Ornamental Plant Research

ARTICLE

Open Access

https://doi.org/10.48130/opr-0026-0003

Ornamental Plant Research **2026**, 6: e013

Genome-wide analysis of MYB transcription factors in *Cymbidium* species reveals their roles in anthocyanin synthesis

Yunzhu Wang[1#], Sujuan Li[1#], Qunhua Zhang[2], Zhenting Shi[2], Guowei Zhai[1], Jie Luo[1] and Chongbo Sun[2*]

[1] *State Key Laboratory for Quality and Safety of Agro-Products, Zhejiang Academy of Agricultural Sciences, Hangzhou 310021, China*
[2] *Institute of Horticulture, Zhejiang Academy of Agricultural Sciences, Hangzhou 310021, China*
[#] Authors contributed equally: Yunzhu Wang, Sujuan Li
[*] Correspondence: chongpo1230@163.com (Sun C)

Abstract

The MYB proteins represent one of the most prevalent transcription factors in plants that play critical roles in various biological processes. The *Cymbidium* species occupy a dominant role in global floriculture markets due to their high ornamental value. In this study, a genome-wide identification and characterization of MYB proteins in *C. sinense*, *C. haematode*, and other closely related *Cymbidium* species were performed. A total of 170 MYB proteins (CysMYBs) were identified in *C. sinense*. Phylogenetic analysis classified these CysMYBs into 33 subfamilies, with 23 members from S4, S5, S6, S7, and S9 subfamilies that are potentially involved in anthocyanin regulation. Chromosomal distribution revealed uneven gene clustering, and 23 pairs of syntenic relationships, indicating gene expansion through fragment or tandem duplication events. Transcriptome and qRT-PCR analyses demonstrated functional divergence and species-specific regulation of anthocyanin synthesis in *C. sinense*, and its close relative *C. haematodes*. In particular, *CysMYB72* and *CysMYB114* exhibited species-specific roles in anthocyanin synthesis in *C. sinense* flowers, whereas *CysMYB131* and *CysMYB88* were proposed as key regulators of anthocyanin accumulation in *C. haematodes*. Cyanidin derivatives were identified as the primary anthocyanins in floral tissues, with the abundance correlating with *CysMYB* expression patterns. In conclusion, this study elucidates the functional divergence of MYB transcription factors in orchids and provides foundational insights into the molecular mechanisms underlying anthocyanin biosynthesis in Orchidaceae species.

Citation: Wang Y, Li S, Zhang Q, Shi Z, Zhai G, et al. 2026. Genome-wide analysis of MYB transcription factors in *Cymbidium* species reveals their roles in anthocyanin synthesis. *Ornamental Plant Research* 6: e013 https://doi.org/10.48130/opr-0026-0003

Introduction

MYB transcription factors (TFs) play diverse roles in regulating plant growth and development, including secondary metabolism, signal transduction, and resistance to biotic and abiotic stresses[1–4]. According to the repeat number, MYB proteins are generally classified into four subfamilies: 1R-MYB (MYB-related), 2R-MYB (R2R3-MYB), 3R-MYB (R1R2R3-MYB), and 4R-MYB[1,5–7]. The first plant MYB was identified in *Zea mays* associated with anthocyanin biosynthesis in aleurone tissues[8]. Since then, a large number of *MYB* superfamily genes has been identified in more than 70 plant genomes[4,5,9].

MYB proteins occupy a dominant role in regulating the biosynthesis and accumulation of anthocyanin. In *Arabidopsis*, MYB members from the S4, S5, S6, S7, and S9 subfamilies are highly correlated with flavonoid biosynthesis[1,10,11]. The early biosynthesis genes (EBGs), such as *CHS*, *CHI*, *F3H*, and *F3'H*, are directly bound with and regulated by the MYB TFs in the anthocyanin biosynthesis pathway (ABP)[12]. In contrast, the late biosynthesis genes (LBGs), including *DFR*, *ANS*, and *UFGT*, are regulated by the MBW (MYB-bHLH-WD40) protein complex that controls the downstream accumulation of anthocyanins[13–15]. Overexpression of the photoreactive transcription factor SmMYB35 promotes anthocyanin accumulation in the stems and petals of transgenic eggplant (*Solanum melongena*)[16]. In addition to transcriptional activators, there are also R2R3-MYB transcriptional inhibitors. For example, LvMYB5 promotes anthocyanin synthesis by activating the *ANS* gene promoter, whereas LvMYB1 inhibits anthocyanin synthesis[17].

The Orchidaceae family comprises over 30,000 species, characterized by their diverse floral organ morphology, and highly abundant corolla-color diversity[18,19]. The *Cymbidium* species have a long cultivation history of more than 2,000 years, and are still dominating the world floriculture markets due to their excellent ornamental value[20,21]. *C. sinense* and *C. haematodes* are terrestrial orchids with various flower colors and strong floral fragrance, making them valuable resources for *Cymbidium* cultivar improvement. *C. haematodes* was initially classified as a subspecies in *C. ensifolium*, and later as a variety of *C. sinense* (*C. sinense* var. haematodes). It is now widely recognized as a distinct species closely related to *C. sinense*[22]. The regulation mechanisms in the anthocyanin biosynthesis pathway have been studied in several Orchidaceae species, mainly focused on tropical orchids, such as *Phalaenopsis*[23], *C. hybridum*[24], and *Catilan*[25]. In *Catilan*, three R2R3-MYB transcription factors (RcPCP1, RcPAP1, and RcPAP2) in the 'KOVA' variety participate in the biosynthesis and accumulation of carotenoids and anthocyanins[25]. However, there are few reports on the regulation of anthocyanin biosynthesis in *Cymbidium*, and none have been reported in *C. haematodes*. Identification and functional characterization of the MYB transcription factors in *Cymbidium* could shed light on the mechanisms underlying anthocyanin biosynthesis and flower color formation in orchids.

In the present study, a genome-wide identification of MYB TFs in *C. sinense* was performed, and 170 high-quality CysMYBs identified. A comprehensive analysis of the molecular characteristics, phylogenetic relationships, and gene/protein structures of the CysMYBs was conducted. Combined with flower color phenotypes, transcriptome sequencing, and qRT-PCR, functional divergence of *CysMYBs* were revealed in anthocyanin biosynthesis regulation between *C. sinense* and *C. haematodes*. Taken together, this study provides a valuable resource for future investigations on the role of MYB proteins in regulating anthocyanin biosynthesis in orchids.

Materials and methods

Plant materials

The tepals (including sepals and petals) were collected from *C. sinense* variety 'Qihei', and six natural *C. haematodes* accessions with distinct flower colors: dark red (DR), red (R), light red (LR), yellow-green (YG), white-green (WG), and white (W), were used as the experimental materials. All plants were cultivated in the greenhouses at the Orchid Germplasm Resources Paddy of Zhejiang Province (Hangzhou, China). Healthy individuals from each *C. sinense* 'Qihei' and *C. haematodes* accessions were selected. The tepals were sampled 10 d after flowering under the same blooming condition. For each sample, three biological replicates were performed. All samples were immediately frozen in liquid nitrogen, and then stored at −80 °C for subsequent analysis.

Identification of CysMYBs

The genome assembly and annotation files of *C. sinense* were downloaded from NCBI (National Center for Biotechnology Information), BioSample ID: SAMN20059972. A total of 197 MYB protein sequences of *Arabidopsis thaliana* (AtMYB) were obtained from The Arabidopsis Information Resource (TAIR) (www.arabidopsis.org). The Hidden Markov Models (HMMs) of MYB conserved domains (PF00249, PF13921, PF14379) were retrieved from the Pfam database (http://pfam.xfam.org). Candidate MYB proteins were initially identified by performing a local BLASTP alignment against the *C. sinense* protein database using the 197 AtMYB protein sequences as query sequences, with an e-value cutoff of < 1e-5. Subsequently, the conserved MYB domain was used as a model to search and identify all candidate MYB proteins in the *C. sinense* genome, using the HMMER 3.0 software based on the HMMs. The parameters of HMMER were set as -E 1e-5 --domE 1e-5. The candidate sequences from both approaches were combined, and the domains were validated using InterProScan (www.ebi.ac.uk/interpro) and the NCBI Conserved Domain Database (CDD) (www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). Redundant and structurally incomplete proteins, and those lacking typical domains were removed, resulting in the final set of *C. sinense* MYB (CysMYB) proteins. The physicochemical properties of the CysMYB proteins were analyzed using ProtParam online tool (https://web.expasy.org/protparam), and the subcellular localization of CysMYB family members was performed by Wolf PSORT (https://wolfpsort.hgc.jp) Prediction.

Phylogenetic analysis

A phylogenetic tree was constructed using the MYB proteins from five species, including *A. thaliana*, *Oryza sativa*, *C. sinense*, *C. ensifolium*, and *C. goeringii*. The MYB protein sequences of *O. sativa* were retrieved from the PFAM database (https://phytozome.jgi.doe.gov). The MYB protein sequences of *C. ensifolium* and *C. goeringii* were obtained by BLAST against the genome data[26,27]. In addition, three separate phylogenetic trees were also constructed for R2R3-MYB, 1R-MYB, and the other MYBs from *A. thaliana* and *C. sinense*. All phylogenetic trees were constructed using MEGA X (V10.0, Tokyo Metropolitan University, Tokyo, Japan), with the Maximum Likelihood (ML) method. The bootstrap value was set to 1,000, and other parameters were retained as defaults. The *C. sinense* MYBs were classified into subgroups, based on the established *A. thaliana* classification system[1,6] and the branching topology of the phylogenetic tree. Finally, the evolutionary trees were refined using MEGA X (V10.0), and the online tool iTOL (https://itol.embl.de).

Multiple sequence alignment and WebLogo

The R2R3-MYB protein sequences from CysMYB were subjected to multiple sequence alignment using the DNAMAN software. After that, the conserved amino acid sequences within the R2 and R3 repeats were extracted and analyzed. These trimmed and aligned sequences were then submitted to WebLogo (https://weblogo.threeplusone.com/create.cgi) to generate a sequence logo, which visually represents the consensus and conservation of the R2 and R3 domains in *C. sinense* MYB proteins.

