



Genome-Wide Identification of R2R3-MYB Transcription Factors: Discovery of a “Dual-Function” Regulator of Gypenoside and Flavonol Biosynthesis in *Gynostemma pentaphyllum*

OPEN ACCESS

Edited by:

Xueqing Fu,
Shanghai Jiao Tong University, China

Reviewed by:

Hexin Tan,
Second Military Medical University,
China

Praveen Awasthi,
Institute of Plant Molecular Biology,
Centre for Biology, Academy
of Sciences of the Czech Republic
(ASCR), Czechia

*Correspondence:

Ding Huang
hdh016@126.com
Yong Tan
xjty321@163.com

† These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Plant Metabolism
and Chemodiversity,
a section of the journal
Frontiers in Plant Science

Received: 16 October 2021

Accepted: 02 December 2021

Published: 05 January 2022

Citation:

Huang D, Ming R, Xu S, Yao S,
Li L, Huang R and Tan Y (2022)
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. 12:796248.
doi: 10.3389/fpls.2021.796248

Ding Huang^{1,2*†}, Ruhong Ming^{1†}, Shiqiang Xu^{3,4†}, Shaochang Yao^{1,2}, Liangbo Li¹,
Rongshao Huang¹ and Yong Tan^{1,2*}

¹ College of Pharmacy, Guangxi University of Chinese Medicine, Nanning, China, ² Guangxi Key Laboratory of Zhuang and Yao Ethnic Medicine, Guangxi University of Chinese Medicine, Nanning, China, ³ Guangdong Provincial Engineering and Technology Research Center for Conservation and Utilization of the Genuine Southern Medicinal Resources, Guangzhou, China, ⁴ Guangdong Provincial Key Laboratory of Crops Genetics and Improvement, Crops Research Institute, Guangdong Academy of Agricultural Sciences, Guangzhou, China

The *R2R3-MYB* gene family participates in several plant physiological processes, especially the regulation of the biosynthesis of secondary metabolites. However, little is known about the functions of *R2R3-MYB* genes in *Gynostemma pentaphyllum* (*G. pentaphyllum*), a traditional Chinese medicinal herb that is an excellent source of gypenosides (a class of triterpenoid saponins) and flavonoids. In this study, a systematic genome-wide analysis of the *R2R3-MYB* gene family was performed using the recently sequenced *G. pentaphyllum* genome. In total, 87 *R2R3-GpMYB* genes were identified and subsequently divided into 32 subgroups based on phylogenetic analysis. The analysis was based on conserved exon–intron structures and motif compositions within the same subgroup. Collinearity analysis demonstrated that segmental duplication events were majorly responsible for the expansion of the *R2R3-GpMYB* gene family, and Ka/Ks analysis indicated that the majority of the duplicated *R2R3-GpMYB* genes underwent purifying selection. A combination of transcriptome analysis and quantitative reverse transcriptase-PCR (qRT-PCR) confirmed that *Gynostemma pentaphyllum* myeloblastosis 81 (*GpMYB81*) along with genes encoding gypenoside and flavonol biosynthetic enzymes exhibited similar expression patterns in different tissues and responses to methyl jasmonate (MeJA). Moreover, *GpMYB81* could bind to the promoters of *Gynostemma pentaphyllum* farnesyl pyrophosphate synthase 1 (*GpFPS1*) and *Gynostemma pentaphyllum* chalcone synthase (*GpCHS*), the key structural genes of gypenoside and flavonol biosynthesis, respectively, and activate their expression. Altogether, this study highlights a novel transcriptional regulatory mechanism that suggests that *GpMYB81* acts as a “dual-function” regulator of gypenoside and flavonol biosynthesis in *G. pentaphyllum*.

Keywords: *R2R3-MYB* gene family, *Gynostemma pentaphyllum*, gypenoside, flavonol, transcription factor

INTRODUCTION

The myeloblastosis (MYB) transcriptional regulators comprise one of the largest and important families in the plant kingdom (Riechmann et al., 2000). A classical characteristic of MYB proteins is that they contain conserved MYB DNA-binding domain repeats in the N-terminus. The conserved MYB domain is composed of one to four continuous and non-redundant imperfect sequence repeats, designated as R1, R2, and R3 according to their similarity to the c-MYB protein. Each conserved MYB DNA-binding domain repeat is approximately 50 amino acids in length and encodes three α -helices (Dubos et al., 2010). Depending on the number of MYB repeats in the MYB domain, MYB transcription factors (TFs) can be subdivided into R2R3-MYB (2R-MYB), R1R2R3-MYB (3R-MYB), 4R-MYB (containing four R1/R2 repeats), and the MYB-related subfamily (containing a single or a partial MYB repeat) (Stracke et al., 2001). In the R2R3-MYB family, the last two α -helices of each MYB repeat form a helix–turn–helix (HTH) structure and the third α -helix of R2 and R3 repeats are essential for DNA binding, allowing direct contact and insertion into the DNA major groove (Gabrielsen et al., 1991; Jia et al., 2004).

The R2R3-MYB TFs constitute the largest MYB subfamily in plants and the majority of them can specifically recognize the MYB-core sequence [(C/T)NGTT(G/T)] and AC-rich element [A/CCC(T/A)A(C/A)C/G] (Zhu et al., 2015; Millard et al., 2019). As important regulatory proteins involved in several crucial biological processes, the number of identified and characterized R2R3-MYB TFs in plants is continuously increasing. R2R3-MYB TFs play vital roles in plant growth and development, respond to various biotic and abiotic stresses, and regulate secondary metabolism, especially those affecting nutrition and medicinal components or appearance and quality traits (Huang D. et al., 2019; Wang et al., 2019; He et al., 2020). For example, *Ruby1* and *Ruby2* encode R2R3-MYB TFs and form a gene cluster that shows a regulatory subfunctionalization in anthocyanin biosynthesis in citrus (Huang et al., 2018). Similarly, kiwifruit R2R3-MYB TF MYB7 acts as a positive regulator to activate the promoter of the key gene lycopene beta-cyclase (*AdLCY-b*) in the carotenoid biosynthetic pathway (Ampomah-Dwamena et al., 2019). The genes *AtMYB11*, *AtMYB12*, and *AtMYB111* from subgroup 7 of the *Arabidopsis thaliana* (*A. thaliana*) R2R3-MYB gene family control flavonol accumulation in different parts of the *A. thaliana* seedling (Stracke et al., 2007).

Extensive study on the R2R3-MYB gene family members in horticultural and crop plants has increased our understanding of their functions and transcriptional regulatory mechanism. However, characteristics of this gene family in *Gynostemma pentaphyllum* (*G. pentaphyllum*), a traditional Chinese medicinal herb named jiaogulan, have not yet been declassified. As an economically valuable medicinal and edible plant, jiaogulan tea has been commercialized globally. Gypenosides are a major class of triterpenoids with a dammarane-type carbon skeleton in *G. pentaphyllum*, which exert beneficial effects on human health (Park et al., 2014; Shen et al., 2020; Wang et al., 2020). In addition, flavonoids can be divided into flavonols, flavones, flavonones, and anthocyanidins that play important roles in medicine and hygiene due to their high antioxidant activity

(Chen and Chen, 2013; Wang et al., 2018). Gypenosides are the primary active components of *G. pentaphyllum*. In contrast, flavonols extracted from *G. pentaphyllum* contain mainly kaempferol and quercetin derivatives, which were considered major contributors to the beneficial properties of *G. pentaphyllum* (Xie et al., 2011). In a recent study, gypenoside biosynthetic genes, including *farnesyl pyrophosphate synthase* (*FPS*), *squalene synthase* (*SS*), *squalene epoxidase* (*SE*), *2,3-oxidosqualene cyclase* (*OSC*), and *cytochrome P450* (*CYP450*), have been well elucidated (Huang et al., 2021). In addition, structural genes of the flavonoid metabolic pathway are best understood at present (Nabavi et al., 2020). However, it remains poorly understood whether the R2R3-MYB gene family members are involved in the regulation of both gypenoside and flavonol biosynthesis in *G. pentaphyllum*.

The recently published *G. pentaphyllum* genome sequence provides a convenient tool to identify and characterize the R2R3-MYB gene family (Huang et al., 2021). In this study, we performed a genome-wide identification of R2R3-MYB genes in *G. pentaphyllum* and screened 87 R2R3-MYB genes. Next, a comprehensive analysis including phylogenetic relationship, gene structure, conserved domains and motifs, chromosomal location, gene duplication, and collinearity was performed. Based on the weighted gene co-expression network analysis (WGCNA) and expression pattern response to methyl jasmonate (MeJA) treatment, GpMYB81 was suggested as a “dual-function” TF that can regulate both gypenoside and flavonol biosynthesis. In addition, GpMYB81 could bind to the promoters of *GpFPS1* and *GpCHS* genes and activate their transcription, thus opening up the possibility for improving the yield of both gypenosides and flavonols in *G. pentaphyllum* through metabolic engineering.

