMYB4 (BAI49721.1), PhMYB27 (AHX24372.1), VcMYB4 (ALP43794.1), EpMYB4 (WCJ25261.1), VvMYB6 (XP_002273328.1).

motif 2 and motif 3. Additionally, members of the MYB family in *H. macrophylla* that reside within the same phylogenetic branch display identical or similar conserved motifs, as exemplified by HmMYB4 and HmMYB19. These proteins contain typical MYB gene family structural domains, indicating a close relationship between the exons and the evolutionary history of the gene family. To further elucidate the structural characteristics of the MYB gene family in *H. macrophylla*, an analysis of the exons in the MYB genes was performed (Fig. 4). Among the 13 MYB genes analyzed, HmMYB3 and HmMYB15 were found to contain only one exon, while the remaining genes possess two to three exons.

Expression analysis of 13 R2R3-MYB transcription factors related to color in hydrangea

R2R3-MYB transcription factors (TFs) play a crucial role in plant metabolism. To further investigate the functions of R2R3-MYB TFs in *H. macrophylla*, samples were collected from the following developmental stages of the infertile flowers of 'Forever Summer': flower bud stage (FSF1), discoloration stage (FSF2), flowering stage (FSF3), flowering stage of 'Emilia' (EMF3), and flowering stage of 'Fendai' (HMF3). Transcriptome sequencing was conducted, resulting in a gene expression heat map (Fig. 5A). HmMYB1, HmMYB4 and HmMYB19 have no difference in expression, so they are not shown in the figure. The expression levels of HmMYB15 and HmMYB54 in Forever Summer increased with the accumulation of anthocyanins during flowering, with expression levels

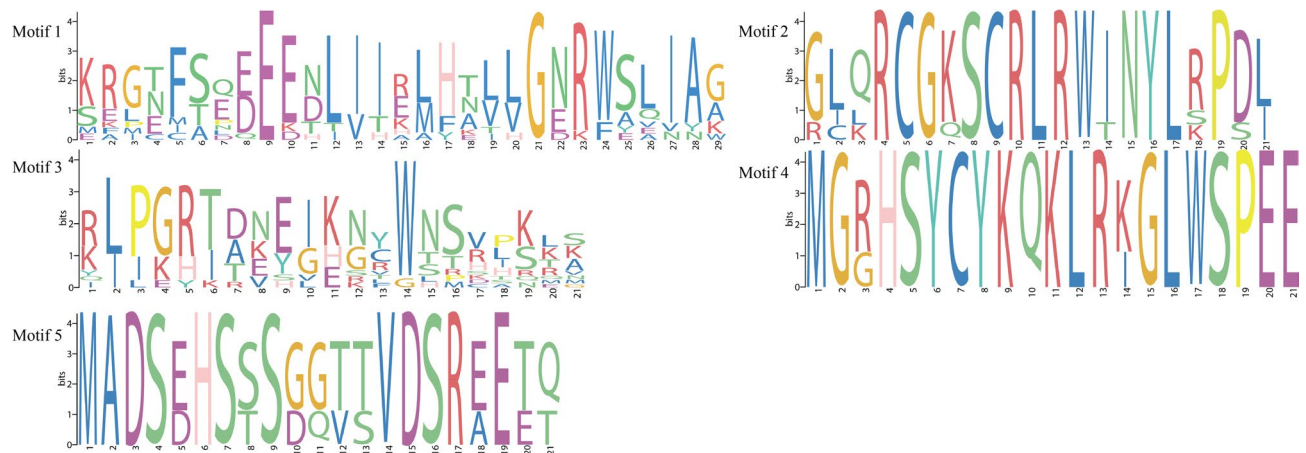


Fig. 3. Motif amino acid sequence of MYB protein.

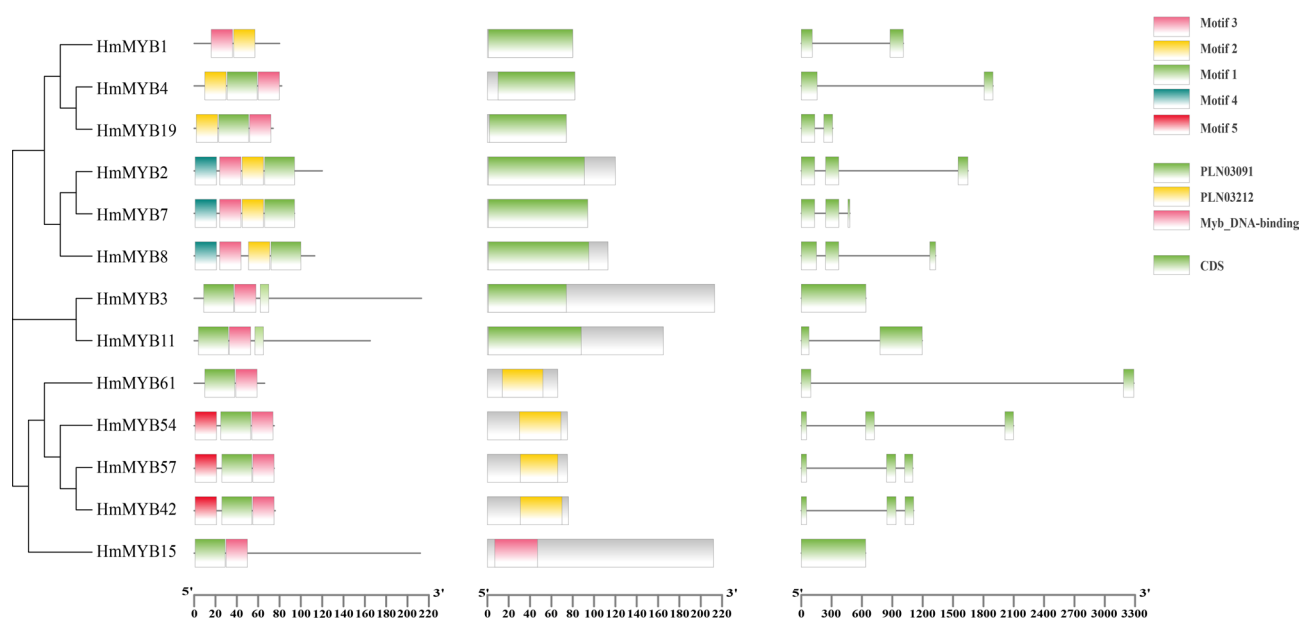


Fig. 4. Protein conserved motif, conserved domain and gene structure of Hydrangea R2R3-MYB.

at FSF3 being higher than those at EMF3 and HMF3. In contrast, the expression level of HmMYB61 in Forever Summer decreased as the flowering period progressed. Additionally, the expression heat map for green stems at an early developmental stage, normal purple stems, and colorless stems subjected to light deprivation treatment with aluminum foil was generated (Fig. 5B). HmMYB61 exhibited higher expression levels in both green and purple stems, gradually declining with the light deprivation treatment. Meanwhile, HmMYB54 showed the highest expression level in the purple stems, with expression decreasing following the light deprivation treatment. To explore the regulatory mechanism of Al^{3+} in altering the flower color of hydrangeas, $AlCl_3$ was applied to the roots and leaves of *H. macrophylla* (Fig. 5C). After treatment, the expression levels of HmMYB42, HmMYB57, and HmMYB61 significantly increased. It can be inferred that HmMYB15, HmMYB54, HmMYB61, HmMYB42 and HmMYB57 may be related to the formation of anthocyanins.

