**ORIGINAL ARTICLE**

# Integrated metabolome and transcriptome analysis provides insight into the leaf color change of *Cymbidium ensifolium*

Huan Mei[1] · Xiaoying Zhang[1] · Fukang Zhao[1] · Ruoxin Ruan[1] · Qiaojuan Fu[1]

## Abstract

*Cymbidium ensifolium* L. is an economically important terrestrial orchid. Leaf color is a key horticultural trait with aesthetic and ornamental significance. However, the molecular basis of leaf color change in *C. ensifolium* remains to be fully understood. Here, different color leaves of *C. ensifolium* were collected and classified into two groups, namely, leaf buds that appear red in the juvenile stage (RL) and those that transform green in the mature stage (GL). Significantly higher anthocyanin content was observed in RL than GL. RNA-seq data showed a total of 3955 differentially expressed unigenes (DEUs) were screened, and 381 differentially accumulated metabolites (DAMs) were detected by metabolome analysis in RL vs GL. Further analysis revealed that the anthocyanin biosynthesis pathway (ko00942) is essential for leaf color change. Cyanidin-3-O-glucoside, petunidin-3-O-(6″-O-p-coumaroyl) rutinoside, and peonidin-3-O-glucoside were identified as key anthocyanins affecting leaf color change. Meanwhile, we identified key structural unigenes associated with leaf color change, including *CHS*, *CHI*, *DFR*, *F3H*, *F3'H*, and *ANS*. Correlation analysis revealed 77 DEUs associated with transcription factors to exhibit strong correlations with both DEUs and DAMs involved in anthocyanin biosynthesis. Our study identifies the important metabolic and functional role of the anthocyanin biosynthesis pathway in *C. ensifolium* leaf color change.

**Keywords** *Cymbidium ensifolium* · Leaf color change · RNA-Seq · Metabolome · Anthocyanin biosynthesis

## Introduction

*Cymbidium ensifolium* L. is a valuable terrestrial orchid with fragrant flowers, graceful leaves, and a delicate fragrance. It is primarily found in southern China, southeastern Asia, northwest India, Korea, northern and eastern Australia, Borneo, New Guinea, and the Philippines (Wang et al. 2011). The horticultural traits of *C. ensifolium*, notably its floral organ, have been the focus of much research, particularly on floral morphology, floral fragrance, flowering time, and flower color (Ai et al. 2021; Yang et al. 2021; Zhang et al. 2021, 2022a; Sun et al. 2021). *C. ensifolium* can be classified into several varieties depending on the color and shape

Huan Mei and Xiaoying Zhang contributed equally to this work.

✉ Qiaojuan Fu
  fqjwxj@126.com

1   Institute of Horticulture, Hangzhou Academy of Agricultural Sciences, Hangzhou 310024, China

of the flower and leaf. 'Hong He' is the most common variety, combining safflower, and lotus petals. The leaves of this variety change color during its growth and development, with red leaves in the juvenile stage (RL) that transform into green leaves in the mature stage (GL). Prior to flowering, *C. ensifolium* generally experiences a 5- to 6-year juvenile phase, with a subsequent flowering time of approximately 1–2 months (Li et al. 2013; Zhang et al. 2022b). The variations in leaf color at the different developmental stages are key in determining the aesthetic and economic values of the orchids as the leaf viewing period is longer than that of flower viewing (Zhang et al. 2022b). However, despite its key significance, leaf color change in *C. ensifolium* is poorly understood. Compared with flower color, the influencing factors of leaf color are more complex (Gao et al. 2020). Thus, in order to overcome the gap in the current literature, it is crucial to investigate the molecular basis of leaf color change in *C. ensifolium*, which will consequently enhance the understanding of its ecological significance and increase its economic.

Variations in the levels of pigment families (e.g., chlorophylls and anthocyanins) are reported to be responsible

for plant pigmentation (Pahlavani et al. 2004; Fernández-García et al. 2012; Wang et al. 2020; Li et al. 2015). Chlorophyll provides green pigmentation and comprises chlorophyll a and chlorophyll b molecules (Wang et al. 2020). For plants, chlorophyll synthesis is robustly affected by the metabolic pathway's genes such as Glutamyl-tRNA reductase (*GluTR*), and Glutamyl-tRNA reductase (*HEMA*), etc. Anthocyanins are water-soluble, derived from phenylalanine, and engaged in the flavonoid biosynthetic pathway. They are synthesized in the cytosol and localized in the vacuoles (Tanaka et al. 2008; Zhao et al. 2022). Anthocyanin biosynthesis is primarily supported by structural genes, including *F3H*, *F3'H*, *F3'5'H*, *DFR*, *ANS*, *UFGT*, etc. (Ferrer et al. 2008; Holton and Cornish 1995; Sunil and Shetty 2022). In addition, genes related to anthocyanin biosynthesis are controlled by transcription factors (TFs) from the MYB, bHLH and WD40 families (Ahmad et al. 2022; Lloyd et al. 2017; Liu et al. 2021). In ornamental plants, bright colors (red, pink, purple, violet, and blue) coloration or fading results are generally attributed to the complex metabolic accumulation or loss in anthocyanin biosynthesis pathways (Gao et al. 2020; Lim et al. 2020; Song et al. 2021). Understanding the underlying mechanisms of color change can help with the effective selection of desirable leaf traits in breeding. However, research on the mechanism of leaf color change in *C. ensifolium* from red to green is limited.

Studies have demonstrated that combining metabolomics and transcriptomics provides an effective approach to identifying trait-related key genes associated with metabolites, and it currently plays a significant role in fundamental plant biology (Wei et al. 2016; Zhang et al. 2023, 2020; Li et al. 2020). In *Crataegus maximowiczii* C. K. Schneid, Zhang et al. integrated metabolome and transcriptome analyses to reveal cyanidin-based anthocyanins as the main pigments responsible for the black coloration of the fruit (Zhang et al. 2023). Illumina transcriptome analysis and widely targeted metabolomics (LC–ESI–MS/MS) were successfully used to identify the B-class *AP3-* and *AGL6*-like MADS-box genes participating in the flower color formation of the *Cattleya* hybrid 'KOVA' (Li et al. 2020). In the absence of a complete and high-quality genome sequence of *C. ensifolium*, the integrative analysis of metabolome and transcriptome provides a successful approach for gaining new insights into the molecular mechanisms of orchids.

In this study, analysis of the pigment contents from RL to GL of *C. ensifolium* 'Hong He' was performed. The molecular basis of leaf color change was explored using a widely targeted metabolomics approach and transcriptome analysis. Moreover, we focused on the differentially accumulated metabolites (DAMs), differentially expressed unigenes (DEUs) and TFs related to anthocyanin biosynthesis. Investigating the molecular mechanism of leaf color change in *C.*

*ensifolium* can determine new directions for future genetic modifications and improvements in orchids.

## Materials and methods

### Plant materials

*C. ensifolium* 'Hong He' was cultivated at a spacing of 30 cm in the multi-span greenhouse of the Hangzhou Academy of Agricultural Sciences (China, 30.16° N, 120.08° E). The substrates used by the plants are 5–10 mm bark and 3–6 mm perlite. The plants were grown at 28/20 °C (day/night), 80% relative humidity, and illumination levels range of 8000 Lux to 10,000 Lux. RL and GL from the 3-year-old plants were collected with three biological replicates, respectively. The color leaves were dropped to liquid nitrogen for rapid freezing after sampling and then stored at − 80 °C. The samples were used to measure the pigment contents (anthocyanin, total flavone and chlorophyll), transcriptome sequencing, and metabolic profiling.

### Measurement of pigment contents

For anthocyanin profiling (Meiling et al. 2018), fresh leaves (0.1 g) were mixed with 4 mL of a hydrochloric acid–ethanol solution (0.1 mol/L) and incubated in a 60 °C water bath for 6 h. The supernatant was measured at 530 nm, 620 nm, and 650 nm, respectively. Total anthocyanin concentration (μg/g) was determined as $[(A_{530} - A_{620}) - 0.1 \times (A_{650} - A_{620})] \div \varepsilon \times V \times N \div M \times 10^3 \times 287.24$ ($\varepsilon$ = molar extinction coefficient, $4.62 \times 10^4$; V = volume of the extract, mL; N = dilution multiple; M = sample fresh quantity, g).

Extraction of flavone was conducted according to SN/T 4592-2016 (Industry Standard for Entry Exit Inspection and Quarantine of the People's Republic of China). 0.1 g of the sample and 3 mL of absolute ethanol were added into a 10 mL centrifuge tube. The solution was shaken well and sonicated at 60 °C for 1 h. Then the mixture was generated by the mixing proportion [sample liquid; aluminum nitrate solution (100 g/L); potassium acetate solution (98 g/L); distilled water = 30:2:2:66]. The mixture was rested at room temperature for 1 h for color reaction. The absorbance value was determined at 420 nm and the standard curve was determined. The total flavone content (μg/g) was estimated as $C \times V \div V_{sample} \times N \div M$ (C = flavone content obtained from the standard curve, μg; V = volume of the total extract, mL; $V_{sample}$ = volume of the sample extract for the standard curve, mL; N = dilution multiple; M = sample quantity, g).