MATERIALS AND METHODS

Plant Materials and Methyl Jasmonate Treatment

Plant materials were cultivated in a fully controlled climate room of Guangxi University of Chinese Medicine (Nanning, China), with a 16-h light/8-h dark cycle at 24°C temperature. *G. pentaphyllum* seedlings were culture in Hoagland's nutrient solutions. For MeJA treatment, 6-week-old *G. pentaphyllum* seedlings were cultured in Hoagland's nutrient solutions with 100 μ M MeJA. For quantification of gene expression using quantitative reverse transcriptase-PCR (qRT-PCR), *G. pentaphyllum* seedlings were collected at 0, 6, 12, and 24 h after MeJA treatment; the leaves of three seedlings were randomly selected to form three biological replicates. All the plant samples were frozen with liquid nitrogen and stored at -80°C .

Identification of *Gynostemma pentaphyllum* R2R3-MYB Family Genes

The Hidden Markov Model (HMM) file of MYB DNA-binding domain (PF00249), obtained from the Pfam database,¹ was used as the query for HMM search using HMMER 3.0 (Finn et al., 2011) to identify MYB genes from *G. pentaphyllum* genome

¹<http://pfam.xfam.org>

with default parameters. To ensure the presence of two MYB DNA-binding domain repeats, candidate MYB protein sequences were further examined using the Simple Modular Architecture Research Tool (SMART) database.² Finally, a manual inspection was performed to confirm the reliability of our results.

Sequence Analysis and Phylogenetic Analysis of R2R3-MYB Genes

The exon/intron structure of all the R2R3-GpMYB genes was displayed using the TBtools software (Chen et al., 2020) based on gene annotation data in general feature format 3 (GFF3) format. The conserved motif of R2R3-GpMYB protein sequences was predicted using a motif-based sequence analysis tool Multiple Expectation maximizations for Motif Elicitation (MEME) version 5.1.1 program (Bailey et al., 2009). The parameters were as follows: maximum motif number of 25; other options were set to default.

Multiple sequence alignments of *G. pentaphyllum* and *A. thaliana* R2R3-MYB protein sequences were performed using molecular evolutionary genetics analysis (MEGA) version 10.1.7. Subsequently, a maximum likelihood (ML) phylogenetic tree was constructed using the FastTree version 2.1.1 (Price et al., 2009). The ML phylogenetic tree was visualized by the Interactive Tree of Life (iTOL) (Letunic and Bork, 2021). Additionally, an ML phylogenetic tree including full length of 87 R2R3-GpMYB protein sequences was constructed using the same methods. Finally, a combination of the phylogenetic tree, conserved domains, gene structures, and conserved motifs of R2R3-GpMYB protein sequences was visualized using the TBtools software (Chen et al., 2020).

Genomic Localization and Gene Duplication of R2R3-GpMYB Genes

The physical positions of the identified R2R3-GpMYB genes were mapped to 11 chromosomes of the *G. pentaphyllum* genome using the TBtools software (Chen et al., 2020). The orthologous MYB genes between *G. pentaphyllum* and *A. thaliana* as well as those between *G. pentaphyllum* and *C. sativus* were identified using OrthoVenn2 (Xu et al., 2019). Multiple Collinearity Scan toolkit (MCScanX) was used to analyze the gene duplication events with default parameters (Wang et al., 2012). Non-synonymous (ka) and synonymous (ks) substitutions of each duplicated R2R3-MYB gene were calculated using the TBtools software (Chen et al., 2020).

Ribonucleic Acid Isolation and Quantitative Reverse Transcriptase-PCR Analysis

Total RNA isolation and qRT-PCR analysis were performed using the methods described by Xu et al. (2020). qRT-PCR was performed using the LightCycler 96 System (Roche, United States). The *GpActin* gene was used for quantitative gene expression normalization (Xu et al., 2020; Huang et al., 2021). The $2^{-\Delta\Delta Ct}$ analysis method was adopted to calculate

the relative gene expression. Primer information is given in **Supplementary Table 1**.

Yeast One-Hybrid Assays

Yeast one-hybrid (Y1H) assays were performed as described previously (Huang et al., 2018). To construct the prey vector, the open reading frame (ORF) of the *GpMYB81* gene was cloned into the pGADT7 plasmid. To construct the bait vectors, the fragments of *GpFPS1* and *GpCHS* promoters (about 1.5 Kb) were cloned into the pAbAi plasmid. Yeast cells were grown for 3 days at 30°C on synthetic dropout (SD)/-Ura/-Leu medium added with or without aureobasidin A (AbA). Primer information is given in **Supplementary Table 1**.

Dual-Luciferase Assays

A dual-luciferase (LUC) reporter assay was conducted in *Nicotiana benthamiana* leaves according to the method described previously (Huang et al., 2018). To construct the effector vector, the ORF of the *GpMYB81* gene was cloned into the pK2GW7 plasmid. An empty vector of pK2GW7 was used as a negative control. To construct the reporter vectors, the fragments of *GpFPS1* and *GpCHS* promoters (about 1.5 Kb) were cloned into the pGreenII 0800-LUC plasmid. Fluorescence was detected using an *in vivo* imaging system (NightShade LB 985, Germany). Primer information is given in **Supplementary Table 1**.

RESULTS

Identification and Characterization of *Gynostemma pentaphyllum* R2R3-MYB Family Genes

In total, 248 candidate genes were originally obtained from the *G. pentaphyllum* genome as encoding proteins that contained MYB domains. After removing the redundant transcripts, all the candidates were further verified *via* Pfam, HMMscan, and SMART. As a result, 87 R2R3-GpMYB genes were identified in *G. pentaphyllum*. Among these, 86 R2R3-GpMYB genes were mapped to 11 chromosomes and renamed from *GpMYB1* to *GpMYB86* according to their location on the chromosomes. However, one exception was observed, an R2R3-GpMYB gene renamed *GpMYB87* was not located on any chromosome.

The amino acid number of R2R3-GpMYB proteins ranged from 126 to 556, with theoretical isoelectric point and molecular weight values ranging from 4.97 (*GpMYB14*) to 9.87 (*GpMYB5*) and 14.68 (*GpMYB5*) to 61.53 (*GpMYB68*) kDa, respectively. To provide possible clues for functional studies, we predicted their subcellular locations. The results indicated that all the R2R3-GpMYB proteins were located in the nucleus. These results are shown in **Supplementary Table 2**.

Phylogenetic Analysis and Classification of R2R3-MYB Genes in *Gynostemma pentaphyllum*