Conserved motif, gene structure, and chromosomal distribution analysis

The conserved motifs of the CysMYB protein were identified using the online tool MEME (http://meme-suite.org/tools/meme), with the maximum number of motifs set to 15, and other parameters set as default. The phylogenetic tree of the CysMYB protein sequences was constructed using MEGA X software (V10.0). The genome annotation file of *CysMYB* genes (GFF format) was obtained from the *C. sinense* genome assembly[28]. Subsequently, by integrating gene annotation, the conserved motifs and gene structures were visualized using TBtools software (v1.120)[29].

The chromosomal location of *CysMYB* gene family members was extracted from the *C. sinense* GFF file, and visualized using TBtools software (v1.120). To analyze segmental and tandem duplication events caused by gene duplication in the *C. sinense* genome, CysMYB protein sequences were analyzed using MCscanX software with default parameters, and TBtools (v1.120). This analysis generated a synteny map illustrating the collinear relationships.

Flavonoid content detection

The flavonoid contents were analyzed by MetWare (www.metware.cn) using an ultra-performance liquid chromatography-tandem mass spectrometry (UPLC-MS/MS) platform. The petal samples of *C. sinense* 'Qihei', and five *C. haematodes* accessions (DR, R, LR, YG, and WG) at the full blooming stage were collected from storage at −80 °C. The samples were ground into a fine powder using a ball mill, and 50 mg of the powder was weighed and transferred into a 2 mL centrifuge tube. Subsequently, 500 μL of extraction solution (50% methanol aqueous solution containing 0.1% hydrochloric acid) was added to the tube. The mixture was vortexed for 5 min using a multi-tube vortex mixer, followed by ultrasonication for 5 min. After overnight incubation at 4 °C, the samples were centrifuged at 12,000 rpm, and 4 °C for 3 min. The supernatant was collected, and the extraction procedure was repeated once. The two supernatants were combined and filtered through a 0.22 μm syringe filter. The resulting filtrate was transferred into a 2 mL sample vial as the final anthocyanin extract for subsequent measurement and analysis. Three biological replicates, and three technical replicates were performed for each sample.

Gene expression analysis of *CysMYB* genes

RNA sequencing was performed on tepals of *C. sinense* 'Qihei', and five *C. haematodes* accessions (DR, R, LR, YG, and WG). Total RNA was extracted from the tepals, and subjected to RNA-seq analysis, following a protocol adapted from Wang et al.[30]. Three biological replicates of each sample were conducted. Hisat2 (http://ccb.jhu.edu/software/hisat2) was used to compare reads with the *C. sinense* genome (version 4.03, http://solanaceae.plantbiology.msu.edu/pgsc_download.shtml). Transcripts were assembled and quantified using StringTie (http://ccb.jhu.edu/software/stringtie), where the

reference guide model was used to estimate transcript profiling. Expression levels of the *CysMYB* genes were calculated as FPKM (Fragments Per Kilobase of exon per million fragments Mapped) value using Cufflinks software (http://cole-trapnell-lab.github.io/cufflinks, v2.2.1). Differential expression analysis between sample groups was performed using DESeq2 to obtain differentially expressed gene sets under different biological conditions. TBtools software (v1.120) was used to draw the heatmaps of the *CysMYB* genes according to the FPKM values (Supplementary Table S1).

qRT-PCR analysis

qRT-PCR analysis was performed to analyze the expression of candidate genes in the tepals of *C. sinense* 'Qihei' (Mol), and six *C. haematodes* accessions with distinct flower colors (DR, R, LR, YG, WG, and W). Total RNA was extracted using the RNA extraction kit according to Wang et al.[30], and qualified RNA was reversely transcribed into cDNA using a cDNA synthesis system (Invitrogen, Shanghai, China). qRT-PCR primers (Supplementary Table S2) were designed by Primer Premier 5 software, and synthesized by Genepioneer Biotechnologies Co., Ltd. (Nanjing, China). The gene *Mol017122*, whose stable expression was experimentally validated, was used as the internal standard gene. Three biological replicates, and three technical replicates were performed for each sample. The relative expression levels (Supplementary Table S3) were calculated using the $2^{-\Delta\Delta CT}$ method, and are presented as mean ± standard deviation (SD) values of three biological replicates.

Subcellular localization

The full-length coding sequences of *CysMYB131* cDNA were cloned into the pCAMBIA1300 vector. An artificial green fluorescent protein (GFP) tag was fused to the C-terminus of CysMYB131, under the control of the 35S promoter. To verify the subcellular localization, the construct was co-transformed with a nuclear marker-DsRed into epidermal cells of *Nicotiana benthamiana* leaves via transient transformation. The primer sequences for vector construction were designed using CE Design software (Supplementary Table S4). Fluorescence signals from the fusion protein, and the DsRed marker were observed using an Olympus FV3000 confocal laser microscope (Hachioji City, Tokyo, Japan).

Statistical analysis

Statistical analysis and plotting were conducted using Graphpad Prism software (v. 8.4.3) using one-way ANOVA ($p < 0.05$). * indicates significant differences in values: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, and **** $p < 0.0001$.

Results

Genome-wide identification of the CysMYB proteins in *C. sinense*

A total of 180 MYB transcription factors were identified in *C. sinense*, and designated CysMYB1 to CysMYB180. The CysMYBs were analyzed with InterPro and CDD database to remove redundant and structurally incomplete sequences. Finally, 170 high-quality CysMYB proteins were retained for subsequent analysis (Supplementary Table S5). In total, we 104 R2R3-MYBs, 43 1R-MYBs, four 3R-MYBs, two atypical MYBs (CysMYB5R and CysMYBCDC), and 17 MYB-CC members were identified containing the conserved LHEQLE motif. Notably, no typical 4R-MYB was identified, whereas CysMYB5R with five MYB repeats was found instead. The CysMYB proteins varied in length from 122 aa (CysMYB105) to 1106 aa (CysMYB122), with molecular weights (MW) ranging from 13,528.72 to 124,716.93 kD. The theoretical isoelectric point (pI) ranged from 4.49 (CysMYB152) to 10.32 (CysMYB175). Subcellular localization prediction indicated that 164 MYB proteins (96.47%) were located in the nucleus. In addition to nuclear proteins, the CysMYBs were also predicted with mitochondrial, cytoplasmic, or chloroplast localizations. Detailed information, including gene/protein name, protein length, number of introns/exons, MYB-domain type, chromosome localization, isoelectric point, molecular weight, and subcellular localization prediction of the 170 CysMYBs wis presented in Supplementary Table S5.

Phylogenetic tree analysis of the MYB transcription factors

To investigate the phylogenetic relationships of MYB proteins, a multiple sequence alignment of 881 protein sequences from five different species, including *A. thaliana*, rice, and three *Cymbidium* species (*C. sinense*, *C. geogerii*, and *C. ensifolium*) were performed. A Maximum-Likelihood (ML) phylogenetic tree was constructed, which classified the proteins into six groups (Fig. 1a). Consistent with findings in other plant species, R2R3-MYBs constituted the predominant account of all proteins. To further elucidate the evolution and potential functions of the CysMYB, three phylogenetic trees were constructed for Cys1R-MYB (Supplementary Fig. S1), CysR2R3-MYB (Fig. 1b), and the remaining CysMYB proteins (Supplementary Fig. S2), respectively. In the present study, the 42 Cys1R-MYB in *C. sinense* and 52 At1R-MYB in *Arabidopsis* are clustered into 16 subgroups (Supplementary Fig. S1). Based on the established classification of *Arabidopsis* R2R3-MYB[1], the *C. sinense* R2R3-MYBs were divided into 33 subfamilies (Fig. 1b). These subfamilies are named continuously with the 25 AtMYBs subgroups (S1−S25), while the remaining CysR2R3-MYB subgroups designated S26−S33. Given that genes within the same phylogenetic clade or group often share similar functions, S4, S5, S6, S7 and S9 were focused on, which are known in *Arabidopsis* to positively or negatively regulate anthocyanin synthesis.

The S4 subfamily contains seven CysMYB genes and six AtMYB genes, suggesting a potential role in the negative regulation of anthocyanin synthesis. Meanwhile, three proteins (CysMYB6, CysMYB131, and CysMYB92) were grouped into the S5 subfamily, two (CysMYB168 and CysMYB107) in the S6 subfamily, and two (CysMYB68 and CysMYB74) in the S7 subfamily. The S9 subfamily comprised five proteins (CysMYB88, CysMYB71, CysMYB72, CysMYB114, and CysMYB134), which were grouped with known AtMYBs that are involved in controlling flower color. In addition to R2R3-S4 subfamily proteins, the four 3R-CysMYB transcription factors clustered with *Arabidopsis* MYB-LIKE2 (MYBL2, AT1G71030)[31,32] were also identified as potential negative regulators of anthocyanin biosynthesis. These 23 CysMYB genes were regarded as candidates with similar functions to the known AtMYBs in regulating anthocyanin biosynthesis and accumulation.