Correlation between R2R3-MYB transcription factors and anthocyanin content in *H. macrophylla*

The expression levels of R2R3-MYB transcription factors in *H. macrophylla* were analyzed for their correlation with anthocyanin content (Fig. 6A). The results indicated significant positive correlations for HmMYB8, HmMYB15, HmMYB42, HmMYB54, and HmMYB57, while HmMYB61 displayed significant correlations with nearly all anthocyanins, except for delphinidin-3-O-glucoside and cyanidin-3-O-(2''-O-glucosyl) glucoside, with the highest correlation coefficient observed for delphinidin-3,5,3'-O-glucoside. HmMYB15 was positively correlated with delphinidin-3-O-glucoside at the level of $P < 0.05$. HmMYB54 showed a highly significant

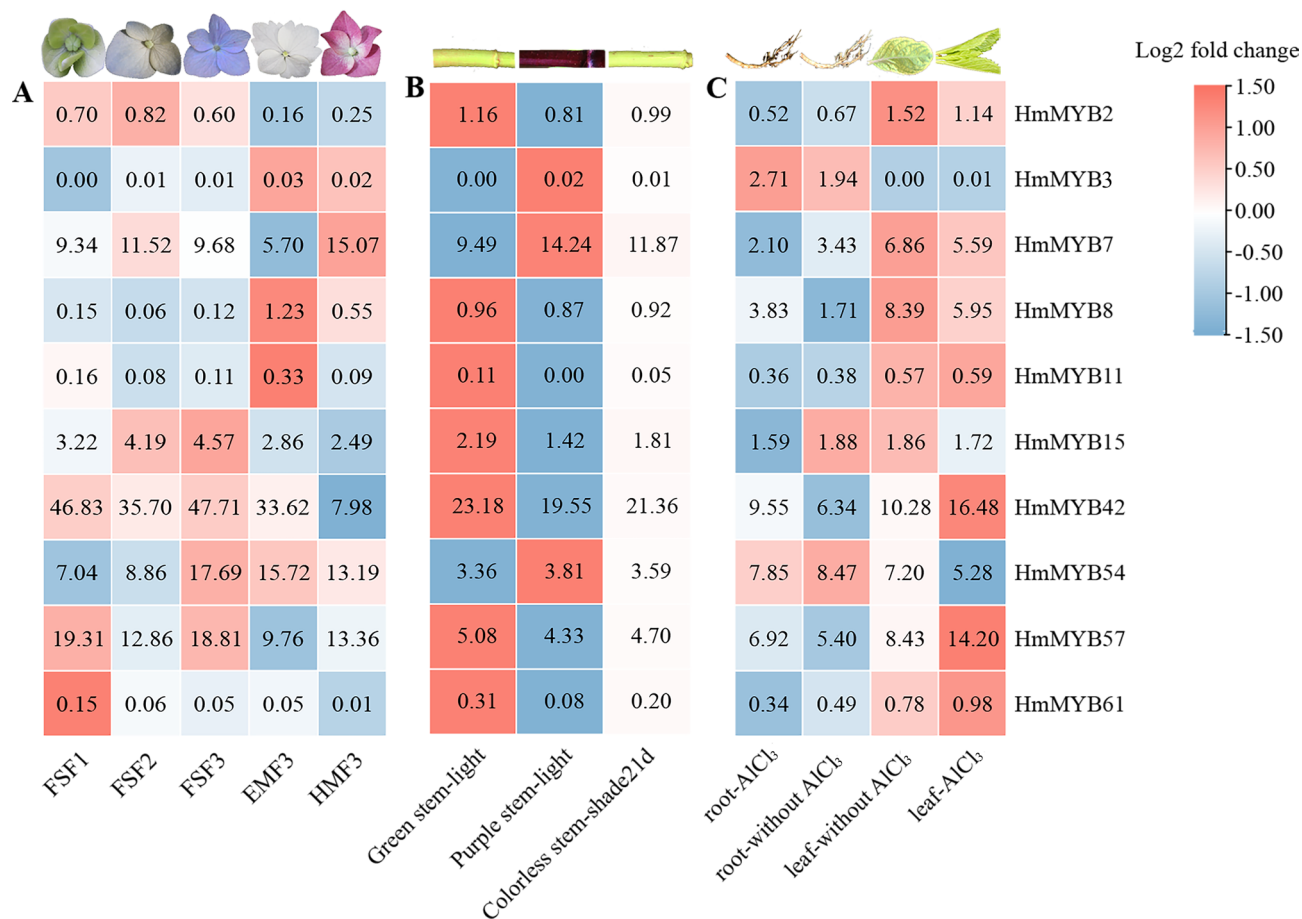


Fig. 5. Expression thermogram of R2R3-MYB transcription factor in hydrangea⁷⁷. Note: A hydrangea 'Forever Summer' bud stage (FSF1), discoloration stage (FSF2), flowering stage (FSF3), 'Emilia' flowering stage (EMF3) and 'FenDai' flowering stage (HMF3); B green stems, purple stems and stems treated in darkness for 21 days; C roots and leaves treated with AlCl₃.

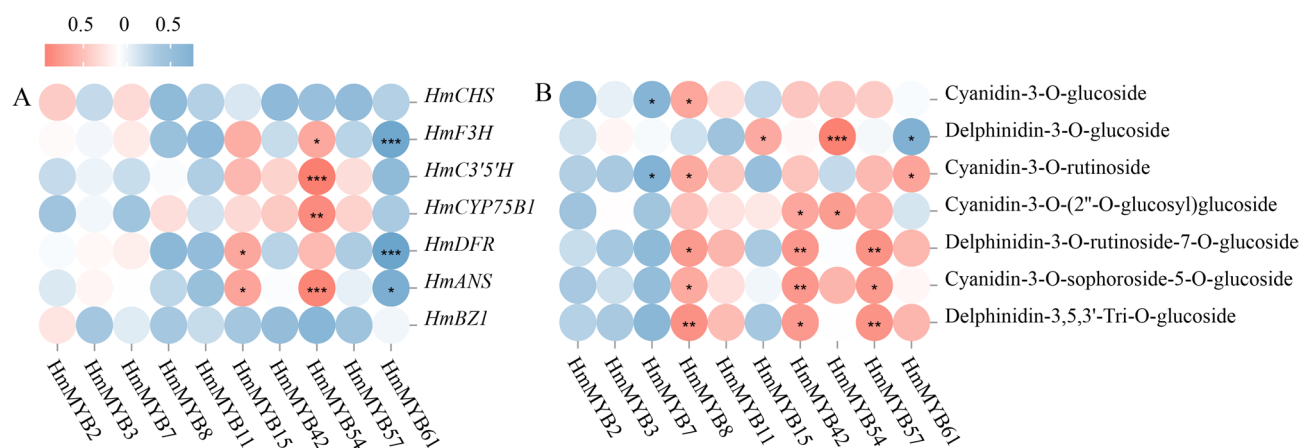


Fig. 6. Correlation thermogram of R2R3-MYB transcription factor with anthocyanins and key structural genes in hydrangea. Note: A key structural genes of anthocyanin synthesis pathway; B delphinidin-3-O-glucoside and other seven anthocyanins. * means $0.01 < P < 0.05$; ** means $0.001 < P < 0.01$; *** means $P \leq 0.001$.

positive correlation with delphinidin-3-O-glucoside, at the level of $P \leq 0.001$. Additionally, HmMYB42 was positively correlated with delphinidin-3-O-rutinoside-7-O-glucoside and cyanidin-3-O-sophoricoside-5-O-glucoside at the level of $P < 0.01$. HmMYB57 was positively correlated with delphinidin-3-O-rutinoside-7-O-glucoside and delphinidin-3,5,3'-tri-O-glucoside. HmMYB7 was negatively correlated with cyanidin-3-O-glucoside and cyanidin-3-O-rutinoside at the level of $P < 0.05$. HmMYB61 was negatively correlated with delphinidin-3-O-glucoside and positively correlated with cyanidin-3-O-rutinoside. In addition, the correlation between the expression levels of R2R3-MYB transcription factors and key structural genes involved in anthocyanin synthesis was analyzed (Fig. 6B). The results showed that HmMYB15 was positively correlated with *HmDFR* and *HmANS* at $P < 0.05$. HmMYB54 was correlated with *HmF3H* at $p < 0.05$, with *HmCYP75B1* at $P < 0.01$, with *HmC3'5'H* and *HmANS* at $P \leq 0.001$, with HmMYB61 negatively correlated with *HmANS* at $P < 0.05$ and with *HmF3H* and *HmDFR* at $P \leq 0.001$.