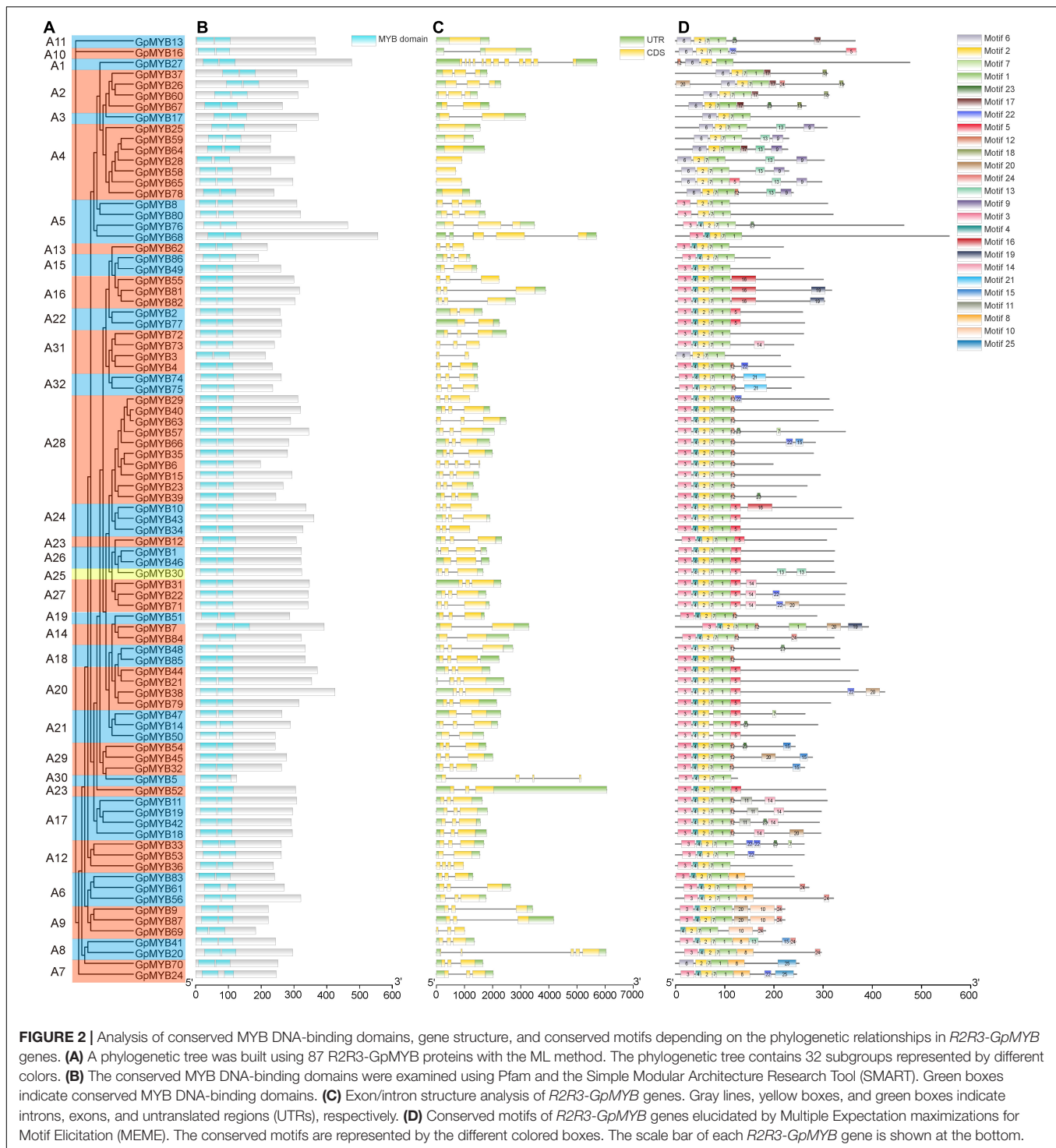
To elucidate the evolutionary relationship and gene function of the R2R3-GpMYB gene family, a ML tree containing 87 R2R3-GpMYB genes and 124 R2R3-AtMYB genes was constructed

²<http://smart.embl-heidelberg.de/>

activate the accumulation of anthocyanin and glucosinolate or contained other special regulated pathways to produce these metabolites.

The typical *R2R3-MYB*-encoded proteins were characterized by R2 and R3 repeats (Dubos et al., 2010). As shown in **Figures 2A,B**, the 87 identified *R2R3-GpMYB* proteins from 32 subgroups contained two conserved MYB repeats and were separated by approximately 108 basic residues. The exon–intron structure analysis indicated that the number of exons in *R2R3-GpMYB* genes varied from 0 to 11, the majority of which contained two

FIGURE 1 | Phylogenetic tree comparison of R2R3-MYB proteins between *Gynostemma pentaphyllum* (*G. pentaphyllum*) and *Arabidopsis thaliana* (*A. thaliana*). A maximum likelihood (ML) phylogenetic tree, including full length of 87 R2R3-GpMYB and 124 R2R3-AtMYB protein sequences, was generated using the Jones–Taylor–Thornton (JTT) algorithm *via* the FastTree software. The genes of the R2R3-GpMYB family were divided into 32 subgroups (designated as A1–A32). In addition, the classification method of *A. thaliana* (Stracke et al., 2001) was adopted.



introns, accounting for about 65.5%. Generally, similar structures of exon/intron were observed among the genes in the same subgroup, especially the number of introns. For example, the *R2R3-GpMYB* genes in the A4 subgroup contained no intron, whereas the A17 subgroup harbored two introns (Figures 2A,C).

The conserved motifs of all the *R2R3-GpMYB* proteins were studied using a motif-based sequence analysis tool

(Supplementary Table 3). As shown in Figure 2D, motif 1, motif 2, motif 3, motif 4, motif 6, and motif 7 in the N-terminus encoded the conserved MYB DNA-binding domain, whereas motifs in the C-terminus were highly variable. The majority of *R2R3-GpMYB* genes belonging to the same subgroup with similar functions exhibited similar motif compositions outside the MYB domain (Figures 2A,D). For example, the A2 subgroup contained

motif 17 and motif 18, which played important roles in the development of axillary meristem (Lee et al., 2009), whereas motif 9 and motif 13 in the A4 subgroup participated in providing resistance to biotic and abiotic stresses (Jung et al., 2008). These results indicated that these motifs were conserved in specific subgroups, and proteins sharing these motifs within a group in the phylogenetic tree likely had similar functions.

Chromosomal Distribution and Synteny Analyses of *Gynostemma pentaphyllum* R2R3-MYB Family

The *G. pentaphyllum* genomic database and genome chromosomal location results revealed that 86 out of 87 R2R3-GpMYB genes were unevenly distributed on 11 chromosomes (Figure 3). In detail, chromosome 11 had 16 R2R3-GpMYB genes, accounting for the largest number of R2R3-GpMYB genes, followed by chromosome 7 (11 R2R3-GpMYB genes), whereas chromosome 8 contained only two genes and had the minimum number of R2R3-GpMYB genes. The majority of R2R3-GpMYB genes were located on both ends of the chromosome. In addition, no correlation was found between the chromosome length and the distribution of R2R3-GpMYB gene family members on the chromosome. According to a previous study, if two or more genes are present within 200 kb, the elements are considered a tandem repeat event (Holub, 2001). In total, six R2R3-GpMYB genes underwent five tandem repeat events (Supplementary Table 4).

We employed basic local alignment search tool for proteins (BLASTP) and MCScanX to construct the collinearity of the R2R3-MYB gene family in *G. pentaphyllum* and identify the possible relationship and potential duplication events between them. Intrachromosomal duplications of the R2R3-MYB gene family were observed in the *G. pentaphyllum* genome (Supplementary Table 4). In detail, 34 pairs of R2R3-GpMYB genes duplicated tandemly on all the 11 chromosomes (Figure 4).

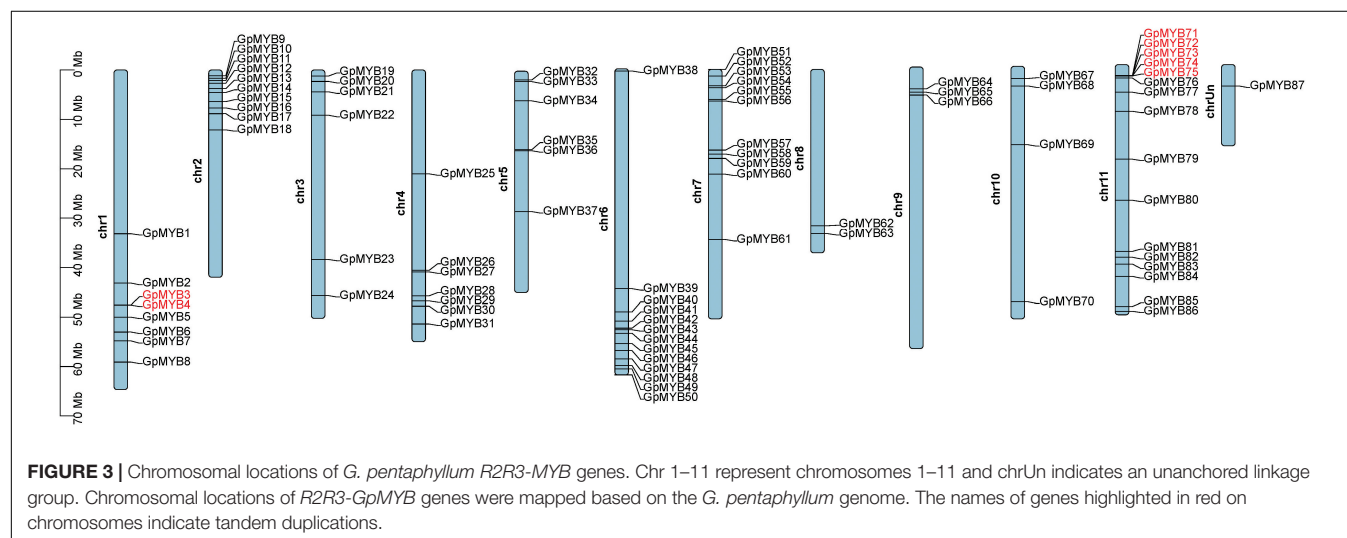
To further illustrate the potential evolutionary patterns of the R2R3-GpMYB gene family, a comparative orthologous analysis was performed between *G. pentaphyllum* and other two representative species, namely, *A. thaliana* and *Cucumis sativus*