The total chlorophyll content was determined by spectrophotometric analyses (Xuekui 2006). Fresh leaves (0.1 g) from RL and GL were homogenized in a glass tube with 5 mL of extracting solution (95% ethanol) and 50 mg of

calcium carbonate powder. The glass tube was maintained in the dark until the tissue was completely whitened. The absorbance was measured at 665 nm and 649 nm. The chlorophyll was measured in three replicates. The total chlorophyll content (μg/g) was calculated as $(18.08 \times A_{649} + 6.63 \times A_{665}) \times V \times N \div M$ (V = volume of the extract, mL; N = dilution multiple; M = sample fresh quantity, g).

## RNA extraction, library construction, and sequencing

According to the TRIzol® reagent (Invitrogen, USA), RNA was extracted from six samples (two tissues, three biological replicates) of RL and GL. Sequencing libraries were constructed using the NEBNext® UltraTM RNA Library Prep Kit for Illumina® (NEB, USA). Transcriptome sequencing was performed on an Illumina Novaseq 6000 platform with a 150-bp paired-end strategy. Raw data in fastq format was further processed using the BMK online platform (http://www.biocloud.net) for quality control and to output the clean reads. Transcriptome assembly was performed using Trinity (version 2.14.0) with non-default parameters (min_contig_length 200, group_pairs_distance 500) (Grabherr MG 2011). Clean reads were aligned to the unigenes library using Bowtie (Langmead et al. 2009), and expression level estimation was performed by RSEM (version 1.2.19) (Li and Dewey 2011). The fragments per kilobase of exon per million fragments mapped (FPKM) value for each unigene was subsequently obtained (Trapnell et al. 2010). Using BLAST alignment (cutoff E-value $\leq 10^{-5}$) and Hmmer alignment (version 3.1b2) (cutoff E-value $\leq 10^{-10}$), the unigenes were annotated based on eight functional annotation databases, including Clusters of Orthologous Groups (COG), euKaryotic Orthologous Groups (KOG), Swiss-Prot, Protein Family (Pfam), Kyoto Encyclopedia of Genes and Genomes (KEGG), Gene Ontology (GO), eggNOG, and NCBI nonredundant protein sequences (NR). The TFs were identified and classified by iTAK and PlantTFDB (http://planttfdb.gao-lab.org/) with E-value $\leq 10^{-5}$.

## Extraction, data preprocessing and annotation of metabolome

The extraction procedure for metabolome analysis (Dunn et al. 2011; Want et al. 2010) was as follows: leaves from *C. ensifolium* were freeze-dried by a vacuum freeze-dryer (Scientz-100F) and crushed using a mixer mill with a zirconia bead for 1.5 min at 30 Hz. A total of 100 mg of lyophilized powder was dissolved in 1.2 mL of a 70% methanol solution. The samples were placed in a refrigerator overnight at 4 °C after being vortexed for 30 s every 30 min. This step was repeated six times. The samples were centrifuged (12,000 rpm, 10 min) and the extracts were filtrated for

subsequent analysis. Widely targeted metabolite analysis was conducted using the UPLC-ESI–MS/MS system, which consisted of the ultra-high-performance liquid chromatograph (UPLC, SHIMADZU Nexera X2, Japan) and a tandem mass spectrometer (MS/MS 4500 QTRAP, Applied Biosystems, USA). The raw data was collected using MassLynx (version 4.2, Waters Corpration) and processed for the identification and quantification of metabolites based on Progenesis QI with the online METLIN database and Biomark's self-built library. The results of the OPLS-DA were plotted with ropls package (version 1.18.8).

## Identification of differentially expressed unigenes

Differential expression analysis was performed using DESeq2 (version 1.30.1) (Love et al. 2014). The resulting *P* values were adjusted using Benjamini and Hochberg's approach for controlling the false discovery rate (FDR). Genes with the threshold of twofold change (FC) in expression [|log$_2$ (FC) | ≥ 1] and an adjusted $P < 0.01$ were assigned as DEUs.

## Identification of differentially accumulated metabolites

The constraints variable importance in projection (VIP) ≥ 1, [log$_2$ (FC) | ≥ 1], and $P < 0.05$ were used to identify metabolites with significant differences in content as DAMs.

## Statistical analysis

PCA and PCC analyses were performed using the Prcomp and Cor functions in R (version 3.6.3), and the heat map was generated using pheatmap (version 1.0.12). GO and KEGG enrichment analyses were conducted using the R package ClusterProfiler (version 3.14.3) (Yu et al. 2012).

## qRT-PCR verification

qRT-PCR was conducted using SYBR® Premix Ex TaqTMII (Takara, Japan), with three replicates for each sample. We employed *EF1a* (GenBank: KC794500.1) as the reference gene. The amplification procedure adopted in the AnalytikjenaqTOWER2.2 was as follows: (i) pre-denaturation at 95 °C for 3 min; (ii) denaturation at 95 °C for 10 s; and (iii) annealing at 60 °C for 30 s. Steps (ii) and (iii) were repeated for 40 cycles. We calculated the relative abundance of transcripts using the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen 2001).

## Results

### Phenotypical and physiological differences between RL and GL

The leaves of *C. ensifolium* were red during the leaf bud stage and finally matured into green (Fig. 1A). In order to determine the pigment components involved in the leaf color change, we measured the contents of anthocyanin, total flavone, and chlorophyll. Anthocyanin content declined sharply ($P = 6.09 \times 10^{-5}$, Student's *t*-test), changing from 432.93 μg/g in RL to 78.94 μg/g in GL (Fig. 1B). A slightly lower total flavone content was observed in RL compared to GL (Fig. 1C). However, no significant differences were observed in the chlorophyll content between RL and GL (Fig. 1D). To further explore the mechanism of leaf color change in *C. ensifolium*, a widely targeted metabolomics approach and transcriptome analysis were conducted.

### De novo transcriptome sequencing of RL and GL

A total of 47.58 Gb of clean data was obtained by de novo transcriptome sequencing on RL and GL in *C. ensifolium*, with 7.54 Gb of clean data per sample and a mean GC content of 46.56%. The results determined that Q20 and Q30 were greater than or equal to 97.99% and 89.20%, respectively, for each sample (Supplementary Table 1). An impressive range of 81.91% to 88.23% was observed for the mapped ratio, highlighting the reliable quality of the RNA-seq data (Supplementary Table 1). Principal component analysis (PCA) revealed that the two principal components explained 59.3% of the total variance (39.1% by PC1, 20.2% by PC2). Each sample tissue (three biological replicates) formed a cluster, indicating a significant difference in gene expression between RL and GL (Fig. 2A). The correlation coefficient among the biological replicates for each sample was high (Pearson's $r > 0.982$ in RL, Pearson's $r > 0.994$ in GL), while the correlation coefficient between RL and GL was low (Pearson's $r < 0.189$) (Fig. 2B). Following the
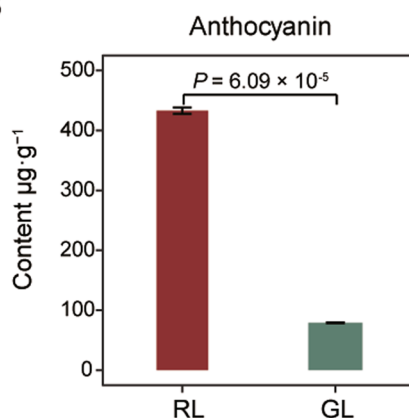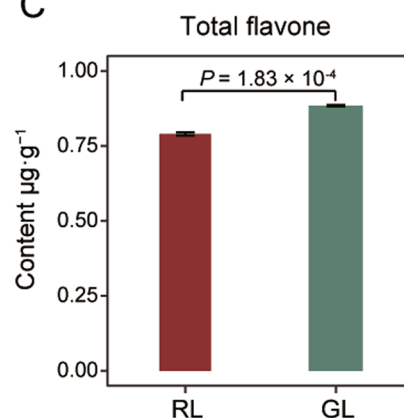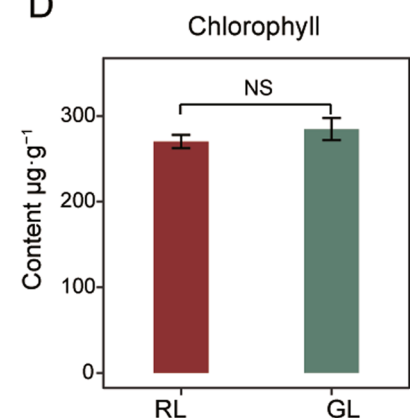