Expression analysis of 6 R2R3-MYB transcription factors related to color in hydrangea

In order to deeply explore the role of S5 subfamily genes in hydrangea flower color regulation, the genes *HmMYB3*, *HmMYB11*, *HmMYB42*, *HmMYB54*, *HmMYB57* and *HmMYB61* expression was verified in Hydrangea 'Forever Summer' bud stage (FSF1), discoloration stage (FSF2), flowering stage (FSF3), 'Emilia' flowering stage (EMF3) and 'Fendai' flowering stage (HMF3) by qPCR experiment. The qPCR detection data and the analysis results of RNAseq were integrated and compared (Fig. 7). The experimental results show that HmMYB11, HmMYB42 and HmMYB54 in the two groups of data have a correlation at $P < 0.05$, while HmMYB61 has a significant correlation at $p < 0.0001$.

From the expression level, six genes at different developmental stages are highly consistent with the previous transcriptome sequencing data. Specifically, the expression levels of qPCR and RNAseq of *HmMYB3* in three hydrangea varieties are very low, with little difference, and the expression of gene *HmMYB11* is the highest at the flowering stage (EMF3) in Emilia. In addition, three genes, *HmMYB42*, *HmMYB54* and *HmMYB57*, have the highest expression at the flowering stage of hydrangea (FSF3). It is worth noting that the expression level of gene *HmMYB54* in Forever Summer sterile flowers showed a significant upward trend with the continuous accumulation of flower color substances, which probably participated in the positive regulation process of flower color synthesis in Forever Summer sterile flowers; On the contrary, the expression of *HmMYB11* and *HmMYB61* decreased gradually with the accumulation of flower color, which may play a negative regulatory role in the process of flower color formation in Forever Summer infertility.

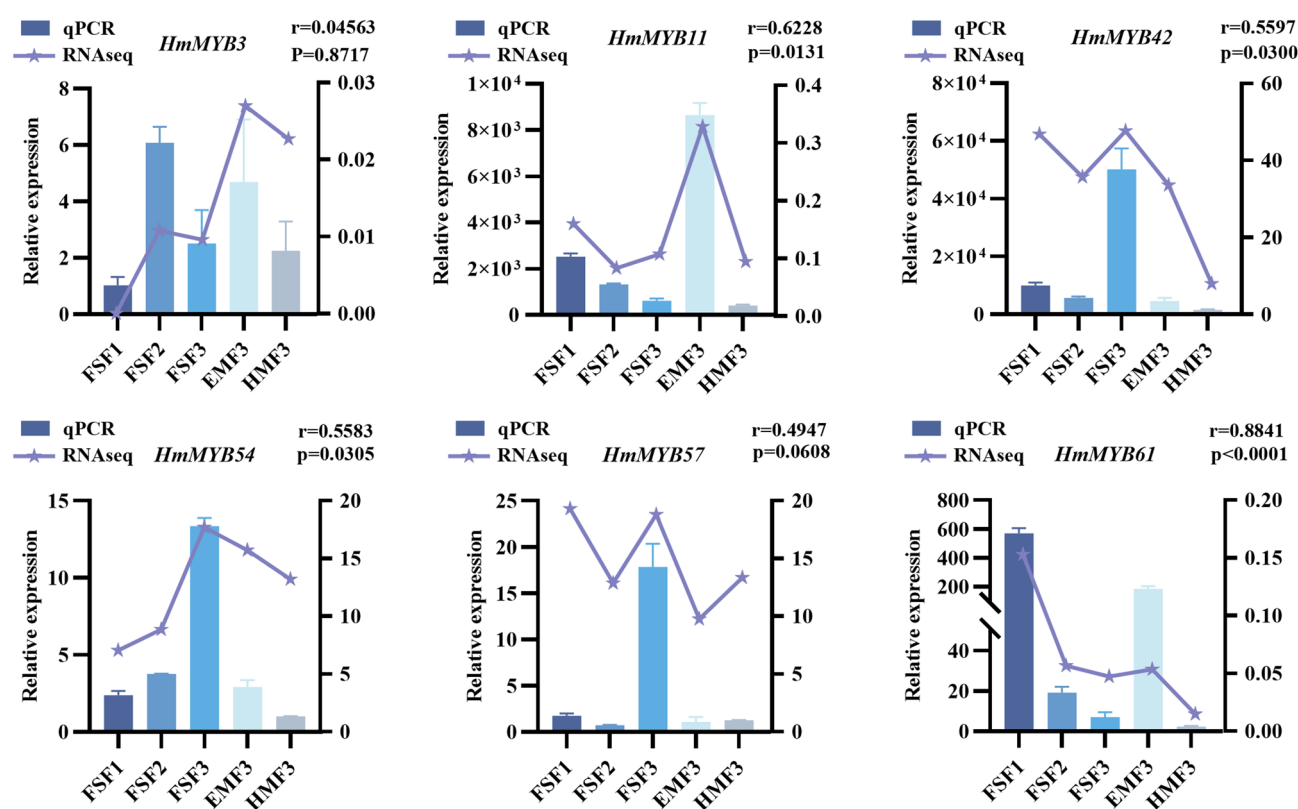


Fig. 7. Gene expression of S5 subfamily in hydrangea. Note: hydrangea Forever Summer bud stage (FSF1), discoloration stage (FSF2), flowering stage (FSF3), Emilia flowering stage (EMF3) and Fendai flowering stage (HMF3). The bar chart represents qPCR, and the line chart represents RNAseq.