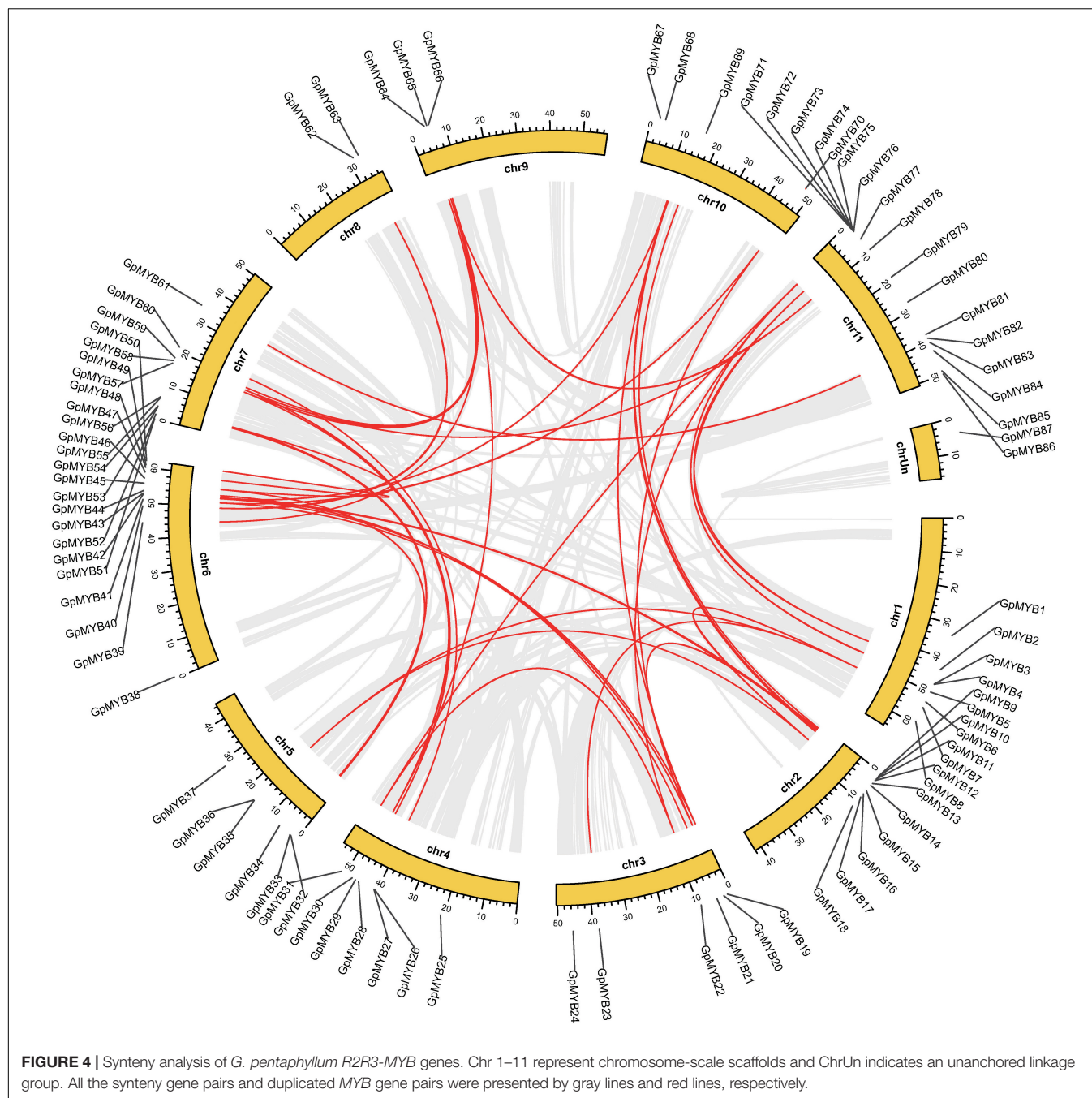
(*C. sativus*), which belong to the Brassicaceae and Cucurbitaceae families, respectively (Figure 5). The orthologous gene pairs between *G. pentaphyllum* and *A. thaliana* and *G. pentaphyllum* and *C. sativus* were 44 and 70, respectively (Supplementary Tables 5, 6). These results revealed that the identified orthologous events of GpMYB-CsMYB were considerably more than those of GpMYB-AtMYB based on the close evolutionary relationship between *G. pentaphyllum* and *C. sativus*. An extensive level of synteny conservation and increased number of orthologous events of GpMYB-CsMYB indicated that R2R3-GpMYB genes in *G. pentaphyllum* shared a similar structure and function with R2R3-CsMYB genes in *C. sativus*.

To further investigate the driving force behind the duplication of R2R3-MYB gene pairs in *G. pentaphyllum*, Ka/Ks (non-synonymous/synonymous substitution ratio) calculation of the duplicated R2R3-MYB gene pairs was performed to determine whether a selective pressure acted on the R2R3-GpMYB genes (Supplementary Tables 5, 6). Interestingly, all the Ka/Ks values of orthologous R2R3-GpMYB gene pairs were less than 1, indicating that these genes were subjected to purifying selection with limited functional divergence during evolution after duplication events.

Identification of R2R3-MYB Was Related to Both Gypenoside and Flavonol Biosynthesis in *Gynostemma pentaphyllum*

Transcriptional activators usually present similar expression patterns to the downstream structural genes of the metabolic pathway, narrowing the scope of screening candidate regulators and providing functional prediction. In a previous study, the early biosynthesis genes (EBGs) and late biosynthesis genes (LBGs) corresponding to gypenoside biosynthesis were elucidated (Huang et al., 2021). Furthermore, based on the gene expression profiles of different tissues (tendrils, young leaf, mature leaf, root, stem, flower, and fruit), a WGCNA was performed to identify the potential upstream regulators of gypenoside biosynthetic pathway genes (Huang et al., 2021).





In the constructed gypenoside biosynthesis regulatory network, five *GpMYB* genes were identified, three of which were R2R3-MYB genes, including *GpMYB60*, *GpMYB80*, and *GpMYB81*. Among these, *GpMYB60* was related to *AtMYB105* and *AtMYB117*, suggesting that the function of *GpMYB60* was related to the development of floral organs and the initiation of ovule outgrowth (Lee et al., 2009). *GpMYB80* was the closest homolog of *AtMYB97* and *AtMYB120*, acting as a transcriptional activator to control the differentiation of the pollen tube required for sperm release (Liang et al., 2013), and *GpMYB81* was highly close to other positive regulators of flavonol

biosynthesis, such as *AtMYB11*, *AtMYB12*, and *AtMYB111* (Tan et al., 2019). In addition, flavonol biosynthetic pathway genes, including *Gynostemma pentaphyllum* 4-coumarate-CoA ligase (*Gp4CL*), *GpCHS*, *Gynostemma pentaphyllum* chalcone isomerase (*GpCHI*), *Gynostemma pentaphyllum* flavanone 3-hydroxylase (*GpF3H*), and *Gynostemma pentaphyllum* flavonol synthase (*GpFLS*), were identified in the gypenoside-related module (**Supplementary Table 7**). As confirmed by qRT-PCR, *GpMYB81*, along with genes encoding gypenoside and flavonol biosynthetic enzymes, showed similar expression patterns, i.e., a high expression in young leaf tissue (**Figure 6A**).