Conserved domain and structural characteristics of CysMYBs

The amino acid frequency within the repeat domains was investigated using 104 R2R3-MYB protein sequences. Multiple sequence alignment revealed that the conserved amino acids of R2 and R3 domains are [-W-(X20)-W-(X20)-W-] and [-F/I-(X19)W-(X19)-W-], respectively, where W represents tryptophane residues, F is phenylalanine, I is isoleucine, and X denotes any amino acid. Furthermore,

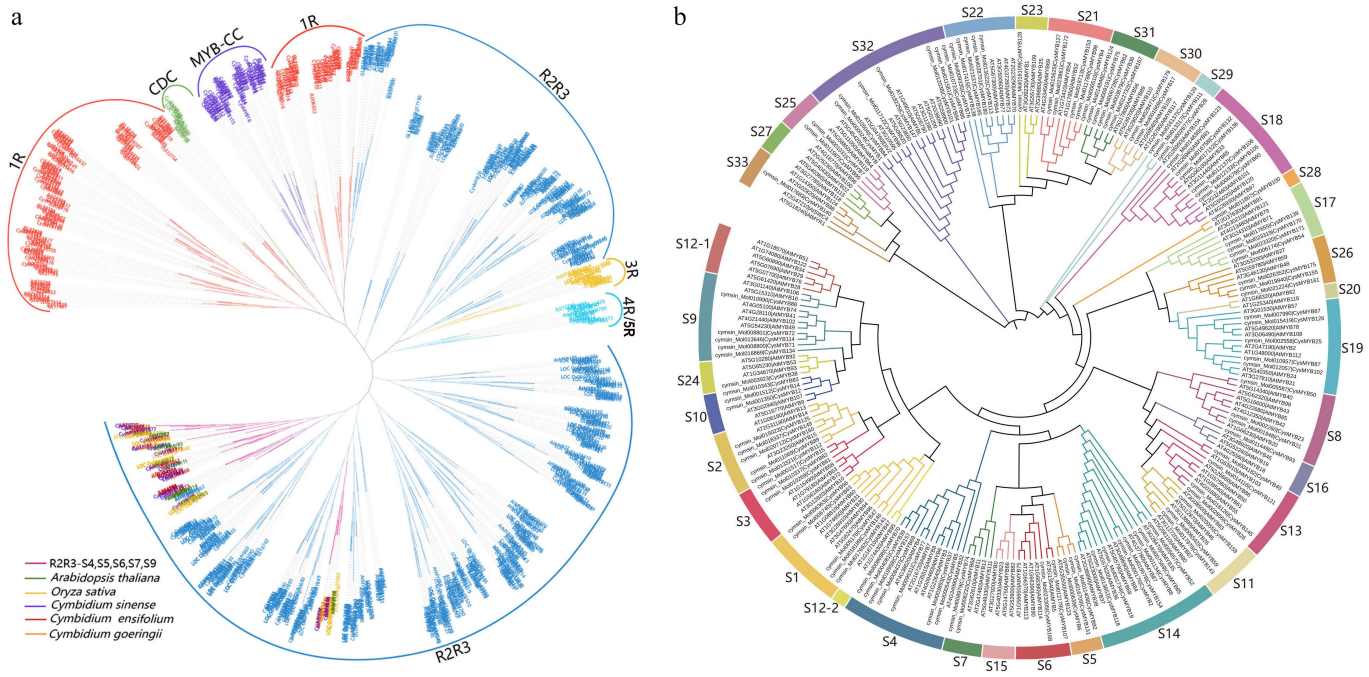

Fig. 1 Phylogenetic analysis of MYB proteins. The phylogenetic trees were constructed using MEGA X with the Maximum-likelihood (ML) method and 1,000 bootstrap replicates. (a) Phylogenetic tree of MYB gene families from five plant species, including *A. thaliana*, *O. sativa*, *C. sinense*, *C. goeringii*, and *C. ensifolium*. The tree was divided into six major groups, highlighted in different colors. Members from R2R3-S4, S5, S6, S7, and S9 subfamilies of the five species are marked with different colors. (b) Phylogenetic tree of R2R3-MYBs proteins of *C. sinense* and *A. thaliana*. The tree was classified into 33 phylogenetic subgroups (S1–S33), each indicated by a different color.

the R3 domain contains a specific binding site, '(D/E)Lx2(R/K)x3Lx6-Lx3R', which interacts with the bHLH protein (Fig. 2).

Gene structure and motif analysis of CysMYBs

The diversity in gene structure and motif composition reflects functional divergence among gene family members. The exon/intron pattern of the *CysMYB* family genes showed that most genes within the same subfamily share similar exon/intron structures, which are closely related to subfamily classification and functional roles. Furthermore, exon/intron structure analysis revealed considerable variations in the exon numbers among *CysMYB* subfamilies, ranging from 2 to 14, indicating that the exons underwent loss or gain during gene evolution of the MYB family (Fig. 3; Supplementary Fig. S3).

The most common exon numbers in *CysMYBs* were two or three. Specifically, 71 *MYB* genes were identified, with three exons and two introns, accounting for 41.76% of all *MYBs*. *CysMYB91* and *CysMYB135* possessed the highest number of exons and introns. To further investigate the sequence features, 15 conserved motifs in the CysMYB proteins (Motif1 to Motif15) were identified using MEME (Supplementary Fig. S3). Motifs 1–4 were present in the vast majority of the CysMYB proteins, and represent the characteristic DNA-binding domains of R2R3-CysMYB proteins. For example, motifs 3, 10, 12, and 13 are specific patterns to R2R3-S32, while motif 19 is specific to R2R3-S27 (Fig. 3; Supplementary Fig. S3). Meanwhile, the motif6 was unique to 1R-MYB, and motif12 was unique to R2R3-MYB. All seven CysMYBs in the S4 subfamily contain the 'EAR repression motif' (PDLNL(D/E)L) in motif 5 (PDLNLEL) (Supplementary Fig. S4). Except for motif1, 2, and 3, CysMYB88 also share motif8 with AtMYB106 and AtMYB16 (Supplementary Fig. S4). Whereas CysMYB114 and CysMYB134 do not contain motif10 or motif8, which may be due to the comparatively shorter length resulting from genome assembly errors (Supplementary Fig. S4).

Genomic distribution and syntenic analysis of the CysMYBs

Chromosomal localization analysis showed that the *CysMYB* genes were unevenly distributed across the 19 chromosomes of *C. sinense*. No *MYB* genes were on chromosome 11, while several chromosomes showed notable gene clustering (Fig. 4a). Chromosome 2 contained the largest number of *CysMYBs* (20 genes), followed by chromosomes 10 (17 genes), and 8 (15 genes). In contrast, chromosomes 15 and 20 contained the fewest, each with only two *CysMYB* members. The positions of several genes on the same chromosome are relatively close, such as *CysMYB103* and *CysMYB104* on chromosome 3, *CysMYB26* and *CysMYB27* on chromosome 9, and *CysMYB110*, *CysMYB111*, and *CysMYB112* on chromosome 10. This indicates that these genes may be homologous genes arising from segmental or tandem duplication. Intra-genomic collinearity analysis further identified 23 syntenic gene pairs within the *CysMYB* family genes, such as *CysMYB55*/*CysMYB73*, *CysMYB167*/*CysMYB124*, and *CysMYB118*/*CysMYB65* (Fig. 4b). These results indicate that the gene duplication has contributed to the expansion of the *CysMYB* family, potentially accompanied by functional divergence among its members.

Analysis of flavonoid components

The composition and content of flavonoids were analyzed in the tepals of *C. sinense* 'Qihei' (Mol), and five *C. haematodes* accessions with distinct flower colors including dark red (DR), red (R), light red (LR), yellow green (YG), and white green (WG) using LC-MS/MS (Supplementary Table S6). Five classes of flavonoid metabolites were detected: proanthocyanidins, anthocyanins, flavonols, dihydroflavonols, and flavanones (Supplementary Fig. S5). The flavonoid composition in *C. sinense* 'Qihei', and five *C. haematodes* accessions exhibited remarkable differences. Flavonols were the most

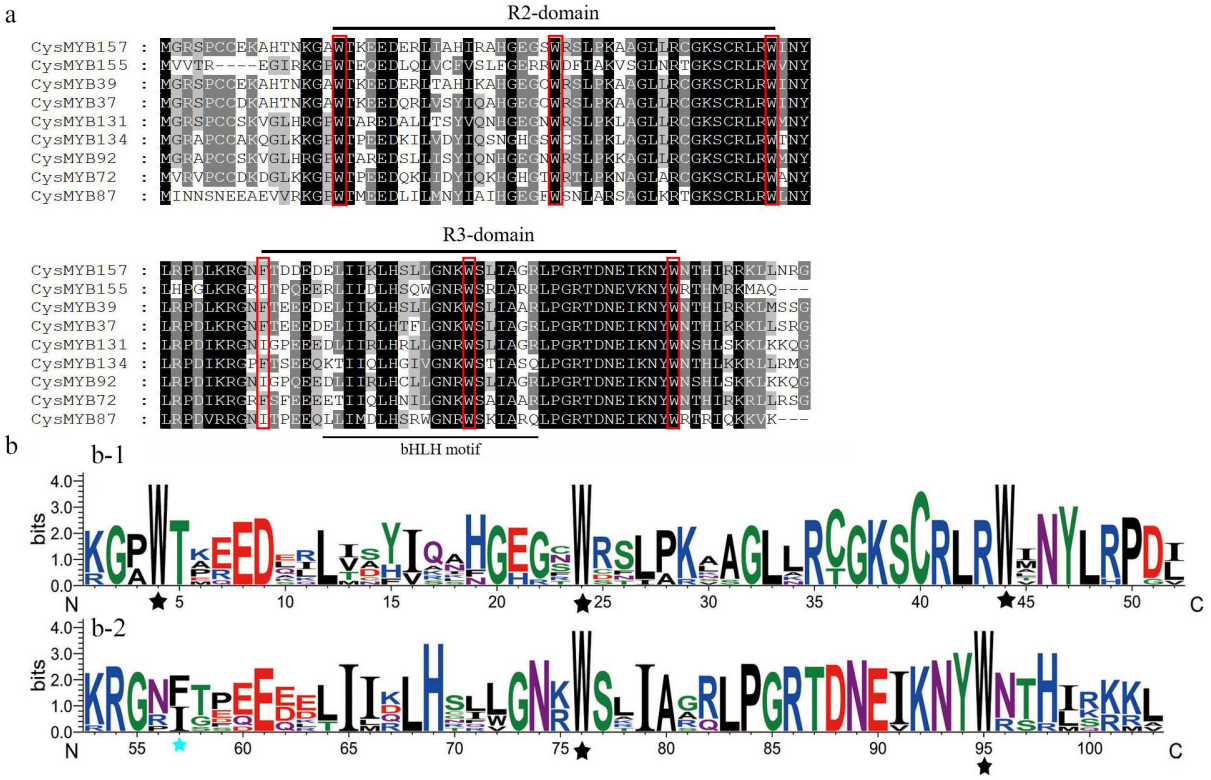

Fig. 2 Consensus sequences, and the level of conservation of R2R3-type and R3-MYB domains from *C. sinense*. (a) R2 and R3 domains, based on nine R2R3-MYB alignments. The conserved tryptophan residues (Trp, W) in the MYB domain and the replaced amino acid residues in the R3 repeat were indicated with red boxes. (b) Sequence logos of the R2 and R3 MYB repeats were obtained based on full-length alignments of 104 R2R3-CysMYB. Black asterisks indicate the W residues in R2 and R3 domains, while the blue asterisk indicates the replaced amino acid residues in the R3 repeat.

abundant class in the five *C. haematodes* accessions, accounting for 86.2%–96.6% of total flavonoids (Supplementary Fig. S5). Anthocyanins were the second most abundant, with significantly higher content in DR (1,557.10 ug/g), R (2,096.26 ug/g), and LR (1,766.88 ug/g) than in the other *C. haematodes* accessions and *C. sinense* (Supplementary Table S6; Supplementary Fig. S5).