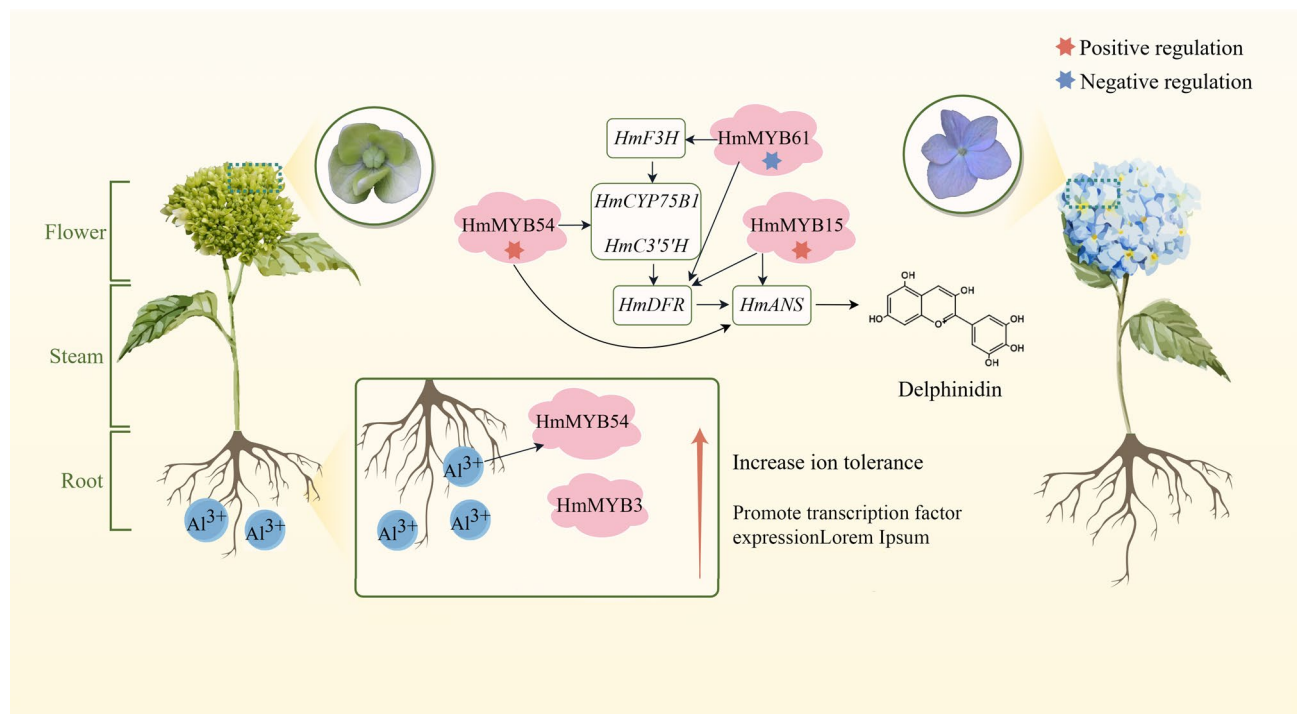


Fig. 8. Formation process of sterile flower color of hydrangea regulation mode. Note: The upper part of the picture shows that MYB transcription factor affects the formation of color by regulating the key genes of anthocyanin synthesis; The second part shows the response mechanism of roots under Al^{3+} stress, which can enhance ion tolerance and promote flower color formation by increasing the expression of MYB transcription factor.

Prediction formation of the regulatory model of R2R3-MYB transcription factors in hydrangea

According to the correlation between transcription factor R2R3-MYB of hydrangea and anthocyanin expression, as well as the correlation analysis of key structural genes of anthocyanin synthesis, the flower color formation mechanism of hydrangea is predicted (Fig. 8). Taking hydrangea Forever Summer as an example, during the flower color formation process, *HmMYB61* negatively regulates *HmF3H* and *HmDFR* to influence anthocyanin synthesis; *HmMYB15* positively regulates the synthesis of *HmDFR* and *HmANS* and affects the accumulation of delphinidin 3-O- glucoside; *HmMYB54* positively regulates the synthesis of *HmF3H*, *HmCYP75B1*, *HmC3'5'H*, and *HmANS*, with the greatest influence on *HmC3'5'H* and *HmANS*, and then affects the accumulation of anthocyanins in hydrangea, especially the accumulation of delphinidin -3-O- glucoside, which makes the sterile flowers of hydrangea appear blue.

Under Al^{3+} stress, the expression of *HmMYB54* and *HmMYB3* in the root of hydrangea increased to adapt to ions. A large number of studies have confirmed that aluminum ions play a key role in the formation of blue pigment in hydrangea sepals, which is closely related to the unique physiological and biochemical characteristics and complex molecular regulatory network of hydrangea. The bluing process of hydrangea sepals is essentially a dynamic process coordinated by environmental factors and gene expression. As an important environmental signal molecule, Al^{3+} can affect the synthesis and stability of anthocyanins through a series of complex physiological and biochemical reactions²⁶. *HmMYB54* is positively regulated in the process of endless summer flower color formation of hydrangea, and Al^{3+} may promote the expression of *HmMYB54* gene through many potential ways, and finally regulate the accumulation of blue anthocyanins in hydrangea.

Discussion

The R2R3-MYB transcription factors, as the largest subfamily within the MYB family, are recognized for their multifaceted roles in plant biology, encompassing diverse functions related to primary and secondary metabolism, growth and development, and responses to environmental stresses. These transcription factors are involved in the biosynthesis of critical metabolites, including glucosinolates (GSL)³¹ anthocyanins^{32–35} and phenylpropanoids^{36,37} as well as the synthesis of epicuticular wax³⁸ carotenoids³⁹ and capsaicinoids⁴⁰. Additionally, they play vital roles in cellular processes such as leaf senescence⁴¹ root development^{42,43} and pollen maturation⁴⁴. Current studies have documented the presence of R2R3-MYB transcription factors in various species, including *Arabidopsis thaliana*⁴⁵ cucumber⁴⁶ Japanese morning glory⁴⁷ and mung bean⁴⁸. Notably, *Arabidopsis* possesses 125 identified R2R3-MYB members⁴⁵ while 55 members exist in cucumber⁴⁶ 127 in Japanese petunia⁴⁷ and 168 in mung bean⁴⁸. These transcription factors are characterized by their unique molecular structures, which include distinct R2 and R3 repeats essential for their DNA-binding capabilities. The

functional properties of R2R3-MYB proteins are governed by specific activation and repression domains, which modulate their regulatory functions across various biosynthetic pathways, particularly those associated with flavonoids and secondary metabolites⁴⁹.

Recent structural studies have elucidated the interactions of R2R3-MYB proteins with target DNA, highlighting specific residues that are critical for binding affinity and regulatory activity. These interactions are influenced by methylation states, which can inhibit DNA binding⁵⁰. Furthermore, R2R3-MYB proteins engage in complex interactions with other transcription factors, such as bHLH and WD40, forming regulatory complexes that intricately fine-tune gene expression patterns through both direct and indirect mechanisms⁵¹. The regulatory effects exerted by R2R3-MYB proteins include direct activation or repression of target genes, as well as indirect modulation via protein-protein interactions^{52,53}. Insights gained from molecular dynamics simulations have further revealed the importance of hydrophobic and hydrogen bonding interactions in maintaining the stability and functional integrity of these complexes during DNA binding⁵⁰. The evolutionary expansion and diversification of the R2R3-MYB family across multiple plant species underscore their significance and suggest that these transcription factors play critical roles in the adaptation and innovation of various plant traits^{51,52}. This synthesis of findings emphasizes the vital contributions of R2R3-MYB transcription factors to plant biology and their potential applications in understanding plant adaptation and trait enhancement.