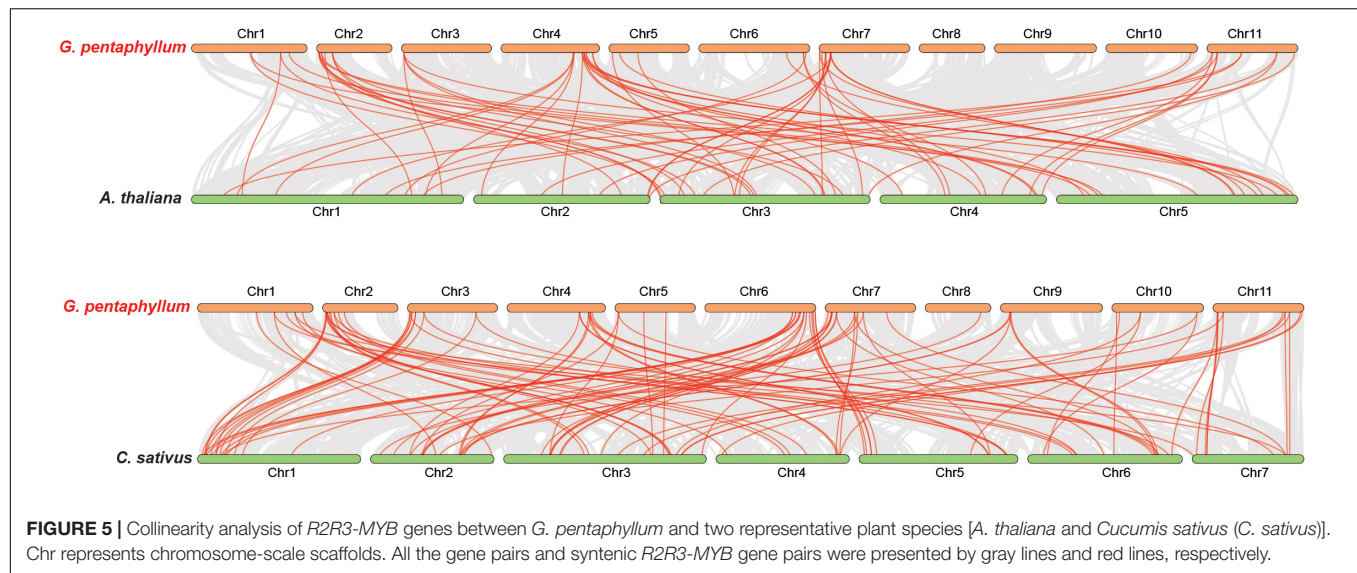


FIGURE 5 | Collinearity analysis of *R2R3-MYB* genes between *G. pentaphyllum* and two representative plant species [*A. thaliana* and *Cucumis sativus* (*C. sativus*)]. Chr represents chromosome-scale scaffolds. All the gene pairs and syntenic *R2R3-MYB* gene pairs were presented by gray lines and red lines, respectively.

As an effective elicitor, MeJA can intensify the accumulation of several secondary metabolites in various medicinal plants (Zhan et al., 2018; Deng et al., 2020). Candidate genes including TFs and biosynthetic pathway genes showed similar expression patterns in response to MeJA, further narrowing down the number of candidate genes. The qRT-PCR results showed that the expression of gypenoside biosynthetic pathway genes comprising *GpFPS1*, *Gynostemma pentaphyllum* squalene synthase 1 (*GpSS1*), *Gynostemma pentaphyllum* squalene epoxidase 2 (*GpSE2*), and *Gynostemma pentaphyllum* 2,3-oxidosqualene cyclases 1 (*GpOSC1*), and flavonol biosynthetic pathway genes, including *Gp4CL*, *GpCHS*, *GpCHI*, *GpF3H*, and *GpFLS*, was significantly increased after the MeJA treatment. Moreover, among the three candidate *R2R3-GpMYB* genes, the expression of *GpMYB81* increased gradually with the increase in the MeJA treatment time (Figure 6B). According to functional predictions and co-expression patterns, we speculated that *GpMYB81* most likely functions as a “dual-function” activator of both gypenoside and flavonol biosynthesis.

Gypenoside and Flavonol Biosynthetic Pathway Genes Were Transcriptionally Activated by *GpMYB81*

To investigate the mechanism underlying similar expression patterns among gypenoside, flavonol biosynthetic pathway genes, and their potential regulator *GpMYB81*, the transcriptional *cis*-elements of *GpFPS1* and *GpCHS* were analyzed (Figure 7A). The results revealed that the promoters of *GpFPS1* and *GpCHS* contained conserved MYB-recognition elements (MREs) or AC-rich elements, suggesting that *GpMYB81* might bind to the promoters of *GpFPS1* and *GpCHS*. To prove this hypothesis, Y1H assays were performed. As shown in Figure 7B, *GpMYB81* could bind to the promoters of *GpFPS1* and *GpCHS* *in vivo*. Moreover, the transient expression of the promoter activity assays revealed that *GpMYB81* could activate the expression of *GpFPS1* and *GpCHS* (Figure 7C). These results confirmed that *GpMYB81*

can simultaneously activate gypenoside and flavonol biosynthetic pathway genes, thereby parallelly promoting the accumulation of gypenosides and flavonols.

DISCUSSION

Secondary metabolism not only endows plants with the ability to adapt to the ecological environment, but also provides abundant pharmaceutical ingredients of considerable health benefits to humans (Ramakrishna and Ravishankar, 2011; Thimmappa et al., 2014; Li et al., 2015). Terpenoid saponins and flavonoids are two important secondary metabolites, several of which have been proven to exhibit antiviral, antifungal, and anticancer properties (Zhang et al., 2018; Bonta, 2020; Russo et al., 2020; Yang et al., 2021). *G. pentaphyllum* is a traditional Chinese medicinal herb and known for its industrial use. It often serves as a natural source of gypenosides and flavonoids in medicinal products. As reported in our previous study (Huang et al., 2021), the genome sequencing of *G. pentaphyllum* was completed in 2021, providing the opportunity to perform a genome-wide analysis of the *R2R3-GpMYB* gene family. However, the *R2R3-GpMYB* gene family has not been comprehensively analyzed and their dual functionality in regulating distinct pathways of synthesis has not been well studied.

In this study, 87 members of the *R2R3-GpMYB* family were identified and characterized in *G. pentaphyllum*. Although the number of *R2R3-GpMYB* was higher than that in *C. sativus* (69) (Cheng et al., 2020), it was significantly less than that in cotton (205) (Huang J. et al., 2019), *Populus trichocarpa* (196) (Wilkins et al., 2009), and banana (285) (Pucker et al., 2020). Whole-genome duplication (WGD) is a crucial event for the rapid expansion and evolution of gene families, e.g., the recent gene duplication in several angiosperms (Pucker et al., 2020). However, WGD events and tandem duplications are rare in the *G. pentaphyllum* genome. Conversely, numerous segmental duplication events were identified in