**Fig. 1** Characteristics of phenotype and pigment content in red leaves (RL) and green leaves (GL) of *C. ensifolium*. **A** Phenotypes of RL and GL. **B** Significant difference in anthocyanin content of RL and GL. **C** Significantly higher total flavone content in GL than RL. **D** No difference in chlorophyll content between RL and GL. Bars represent the mean ± SD ($n = 3$). *NS* no significant difference
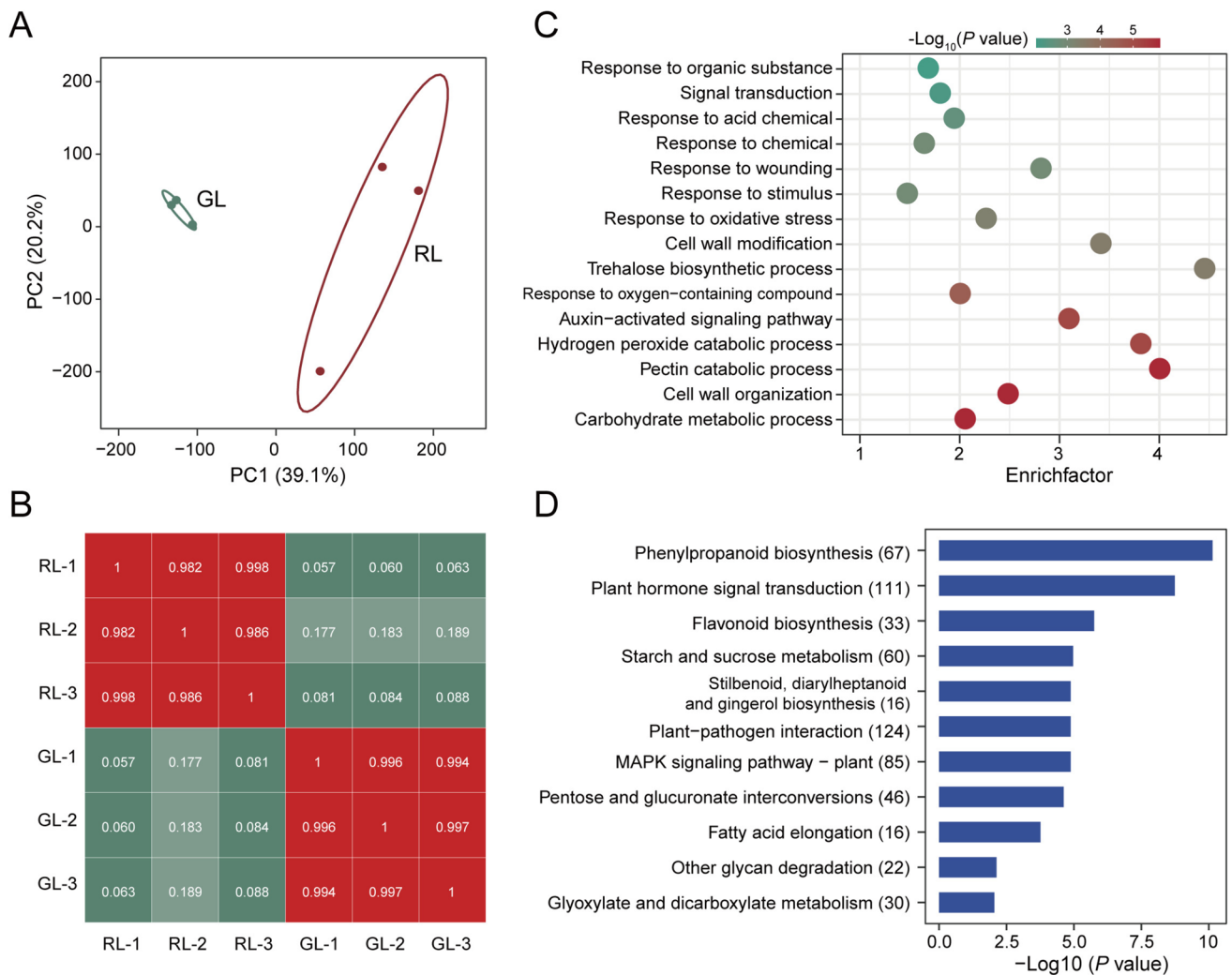
**Fig. 2** Information of de novo assembled unigenes in *C. ensifolium* determined from transcriptomic analysis. **A** Principal Component Analysis of unigene expression derived from the transcriptome. **B** Pearson's correlation coefficient of unigene expression. Numbers 1, 2, and 3 correspond to each replication. **C** Bubble chart of significant GO biological process terms (FDR < 0.01) among the DEUs. The redder the bubble color, the higher the enrichment degree. **D** Bar diagram showing the KEGG analysis of the DEUs. *X*-axis represents − log10 (*P* value), with greater values indicating a higher enrichment degree. *Y*-axis plots KEGG terms; numbers represent the DEUs included in each term. FDR < 0.01 was set as the cutoff

assembly of the sequences, 59,445 unigenes were obtained with average lengths of 1013 bp and N50 lengths of 1,890 bp (Supplementary Table 2). A total of 25,856 unigenes were functionally annotated in the eight public databases (COG, KOG, SwissProt, Pfam, KEGG, GO, eggNOG, and NR) (Figure S1), with 3274 annotated in all databases and accounting for 5.51% of the total unigenes. In addition, 3,955 DEUs were identified, 1,338 of which were significantly upregulated and 2617 were downregulated in GL compared to RL (Figure S2, Supplementary Table 3).

To get a deeper understanding of major DEG functions throughout changing leaf color, enrichment analysis was performed. The biological process revealed that the DEUs were significantly enriched for 15 GO terms, containing the carbohydrate metabolic process (GO:0005975), cell wall organization (GO:0071555), pectin catabolic process (GO:0045490), and auxin-activated signaling pathway (GO:0009734) (Fig. 2C). In the molecular function and cellular component categories, the terms cell wall (GO:0005618), extracellular region (GO:0005576), heme binding (GO:0020037), and protein kinase activity (GO:0020037) were most enriched (Figure S3). Furthermore, the phenylpropanoid biosynthesis (ko00940), starch and sucrose metabolism (ko00500), and flavonoid biosynthesis (ko00941) pathways exhibited the better significance and richer factors based on KEGG analysis (Fig. 2D).

## Metabolite identification and quantification

In order to explore the metabolic networks occurring during leaf color change in *C. ensifolium*, widely targeted metabolite analysis was used to detect the metabolites based on the UPLC-ESI–MS/MS system. A total of 897 metabolites were identified (Supplementary Table 4). The PCA analysis revealed significant differences between RL and GL, explaining 89.0% of the total variance (64.1% by PC1, 24.9% by PC2) (Fig. 3A). The Pearson's correlation coefficient (PCC) was high among biological replicates (Pearson's $r > 0.816$ in RL, Pearson's $r > 0.993$ in GL) (Fig. 3B). Moreover, the orthogonal projections to latent structures-discriminant

analysis (OPLS-DA) was stable and reliable ($R^2X = 0.936$, $R^2Y = 1$, $Q^2Y = 0.996$), and can be used to identify DAMs according to the VIP value (Fig. 3C). We performed 200 permutation experiments for the alignment verification of OPLS-DA (Figure S4). Compared with RL, 381 DAMs were identified, 309 of which were upregulated and 72 were downregulated in GL (Fig. 3D, Supplementary Table 4).