The investigation into the evolutionary relationships of R2R3-MYB transcription factors in *H. macrophylla* and Arabidopsis was conducted through the analysis of assembled and annotated genomic data from *H. macrophylla*. This study identified a total of 72 MYB transcription factors, which were subsequently compared to the 133 MYB members identified in Arabidopsis to construct a comprehensive phylogenetic tree. Based on the established classification criteria for R2R3-MYB transcription factors in Arabidopsis, 13 distinct R2R3-MYB transcription factors were identified within *H. macrophylla*, distributed among subfamilies 2, 5, 13, 15, and 23. Notably, the R2R3-MYB family in *H. macrophylla* exhibits a reduced number of members relative to other plant species, suggesting a unique evolutionary trajectory. Extensive research has explored the evolutionary dynamics of R2R3-MYB transcription factors, underscoring the intricate relationships established through gene duplication and functional diversification across a variety of plant taxa. It has been demonstrated that significant expansions of R2R3-MYB genes are primarily attributed to whole-genome duplications, in addition to subsequent smaller duplicative events, which have collectively contributed to their functional diversity⁵⁴. A pertinent example of this diversification can be observed in Arabidopsis, where paralogous MYB genes, such as WEREWOLF and GLABROUS1, have been shown to define distinct cell types while maintaining functionally equivalent roles. This illustrates that alterations in cis-regulatory sequences can lead to novel functions without necessitating changes in protein structure⁵⁵. Moreover, the comparative analysis of MYB gene families across different taxa, including bHLH genes, has revealed that these expansions are closely associated with the morphological and metabolic complexities of higher plants. This association suggests that the evolution of these transcription factors is a direct response to selective pressures encountered in their respective environments⁵⁶. Further studies examining the phylogenetic relationships among various MYB subfamilies, including 1R-MYB, 3R-MYB, and 4R-MYB, have provided insights into the conservation and divergence of these transcription factors⁵⁷. The evolutionary pathways of these subfamilies imply that they have been shaped by specific adaptive radiations, enabling them to assume distinct functional roles in processes such as trichome morphogenesis and stress responses⁵⁸. These findings collectively emphasize the significance of gene duplication events and the functional plasticity of MYB transcription factors, which enhance the capacity of plants to adapt to diverse environmental conditions. Ultimately, synthesizing data from an evolutionary perspective on R2R3-MYB transcription factors reveals a consistent pattern of expansion and functional diversification, facilitating the emergence of specialized roles in plant biology. Understanding this ongoing evolutionary process is crucial for elucidating how plants have adapted to their environments over time.

The intricate functions of R2R3-MYB transcription factors, particularly their roles in regulating anthocyanin biosynthesis and other secondary metabolites, highlight their significance within various metabolic pathways. Genes within the same subfamily generally exhibit similar functions⁵⁹. For instance, AtMYB123 (TT2), the sole representative of subfamily 5 among R2R3-MYB transcription factors, encodes a protein crucial for proanthocyanidin biosynthesis in the seed coat of *Arabidopsis*⁶⁰ and is known to enhance flavonoid production in peaches⁶¹. Within subfamily 5, *H. macrophylla* genes HmMYB3, HmMYB11, HmMYB42, HmMYB61, HmMYB54, and HmMYB57 cluster together, suggesting their potential roles in anthocyanin or proanthocyanidin synthesis pathways⁶². Additionally, HmMYB4 and HmMYB19, alongside HmMYB2, HmMYB7, and HmMYB8, are positioned in adjacent branches, possessing similar conserved motifs that indicate possible functional similarities.

Research on flower color formation in *H. macrophylla* suggests that the transcription factor HmWER-like exerts a negative regulatory effect on the coloration of sterile flowers in the blue hydrangea cultivar 'Forever Summer'¹. Notably, genes such as HmDFR⁶³ and HmF3H⁶⁴ are thought to play pivotal roles in the color variations observed among different *H. macrophylla* cultivars. Under specific cultivation conditions, the interaction of aluminum ions (Al³⁺), delphinidin-3-glucoside, and copigments has been shown to facilitate a color transformation in the hydrangea sepals from red to blue⁶⁵. Despite these findings, there remains a scarcity of research focused on the identification and functional analysis of R2R3-MYB genes in *H. macrophylla*. This study aimed to bridge this gap by conducting heatmap analyses on transcriptomic data obtained from sterile flowers across various cultivars and developmental stages. The clustering of HmMYB54, HmMYB61, HmMYB42, and HmMYB57 within subfamily 5 indicates their potential association with anthocyanin formation. The investigation sought to elucidate the correlation between *H. macrophylla* MYB transcription factors, anthocyanin content, and key genes implicated in anthocyanin synthesis, ultimately aiming to predict the regulatory mechanisms that underpin flower color formation.

A comprehensive examination of R2R3-MYB transcription factors reveals their integral role in anthocyanin biosynthesis regulation, emphasizing their interaction with various metabolic pathways. The mechanisms governing anthocyanin production have been characterized extensively, highlighting the activation of R2R3-MYB factors modulated by environmental stimuli and internal signals, thereby influencing color variations in horticultural plants⁶⁶. These transcription factors engage with other regulatory proteins, such as bHLH transcription factors, to establish a complex regulatory network that balances the accumulation of anthocyanins and proanthocyanidins⁶⁷. Furthermore, investigations have demonstrated the involvement of MYB transcription factors in the biosynthesis of flavonoids and lignins, elucidating how these secondary metabolic pathways integrate with primary metabolism⁶⁸.

The functional repertoire of R2R3-MYB transcription factors extends to developmental processes, where they participate in cell cycle regulation and organ-specific developmental functions, interfacing with hormone signaling pathways to modulate growth responses⁶⁹. Additionally, these transcription factors are essential for enhancing plant resilience against abiotic stressors and mediating plant-pathogen interactions, showcasing their dual capacity to respond to environmental challenges and biotic threats^{68,70}. Comparative analyses among different species illustrate the unique adaptations of R2R3-MYB functions in crop plants versus model organisms, underlining their evolutionary development in response to specific ecological niches⁷¹. The complexity of transcriptional and post-transcriptional regulation of R2R3-MYB genes encompasses various mechanisms, including epigenetic modifications that influence protein stability and turnover, thereby affecting metabolic outcomes⁶⁹. In the context of biotechnological applications, R2R3-MYB transcription factors have emerged as crucial tools in genetic engineering aimed at enhancing crop traits such as stress tolerance, nutritional quality, and ornamental characteristics, signifying substantial potential for advancements in synthetic biology and molecular breeding strategies^{68,72}.

To date, research on the MYB gene family, particularly concerning its functional characterization, has predominantly focused on select members, especially within the model organism *Arabidopsis thaliana*. Studies have illustrated the significant roles of MYB transcription factors in regulating anthocyanin biosynthesis across various plant species, providing insights applicable to non-model flowering plants such as *H. macrophylla*. For instance, R2R3-MYB transcription factors are highlighted as crucial regulators of anthocyanin synthesis, playing pivotal roles in the transcriptional regulation of biosynthetic genes⁶⁹. In the Rosaceae family, MYB genes, particularly MYB10 and MYBA, have been shown to control anthocyanin levels in apple fruits, with functional studies indicating that these transcription factors can induce anthocyanin pathways in heterologous systems, such as tobacco and strawberries⁷³. Similarly, other investigations revealed that MYB112 in *Arabidopsis* enhances anthocyanin accumulation under stress conditions, further emphasizing the complex regulatory networks involving MYB factors⁷⁴. Moreover, the investigation into petunia has revealed that members like DEEP PURPLE and PURPLE HAZE are instrumental in both floral and vegetative pigmentation, demonstrating the adaptability of MYB regulators in different contexts⁷⁵. These findings underscore the diverse functions and evolutionary significance of MYB transcription factors across flowering plants. Functional analyses conducted in various species suggest that the modulation of anthocyanin biosynthesis by MYBs can be leveraged for enhancing color quality in ornamental and food crops, as seen in the manipulation of MYB119 in poplar to boost anthocyanin levels without detriment to plant growth⁷⁶. The ongoing exploration of these transcription factors not only sheds light on the genetic basis of pigmentation but also paves the way for future research into the regulatory mechanisms governing anthocyanin synthesis in non-model plants like *H. macrophylla*, which remains underexplored in the context of MYB functionality.