Cyanidin-based anthocyanins were identified as the major pigments responsible for tepal coloration in both *C. haematodes* and *C. sinense* (Fig. 5). Total contents of anthocyanidins in the DR, R, LR, YG, and WG accessions were 1,483.44 μg/g (96.8%), 1,985.37 μg/g (95.9%), 1,622.87 μg/g (92.3%), 297.51 μg/g (67.3%), and 438.43 μg/g (93.9%), respectively (Fig. 5b). Peonidin ranked the third most abundant anthocyanidin in the five *C. haematodes* accessions, whereas delphinidin content was relatively higher in *C. sinense* (13.91 ug/g) than in *C. haematodes* (1.55–5.14 ug/g). Cyanidin levels were significantly elevated in the red-colored accessions (DR, R, and LR) compared to light-colored accessions (YG and WG). These results demonstrate that cyanidin derivatives are the key anthocyanin contributing to color formation in the tepals of *C. haematodes* and *C. sinense* orchids, despite the variations in the composition of other anthocyanins.

Diverse expression patterns of the *CysMYB* genes

To investigate the roles of *CysMYB* genes in anthocyanin synthesis, the expression patterns in *C. sinense* and the closely related species *C. haematodes* were analyzed. Transcriptome sequencing was performed using tepals of *C. sinense* 'Qihei', and five *C. haematodes* accessions (designated as *C. haematodes*-DR, R, LR, YG, WG). A total of 148 *CysMYB*s were expressed across all 18 samples, exhibiting species-specific expression patterns (Supplementary Fig. S6; Supplementary Table S1). Genes within the same phylogenetic subfamily generally exhibit similar expression patterns. Based on the classification of *CysMYB* and *AtMYB*, 23 *R2R3-MYB* genes in subfamilies S4, S5, S6, S7, and S9 may be involved in anthocyanin synthesis. The expression levels of 31 selected *CysMYB* genes are presented in the heatmap in Fig. 6. Among the positively regulated genes, *cymsin_Mol011408* (*CysMYB92*) and *cymsin_Mol016709* (*CysMYB131*) were highly expressed in R and DR tepals of *C. haematodes*, but showed relatively low expression in other samples. Considering that cyanidin-based derivatives were the predominant anthocyanins (Fig. 5), these genes may participate in the regulation of cyanidin biosynthesis in *C. haematodes* and *C. sinense* tepals. *CysMYB79* that belongs to MYB-CC also has relative high expression levels in R and DR.

In addition, several genes showed distinct sample-specific expression. For instance, *CysMYB87* was highly expressed in R tepals. In contrast, *CysMYB114* (*cymsin_Mol013646*) was strongly expressed in *C. sinense* tepals, indicating a species-specific function in anthocyanin synthesis. Similarly, *CysMYB30*, *CysMYB72*, *CysMYB99*, and *CysMYB157* were all highly expressed in *C. sinense*, but lowly expressed across *C. haematodes* accessions, suggesting that they may be *C. sinense*-specific regulators. Moreover, the expression of these genes arepositively correlated with the relatively high level of pelargonidins and delphinidins specifically detected in *C. sinense* tepals. Six genes, including *CysMYB107*, *CysMYB102*, *CysMYB39*, *CysMYB70*, *CysMYB77*, and *CysMYB6*, were exclusively highly expressed in *C. haematodes*-YG, and low in other samples, suggesting a potential role in flavonoid synthesis and accumulation in YG. In conclusion, *CysMYB* genes exhibit functional divergence, and species-specific regulatory roles in anthocyanin synthesis between *C. sinense*, and the close relative *C. haematodes*.

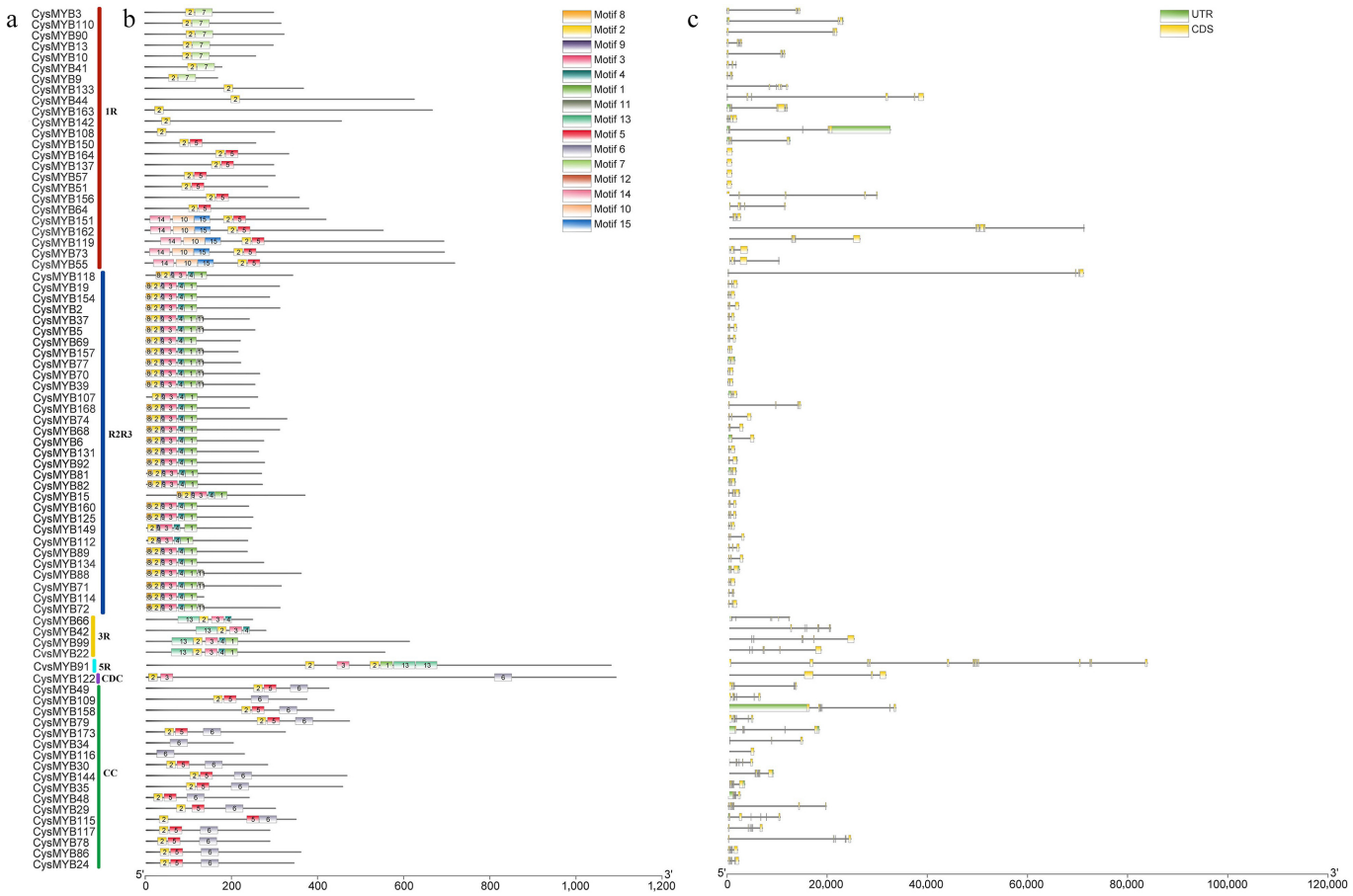

Fig. 3 Conserved motif and gene structure analysis of 78 selected *CysMYBs*, including 24 1R-MYBs (red bar), 31 R2R3-MYBs (blue bar), four 3R-MYBs (yellow bar), one 5R-MYB (cyan bar), one MYBCDC (purple bar), and 17 MYB-CC (green bar). (a) Protein names and subgroups. (b) Schematic representation of conserved motifs in CysMYB proteins. Different motifs are indicated with of different color boxes, and non-conserved sequences are represented by black lines. (c) Exon/intron structures of the *CysMYB* genes. Yellow boxes indicate exons, green boxes indicate UTR, and the lines between boxes indicate introns. The scale bars below (b) and (c) illustrate the length of the genes and motifs, respectively.