To further understand the category of metabolites, we annotated the DAMs based on the HMDB and KEGG databases. A total of 60 DAMs were annotated into 19 classification entries, such as carboxylic acids and derivatives, organooxygen compounds, and flavonoids (Fig. 3E). Based on the rich factor ≥ 1.5, 40 DAMs were assigned to 22 KEGG
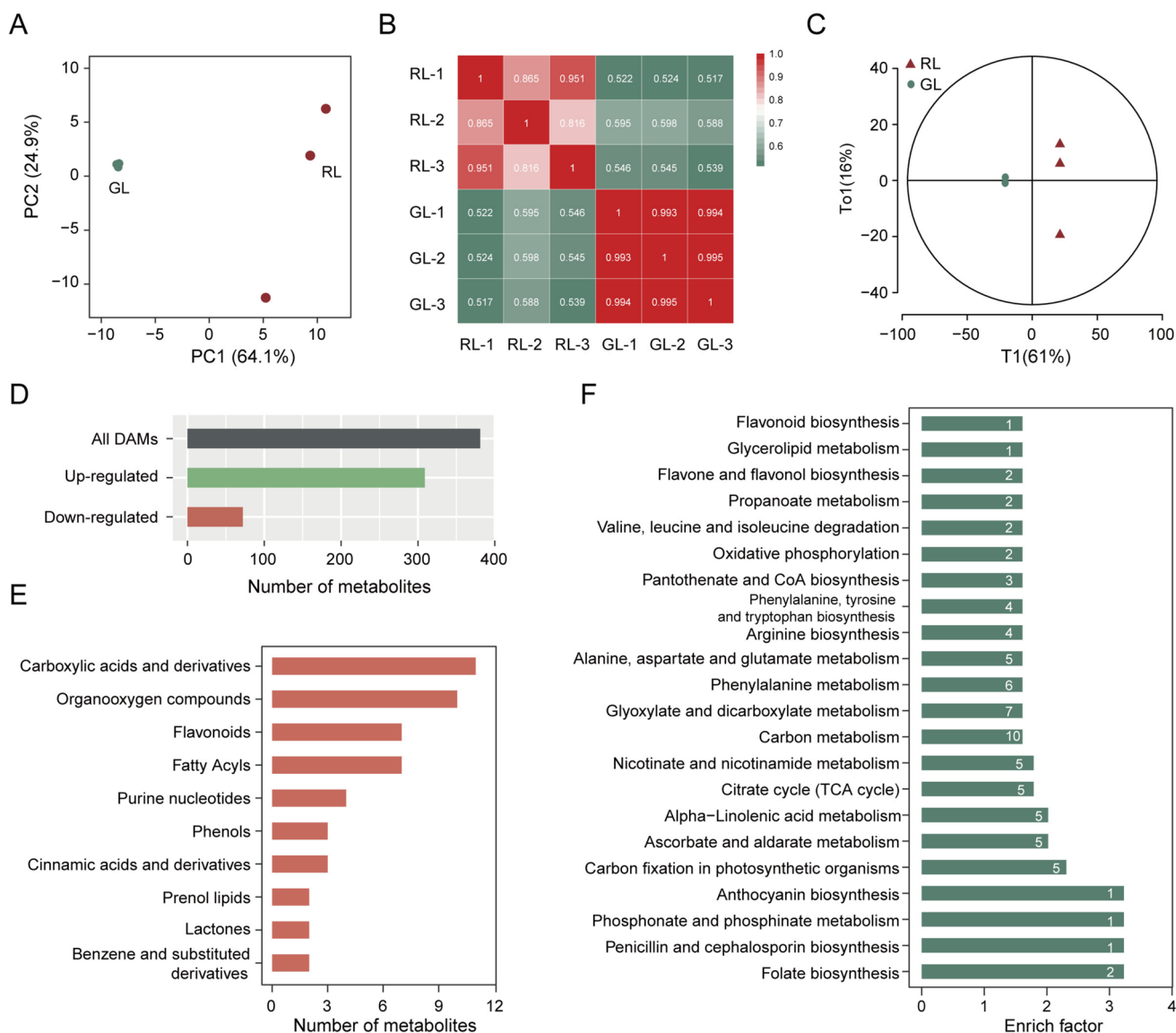


**Fig. 3** Metabolomic UPLC-ESI–MS/MS-based analyses of the *C. ensifolium* leaves. **A** Principal Component Analysis of metabolites. **B** Pearson's correlation coefficient of metabolites. **C** OPLS-DA model of metabolites. **D** Bar chart depicting the DAMs between RL and GL. **E** Annotation of DAMs against HMDB databases. **F** Bar chart depicting the KEGG analysis of the DAMs

pathways. A total of 7 DAMs were mapped to glyoxylate and dicarboxylate metabolism (ko00630), 1 DAM was mapped to anthocyanin biosynthesis (ko00942), and 2 DAMs were mapped to flavone and flavonol biosynthesis (ko00944) (Fig. 3F).

## Comprehensive analysis of the metabolome and transcriptome of *C. ensifolium* leaf

We detected many DEUs and DAMs were simultaneously annotated in 57 different KEGG pathways, such as flavone and flavonol biosynthesis associated with 2 DAMs and 6 DEUs, the citrate cycle (TCA cycle) associated with 5 DAMs and 4 DEUs, and the phenylalanine metabolism associated with 6 DAMs and 9 DEUs (Supplementary Table 5). In order to investigate the deeper association between DEUs and DAMs, a comprehensive analysis was performed using PCC. Supplementary Table 6 reported the correlations based on $|r| > 0.90$ and $P < 1.0 \times 10^{-2}$, revealing strong correlations between the 3,723 DEUs and 381 DAMs. *TRINITY_DN703_c0_g1* (encoding kinesin-like protein KIN-5A) exhibited the strongest negative correlation ($r = -0.999$, $P = 3.68 \times 10^{-9}$) with icariside B2, and *TRINITY_DN6894_c0_g1* (encoding MYB) showed the strongest positive correlation quercetagetin-7-O-glucoside ($r = 0.999$, $P = 5.81 \times 10^{-9}$) (Supplementary Table 6). The highly correlated DEUs and DAMs, which were involved in multiple metabolic pathways, may play important roles in the leaf color change of *C. ensifolium*.

## Regulatory network of anthocyanin biosynthesis

The anthocyanin content dropped significantly from RL to GL (Fig. 1B). The transcriptome results demonstrated the enrichment of DEUs in flavonoid biosynthesis, and the investigation of metabolomes revealed a superior rich factor in anthocyanins biosynthesis (ko00942, an important branch of flavonoid synthesis) during the leaf color change of *C. ensifolium*. Thus, we focused on anthocyanin biosynthesis and analyzed the key DEUs and DAMs implicated in this pathway. A total of 10 anthocyanin synthesis-related DEUs were identified, containing *CHS* (*TRINITY_DN316_c0_g1*, *TRINITY_DN1786_c0_g1*), *CHI* (*TRINITY_DN10644_c0_g1*, *TRINITY_DN143_c0_g1*), *DFR* (*TRINITY_DN2247_c1_g1*), *F3H* (*TRINITY_DN76_c0_g3*, *TRINITY_DN3316_c1_g1*), *F3'H* (*TRINITY_DN3907_c0_g1*, *TRINITY_DN9664_c0_g1*), and *ANS* (*TRINITY_DN852_c0_g1*). Interestingly, all 10 major DEUs in the anthocyanin biosynthesis pathway were identified to be downregulated in RL vs GL, with FC varying from 3.49 to 79.17 (Supplementary Table 7). Metabolome analysis revealed that 3 downregulated and 4 upregulated DAMs were involved in the anthocyanin biosynthesis pathway. Cyanidin-3-O-glucoside (down by 6.47

FC), petunidin-3-O-(6″-O-p-coumaroyl) rutinoside (down by 7.26 FC), and peonidin-3-O-glucoside (down by 15.70 FC) depicted a down accumulation pattern in GL vs RL. Although petunidin-3-O-glucoside (up by 17.78 FC), delphinidin-3-O-rutinoside-7-O-glucoside (up by 18.54 FC), malvidin-3-O-(6″-O-caffeoyl) glucoside (up by 21.20 FC) and pelargonidin-3,5-O-diglucoside (up by 20.71 FC) were identified with up accumulation in GL vs RL. Figure 4A presents the anthocyanin synthesis pathway determined by combining transcriptomic and metabolomic data. The results signify the important role of cyanidin-3-O-glucoside, petunidin-3-O-(6″-O-p-coumaroyl) rutinoside, and peonidin-3-O-glucoside in coloration in *C. ensifolium*. Furthermore, *CHI*, *DFR*, *F3H*, and *ANS* were positively correlated with cyanidin-3-O-glucoside ($r > 0.932$, $P < 6.83 \times 10^{-3}$), while *CHI* exhibited a negative correlation with peonidin-3-O-glucoside ($r = 0.927$, $P = 7.83 \times 10^{-3}$) (Fig. 4B).
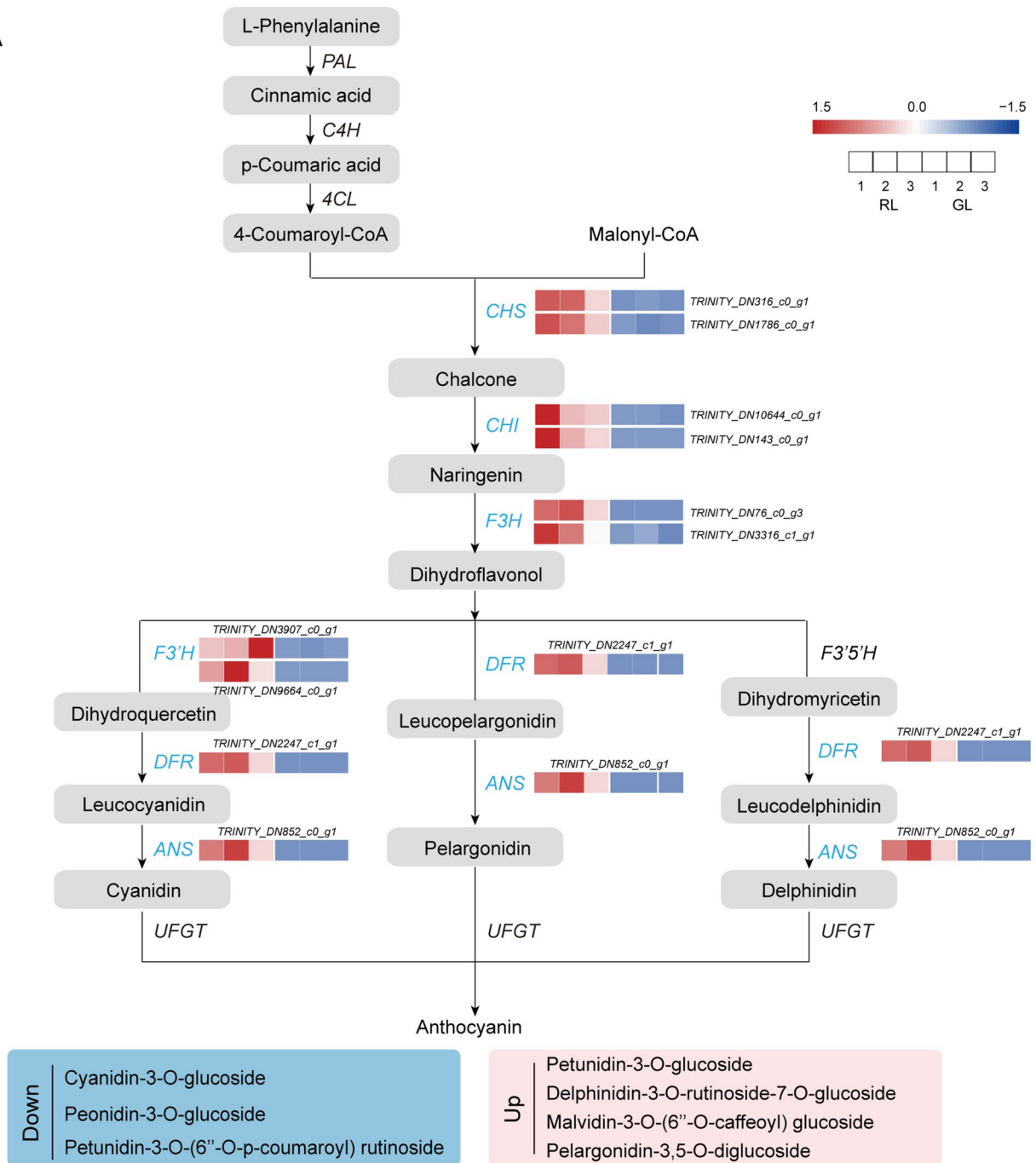
## Identification of transcription factors