Conclusion

The findings from this study significantly advance the understanding of the genetic mechanisms underlying flower color variation in *H. macrophylla*. The identification of 72 MYB genes, particularly the roles of R2R3-MYB transcription factors such as HmMYB15 and HmMYB54, provides a robust framework for further exploration of anthocyanin biosynthesis pathways. These transcription factors not only correlate with anthocyanin accumulation during flowering but also exhibit responsiveness to environmental stressors, suggesting that they play a dual role in both ornamental traits and adaptive strategies. The correlation between MYB gene expression and anthocyanin levels points to a complex regulatory network that governs flower coloration, integrating genetic, environmental, and developmental factors. By situating these results within the context of existing literature, it becomes evident that MYB transcription factors are critical not just for *H. macrophylla*, but potentially for a wide range of plant species, as they contribute to phenotypic plasticity in response to changing environments. This adaptive capacity is essential for survival and reproduction, particularly in the face of climate change and habitat alterations. Additionally, the proposed regulatory model highlights the intricate interactions between MYB genes and structural genes involved in anthocyanin biosynthesis, thereby offering potential pathways for targeted genetic manipulation to enhance desirable traits in ornamental horticulture. Future research should aim to further dissect these interactions, perhaps exploring how external factors modulate MYB gene expression and flower color dynamics. In summary, this study not only enhances our comprehension of flower color genetics in *H. macrophylla* but also underscores the broader ecological significance of MYB transcription factors in plant adaptation and resilience.

Data availability

The data supporting this study's findings are available from the corresponding author upon reasonable request.

Received: 2 January 2025; Accepted: 30 July 2025

Published online: 09 August 2025

References

- Peng, J. Q. Study on the regulation mechanism of pigment formation in sterile flowers of hydrangea. Dissertation. Central South University of Forestry and Technology. (2021). <https://doi.org/10.27662/d.cnki.gznlc.2021.000858>
- Shi, C. C. & Mining R2R3-MYBs, a key gene for the formation of hyacinthus orientalis blue color in grapes, based on omics analysis and its function research. Henan Agricultural University. (2024). <https://doi.org/10.27117/d.cnki.ghenu.2024.000482>
- Meng, Y. Analysis of Petal Pigment Composition of Genus Eleocharis and Its Influence on Flower Color. Zhejiang A&F University. (2024). <https://doi.org/10.27756/d.cnki.gzjlx.2024.000230>
- Chen, K. L., Ban, W. Z., Du, L. J., Zhou, X. H. & Zhang, X. F. Flavonoids analysis and identification of key genes of flower color formation in petals of blue star water Lily. *J. Southwest. Univ. (Natural Sci. Edition)*. **46**, 71–80. <https://doi.org/10.13718/j.cnki.xdzk.2024.02.008> (2024).
- Zhuang, W. B. et al. Molecular regulation mechanism of anthocyanin biosynthesis and coloration in plants. *Plant Physiol.* **54**, 1630–1644. <https://doi.org/10.13592/j.cnki.ppj.2018.0231> (2018).
- Akashi, T. et al. Molecular cloning and biochemical characterization of a novel cytochrome P450, flavone synthase II, that catalyzes direct conversion of flavanones to flavones. *Plant. Cell. Physiol.* **40**, 1182–1187. <https://doi.org/10.1093/oxfordjournals.pcp.a029505> (1999).
- Yang, J. et al. Genome-wide identification and expression analysis of the R2R3-MYB gene family in tobacco (*Nicotiana tabacum* L.). *BMC Genom.* **23**, 432. <https://doi.org/10.1186/s12864-022-08658-7> (2022).
- Li, X. J. et al. AcMYB96 promotes anthocyanin accumulation in onion (*Allium Cepa* L.) without forming the MBW complex. *Plant. Physiol. Biochem.* **214**, 108965. <https://doi.org/10.1016/j.plaphy.2024.108965> (2024).
- Martinez-Sanchez, M. et al. SmuMYB113 is the determinant of fruit color in Pepino (*Solanum muricatum*). *Front. Plant. Sci.* **15**, 1408202. <https://doi.org/10.3389/fpls.2024.1408202> (2024).
- Song, G. M. et al. Identification and expression analysis of R2R3-MYB transcription factors associated with flavonoid biosynthesis in *Panax quinquefolius*. *Int. J. Mol. Sci.* **25**, 3709. <https://doi.org/10.3390/ijms25073709> (2024).
- Dubos, C. et al. MYB transcription factors in Arabidopsis. *Trends Plant. Sci.* **15** <https://doi.org/10.1016/j.tplants.2010.06.005> (2010). 573–81.
- Wang, X. C. et al. MYB4 plays dual roles in flavonoid biosynthesis. *Plant. J.* **101**, 637–652. <https://doi.org/10.1111/tpj.14570> (2020).
- Qian, J. Y. et al. Screening of MYB transcription factors related to anthocyanin synthesis in Zinnia and study on the function of ZeMYB9. *J. Hort.* **49**, 1505–1518. <https://doi.org/10.16420/j.issn.0513-353x.2021-0474> (2022).
- Xie, N. et al. Integrated transcriptomic and WGCNA analyses reveal candidate genes regulating mainly flavonoid biosynthesis in *litsea Coreana* var. *Sinensis*. *BMC Plant. Biol.* **24**, 231. <https://doi.org/10.1186/s12870-024-04949-1> (2024).
- Li, M. Z. et al. Functional characterization of tea (*Camellia sinensis*) MYB4a transcription factor using an integrative approach. *Front. Plant. Sci.* **8**, 943. <https://doi.org/10.3389/fpls.00943> (2017).