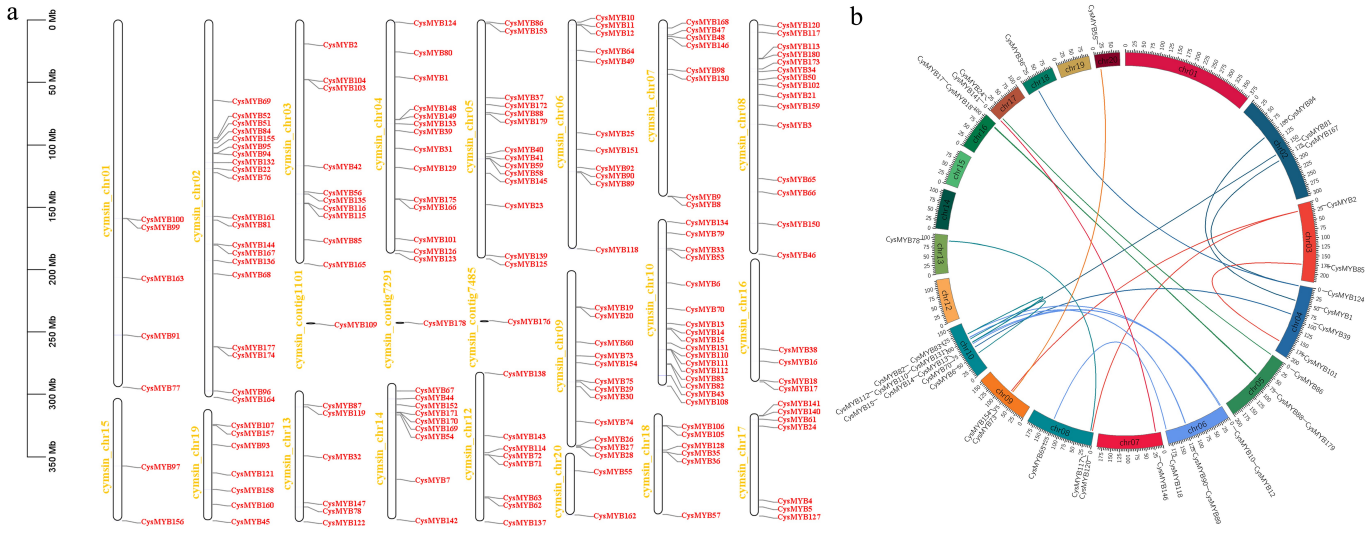

Fig. 4 Chromosomal distribution and gene duplications of the *C. sinense MYBs* genes. (a) Chromosomal locations of the CysMYBs in *C. sinense* that are mapped based on the of the genome annotation of *C. sinense*. The chromosome/contig ID is indicated on the left side of the chromosomes/contigs. The scale is indicated in megabase (Mb). (b) Collinearity analysis of the *C. sinense MYBs* genes. The syntenic *CysMYB* gene pairs are indicated by different color lines and marked on the out layer of the chromosomes.

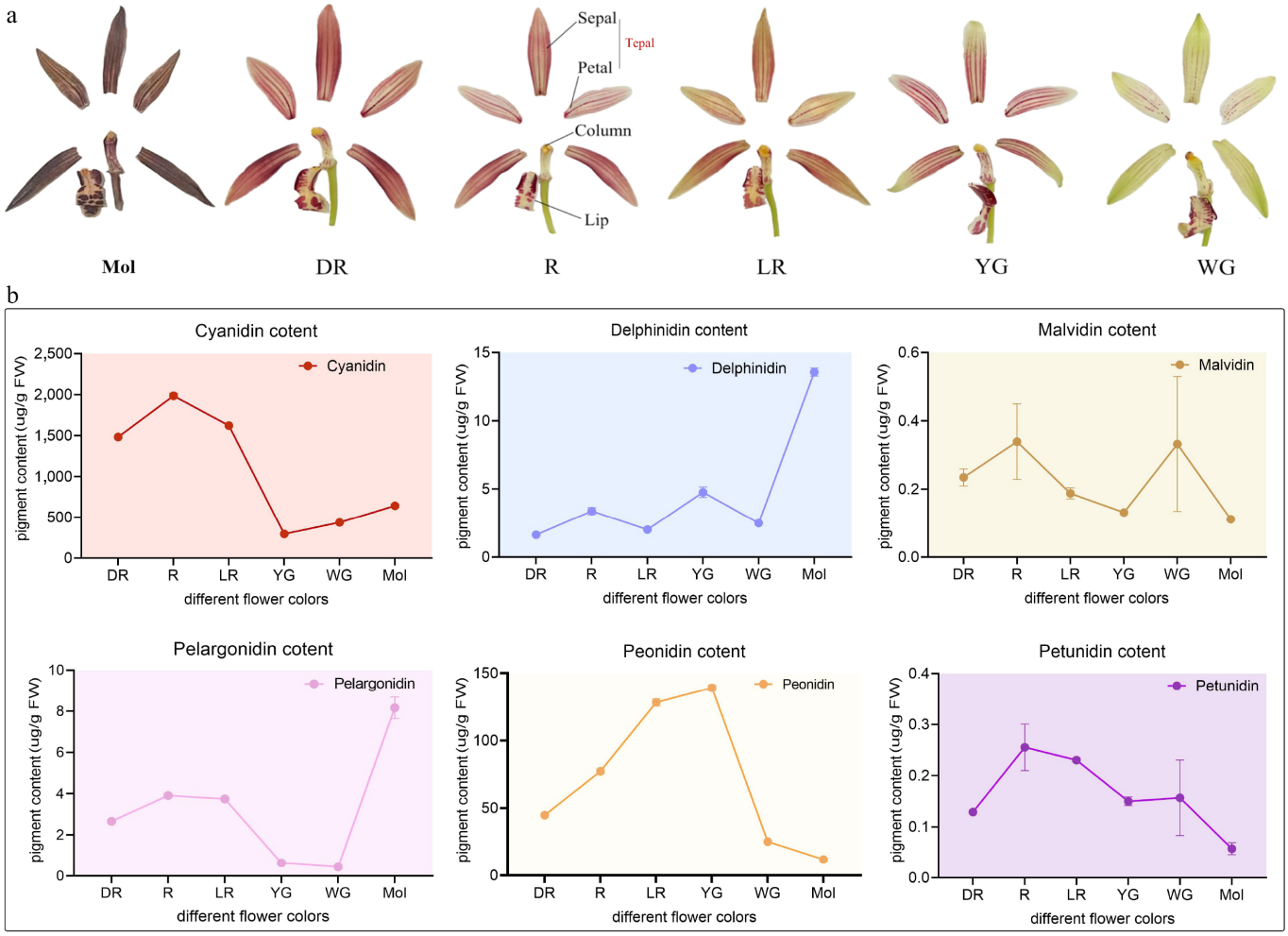

Fig. 5 The major anthocyanin composition and floral organ structure of *C. sinense* (Mol) and *C. haematodes*-DR, R, LR, YG, and WG. (a) Foral organ structure of the *Cymbidium* species for anthocyanin content detection. DR: dark-red tepals; R: red tepals; LR: light-red tepals; YG: yellow-green tepals; WG: white-green tepals. (b) Anthocyanin composition and content among different samples. The x-axis represents samples with different flower colors, *C. sinense* (Mol) and *C. haematodes* accessions with different flower colors (DR, R, LR, YG, and WG). The y-axis represents the content of pigments (ug/g FW). The pigments are indicated in different colors of dots, lines, and shades. Cyanidin: red; Delphinidin: blue; Malvidin: brown; Pelargonidin: pink; Peonidin: yellow; Petunidin: purple.

qPCR analyses of *CysMYB* genes in different colors of tepals

To verify the function of *CysMYB* genes in anthocyanin biosynthesis, qPCR was performed to detect the expression levels in the tepals of *C. sinense*, and six *C. haematodes* accessions with different flower colors (DR, R, LR, YG, WG, and W). Nine candidate *CysMYBs* were selected and verified using qPCR according to phylogenetic analysis and gene expression abundance in RNA-seq data (Fig. 7; Supplementary Table S3). The results revealed significant differences in the expression levels of the nine genes across the tepals of different colors (Fig. 7). Among them, *CysMYB72* showed a significantly high expression level in *C. sinense*, but low or nearly undetectable expression in *C. haematodes*, suggesting that *CysMYB72* may be a species-specific regulator of anthocyanin synthesis in *C. sinense*. Similarly, *CysMYB114* was also highly expressed in *C. sinense*. Different from the RNA-seq expression, *CysMYB88* exhibits high expression levels in red spectrum samples (Mol, DR, R and LR), but low expression level in light-colored samples YG, WG, and W. These results suggest that *CysMYB72* and *CysMYB114* may play a function in flower color determination in *C. sinense*, whereas *CysMYB131* and *CysMYB88* are potential regulators of red coloration in *C. haematodes*. However, their specific roles in anthocyanin synthesis require further experimental validation.

Subcellular localization analysis

According to online prediction, the majority of the CysMYB proteins (96.47%) were localized to the nucleus (Supplementary Table S5). To verify this prediction, CysMYB131 was selected for experimental subcellular localization analysis. The CysMYB131-GFP protein was transiently expressed in the epidermal cells of *Nicotiana benthamiana* leaves, which showed that CysMYB131 was localized to the nucleus (Fig. 8). Moreover, the nucleus marker DsRed was used to co-transfect with the CysMYB131 protein. The results confirmed that CysMYB131 is a nucleus-localized transcription factor (Fig. 8).