To promote functional gene research in *C. ensifolium*, the identification of TFs was performed. A total of 242 DEUs were identified to be associated with TFs (Figure S5, Supplementary Table 8). In particular, the bHLHs and MYBs TF families, which are considered to regulate the biosynthesis of anthocyanins (Ahmad et al. 2022; Lloyd et al. 2017), accounted for the largest proportion, with 23 (9.50%) and 18 (7.44%) members, respectively. The correlation coefficients between 242 DEUs associated with TFs and key anthocyanin synthesis-related DEUs was presented in Supplementary Table S9. In addition, we identified 93 DEUs associated with TFs to be significantly correlated with key anthocyanin synthesis-related DAMs (Supplementary Table S10). Further investigation determined a set of 77 DEUs associated with TFs to be strongly correlated with both the DEUs and DAMs implicated in anthocyanin biosynthesis (Figure S6). The 77 DEUs were assigned to 31 different TF families, with the top 5 TF families (in terms of quantity) determined as bHLH ($N = 13$), NAC ($N = 9$), AP2/ERF-ERF ($N = 6$), MYB ($N = 5$), and MYB-related ($N = 4$). Interestingly, a general downregulation tendency was observed in the MYBs and bHLHs families from RL to GL among these 77 TFs. For example, *TRINITY_DN1376_c0_g1* (R2R3-MYB) exhibited a positive correlation with 1 key anthocyanin synthesis-related DAMs and 5 anthocyanin synthesis-related DEUs from RL to GL. This implies the role of these TF families in regulating the biosynthesis and transport of anthocyanins during leaf color change, especially MYBs and bHLHs.

## Real-time polymerase chain reaction (PCR) validation

To assess the accuracy and reproducibility of the RNA-Seq data, the relative expression levels of key structural

## A



L-Phenylalanine

↓ *PAL*

Cinnamic acid

↓ *C4H*

p-Coumaric acid

↓ *4CL*

4-Coumaroyl-CoA          Malonyl-CoA

*CHS*   TRINITY_DN316_c0_g1
        TRINITY_DN1786_c0_g1

Chalcone

*CHI*   TRINITY_DN10644_c0_g1
        TRINITY_DN143_c0_g1

Naringenin

*F3H*   TRINITY_DN76_c0_g3
        TRINITY_DN3316_c1_g1

Dihydroflavonol

*F3'H*   TRINITY_DN3907_c0_g1
         TRINITY_DN9664_c0_g1

*DFR*   TRINITY_DN2247_c1_g1

*F3'5'H*

Dihydroquercetin          Leucopelargonidin          Dihydromyricetin

*DFR*   TRINITY_DN2247_c1_g1         *DFR*   TRINITY_DN2247_c1_g1

Leucocyanidin          *ANS*   TRINITY_DN852_c0_g1          Leucodelphinidin

*ANS*   TRINITY_DN852_c0_g1          Pelargonidin          *ANS*   TRINITY_DN852_c0_g1

Cyanidin          Delphinidin

*UFGT*          *UFGT*          *UFGT*

Anthocyanin

**Down**
Cyanidin-3-O-glucoside
Peonidin-3-O-glucoside
Petunidin-3-O-(6''-O-p-coumaroyl) rutinoside

**Up**
Petunidin-3-O-glucoside
Delphinidin-3-O-rutinoside-7-O-glucoside
Malvidin-3-O-(6''-O-caffeoyl) glucoside
Pelargonidin-3,5-O-diglucoside

## B



CHI
*TRINITY_DN10644_c0_g1*

DFR
*TRINITY_DN2247_c1_g1*

Peonidin-3-O-glucoside

$r = 0.932$   $P = 6.83 \times 10^{-3}$

$r = 0.989$   $P = 1.78 \times 10^{-4}$

Cyanidin-3-O-glucoside

$r = 0.986$   $P = 2.77 \times 10^{-4}$

$r = 0.978$   $P = 7.37 \times 10^{-4}$

F3H
*TRINITY_DN76_c0_g3*

ANS
*TRINITY_DN852_c0_g1*

$r = 0.927$   $P = 7.83 \times 10^{-3}$
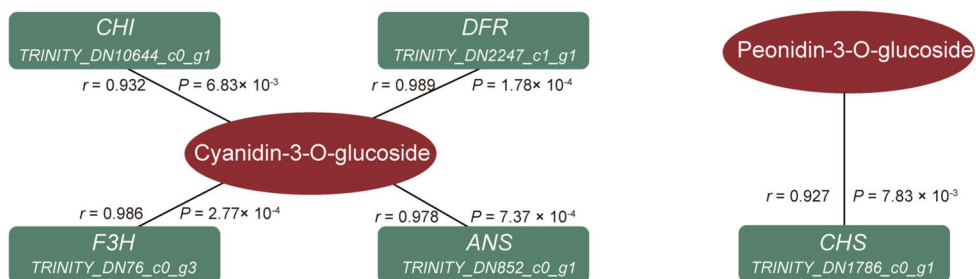
CHS
*TRINITY_DN1786_c0_g1*

◂**Fig. 4** DEUs and DAMs in the synthesis pathway of anthocyanins. **A** Regulatory network of anthocyanin biosynthesis. Red indicates higher levels of unigene expression in RL relative to GL and blue indicates lower levels of unigene expression in RL relative to GL. Blue rectangular boxes represent the DAMs that are significantly lower in GL than RL. *PAL* phenylalanine lyase, *C4H* cinnamate 4-hydroxylase, *4CL* 4-coumadin CoA ligase, *CHI* chalcone isomerase, *CHS* chalcone synthase, *F3H* flavanone 3-hydroxylase, *F3'H* flavonoid 3'-hydroxylase, *F3'5'H* flavonoid 3',5'-hydroxylase, *DFR* dihydroflavonol 4-reductase, *ANS* anthocyanidin synthase, and *UFGT* flavonoid 3-O-glucosyltransferase. **B** Correlations between the anthocyanin synthesis-related DAMs and DEUs

unigenes were selected and examined using quantitative real-time PCR (qRT-PCR). Supplementary Table 11 reports the enumeration of the specific primers. The qRT-PCR and RNA-seq correlation results are represented by $R^2$. The $R^2 > = 0.767$ indicates that the expression patterns of the DEUs obtained from the RNA-seq and qRT-PCR are consistent, confirming that our RNA-seq data are authentic and reliable (Fig. 5). Thus, the acquired data verified the importance of anthocyanin biosynthesis-related DEUs in the change of leaf color in *C. ensifolium*.