- Zhou, H. et al. R2R3-MYB transcription factor CaMYB5 regulates anthocyanin biosynthesis in pepper fruits. *Int. J. Biol. Macromol.* **28**, 308. <https://doi.org/10.1016/j.ijbiomac.2025.142450> (2025).
- Li, M. et al. The mutations in RcMYB114 affect anthocyanin glycoside accumulation in Rose. *Biology (Basel)*. **14** <https://doi.org/10.3390/biology14030258> (2025).
- Huang, R. et al. Integrative analysis of the R2R3-MYB gene family revealed that BsMYB36 and BsMYB51 significantly regulate the accumulation of flavonoids in *Bletilla striata* (Orchidaceae). *Plant. Physiol. Biochem.* **222**, 109733. <https://doi.org/10.1016/j.plaphy.2025.109733> (2025).
- Xiong, L. et al. Cloning and expression analysis of color-related gene HmC3'5'H in hydrangea. *Mol. Plant. Breed.* **10**, 1–12 (2024). <http://kns.cnki.net/kcms/detail/46.1068.S.20230718.1756.019.html>
- Tang, Q. Q. et al. Hydrangea genome-wide association study. *J. Plant. Genetic Resour.* **24**, 1174–1185. <https://doi.org/10.13430/j.cnki.jpgr.20230208004> (2023).
- Long, S. Y. Effects of ALA on physiological characteristics of hydrangea under high temperature and drought stress. *Cent. South. Univ. Forestry Technol.* **10**, 63–69. <https://doi.org/10.11937/bfyy.20214364> (2022).
- Peng, J. Q., Dong, X. J., Xue, C., Liu, Z. M. & Cao, F. X. Exploring the molecular mechanism of blue flower color formation in *H. macrophylla* cv. *Forever summer*. *Front. Plant Sci.* **12**, 585665. <https://doi.org/10.3389/fpls.2021.585665> (2021).
- Chen, H. et al. Integrative transcriptomics and proteomics elucidate the regulatory mechanism of *H. macrophylla* Flower-Color changes induced by exogenous aluminum. *Agronomy* **12**, 969. <https://doi.org/10.3390/agronomy12040969> (2022).
- Rahmati, R. et al. Comparative transcriptome analysis unveils the molecular mechanism underlying sepal colour changes under acidic pH substratum in *H. macrophylla*. *Int. J. Mol. Sci.* **23**, 15428. <https://doi.org/10.3390/ijms232315428> (2022).
- Anaya-Covarrubias, J. Y. et al. Hydrangea DNA methylation caused by pH substrate changes to modify sepal colour is detected by MSAP and ISSR markers. *Agronomy* **9**, 871. <https://doi.org/10.3390/agronomy9120871> (2019).
- Yoshida, K., Oyama, K. I. & Kondo, T. Insight into chemical mechanisms of sepal color development and variation in hydrangea. *Proc. Jpn Acad. Ser. B Phys. Biol. Sci.* **97**, 51–68. <https://doi.org/10.2183/pjab.97.003> (2021).
- Wang, Q., Lyu, T. & Lyu, Y. M. Exploring the molecular mechanism of sepal formation in the decorative flowers of *H. macrophylla* 'forever summer' based on the ABCDE model. *Int. J. Mol. Sci.* **23**, 14112. <https://doi.org/10.3390/ijms232214112> (2022).
- Zhang, X. et al. Identification of two novel R2R3-MYB transcription factors, PsMYB114L and PsMYB12L, related to anthocyanin biosynthesis in *paeonia suffruticosa*. *Int. J. Mol. Sci.* **20**, 1055. <https://doi.org/10.3390/ijms20051055> (2019).
- Akagi, T. et al. DkMyb4 is a Myb transcription factor involved in Proanthocyanidin biosynthesis in persimmon fruit. *Plant. Physiol.* **151**, 2028–2045. <https://doi.org/10.1104/pp.109.146985> (2009).
- Huang, H. Cloning and functional study of color-related transcription factors CjMYB1 and CjUGT94E5 in camellia. *Chin. Acad. Forestry.* <https://doi.org/10.27625/d.cnki.gzlk.20010.000000000625> (2022).
- Seo, M. S. & Kim, J. S. Understanding of MYB transcription factors involved in glucosinolate biosynthesis in brassicaceae. *Molecules* **22**, 1549. <https://doi.org/10.3390/molecules22091549> (2017).
- Yang, Y. T. et al. The pea R2R3-MYB gene family and its role in anthocyanin biosynthesis in flowers. *Front. Genet.* **13**, 936051. <https://doi.org/10.3389/fgene.2022.936051> (2022).
- Sumbur, B. et al. The characterization of R2R3-MYB genes in *ammopiptanthus nanus* uncovers that the miR858-AnaMYB87 module mediates the accumulation of anthocyanin under osmotic stress. *Biomolecules* **13**, 1721. <https://doi.org/10.3390/biom13121721> (2023).
- Chen, D. Z. et al. Genome-wide identification of R2R3-MYB gene family and association with anthocyanin biosynthesis in brassica species. *BMC Genom.* **23**, 441. <https://doi.org/10.1186/s12864-022-08666-7> (2022).
- Duan, A. Q., Tan, S. S., Deng, Y. J., Xu, Z. S. & Xiong, A. S. Genome-Wide identification and evolution analysis of R2R3-MYB gene family reveals S6 subfamily R2R3-MYB transcription factors involved in anthocyanin biosynthesis in Carrot. *Int. J. Mol. Sci.* **23**, 11859. <https://doi.org/10.3390/ijms231911859> (2022).
- Li, J. Y. et al. Systematic analysis of the R2R3-MYB family in camellia sinensis: evidence for galloylated catechins biosynthesis regulation. *Front. Plant. Sci.* **12**, 782220. <https://doi.org/10.3389/fpls.2021.782220> (2022).
- Singh, V., Kumar, N., Dwivedi, A. K., Sharma, R. & Sharma, M. K. Phylogenomic analysis of R2R3 MYB transcription factors in sorghum and their role in conditioning biofuel syndrome. *Curr. Genomics.* **21**, 138–154. <https://doi.org/10.2174/1389202921666200326152119> (2020).
- Ahmad, H. M. et al. Comparative genomic analysis of MYB transcription factors for cuticular wax biosynthesis and drought stress tolerance in *Helianthus annuus* L. *Saudi J. Biol. Sci.* **28**, 5693–5703. <https://doi.org/10.1016/j.sjbs.2021.06.009> (2021).