Discussion

The MYB transcription factor family is among the largest transcription factor superfamily in all eukaryotes[1,3], which has been identified in numerous plant species. *C. sinense* and its close relative

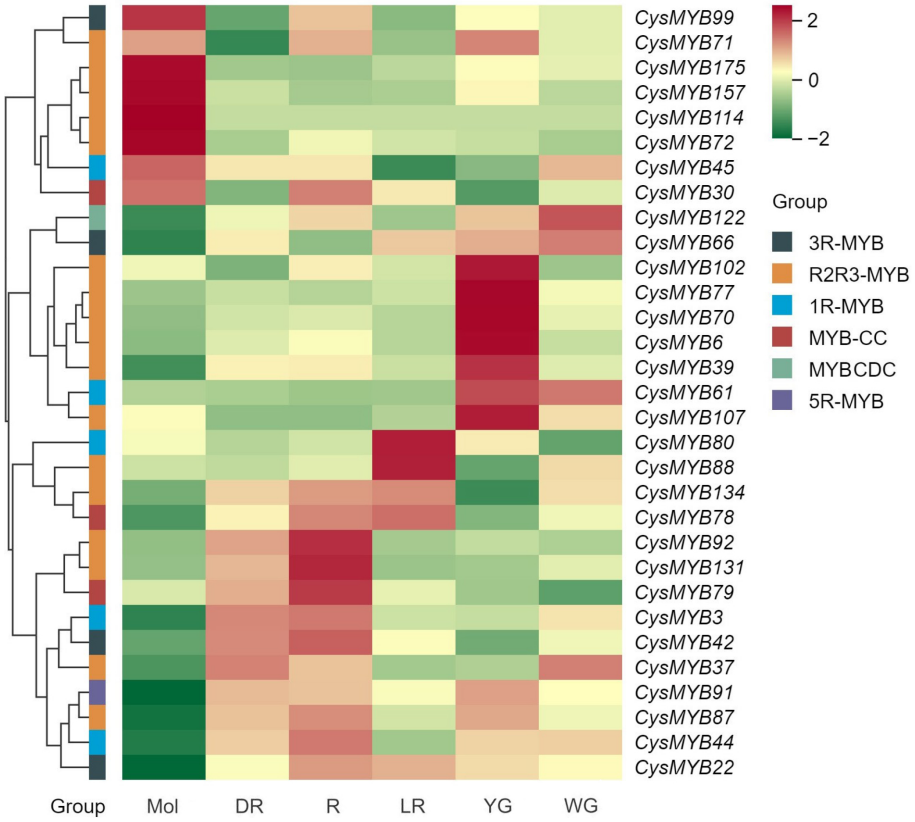

Fig. 6 Expression patterns of 31 *CysMYBs* in tepals of *C. sinense* (Mol) and *C. haematodes* accessions with different flower colors (DR, R, LR, YG, and WG), according to the FPKM values. The CysMYB groups are indicated with different color blocks: 1R-MYB (black), R2R3-MYB (yellow), 3R-MYB (red), 5R-MYB (green), MYB-CC (blue), MYBCDC (purple). DR: dark-red tepals; R: red tepals; LR: light-red tepals; YG: yellow-green tepals; WG: white-green tepals.

C. haematodes are valuable resources for orchid breeding due to their unique floral structures and rich coloration. However, research on the molecular mechanisms underlying floral pigmentation in *Cymbidium* species remains limited. The recent release of the whole-genome sequence of *C. sinense*[28] has enabled genome-wide analysis of the *MYB* genes. In the present study, 170 high-quality CysMYBs were identified in the *C. sinense* genome, which is larger than that of *Solanum lycopersicum* (127 SlMYBs)[33], but lower than that in *Arabidopsis* (197 AtMYBs)[34], *O. sativa* (239 OsMYBs)[35], and *S. tuberosum* (253 StMYBs)[36]. R2R3-MYB proteins represent the most predominant and prevalent MYB in higher plants, with functions in responses to stresses[37–39], specialized metabolism[1,10,40], cell differentiation[41], and development[42]. One hundred and four R2R3-CysMYBs were identified, accounting for 61.18% of all CysMYBs (Fig. 1). The number was similar to that in other Orchidaceae species, such as *C. goeringii*[43] (104 R2R3-MYBs), *C. ensifolium*[44] (102 R2R3-MYBs), *Dendrobium officinale*[45] (101 R2R3-MYBs), and *P. equestris* (96 R2R3-MYBs)[46]. The predominance of R2R3-MYBs aligns with findings in other plant species, reinforcing their critical role in regulating various biological processes. Moreover, no 4R-CysMYB members were identified, while a unique 5R-MYB (CysMYB5R) was discovered in *C. sinense*, suggesting evolutionary divergence in the MYB family. Conserved tryptophan residues in R2 and R3 domains (Fig. 2) are critical for DNA binding and interaction with bHLH proteins, forming the MBW complex that regulates anthocyanin biosynthesis. Structural analysis of the CysMYBs, including conserved motifs and gene architectures (Fig. 3), support functional divergence and specialization among MYB subfamilies. Meanwhile, the variability in exon numbers and motif distributions among subfamilies underscores the evolutionary plasticity of MYB genes, which may drive their functional diversification.

Anthocyanins comprise various anthocyanidin pigments that impart floral organs with hues ranging from red, orange, to blue and purple[47]. Previous studies have shown that cyanidin, pelargonidin, and peonidin are the predominant anthocyanidins in Orchidaceae plants[21,48,49]. Cyanidins are identified as the main contributors to the coloration of perianths in certain cultivars of *Pleione* spp.[49], *Oncidium*[50], and *Cattleya hybrid*[25]. Similarly, in *C. hybrids*, cyanidin and peonidin have been reported as the basis for flower color variation[51]. In agreement with these findings, it was found that the anthocyanins in the tepals of *C. sinense* and *C. haematodes* are primarily cyanidin derivatives (Fig. 5; Supplementary Fig. S5). Interestingly, the anthocyanin content in *C. sinense* exhibited a significant reduction compared to *C. haematodes* accessions. Moreover, although DR accession exhibited a darker floral hue than R, the contents of total flavonoid, anthocyanin, and cyanidin all peaked in R, followed by LR and DR. These results suggested the participation of other pigments (i.e., flavonols and carotenoids) in determining the flower color of *C. sinense* and *C. haematodes*.

Studies in many plant species have established that the S4, S6, and S7 subgroups of the R2R3-MYB subfamily play key roles in regulating anthocyanin biosynthesis. The mechanisms underlying anthocyanin biosynthesis pathway have been studied in several Orchidaceae species, mainly focused on the tropical orchids. For example, in *Phalaenopsis*, *PeMYB2*, *PeMYB11*, and *PeMYB12* regulate the formation of red background coloration, spot patterns, and red veins in sepals and petals[23]. The coloration of the base color and spots in the lip flap is determined by *PeMYB12* and *PeMYB11*, respectively[23].

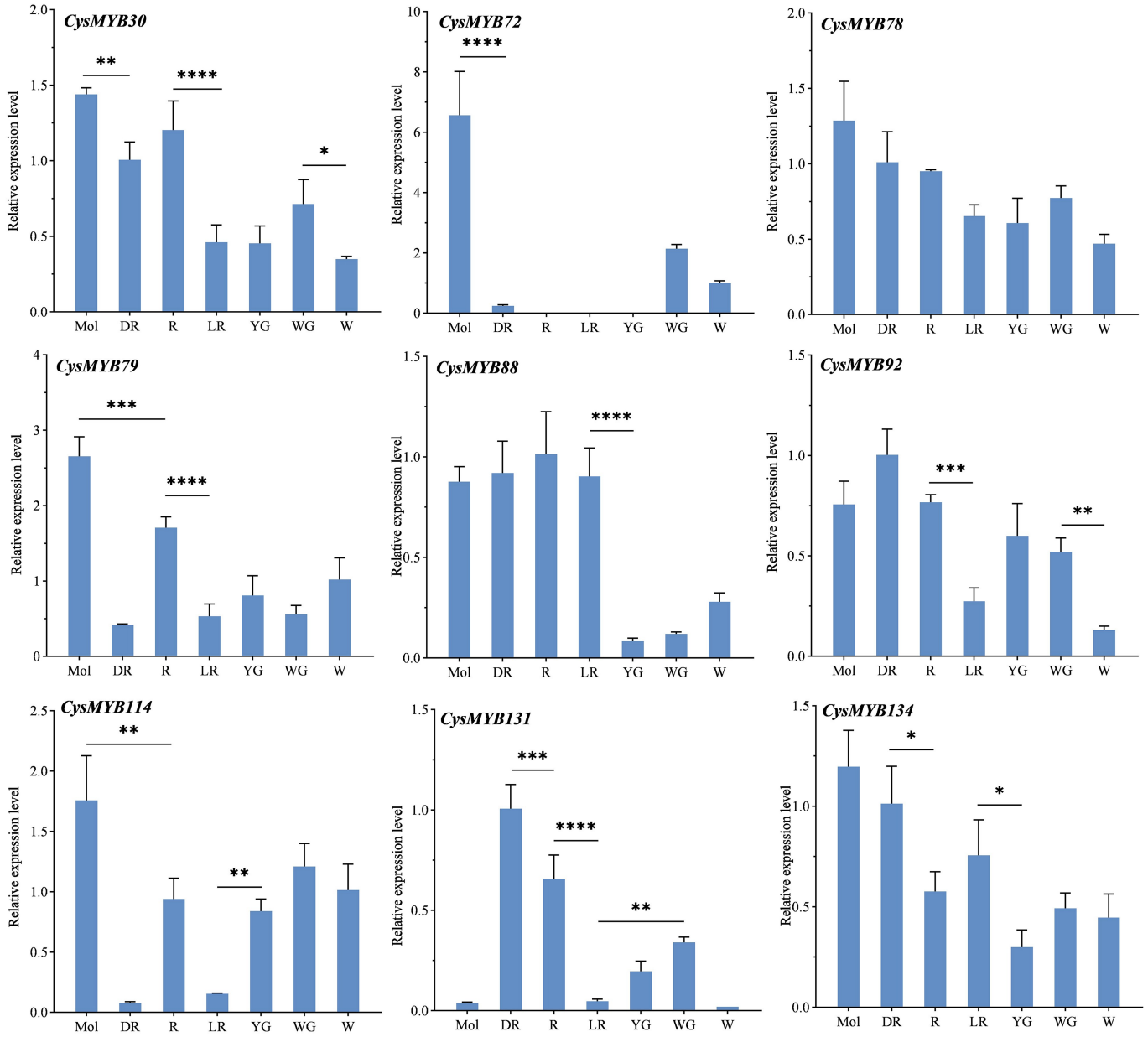

Fig. 7 The expression of nine *CysMYBs* genes among different flower colors of *C. sinense* and *C. haematodes* using qRCR. Values are mean ± standard deviation (SD) of three biological replicates. Statistics were analyzed using one-way ANOVA ($p < 0.05$), and * indicates significant differences in values. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, and **** $p < 0.0001$. Mol, *C. sinense* 'Qihei'; DR, R, LR, YG, WG, and W indicate *C. haematodes* accessions with different colors. DR: dark-red tepals; R: red tepals; LR: light-red tepals; YG: yellow-green tepals; WG: white-green tepals; W: white tepals.