## Discussion

The color of leaves frequently varies as a result of a plant's growth and development. Examples include *Alternanthera bettzickiana* (Li et al. 2022), *Acer pictum* subsp. *mono* (Ge et al. 2019), *Padus virginiana* (Li et al. 2021). To understand the mechanism of leaf color change, earlier studies concentrated on physiology, cytology, and molecular biology techniques (Qi et al. 2022; Wang et al. 2021). However, the analysis of the regulatory mechanisms and abundant metabolites using multi-omics methods to investigate leaf color change is lacking. The integrated analysis of the transcriptome and metabolome can be employed to identify secondary metabolites or related genes that affect leaf color (Ge et al. 2019; Feng et al. 2021; Qi et al. 2022; Qiu et al. 2020). In our study, the combined analysis of metabolome and transcriptome data was performed to elucidate mechanism-based information for the genetic improvement of *C. ensifolium*.

### Pigment accumulation between RL and GL

The aesthetic and ornamental worth of *C. ensifolium* is greatly influenced by the leaf color, a key horticultural characteristic. Studies have shown that the formation of leaf color in colored-leaf plants is complex and based on the biosynthesis, distribution, type, and content of pigments (Tang et al. 2020). This study revealed that the anthocyanin content in GL is 5.48 times lower than that in RL, with a greater degree of change amplitude than the total flavone and chlorophyll content (Fig. 1B–D). We hypothesized that

anthocyanins might be a key cause of color change in *C. ensifolium* leaves. Anthocyanidins are important secondary metabolites and have the ability to safeguard photosynthesis, resist oxidation, moderate stress, and postpone senescence. More significantly, numerous studies have demonstrated that anthocyanins are crucial for the development of pigments in key plant organs including leaves, flowers, fruits, and seeds (Gao et al. 2020; Song et al. 2021; Lim et al. 2020; Zhang et al. 2023). Cyanidin controls the change from the pink to purple-red color spectrum (Kawase et al. 1970; Veberic et al. 2015; Yi et al. 2021; Han et al. 2021; Bureau et al. 2009; Robinson and Robinson 1932) and is the dominant anthocyanin in the scarlet skin of *Crataegus pinnatifida* Bunge and *Crataegus laevigata* (Poir.) DC. fruits (Veberic et al. 2015). Longan's pericarp has a bright red color owing to the accumulation of cyanidin derivatives (cyanidin 3-O-glucoside, cyanidin 3-O-6″-malonyl-glucoside, and cyanidin O-syringic acid) (Yi et al. 2021). The mature fruit pericarp of sweet osmanthus is purple-black due to elevated levels of cyanidin-3-O-rutinoside and peonidin-3-O-rutinoside (Han et al. 2021). In this work, cyanidin-3-O-glucoside, petunidin-3-O-(6″-O-p-coumaroyl) rutinoside, and peonidin-3-O-glucoside all exhibited a significant down accumulation pattern in GL. The reduction of the three key anthocyanins was thought to be one of the primary causes of the *C. ensifolium* leaf color shift from red to green.

### Anthocyanin synthesis-related DEUs of *C. ensifolium* leaves

The anthocyanin biosynthetic route in plants has been extensively researched. There are numerous enzymes that can catalyze anthocyanin synthesis, such as *CHS*, *CHI*, *DFR*, *ANS*, and *F3'H* (Ferrer et al. 2008; Holton and Cornish 1995; Sunil and Shetty 2022). *CHS* catalyzes the key branch point from monolignol biosynthesis to the anthocyanins, while *F3'H* determines whether the cyanidin-based anthocyanins will be synthesized (Jaakola 2013). Genes related to anthocyanin synthesis are vital for the leaf color. In *Paeonia qiui* found that *F3H*, *F3'H*, *DFR* and *ANS* were upregulated during spring leaf red color fading (Luo et al. 2017). An increased expression of the *UGFT* and *BZ1* genes contributes to the purple-red color change of *Padus virginiana* (Li et al. 2021). From our results, unigenes annotated as *CHS*, *CHI*, *DFR*, *F3H*, *F3'H*, and *ANS* exhibited a consistent downregulation from RL to GL in *C. ensifolium* leaves (Fig. 5, Supplementary Table 7). We inferred that the 10 anthocyanin-related DEUs were identified as key genes affecting leaf color change. Further functional verification of identified DEUs in the near future can shed light on the regulatory processes underlying anthocyanin in *C. ensifolium*. Furthermore, *C. ensifolium* leaves can be utilized for natural pigmentation.
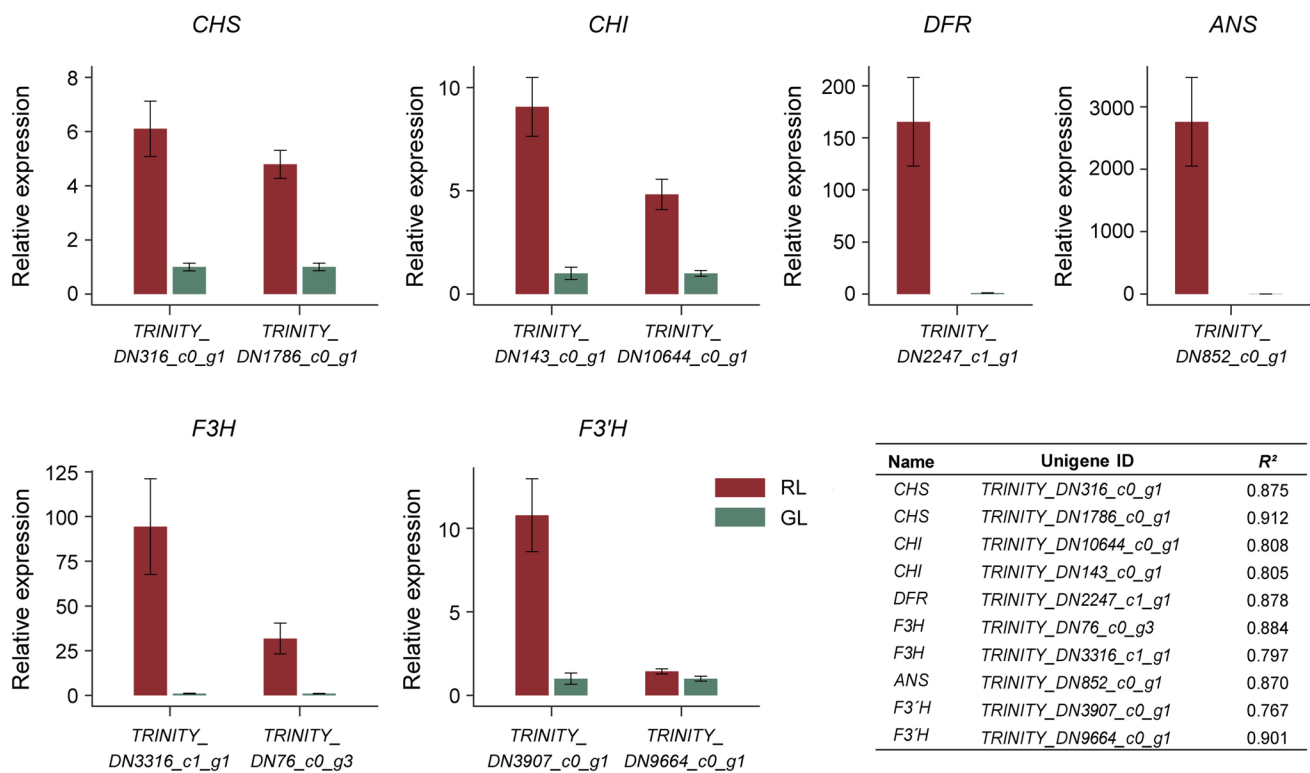
**Fig. 5** Validation of qRT-PCR. Bars represent the mean $\pm$ SD ($n=3$). $R^2$ represents the correlation analysis of the qRT-PCR and RNA-seq results

## Transcription factors related to anthocyanin biosynthesis

The biosynthesis of anthocyanins is strictly regulated by TFs, particularly MYBs and bHLHs (Ahmad et al. 2022; Lloyd et al. 2017; Liu et al. 2021). Anthocyanin synthesis activation can be restored by simultaneously introducing MYB and bHLH anthocyanin regulators in *Cymbidium* orchids (Albert et al. 2010). In numerous studies on model plants and horticultural crops, the regulation of the anthocyanin-promoting R2R3-MYB TF can control anthocyanin accumulation (Yang et al. 2022). Among the 77 DEUs associated with TFs, we found that a R2R3 MYB (*TRINITY_DN1376_c0_g1*) might potentially have a significant impact on the color change of leaves, exhibiting positive correlations with *CHS*, *CHI*, *F3H* and Petunidin-3-O-(6″-O-p-coumaroyl) rutinoside. Its expression decreased in GL compared to RL by 10.26 FC. Furthermore, *TRINITY_DN6149_c2_g1* (encoding bHLH25) demonstrated a strong positive association with 2 key anthocyanin synthesis-related DAMs and 8 anthocyanin synthesis-related DEUs. Previous research attributed the coordinated regulation of the MYBs and bHLHs to the purple-red leaf stage of *Padus virginiana* (Li et al. 2021). According to our research, MYBs and bHLHs may positively or cooperatively regulate the expression of several structural genes involved in the production of

anthocyanins, which is what causes the color shift in *C. ensifolium* leaves. Several additional types of TFs with differential expressions, such as bZIPs, MADs, WRKYs, NACs, and AP2/ERF-ERFs, were also discovered to be connected to the expression of anthocyanin synthesis-related DEUs and DAMs. *TRINITY_DN1116_c0_g1* (which encodes for the ethylene-responsive TF ERF039) was reduced from RL to GL by 12.89 FC and was positively correlated with *CHI* and peonidin-3-O-glucoside. *TRINITY_DN1013_c0_g2* (WRKY44) increased in expression by 3.65 FC from RL to GL and was negatively correlated with peonidin-3-O-glucoside, *CHI*, and *CHS*. TF such as ERF, WRKY, bZIPs, MADs may synergistically regulate the expression of unigenes in anthocyanin biosynthesis. Similar results were found in *Artemisia annua* L (Hassani et al. 2020).