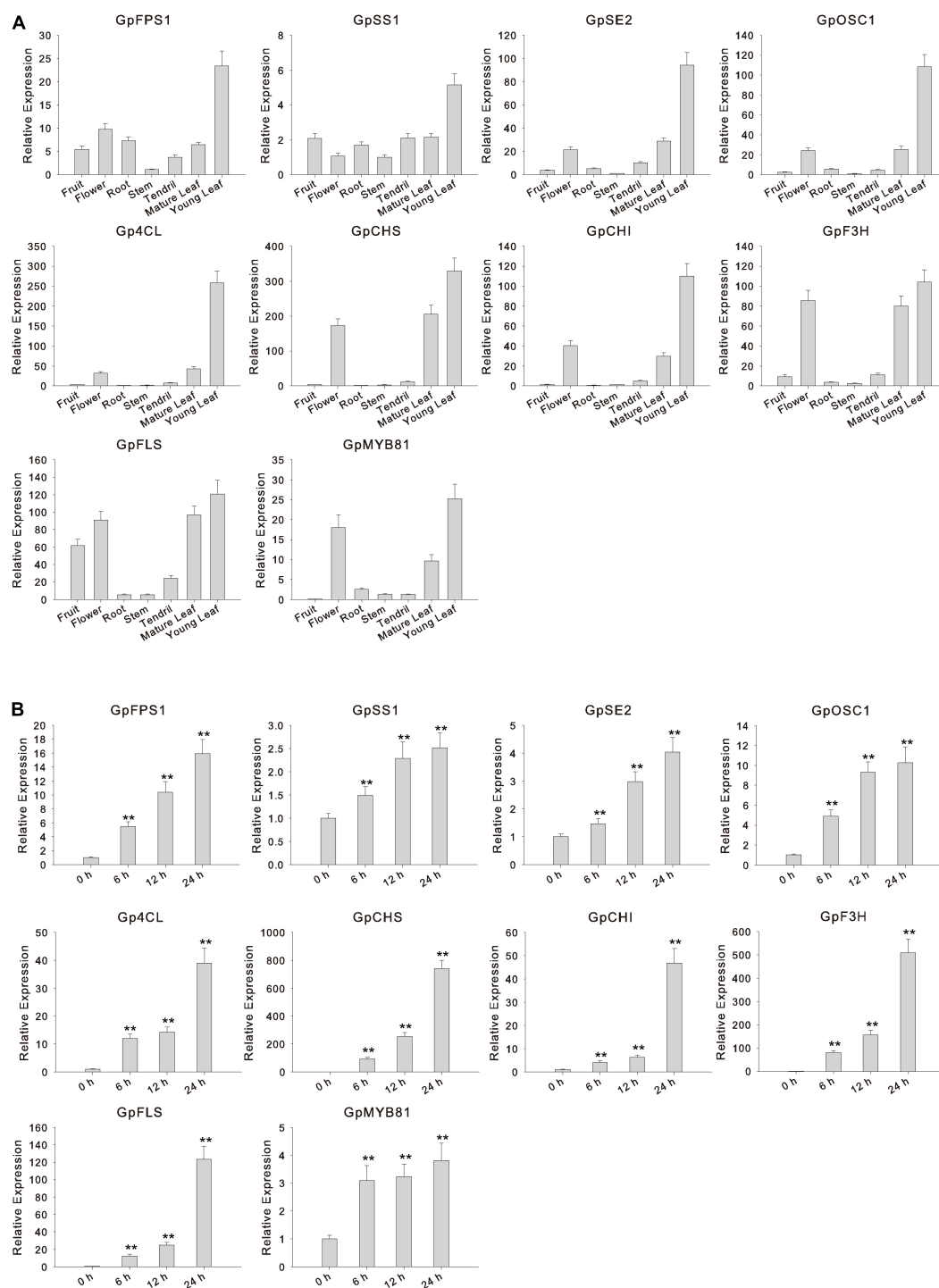
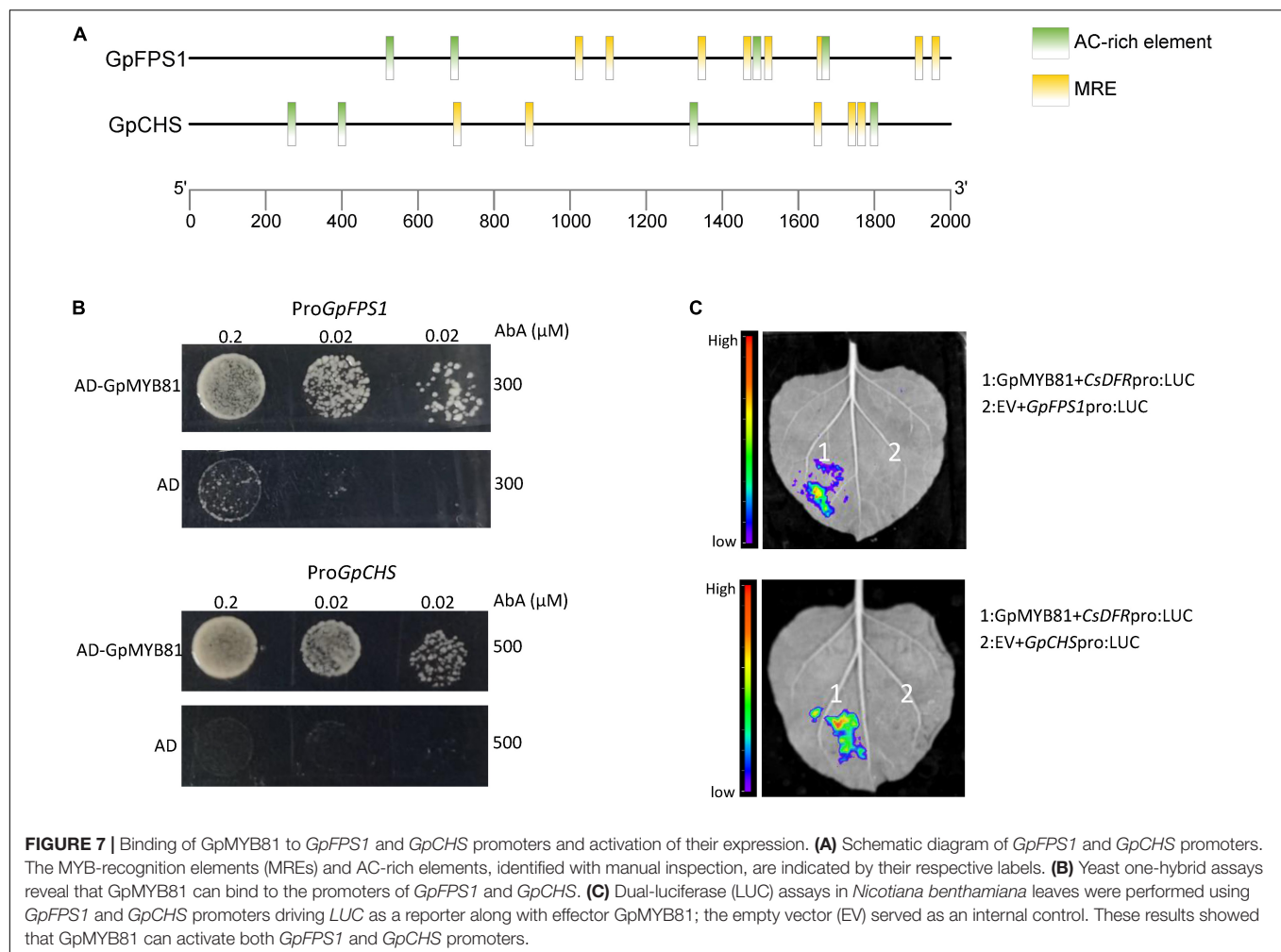


FIGURE 6 | Expression profile of 10 selected candidate genes in different tissues and response to methyl jasmonate (MeJA) treatment. **(A)** Expression analysis of gypenoside and flavonol biosynthetic genes and candidate regulator *GpMYB81* in different tissues of *G. pentaphyllum*. Values represent the mean \pm SE ($n = 3$ biological replicates). **(B)** Expression analysis of gypenoside and flavonol biosynthetic genes and candidate regulator *GpMYB81* with 0 (negative control), 6, 12, and 24 h treatment with MeJA. ** $p < 0.01$ (two-tailed Student's *t*-test). Values represent the mean \pm SE ($n = 3$ biological replicates).

G. pentaphyllum, indicating that segmental duplication events were the most important factor in the expansion of the R2R3-*GpMYB* gene family, which exhibited an evolutionary

pattern similar to that of *MYB* genes in other plants (Cannon et al., 2004). In addition, certain R2R3-*GpMYB* genes were not associated with syntenic gene pairs in *A. thaliana* or



C. sativus, indicating their possible specificity to *G. pentaphyllum* during the course of evolution, providing new insight or interest to explore those special R2R3-GpMYB genes in the future.

Flavonoids and terpenoids, the two largest groups of specialized plant metabolites, are derived from two distinct pathways. Notably, there are several examples where MYB proteins act to regulate terpenoid and flavonoid biosynthesis. In *A. thaliana*, the R2R3-MYB TFs that are currently known to regulate flavonol biosynthesis belong to subgroup 7 (Stracke et al., 2007). Some MYB TFs involved in the regulation of triterpene saponins have also been characterized. For instance, *Panax ginseng* myeloblastosis 2 (PgMYB2) in *Panax ginseng* was characterized as a positive regulator of ginsenoside metabolism (Liu et al., 2019). However, TFs act in a coordinated manner to simultaneously regulate different pathways of specialized metabolism in a novel regulatory mechanism. In this study, GpMYB81 belongs to subgroup 7 of the R2R3-MYB gene family from *A. thaliana*, which was identified in the gypenoside-related module. In addition, the expression patterns of GpMYB81 were highly correlated with gypenoside and flavonol biosynthesis genes and with those involved in response to MeJA, further supporting the finding that GpMYB81 can regulate the accumulation of both gypenosides and flavonols. The chalcone

synthase (*CHS*) gene plays a vital role in flavonol biosynthesis and greatly impacts the content of flavonols, whereas the overexpression or RNA interference of the *FPS* gene can significantly increase or decrease the biosynthesis of gypenoside (Kim et al., 2014; Zhang et al., 2017). In this study, GpMYB81 could bind to the promoters of *GpFPS1* and *GpCHS* and increase the transcriptional activities of these promoters. These findings suggested that GpMYB81 acts as a “dual-function” TF that regulates both gypenoside and flavonol pathways in *G. pentaphyllum*.

CONCLUSION

In summary, this study presented a detailed genome-wide analysis of the R2R3-GpMYB gene family. A total of 87 R2R3-GpMYB genes were identified in *G. pentaphyllum* and divided into 32 subgroups, with an uneven distribution on 11 chromosomes. Similar exon-intron structures and conserved motif compositions of R2R3-GpMYB genes were observed in the same subgroup, which provided additional support for phylogenetic analysis. Synteny analysis indicated that segmental duplication events primarily contributed to the expansion of the R2R3-GpMYB gene family. The Ka/Ks analysis suggested that

the *R2R3-GpMYB* gene family underwent purifying selection. A combination of similar gene expression patterns, Y1H, and dual-LUC assay results verified that *GpMYB81* acted as a “dual-function” activator in gypenoside and flavonol biosynthesis by directly binding to the promoters of *GpFPS1* and *GpCHS*. These results provide novel insights into the parallel transcriptional regulation of gypenoside and flavonol biosynthesis in *G. pentaphyllum*.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: the transcriptome sequencing data can be found in NCBI under accession codes PRJNA720501 and PRJNA631355.