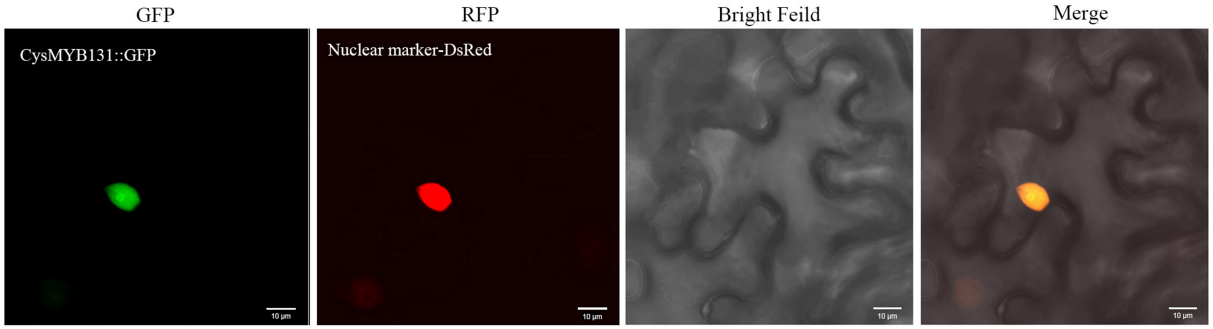

Fig. 8 Subcellular localization of the CysMYB131 protein. Gene-GFP fusion construct was transiently expressed in *Nicotiana benthamiana*. From left to right: fluorescence of the gene-GFP fusion construct (green), fluorescence of the nuclear marker DsRed (red), bright-field, and merged images.

In the present study, among the 104 R2R3-CysMYBs identified in *C. sinense*, seven R2R3-CysMYBs were from the S4 subgroup, two from the S6 subgroup, and two from the S7 subgroup. Previous studies in related orchids have implicated S6 subfamily members in flower color variation. In *C. ensifolium*, *CeMYB52* and *CeMYB104* (S6 subfamily) are proposed as the key genes controlling flower color variation[44]. Similarly, the *CgMYB91* of the S6 subfamily was also associated with anthocyanin formation in *C. goeringii*, as well as *CgMYB32* of S4 subfamily[43]. However, the present expression profile revealed that most *CysMYBs* associated with anthocyanin biosynthesis from subfamilies S4, S6, and S7 exhibited relatively low transcript abundance across red-flowered accessions of both *C. haematodes* and *C. sinense*. In contrast, two S5 subfamily members, *CysMYB131* and *CysMYB92*, were identified as positive regulators of anthocyanin accumulation in red-flowered *C. haematodes* accessions (DR and R). Their expression patterns closely mirrored the variation in cyanidin content across different color accessions of *C. haematodes*.

Phylogenetic analysis classified *CysMYB131* as a TT2-type regulatory gene. Previous studies reported that *AtMYB123* (TT2), a member of the S5 subfamily in *A. thaliana*, promotes proanthocyanidin accumulation in seed coats[52]. However, heterologous overexpression of the TT2 ortholog *FcMYB123* from *Ficus carica* significantly increased anthocyanin levels in apple fruits and calli[53]. Similarly, the peach *TT2-like* gene influences floral coloration by promoting red speckle formation in white petal regions[54]. Collectively, these findings suggest that *CysMYB131* may also function in anthocyanin biosynthesis, and accumulation in *C. haematodes*. Given that cyanidin-based derivatives are the major anthocyanins in *C. haematodes* petals (Fig. 5), *CysMYB131* is proposed as the potential gene in regulating cyanidin biosynthesis in *C. haematodes*. Conversely, the species-specific expression of *CysMYB72* and *CysMYB114* (both from subfamily S9) in *C. sinense* suggests divergent regulatory mechanisms between closely related species. The relative high levels of pelargonidin and delphinidin in *C. sinense* tepals imply that these species-specific expressed genes may influence flower coloration by regulating the biosynthesis of pelargonidin and delphinidin. Recent studies have revealed that *AtMYB106* modulates floral coloration by suppressing petal epidermal cell morphogenesis[55], and also frequently forms the MBW complex to promote anthocyanin biosynthesis in plants[56]. Similarly, *CysMYB88* that also from the S9 subfamily is hypothesized to be a potential MYB transcription factor regulating anthocyanin synthesis in *C. sinense* and *C. haematodes* (Fig. 7).

Further research is needed to investigate the regulatory effects of these candidate MYB transcription factors on the expression of upstream structural genes, and to validate the functions of the MYB TFs along with those of their target genes. These findings highlight the functional diversification of MYB TFs in Orchidaceae, contributing to the vast floral color diversity observed in this family.

Conclusions

This study presents a comprehensive genome-wide identification and functional characterization of MYB transcription factors in *C. sinense*, and its close relative *C. haematodes*, revealing their critical roles in anthocyanin biosynthesis and floral coloration. Transcriptome and qRT-PCR analyses highlighted species-specific expression patterns, such as the high expression of *CysMYB72* and *CysMYB114* in *C. sinense*. On the other hand, *CysMYB131* and *CysMYB88* were strongly associated with cyanidin accumulation in *C. haematodes*. These findings underscore the evolutionary diversification of MYB TFs in orchids and their pivotal role in regulating the biosynthesis of distinct anthocyanins. The findings provide valuable genetic resources for future studies on floral pigmentation. Future research should focus on functional validation of these candidates and their interactions within the anthocyanin regulatory network.

Author contributions

The authors confirm their contributions to the paper as follows: study conception and design: Wang Y, Sun C; data collection: Zhang Q, Zhai G; analysis and interpretation of results: Li S, Luo J, Shi Z; draft manuscript preparation: Wang Y, Li S. All authors reviewed the results and approved the final version of the manuscript.

Data availability

The raw data of the RNA-seq experiment is deposited in the Sequence Read Archive (NCBI): PRJNA1281345 (https://dataview.ncbi.nlm.nih.gov/object/PRJNA1281345). All data and material used in this study are available from the corresponding author upon reasonable request.

Acknowledgments

This research was supported by the 'Pioneer' and 'Leading Goose' R&D Program of Zhejiang (Grant No.2023C02028-1), and the New Variety Breeding Project of the Major Science and Technology Projects of Zhejiang (Grant No. 2021C02071-5).

Conflict of interest

The authors declare that they have no conflict of interest.

Supplementary information accompanies this paper online at: https://doi.org/10.48130/opr-0026-0003.

Dates

Received 20 August 2025; Revised 3 December 2025; Accepted 10 December 2025; Published online 25 March 2026

References

[1] Dubos C, Stracke R, Grotewold E, Weisshaar B, Martin C, et al. 2010. MYB transcription factors in *Arabidopsis*. *Trends in Plant Science* 15:573–581

[2] Li W, Zhu Z, Chern M, Yin J, Yang C, et al. 2017. A natural allele of a transcription factor in rice confers broad-spectrum blast resistance. *Cell* 170:114–126.e15

[3] Jiang CK, Rao GY. 2020. Insights into the diversification and evolution of R2R3-MYB transcription factors in plants. *Plant Physiology* 183:637–655

[4] Cao Y, Jia H, Xing M, Jin R, Grierson D, et al. 2021. Genome-wide analysis of *MYB* gene family in Chinese bayberry (*Morella rubra*) and identification of members regulating flavonoid biosynthesis. *Frontiers in Plant Science* 12:691384

[5] Wang X, Zhao S, Zhou R, Liu Y, Guo L, et al. 2023. Identification of *Vitis vinifera* MYB transcription factors and their response against grapevine berry inner necrosis virus. *BMC Plant Biology* 23:279

[6] Stracke R, Werber M, Weisshaar B. 2001. The *R2R3-MYB* gene family in *Arabidopsis thaliana*. *Current Opinion in Plant Biology* 4:447–456

[7] Yu X, Tang L, Tang X, Mao Y. 2023. Genome-wide identification and analysis of *MYB* transcription factors in *Pyropia yezoensis*. *Plants* 12:3613

[8] Paz-Ares J, Ghosal D, Wienand U, Peterson PA, Saedler H. 1987. The regulatory *c1* locus of *Zea mays* encodes a protein with homology to

myb proto-oncogene products and with structural similarities to transcriptional activators. *The EMBO Journal* 6:3553−3558

[9] Wu Y, Wen J, Xia Y, Zhang L, Du H. 2022. Evolution and functional diversification of R2R3-MYB transcription factors in plants. *Horticulture Research* 9:uhac058

[10] Stracke R, Ishihara H, Huep G, Barsch A, Mehrtens F, et al. 2007. Differential regulation of closely related R2R3-MYB transcription factors controls flavonol accumulation in different parts of the *Arabidopsis thaliana* seedling. *The Plant Journal* 50:660−677

[11] Li C, Yu W, Xu J, Lu X, Liu Y. 2022. Anthocyanin biosynthesis induced by MYB transcription factors in plants. *International Journal of Molecular Sciences* 23:11701

[12] Pelletier MK, Burbulis IE, Winkel-Shirley B. 1999. Disruption of specific flavonoid genes enhances the accumulation of flavonoid enzymes and end-products in *Arabidopsis* seedlings. *Plant Molecular Biology* 40:45−54

[13] Ilk N, Ding J, Ihnatowicz A, Koornneef M, Reymond M. 2015. Natural variation for anthocyanin accumulation under high-light and low-temperature stress is attributable to the *ENHANCER OF AG-4 2 (HUA2)* locus in combination with *PRODUCTION OF ANTHOCYANIN PIGMENT 1 (PAP1)* and *PAP2*. *New Phytologist* 206:422−435

[14] Yan H, Pei X, Zhang H, Li X, Zhang X, et al. 2021. MYB-mediated regulation of anthocyanin biosynthesis. *International Journal of Molecular Sciences* 22:3103

[15] Zhou H, Kui LW, Wang F, Espley RV, Ren F, et al. 2019. Activator-type R2R3-MYB genes induce a repressor-type R2R3-MYB gene to balance anthocyanin and proanthocyanidin accumulation. *New Phytologist* 221:1919−1934

[16] Li L, Li S, Ge H, Shi S, Li D, et al. 2021. A light-responsive transcription factor SmMYB35 enhances anthocyanin biosynthesis in eggplant (*Solanum melongena* L.). *Planta* 255:12