## Conclusion

In this study, the molecular regulatory mechanisms of leaf color change in *C. ensifolium* were characterized by combining metabolome and transcriptome data. We found anthocyanin, whose biosynthesis pathway is essential for leaf color change, was markedly lower in GL than RL. The key anthocyanin-related DEUs (*CHS*, *CHI*, *DFR*, *F3H*, *F3'H*, and *ANS*) were identified with down accumulation pattern

in GL vs RL, implying a significant role in coloration. The reduction of cyanidin-3-O-glucoside, petunidin-3-O-(6″-O-p-coumaroyl) rutinoside, and peonidin-3-O-glucoside was considered as one of the main reasons for leaf color change from red to green in *C. ensifolium*. 77 DEUs associated with TFs were detected to exhibit strong correlations with both DEUs and DAMs involved in anthocyanin biosynthesis, such as *TRINITY_DN1376_c0_g1* (R2R3 MYB) and *TRINITY_DN6149_c2_g1* (encoding bHLH25). In the future, the application of biotechnology is expected to improve the expression of identified DEUs, achieving the goal of prolonging the red color of the leaves in *C. ensifolium*. Taken together, the comprehensive analyses explained the mechanism of color change in *C. ensifolium* leaves, providing an important reference for colored-leaf plants.

**Author contributions** QF conceptualized the project. HM and XZ designed and performed the experiments. HM and XZ performed the bioinformatics analysis. RR and FZ contributed reagents/materials/analysis. HM, XZ, and QF prepared the manuscript. All authors read and approved the final manuscript.

**Data availability** All the raw data are available in the China National Genebank (CNGB, https://db.cngb.org/cnsa/) with the accession number CNP0004200.

## Declarations

**Conflict of interest** The authors declare no competing financial interests.

## References

Ahmad S, Chen JL, Chen GZ, Huang J, Zhou YC, Zhao K, Lan SR, Liu ZJ, Peng DH (2022) Why black flowers? An extreme environment and molecular perspective of black color accumulation in the ornamental and food crops. Front Plant Sci 13:885176–885187. https://doi.org/10.3389/fpls.2022.885176

Ai Y, Li Z, Sun WH, Chen J, Zhang DY, Ma L, Zhang QH, Chen MK, Zheng QD, Liu JF, Jiang YT, Li BJ, Liu X, Xu XY, Yu X, Zheng Y, Liao XY, Zhou Z, Wang JY, Wang ZW, Xie TX, Ma SH, Zhou J, Ke YJ, Zhou YZ, Lu HC, Liu KW, Yang FX, Zhu GF, Huang L, Peng DH, Chen SP, Lan S, Van de Peer Y, Liu ZJ (2021) The *Cymbidium* genome reveals the evolution of unique morphological traits. Hortic Res 8(1):255–269. https://doi.org/10.1038/s41438-021-00683-z

Albert NW, Arathoon S, Collette VE, Schwinn KE, Jameson PE, Lewis DH, Zhang H, Davies KM (2010) Activation of anthocyanin synthesis in *Cymbidium* orchids: variability between known regulators. Plant Cell Tissue Organ Culture (PCTOC) 100(3):355–360. https://doi.org/10.1007/s11240-009-9649-0

Bureau S, Renard CMGC, Reich M, Ginies C, Audergon J-M (2009) Change in anthocyanin concentrations in red apricot fruits during ripening. LWT Food Sci Technol 42(1):372–377. https://doi.org/10.1016/j.lwt.2008.03.010

Dunn WB, Broadhurst D, Begley P, Zelena E, Francis-McIntyre S, Anderson N, Brown M, Knowles JD, Halsall A, Haselden JN, Nicholls AW, Wilson ID, Kell DB, Goodacre R, Human Serum Metabolome C (2011) Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. Nat Protoc 6(7):1060–1083. https://doi.org/10.1038/nprot.2011.335

Feng XK, Gao G, Yu CM, Zhu AG, Chen JK, Chen KM, Wang XF, Abubakar AS, Chen P (2021) Transcriptome and metabolome analysis reveals anthocyanin biosynthesis pathway associated with ramie (*Boehmeria nivea* (L.) Gaud.) leaf color formation. BMC Genom 22(1):684–698. https://doi.org/10.1186/s12864-021-08007-0

Fernández-García E, Carvajal-Lérida I, Jarén-Galán M, Garrido-Fernández J, Pérez-Gálvez A, Hornero-Méndez D (2012) Carotenoids bioavailability from foods: from plant pigments to efficient biological activities. Food Res Int 46(2):438–450. https://doi.org/10.1016/j.foodres.2011.06.007

Ferrer JL, Austin MB, Stewart C Jr, Noel JP (2008) Structure and function of enzymes involved in the biosynthesis of phenylpropanoids. Plant Physiol Biochem 46(3):356–370. https://doi.org/10.1016/j.plaphy.2007.12.009

Gao J, Ren R, Wei YL, Jin JP, Ahmad S, Lu CQ, Wu JQ, Zheng CY, Yang FX, Zhu GF (2020) Comparative metabolomic analysis reveals distinct flavonoid biosynthesis regulation for leaf color development of *Cymbidium sinense* "Red Sun." Int J Mol Sci 21(5):1869–1885. https://doi.org/10.3390/ijms21051869

Ge W, Wang XX, Li JY, Zhu WP, Cui JT, Zhang KZ (2019) Regulatory mechanisms of leaf color change in *Acer pictum* subsp. mono. Front Plant Sci 62(12):793–805. https://doi.org/10.1139/gen-2019-0115

Grabherr MGHB, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol 29(7):644–652. https://doi.org/10.1038/nbt.1883

Han YJ, Lu MM, Yue SM, Li K, Shang FD (2021) Transcriptome and metabolome profiling revealing anthocyanin and phenolic acid biosynthetic mechanisms in sweet osmanthus pericarp. Sci Hortic 289:110489–110498. https://doi.org/10.1016/j.scienta.2021.110489

Hassani D, Fu XQ, Shen Q, Khalid M, Rose JKC, Tang KX (2020) Parallel transcriptional regulation of artemisinin and flavonoid biosynthesis. Trends Plant Sci 25(5):466–476. https://doi.org/10.1016/j.tplants.2020.01.001

Holton TA, Cornish EC (1995) Genetics and biochemistry of anthocyanin biosynthesis. Plant Cell 7(7):1071–1083. https://doi.org/10.1105/tpc.7.7.1071

Jaakola L (2013) New insights into the regulation of anthocyanin biosynthesis in fruits. Trends Plant Sci 18(9):477–483. https://doi.org/10.1016/j.tplants.2013.06.003

Kawase K, Tsukamoto Y, Saito N, Osawa Y (1970) Studies on flower color in *Chrysanthemum morifolium* RAMAT. I. Anthocyanins. Plant Cell Physiol 11(2):349–353. https://doi.org/10.1093/oxfordjournals.pcp.a074516

Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10(3):R25–R34. https://doi.org/10.1186/gb-2009-10-3-r25

Li B, Dewey CN (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinform 12:323–338. https://doi.org/10.1186/1471-2105-12-323

Li XB, Xiang L, Wang Y, Luo J, Wu C, Sun CB, Ming X (2013) Genetic diversity, population structure, pollen morphology and cross-compatibility among Chinese *Cymbidiums*. Plant Breed 133(1):145–152. https://doi.org/10.1111/pbr.12125

Li Y, Zhang ZY, Wang P, Wang SA, Ma LL, Li LF, Yang RT, Ma YZ, Wang Q (2015) Comprehensive transcriptome analysis discovers novel candidate genes related to leaf color in a *Lagerstroemia indica* yellow leaf mutant. Genes Genom 37(10):851–863. https://doi.org/10.1007/s13258-015-0317-y

Li BJ, Zheng BQ, Wang JY, Tsai WC, Lu HC, Zou LH, Wan X, Zhang DY, Qiao HJ, Liu ZJ, Wang Y (2020) New insight into the molecular mechanism of colour differentiation among floral segments in orchids. Commun Biol 3(1):89–101. https://doi.org/10.1038/s42003-020-0821-8