AUTHOR CONTRIBUTIONS

DH and YT conceived this project. DH and RM designed the experiments. DH and SX prepared the samples and wrote the manuscript. DH, RM, and SX analyzed the bioinformatics data. SY, LL, RH, and YT provided valuable suggestions on the research design and the improvement of the manuscript. All authors contributed to the article and approved the submitted version.

REFERENCES

- Ampomah-Dwamena, C., Thrimawithana, A. H., Dejnopratt, S., Lewis, D., Espley, R. V., and Allan, A. C. (2019). A kiwifruit (*Actinidia deliciosa*) R2R3-MYB transcription factor modulates chlorophyll and carotenoid accumulation. *New Phytol.* 221, 309–325. doi: 10.1111/nph.15362
- Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., et al. (2009). MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.* 37, W202–W208. doi: 10.1093/nar/gkp335
- Bonta, R. K. (2020). Dietary phenolic acids and flavonoids as potential anti-cancer agents: current state of the art and future perspectives. *Anticancer Agents Med. Chem.* 20, 29–48. doi: 10.2174/1871520619666191019112712
- Cannon, S. B., Mitra, A., Baumgarten, A., Young, N. D., and May, G. (2004). The roles of segmental and tandem gene duplication in the evolution of large gene families in *Arabidopsis thaliana*. *BMC Plant Biol.* 4:10. doi: 10.1186/1471-2229-4-10
- Chen, A. Y., and Chen, Y. C. (2013). A review of the dietary flavonoid, kaempferol on human health and cancer chemoprevention. *Food Chem.* 138, 2099–2107. doi: 10.1016/j.foodchem.2012.11.139
- Chen, C., Chen, H., Zhang, Y., Thomas, H. R., Frank, M. H., He, Y., et al. (2020). TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant.* 13, 1194–1202. doi: 10.1016/j.molp.2020.06.009
- Cheng, C., Li, Q., Wang, X., Li, Y., and Qian, C. (2020). Identification and expression analysis of the CsMYB gene family in root knot nematode-resistant and susceptible cucumbers. *Front. Genet.* 11:550677. doi: 10.3389/fgene.2020.550677
- Deng, C., Wang, Y., Huang, F., Lu, S., Zhao, L., Ma, X., et al. (2020). SmMYB2 promotes salvianolic acid biosynthesis in the medicinal herb *Salvia miltiorrhiza*. *J. Integr. Plant Biol.* 62, 50–64. doi: 10.1111/jipb.12943
- Dubos, C., Stracke, R., Grotewold, E., Weissshaar, B., Martin, C., and Lepiniec, L. (2010). MYB transcription factors in *Arabidopsis*. *Trends Plant Sci.* 15, 573–581. doi: 10.1016/j.tplants.2010.06.005
- Finn, R. D., Clements, J., and Eddy, S. R. (2011). HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res.* 39, W29–W37. doi: 10.1093/nar/gkr367

FUNDING

This study was supported by the Natural Science Foundation of Guangxi Zhuang Autonomous Region (2020GXNSFBA297025), the Guangxi Middle-aged and Young Teachers' Basic Ability Promotion Project (2020KY07039), the Guangxi University of Chinese Medicine Scientific Research Fund (2019BS007), the Guangxi Key Laboratory of Zhuang and Yao Ethnic Medicine Open Project Fund (20-065-14), and the Special Fund for Introducing Scientific and Technological Talents of Guangdong Academy of Agricultural Sciences (R2020YJ-YB3003).

ACKNOWLEDGMENTS

The authors would like to thank TopEdit (www.topeditsci.com) for its linguistic assistance during the preparation of this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.796248/full#supplementary-material>

- Gabrielsen, O., Sentenac, A., and Fromageot, P. (1991). Specific DNA binding by c-Myb: evidence for a double helix-turn-helix-related motif. *Science* 253, 1140–1143. doi: 10.1126/science.1887237
- He, J., Liu, Y., Yuan, D., Duan, M., Liu, Y., Shen, Z., et al. (2020). An R2R3 MYB transcription factor confers brown planthopper resistance by regulating the phenylalanine ammonia-lyase pathway in rice. *Proc. Natl. Acad. Sci. U.S.A.* 117, 271–277. doi: 10.1073/pnas.1902771116
- Holub, E. B. (2001). The arms race is ancient history in *Arabidopsis*, the wildflower. *Nat. Rev. Genet.* 2, 516–527. doi: 10.1038/35080508
- Huang, D., Ming, R., Xu, S., Wang, J., Yao, S., Li, L., et al. (2021). Chromosome-level genome assembly of *Gynostemma pentaphyllum* provides insights into gypenoside biosynthesis. *DNA Res.* 28:dsab018. doi: 10.1093/dnares/dsab018
- Huang, D., Tang, Z., Fu, J., Yuan, Y., Deng, X., and Xu, Q. (2019). CsMYB3 and CsRuby1 form an ‘activator-and-repressor’ loop for regulation of anthocyanin biosynthesis in citrus. *Plant Cell Physiol.* 61, 318–330. doi: 10.1093/pcp/pcz198
- Huang, J., Guo, Y., Sun, Q., Zeng, W., Li, J., Li, X., et al. (2019). Genome-wide identification of R2R3-MYB transcription factors regulating secondary cell wall thickening in cotton fiber development. *Plant Cell Physiol.* 60, 687–701. doi: 10.1093/pcp/pcy238
- Huang, D., Wang, X., Tang, Z., Yuan, Y., Xu, Y., He, J., et al. (2018). Subfunctionalization of the Ruby2-Ruby1 gene cluster during the domestication of citrus. *Nat. Plants* 4, 930–941. doi: 10.1038/s41477-018-0287-6
- Jia, L., Clegg, M. T., and Jiang, T. (2004). Evolutionary dynamics of the DNA-binding domains in putative R2R3-MYB genes identified from rice subspecies indica and japonica genomes. *Plant Physiol.* 134, 575–585. doi: 10.1104/pp.103.027201
- Jung, C., Seo, J. S., Han, S. W., Koo, Y. J., Kim, C. H., Song, S. I., et al. (2008). Overexpression of AtMYB44 enhances stomatal closure to confer abiotic stress tolerance in transgenic *Arabidopsis*. *Plant Physiol.* 146, 623–635. doi: 10.1104/pp.107.110981
- Kim, Y. K., Kim, Y. B., Uddin, M. R., Lee, S., Kim, S. U., and Park, S. U. (2014). Enhanced triterpene accumulation in *Panax ginseng* hairy roots overexpressing mevalonate-5-pyrophosphate decarboxylase and farnesyl pyrophosphate synthase. *ACS Synth. Biol.* 3, 773–779. doi: 10.1021/sb400194g