39. Yin, Y. et al. Genome-Wide comparative analysis of the R2R3-MYB gene family in five Solanaceae species and identification of members regulating carotenoid biosynthesis in wolfberry. *Int. J. Mol. Sci.* **23**, 2259. <https://doi.org/10.3390/ijms23042259> (2022).
40. Wang, J. et al. Genome-Wide identification and capsacinoid Biosynthesis-Related expression analysis of the R2R3-MYB gene family in capsicum annum L. *Front. Genet.* **11**, 598183. <https://doi.org/10.3389/fgene.2020.598183> (2020).
41. Li, X. X. et al. Systematic analysis of MYB family genes in potato and their multiple roles in development and stress responses. *Biomolecules* **9**, 317. <https://doi.org/10.3390/biom9080317> (2019).
42. Ahmad, I., Soni, S. K., Muthukumar, M. & Pandey, D. In-silico mining and characterization of MYB family genes in wilt-resistant hybrid guava (Psidium Guajava × psidium molle). *J. Genet. Eng. Biotechnol.* **21**, 74. <https://doi.org/10.1186/s43141-023-00528-3> (2023).
43. Lv, X. M. et al. Genome-wide identification of the Longan R2R3-MYB gene family and its role in primary and lateral root. *BMC Plant Biol.* **23**, 448. <https://doi.org/10.1186/s12870-023-04464-9> (2023).
44. Zhou, F. W., Chen, Y. N., Wu, H. T. & Yin, T. M. Genome-Wide comparative analysis of R2R3 MYB gene family in Populus and Salix and identification of male flower bud Development-Related genes. *Front. Plant Sci.* **12**, 721558. <https://doi.org/10.3389/fpls.2021.721558> (2021).
45. Stracke, R., Werber, M. & Weissshaar, B. The R2R3-MYB gene family in Arabidopsis Thaliana. *Curr. Opin. Plant Biol.* **4**, 447–456. [https://doi.org/10.1016/s1369-5266\(00\)00199-0](https://doi.org/10.1016/s1369-5266(00)00199-0) (2001).
46. Li, Q., Zhang, C. J., Li, J., Wang, L. N. & Ren, Z. H. Genome-wide identification and characterization of R2R3MYB family in cucumis sativus. *PLoS One*. **7**, 47576. <https://doi.org/10.1371/journal.pone.0047576> (2012).
47. Komatsuzaki, A., Hoshino, A., Otagaki, S., Matsumoto, S. & Shiratake, K. Genome-wide analysis of R2R3-MYB transcription factors in Japanese morning glory. *PLoS One*. **17** (0271012). <https://doi.org/10.1371/journal.pone.0271012> (2022).
48. Guo, F. X. et al. Identification of transcription factor family of mung bean R2R3-MYB and screening of flavonoid synthesis regulation genes. *Acta Agron. Sinica*. **10**, 1–26 (2024). <http://kns.cnki.net/kcms/detail/11.1809.s.20240826.1614.004.html>
49. Heine, G. F., Hernandez, J. M. & Grotewold, E. Two cysteines in plant R2R3 MYB domains participate in REDOX-dependent DNA binding. *J. Biol. Chem.* **279**, 37878–37885. <https://doi.org/10.1074/jbc.M405166200> (2004).
50. Wang, B. H. et al. Structural insights into target DNA recognition by R2R3-MYB transcription factors. *Nucleic Acids Res.* **48**, 460–471. <https://doi.org/10.1093/nar/gkz1081> (2019).
51. Wilkins, O., Nahal, H., Foong, J., Provart, N. J. & Campbell, M. M. Expansion and diversification of the Populus R2R3-MYB family of transcription factors. *Plant Physiol.* **149**, 981–993. <https://doi.org/10.1104/pp.108.132795> (2008).
52. Jiang, C. K. & Rao, G. Y. Insights into the diversification and evolution of R2R3-MYB transcription factors in plants. *Plant Physiol.* **183**, 637–655. <https://doi.org/10.1104/pp.19.01082> (2020).
53. Hai, D. et al. The evolutionary history of R2R3-MYB proteins across 50 eukaryotes: new insights into subfamily classification and expansion. *Sci. Rep.* **5**, 11037. <https://doi.org/10.1038/srep11037> (2015).
54. Wu, Y., Wen, J., Xia, Y. P. & Hai, D. Evolution and functional diversification of R2R3-MYB transcription factors in plants. *Hortic. Res.* **9**, 58. <https://doi.org/10.1093/hr/uhac058> (2022).
55. Lee, M. M. & Schiefelbein, J. Developmentally distinct MYB genes encode functionally equivalent proteins in Arabidopsis. *Development* **128**, 1539–1546. <https://doi.org/10.1242/dev.128.9.1539> (2001).
56. Feller, A., Machemer, K., Braun, E. L. & Grotewold, E. Evolutionary and comparative analysis of MYB and bHLH plant transcription factors. *Plant J.* **66**, 94–116. <https://doi.org/10.1111/j.1365-313X.2010.04459.x> (2011).
57. Kirik, V. et al. Functional diversification of MYB23 and GL1 genes in trichome morphogenesis and initiation. *Development* **132**, 1477–1485. <https://doi.org/10.1242/dev.01708> (2005).
58. Zapata, F. & Fine, P. V. Diversification of the monoterpene synthase gene family (TPSb) in protium, a highly diverse genus of tropical trees. *Mol. Phylogenet. Evol.* **68**, 432–442. <https://doi.org/10.1016/j.ympev.2013.04.024> (2013).
59. Wu, S. et al. Identification and expression pattern analysis of MYB gene family in Luffa. *Chin. Herb. Med.* **55**, 4159–4170. <https://doi.org/10.7501/j.issn.0253-2670.2024.12.023> (2024).
60. Baudry, A. et al. TT2, TT8, and TTG1 synergistically specify the expression of BANYULS and Proanthocyanidin biosynthesis in Arabidopsis Thaliana. *Plant J.* **39** <https://doi.org/10.1111/j.1365-313X.2004.02138.x> (2004). 366–80.
61. Ravaglia, D. et al. Transcriptional regulation of flavonoid biosynthesis in nectarine (Prunus persica) by a set of R2R3 MYB transcription factors. *BMC Plant Biol.* **13**, 68. <https://doi.org/10.1186/1471-2229-13-68> (2013).
62. Hai, D. et al. Biochemical and molecular characterization of plant MYB transcription factor family. *Biochem. (Mosc)*. **74**, 1–11. <https://doi.org/10.1134/s0006297909010015> (2009).
63. Chen, D. D. & Cloning Identification and Expression Analysis of HmDFR Gene from Hydrangea. Dissertation, Central South University of Forestry and Technology. (2020). <https://doi.org/10.27662/d.cnki.gznlc.2020.00027>
64. Xue, C. et al. Cloning and expression analysis of color-related gene HmF3H in hydrangea. *Mol. Plant. Breed.* **19**, 59–64. <https://doi.org/10.13271/j.mpb.019.000059> (2021).
65. Qi, H., Zhang, G. T., Zhu, Z. Y., Liu, C. & Yuan, S. X. Identification of seven key structural genes in the anthocyanin biosynthesis pathway in sepals of H macrophylla. *Curr. Issues Mol. Biol.* **44**, 4167–4180. <https://doi.org/10.3390/cimb44090286> (2022).
66. Yang, J. F. et al. Multilevel regulation of anthocyanin-promoting R2R3-MYB transcription factors in plants. *Front. Plant Sci.* **13**, 1008829. <https://doi.org/10.3389/fpls.2022.1008829> (2022).
67. Zhou, H. et al. Activator-type R2R3-MYB genes induce a repressor-type R2R3-MYB gene to balance anthocyanin and Proanthocyanidin accumulation. *New Phytol.* **221**, 1919–1934. <https://doi.org/10.1111/nph.15486> (2018).
68. Wang, L. J. et al. R2R3-MYB transcription factor MYB6 promotes anthocyanin and Proanthocyanidin biosynthesis but inhibits secondary cell wall formation in Populus tomentosa. *Plant Sci.* **99**, 733–751. <https://doi.org/10.1111/tpj.14364> (2020).
69. Naing, A. H. & Kim, C. Roles of R2R3-MYB transcription factors in transcriptional regulation of anthocyanin biosynthesis in horticultural plants. *Plant Mol. Biol.* **98**, 1–18. <https://doi.org/10.1007/s11103-018-0771-4> (2018).
70. Zhang, J. Y. et al. Transcriptome-Wide analysis reveals key DEGs in flower color regulation of hosta plantaginea (Lam.) aschers. *Genes* **11** (31). <https://doi.org/10.3390/genes11010031> (2020).
71. Anwar, M. et al. Ectopic overexpression of a novel R2R3-MYB, NtMYB2 from Chinese narcissus represses anthocyanin biosynthesis in tobacco. *Molecules* **23**, 781. <https://doi.org/10.3390/molecules23040781> (2018).
72. Yang, J. et al. Multilevel regulation of anthocyanin-promoting R2R3-MYB transcription factors in plants. *Front. Plant Sci.* **13**, 1008829. <https://doi.org/10.3389/fpls.2022.1008829> (2022).
73. Lin-Wang, K. et al. An R2R3 MYB transcription factor associated with regulation of the anthocyanin biosynthetic pathway in rosaceae. *BMC Plant Biol.* **10**, 50. <https://doi.org/10.1186/1471-2229-10-50> (2010).
74. Yan, H. et al. MYB-Mediated regulation of anthocyanin biosynthesis. *Int. J. Mol. Sci.* **22**, 3103. <https://doi.org/10.3390/ijms22063103> (2021).
75. Albert, N. et al. Members of an R2R3-MYB transcription factor family in Petunia are developmentally and environmentally regulated to control complex floral and vegetative pigmentation patterning. *Plant. Journal: Cell. Mol. Biology.* **65** <https://doi.org/10.1111/j.1365-313X.2010.04465.x> (2011). 771–84.
76. Ko, J. H. et al. Overexpression of PtrMYB119, a R2R3-MYB transcription factor from Populus trichocarpa, promotes anthocyanin production in hybrid Poplar. *Tree Physiol.* **36**, 1162–1176. <https://doi.org/10.1093/treephys/tpw046> (2016).
77. Chen, C. J. Wu, Y., Li, J. W. et al. TBtools-IIv2.225 [software]. (2023). Available from <https://github.com/CJ-Chen/TBtools-II>

Author contributions

YYL and WWZ contributed equally to this manuscript. YYL used software to analyze and verify the data. YW and SMQZ sorted out some datas and modified the format. SQG made English corrections. CZ collected transcriptome data. YYL wrote and edited the manuscript. JQP, SS, SW reviewed the manuscript.

Funding

This research was funded by Natural Science Foundation of Hunan Province (2025JJ50137), Central South University of Forestry and Technology Postgraduate Science and Technology Innovation Fund Funded Project (2023CX02042) and Scientific Research Project of Hunan Provincial Department of Education - Outstanding Youth Project (22B0261).

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-14216-4>.

Correspondence and requests for materials should be addressed to J.P., S.S. or S.W.

Reprints and permissions information is available at www.nature.com/reprints.

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

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. 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-nc-nd/4.0/>.

© The Author(s) 2025