[17] Yin Y, Guo C, Shi H, Zhao J, Ma F, et al. 2022. Genome-wide comparative analysis of the *R2R3-MYB* gene family in five Solanaceae species and identification of members regulating carotenoid biosynthesis in wolfberry. *International Journal of Molecular Sciences* 23:2259

[18] Aguiar JMRBV, Giurfa M, Sazima M. 2020. A cognitive analysis of deceptive pollination: associative mechanisms underlying pollinators' choices in non-rewarding colour polymorphic scenarios. *Scientific Reports* 10:9476

[19] Basist G, Dyer AG, Garcia JE, Raleigh RE, Lawrie AC. 2021. Why variation in flower color may help reproductive success in the endangered Australian orchid *Caladenia fulva*. *Frontiers in Plant Science* 12:599874

[20] Singh DR, Pamarthi RK, Kumar R, Rai D, Meitei AL, et al. 2019. Traditional artifacts from dried leaves of *Cymbidium species* (orchidaceae) in Indian state of Sikkim. *Indian Journal of Traditional Knowledge* 18:390−394

[21] Zhou Z, Ying Z, Wu Z, Yang Y, Fu S, et al. 2021. Anthocyanin genes involved in the flower coloration mechanisms of *Cymbidium kanran*. *Frontiers in Plant Science* 12:737815

[22] Wu Y, Raven P. 2013. *Floral of China*. Beijing: CSPM China; London: HarperCollins UK. pp. 274−275 https://www.iplant.cn/foc

[23] Hsu CC, Chen YY, Tsai WC, Chen WH, Chen HH. 2015. Three R2R3-MYB transcription factors regulate distinct floral pigmentation patterning in *Phalaenopsis* spp. *Plant Physiology* 168:175−191

[24] Nakatsuka T, Suzuki T, Harada K, Kobayashi Y, Dohra H, et al. 2019. Floral organ- and temperature-dependent regulation of anthocyanin biosynthesis in *Cymbidium* hybrid flowers. *Plant Science* 287:110173

[25] Li BJ, Zheng BQ, Wang JY, Tsai WC, Lu HC, et al. 2020. New insight into the molecular mechanism of colour differentiation among floral segments in orchids. *Communications Biology* 3:89

[26] Ai Y, Li Z, Sun WH, Chen J, Zhang D, et al. 2021. The *Cymbidium* genome reveals the evolution of unique morphological traits. *Horticulture Research* 8:255

[27] Sun Y, Chen GZ, Huang J, Liu DK, Xue F, et al. 2021. The *Cymbidium goeringii* genome provides insight into organ development and adaptive evolution in orchids. *Ornamental Plant Research* 1:1−13

[28] Yang FX, Gao J, Wei YL, Ren R, Zhang GQ, et al. 2021. The genome of *Cymbidium sinense* revealed the evolution of orchid traits. *Plant Biotechnology Journal* 19:2501−2516

[29] Chen C, Wu Y, Li J, Wang X, Zeng Z, et al. 2023. TBtools-II: a "one for all, all for one" bioinformatics platform for biological big-data mining. *Molecular Plant* 16:1733−1742

[30] Wang Y, Zhao K, Chen Y, Wei Q, Chen X, et al. 2022. Species-specific gene expansion of the *Cellulose synthase* gene superfamily in the Orchidaceae family and functional divergence of mannan synthesis-related genes in *Dendrobium officinale*. *Frontiers in Plant Science* 13:777332

[31] Nguyen NH, Jeong CY, Kang GH, Yoo SD, Hong SW, et al. 2015. MYBD employed by HY5 increases anthocyanin accumulation via repression of *MYBL2* in *Arabidopsis*. *The Plant Journal* 84:1192−1205

[32] Jiang W, Yin Q, Liu J, Su X, Han X, et al. 2024. The APETALA2–MYBL2 module represses proanthocyanidin biosynthesis by affecting the formation of the MBW complex in seeds of *Arabidopsis thaliana*. *Plant Communications* 5:100777

[33] Li Z, Peng R, Tian Y, Han H, Xu J, et al. 2016. Genome-wide identification and analysis of the MYB transcription factor superfamily in *Solanum lycopersicum*. *Plant and Cell Physiology* 57:1657−1677

[34] Chen Y, Yang X, He K, Liu M, Li J, et al. 2006. The MYB transcription factor superfamily of *Arabidopsis*: expression analysis and phylogenetic comparison with the rice MYB family. *Plant Molecular Biology* 60:107−124

[35] Muthuramalingam P, Jeyasri R, Selvaraj A, Shin H, Chen JT, et al. 2022. Global integrated genomic and transcriptomic analyses of MYB transcription factor superfamily in C3 model plant *Oryza sativa* (L.) unravel potential candidates involved in abiotic stress signaling. *Frontiers in Genetics* 13:946834

[36] Li Y, Liang J, Zeng X, Guo H, Luo Y, et al. 2021. Genome-wide analysis of MYB gene family in potato provides insights into tissue-specific regulation of anthocyanin biosynthesis. *Horticultural Plant Journal* 7:129−141

[37] Jung C, Seo JS, Han SW, Koo YJ, Kim CH, et al. 2008. Overexpression of *AtMYB44* enhances stomatal closure to confer abiotic stress tolerance in transgenic *Arabidopsis*. *Plant Physiology* 146:623−635

[38] Gong Q, Li S, Zheng Y, Duan H, Xiao F, et al. 2020. SUMOylation of MYB30 enhances salt tolerance by elevating alternative respiration via transcriptionally upregulating AOX1a in *Arabidopsis*. *The Plant Journal* 102:1157−1171

[39] He J, Liu Y, Yuan D, Duan M, Liu Y, et al. 2020. An R2R3 MYB transcription factor confers brown planthopper resistance by regulating the phenylalanine ammonia-lyase pathway in rice. *Proceedings of the National Academy of Sciences of the United States of America* 117:271−277

[40] Mandaokar A, Browse J. 2009. MYB108 acts together with MYB24 to regulate jasmonate-mediated stamen maturation in *Arabidopsis*. *Plant Physiology* 149:851−862

[41] Yan Y, Li C, Dong X, Li H, Zhang D, et al. 2020. MYB30 is a key negative regulator of *Arabidopsis* photomorphogenic development that promotes PIF4 and PIF5 protein accumulation in the light. *The Plant Cell* 32:2196−2215

[42] Sun W, Gao Z, Wang J, Huang Y, Chen Y, et al. 2019. Cotton fiber elongation requires the transcription factor GhMYB212 to regulate sucrose transportation into expanding fibers. *New Phytologist* 222:864−881

[43] Chen J, Bi YY, Wang QQ, Liu DK, Zhang D, et al. 2022. Genome-wide identification and analysis of anthocyanin synthesis-related R2R3-MYB genes in *Cymbidium goeringii*. *Frontiers in Plant Science* 13:1002043

[44] Ke YJ, Zheng QD, Yao YH, Ou Y, Chen JY, et al. 2021. Genome-wide identification of the MYB gene family in *Cymbidiumensifolium* and its expression analysis in different flower colors. *International Journal of Molecular Sciences* 22:13245

[45] Fan H, Cui M, Li N, Li X, Liang Y, Liu L, et al. 2020. Genome-wide identification and expression analyses of *R2R3-MYB* transcription factor genes from two Orchid species. *PeerJ* 8:e9781

[46] Wang J, Liang X, Li B, Wang T. 2018. Genome-wide analysis of R2R3-MYB transcription factors in *Phalaenopsis equestris*. *Forest Research* 31:104−113

[47] Tanaka Y, Sasaki N, Ohmiya A. 2008. Biosynthesis of plant pigments: anthocyanins, betalains and carotenoids. *The Plant Journal* 54:733–749

[48] Tatsuzawa F, Ichihara K, Shinoda K, Miyoshi K. 2010. Flower colours and pigments in *Disa* hybrid (Orchidaceae). *South African Journal of Botany* 76:49–53

[49] Zhang Y, Zhou T, Dai Z, Dai X, Li W, et al. 2020. Comparative transcriptomics provides insight into floral color polymorphism in a *Pleione limprichtii* orchid population. *International Journal of Molecular Sciences* 21:247

[50] Liu XJ, Chuang YN, Chiou CY, Chin DC, Shen FQ, et al. 2012. Methylation effect on chalcone synthase gene expression determines anthocyanin pigmentation in floral tissues of two *Oncidium* orchid cultivars. *Planta* 236:401–409

[51] Wang L, Albert NW, Zhang H, Arathoon S, Boase MR, et al. 2014. Temporal and spatial regulation of anthocyanin biosynthesis provide diverse flower colour intensities and patterning in *Cymbidium* orchid. *Planta* 240:983–1002

[52] Nesi N, Jond C, Debeaujon I, Caboche M, Lepiniec L. 2001. The Arabidopsis *TT2* gene encodes an R2R3 MYB domain protein that acts as a key determinant for proanthocyanidin accumulation in developing seed. *The Plant Cell* 13:2099–2114

[53] Li J, An Y, Wang L. 2020. Transcriptomic analysis of *Ficus carica* peels with a focus on the key genes for anthocyanin biosynthesis. *International Journal of Molecular Sciences* 21:1245

[54] Uematsu C, Katayama H, Makino I, Inagaki A, Arakawa O, et al. 2014. Peace, a MYB-like transcription factor, regulates petal pigmentation in flowering peach 'Genpei' bearing variegated and fully pigmented flowers. *Journal of Experimental Botany* 65:1081–1094

[55] Jakoby MJ, Falkenhan D, Mader MT, Brininstool G, Wischnitzki E, et al. 2008. Transcriptional profiling of mature Arabidopsis trichomes reveals that *NOECK* encodes the MIXTA-like transcriptional regulator *MYB106*. *Plant Physiology* 148:1583–1602

[56] Wang X, Wang X, Hu Q, Dai X, Tian H, et al. 2015. Characterization of an activation-tagged mutant uncovers a role of *GLABRA2* in anthocyanin biosynthesis in *Arabidopsis*. *The Plant Journal* 83:300–311