- Lee, D. K., Geisler, M., and Springer, P. S. (2009). LATERAL ORGAN FUSION1 and LATERAL ORGAN FUSION2 function in lateral organ separation and axillary meristem formation in *Arabidopsis*. *Development* 136, 2423–2432. doi: 10.1242/dev.031971
- Letunic, I., and Bork, P. (2021). Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* 49, W293–W296. doi: 10.1093/nar/gkab301
- Li, C., Ng, K. Y., and Fan, L. M. (2015). MYB transcription factors, active players in abiotic stress signaling. *Environ. Exp. Bot.* 114, 80–91. doi: 10.1016/j.envexpbot.2014.06.014
- Liang, Y., Tan, Z. M., Zhu, L., Niu, Q. K., Zhou, J. J., Li, M., et al. (2013). MYB97, MYB101 and MYB120 function as male factors that control pollen tube-synergid interaction in *Arabidopsis thaliana* fertilization. *PLoS Genet.* 9:e1003933. doi: 10.1371/journal.pgen.1003933
- Lin-Wang, K., Bolitho, K., Grafton, K., Kortstee, A., Karunairatnam, S., McGhie, T. K., et al. (2010). An R2R3 MYB transcription factor associated with regulation of the anthocyanin biosynthetic pathway in *Rosaceae*. *BMC Plant Biol.* 10:50. doi: 10.1186/1471-2229-10-50
- Liu, T., Luo, T., Guo, X., Zou, X., Zhou, D., Afrin, S., et al. (2019). PgMYB2, a MeJA-responsive transcription factor, positively regulates the dammaradiol synthase gene expression in *Panax Ginseng*. *Int. J. Mol. Sci.* 20:2219. doi: 10.3390/ijms20092219
- Millard, P. S., Kragelund, B. B., and Burow, M. (2019). R2R3 MYB transcription factors - functions outside the DNA-binding domain. *Trends Plant Sci.* 24, 934–946. doi: 10.1016/j.tplants.2019.07.003
- Nabavi, S. M., Šamec, D., Tomczyk, M., Milella, L., Russo, D., Habtemariam, S., et al. (2020). Flavonoid biosynthetic pathways in plants: versatile targets for metabolic engineering. *Biotechnol. Adv.* 38:107316. doi: 10.1016/j.biotechadv.2018.11.005
- Park, S. H., Huh, T. L., Kim, S. Y., Oh, M. R., Tirupathi Pichiah, P. B., Chae, S. W., et al. (2014). Antiobesity effect of *Gynostemma pentaphyllum* extract (actiponin): a randomized, double-blind, placebo-controlled trial. *Obesity* 22, 63–71. doi: 10.1002/oby.20539
- Price, M. N., Dehal, P. S., and Arkin, A. P. (2009). FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol. Biol. Evol.* 26, 1641–1650. doi: 10.1093/molbev/msp077
- Pucker, B., Pandey, A., Weisshaar, B., and Stracke, R. (2020). The R2R3-MYB gene family in banana (*Musa acuminata*): genome-wide identification, classification and expression patterns. *PLoS One* 15:e0239275. doi: 10.1371/journal.pone.0239275
- Ramakrishna, A., and Ravishankar, G. A. (2011). Influence of abiotic stress signals on secondary metabolites in plants. *Plant Signal. Behav.* 6, 1720–1731. doi: 10.4161/psb.6.11.17613
- Riechmann, J. L., Heard, J., Martin, G., Reuber, L., Jiang, C., Keddie, J., et al. (2000). *Arabidopsis* transcription factors: genome-wide comparative analysis among eukaryotes. *Science* 290, 2105–2110. doi: 10.1126/science.290.5499.2105
- Russo, M., Moccia, S., Spagnuolo, C., Tedesco, I., and Russo, G. L. (2020). Roles of flavonoids against coronavirus infection. *Chem. Biol. Interact.* 328:109211. doi: 10.1016/j.cbi.2020.109211
- Shen, C. Y., Ma, P. Y., Zhu, J. J., Jiang, J. G., Liu, L., Yi, Y. K., et al. (2020). Saponin extracts from *Gynostemma pentaphyllum* (Thunb.) Makino display sedative-hypnotic and anxiolytic effects. *Ind. Crops Prod.* 157:112893. doi: 10.1016/j.indcrop.2020.112893
- Stracke, R., Ishihara, H., Huep, G., Barsch, A., Mehrtens, F., Niehaus, K., et al. (2007). Differential regulation of closely related R2R3-MYB transcription factors controls flavonol accumulation in different parts of the *Arabidopsis thaliana* seedling. *Plant J.* 50, 660–677. doi: 10.1111/j.1365-3113X.2007.03078.x
- Stracke, R., Werber, M., and Weisshaar, B. (2001). The R2R3-MYB gene family in *Arabidopsis thaliana*. *Curr. Opin. Plant Biol.* 4, 447–456. doi: 10.1016/s1369-5266(00)00199-0
- Tan, H., Man, C., Xie, Y., Yan, J., Chu, J., and Huang, J. (2019). A crucial role of GA-regulated flavonol biosynthesis in root growth of *Arabidopsis*. *Mol. Plant.* 12, 521–537. doi: 10.1016/j.molp.2018.12.021
- Thimmappa, R., Geisler, K., Louveau, T., O'Maille, P., and Osbourn, A. (2014). Triterpene biosynthesis in plants. *Annu. Rev. Plant Biol.* 65, 225–257. doi: 10.1146/annurev-arplant-050312-120229
- Wang, J., Zhao, M., Cheng, X., Han, Y., Zhao, T., Fan, M., et al. (2020). Dammarane-type saponins from *Gynostemma pentaphyllum* prevent hypoxia-induced neural injury through activation of ERK, Akt, and CREB pathways. *J. Agric. Food Chem.* 68, 193–205. doi: 10.1021/acs.jafc.9b06659
- Wang, L., Lu, W., Ran, L., Dou, L., Yao, S., Hu, J., et al. (2019). R2R3-MYB transcription factor MYB6 promotes anthocyanin and proanthocyanidin biosynthesis but inhibits secondary cell wall formation in *Populus tomentosa*. *Plant J.* 99, 733–751. doi: 10.1111/tpj.14364
- Wang, Y., Tang, H., Debarry, J. D., Tan, X., Li, J., Wang, X., et al. (2012). MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* 40:e49. doi: 10.1093/nar/gkr1293
- Wang, Z. L., Wang, S., Kuang, Y., Hu, Z. M., Qiao, X., and Ye, M. (2018). A comprehensive review on phytochemistry, pharmacology, and flavonoid biosynthesis of *Scutellaria baicalensis*. *Pharm. Biol.* 56, 465–484. doi: 10.1080/13880209.2018.1492620
- Wilkins, O., Nahal, H., Foong, J., Provart, N. J., and Campbell, M. M. (2009). Expansion and diversification of the Populus R2R3-MYB family of transcription factors. *Plant Physiol.* 149, 981–993. doi: 10.1104/pp.108.132795
- Xie, Z., Zhao, Y., Chen, P., Jing, P., Yue, J., and Yu, L. L. (2011). Chromatographic fingerprint analysis and rutin and quercetin compositions in the leaf and whole-plant samples of di- and tetraploid *Gynostemma pentaphyllum*. *J. Agric. Food Chem.* 59, 3042–3049. doi: 10.1021/jf104329v
- Xu, L., Dong, Z., Fang, L., Luo, Y., Wei, Z., Guo, H., et al. (2019). OrthoVenn2: a web server for whole-genome comparison and annotation of orthologous clusters across multiple species. *Nucleic Acids Res.* 47, W52–W58. doi: 10.1093/nar/gkz333
- Xu, S., Yao, S., Huang, R., Tan, Y., and Huang, D. (2020). Transcriptome-wide analysis of the AP2/ERF transcription factor gene family involved in the regulation of gypenoside biosynthesis in *Gynostemma pentaphyllum*. *Plant Physiol. Biochem.* 154, 238–247. doi: 10.1016/j.plaphy.2020.05.040
- Yang, G., Zhang, J., Wang, J., Wang, J., Wang, J., Zhu, Y., et al. (2021). Gypenoside inhibits bovine viral diarrhoea virus replication by interfering with viral attachment and internalization and activating apoptosis of infected cells. *Viruses* 13:1810. doi: 10.3390/v13091810
- Yu, Y., Xu, M., Ding, X., Chu, Z., and Liu, H. (2021). Activating the MYB51 and MYB122 to upregulate the transcription of glucosinolates biosynthesis genes by copper ions in *Arabidopsis*. *Plant Physiol. Biochem.* 162, 496–505. doi: 10.1016/j.plaphy.2021.03.025
- Zhan, X., Liao, X., Luo, X., Zhu, Y., Feng, S., Yu, C., et al. (2018). Comparative metabolomic and proteomic analyses reveal the regulation mechanism underlying MeJA-induced bioactive compound accumulation in cutleaf groundcherry (*Physalis angulata* L.) hairy roots. *J. Agric. Food Chem.* 66, 6336–6347. doi: 10.1021/acs.jafc.8b02502
- Zhang, X., Abrahams, C., Colquhoun, T. A., and Liu, C. J. (2017). A proteolytic regulator controlling chalcone synthase stability and flavonoid biosynthesis in *Arabidopsis*. *Plant Cell* 29, 1157–1174. doi: 10.1105/tpc.16.00855
- Zhang, X., Shi, G., Wu, X., and Zhao, Y. (2018). Gypenosapogenin H from hydrolyzate of total *Gynostemma pentaphyllum* saponins induces apoptosis in human breast carcinoma cells. *Nat. Prod. Res.* 34, 1642–1646. doi: 10.1080/14786419.2018.1525370
- Zhu, Z., Wang, H., Wang, Y., Guan, S., Wang, F., Tang, J., et al. (2015). Characterization of the cis elements in the proximal promoter regions of the anthocyanin pathway genes reveals a common regulatory logic that governs pathway regulation. *J. Exp. Bot.* 66, 3775–3789. doi: 10.1093/jxb/erv173

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Huang, Ming, Xu, Yao, Li, Huang and Tan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.