Li X, Li Y, Zhao MH, Hu YB, Meng FJ, Song XS, Tigabu M, Chiang VL, Sederoff R, Ma WJ, Zhao XY (2021) Molecular and metabolic insights into anthocyanin biosynthesis for leaf color change in Chokecherry (*Padus virginiana*). Int J Mol Sci 22(19):10697–10716. https://doi.org/10.3390/ijms221910697

Li WJ, Li HG, Shi LS, Shen P, Li YR (2022) Leaf color formation mechanisms in *Alternanthera bettzickiana* elucidated by metabolite and transcriptome analyses. Planta 255(3):59–78. https://doi.org/10.1007/s00425-022-03840-3

Lim GH, Kim SW, Ryu JA-O, Kang SY, Kim JA-O, Kim SA-O (2020) Upregulation of the MYB2 transcription factor is associated with increased accumulation of anthocyanin in the leaves of *Dendrobium bigibbum*. Int J Mol Sci 21(16):5653–5670. https://doi.org/10.3390/ijms21165653

Liu YF, Ma KX, Qi YW, Lv GW, Ren XL, Liu ZD, Ma FW (2021) Transcriptional regulation of anthocyanin synthesis by MYB-bHLH-WDR complexes in Kiwifruit (*Actinidia chinensis*). J Agric Food Chem 69(12):3677–3691. https://doi.org/10.1021/acs.jafc.0c07037

Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25(4):402–408. https://doi.org/10.1006/meth.2001.1262

Lloyd A, Brockman A, Aguirre L, Campbell A, Bean A, Cantero A, Gonzalez A (2017) Advances in the MYB-bHLH-WD Repeat (MBW) pigment regulatory model: addition of a WRKY factor and co-option of an anthocyanin MYB for betalain regulation. Plant Cell Physiol 58(9):1431–1441. https://doi.org/10.1093/pcp/pcx075

Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15(12):550–570. https://doi.org/10.1186/s13059-014-0550-8

Luo JR, Duan JJ, Huo D, Shi QQ, Niu LX, Zhang YL (2017) Transcriptomic analysis reveals transcription factors related to leaf anthocyanin biosynthesis in *Paeonia qiui*. Molecules 22(12):2186–2199. https://doi.org/10.3390/molecules22122186

Meiling S, Yueyang L, Hongduo W, Zhongming Z (2018) Extraction process of anthocyanins from husk of local-grown rice *Oryza sativa* L. nanyangensis in Nanyang, Henan. Chin Agric Sci Bull 34(18):71–76

Pahlavani MH, Mirlohi AF, Saeidi G (2004) Inheritance of flower color and spininess in safflower (*Carthamus tinctorius* L.). J Hered 95(3):265–267. https://doi.org/10.1093/jhered/esh030

Qi XY, Chen SS, Wang HD, Feng J, Chen HJ, Qin ZY, Deng YM (2022) Comparative physiology and transcriptome analysis reveals that chloroplast development influences silver-white leaf color formation in *Hydrangea macrophylla* var. maculata. BMC Plant Biol 22(1):345–358. https://doi.org/10.1186/s12870-022-03727-1

Qiu WW, Su WQ, Cai ZY, Dong L, Li CB, Xin M, Fang WK, Liu YQ, Wang XM, Huang ZB, Ren H, Wu ZJ (2020) Combined analysis of transcriptome and metabolome reveals the potential mechanism of coloration and fruit quality in yellow and purple *Passiflora edulis* Sims. J Agric Food Chem 68(43):12096–12106. https://doi.org/10.1021/acs.jafc.0c03619

Robinson GM, Robinson R (1932) A survey of anthocyanins. II Biochem J 26(5):1647–1664. https://doi.org/10.1021/acs.jafc.0c03619

Song XH, Yang QS, Liu Y, Li JJ, Chang XC, Xian LH, Zhang J (2021) Genome-wide identification of *Pistacia* R2R3-MYB gene family and function characterization of PcMYB113 during autumn leaf coloration in *Pistacia chinensis*. Int J Biol Macromol 192:16–27. https://doi.org/10.1016/j.ijbiomac.2021.09.092

Sun Y, Chen GZ, Huang J, Liu DK, Xue F, Chen XL, Chen SQ, Liu CG, Liu H, Ma H, Yuan Y, Zhang DY, Sun WH, Peng DH, Wang ZW, Lan SR, Zhao GQ, Li FT, Liu ZJ (2021) The *Cymbidium goeringii* genome provides insight into organ development and adaptive evolution in orchids. Ornam Plant Res 1(1):1–13. https://doi.org/10.48130/opr-2021-0010

Sunil L, Shetty NP (2022) Biosynthesis and regulation of anthocyanin pathway genes. Appl Microbiol Biotechnol 106(5–6):1783–1798. https://doi.org/10.1007/s00253-022-11835-z

Tanaka Y, Sasaki N, Ohmiya A (2008) Biosynthesis of plant pigments: anthocyanins, betalains and carotenoids. Plant J 54(4):733–749. https://doi.org/10.1111/j.1365-313X.2008.03447.x

Tang YH, Fang ZW, Liu M, Zhao DQ, Tao J (2020) Color characteristics, pigment accumulation and biosynthetic analyses of leaf color variation in herbaceous peony (*Paeonia lactiflora* Pall.). 3 Biotech 10(2):76–85. https://doi.org/10.1007/s13205-020-2063-3

Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol 28(5):511–515. https://doi.org/10.1038/nbt.1621

Veberic R, Slatnar A, Bizjak J, Stampar F, Mikulic-Petkovsek M (2015) Anthocyanin composition of different wild and cultivated berry species. LWT Food Sci Technol 60(1):509–517. https://doi.org/10.1016/j.lwt.2014.08.033

Wang HZ, Lu JJ, Hu X, Liu JJ (2011) Genetic variation and cultivar identification in *Cymbidium ensifolium*. Plant Syst Evol 293(1):101–110. https://doi.org/10.1007/s00606-011-0429-z

Wang M, Chen L, Liang ZJ, He XM, Liu WR, Jiang B, Yan JQ, Sun PY, Cao ZQ, Peng QW, Lin YE (2020) Metabolome and transcriptome analyses reveal chlorophyll and anthocyanin metabolism pathway associated with cucumber fruit skin color. BMC Plant Biol 20(1):386–398. https://doi.org/10.1186/s12870-020-02597-9

Wang FF, Chen NZ, Shen SH (2021) iTRAQ-based quantitative proteomics analysis reveals the mechanism of golden-yellow leaf mutant in hybrid paper mulberry. Int J Mol Sci 23(1):127–144. https://doi.org/10.3390/ijms23010127

Want EJ, Wilson ID, Gika H, Theodoridis G, Plumb RS, Shockcor J, Holmes E, Nicholson JK (2010) Global metabolic profiling procedures for urine using UPLC-MS. Nat Protoc 5(6):1005–1018. https://doi.org/10.1038/nprot.2010.50

Wei G, Tian P, Zhang FX, Qin H, Miao H, Chen QW, Hu ZY, Cao L, Wang MJ, Gu XF, Huang SW, Chen MS, Wang GD (2016) Integrative analyses of nontargeted volatile profiling and transcriptome data provide molecular insight into VOC diversity in cucumber plants (*Cucumis sativus*). Plant Physiol 172(1):603–618. https://doi.org/10.1104/pp.16.01051

Xuekui W (2006) Principles and techniques of plant physiological and biochemical experiments. Higher Education Press, Beijing, pp 134–136

Yang FX, Gao J, Wei YL, Ren R, Zhang GQ, Lu CQ, Jin JP, Ai Y, Wang YQ, Chen LJ, Ahmad S, Zhang DY, Sun WH, Tsai WC, Liu ZJ, Zhu GF (2021) The genome of *Cymbidium sinense* revealed the evolution of orchid traits. Plant Biotechnol J 19(12):2501–2516. https://doi.org/10.1111/pbi.13676

Yang JF, Chen YZ, Xiao ZH, Shen HL, Li YH, Wang Y (2022) Multilevel regulation of anthocyanin-promoting R2R3-MYB transcription factors in plants. Front Plant Sci 13:1008829–1008845. https://doi.org/10.3389/fpls.2022.1008829

Yi DB, Zhang HN, Lai B, Liu LQ, Pan XL, Ma ZL, Wang YC, Xie JH, Shi SY, Wei YZ (2021) Integrative analysis of the coloring mechanism of red longan pericarp through metabolome and transcriptome analyses. J Agric Food Chem 69(6):1806–1815. https://doi.org/10.1021/acs.jafc.0c05023

Yu GY, Wang LG, Han YY, He QY (2012) ClusterProfiler: an R package for comparing biological themes among gene clusters. OMICS 16(5):284–287. https://doi.org/10.1089/omi.2011.0118

Zhang HL, Zhang SY, Zhang H, Chen X, Liang F, Qin HL, Zhang Y, Cong RC, Xin HB, Zhang Z (2020) Carotenoid metabolite and transcriptome dynamics underlying flower color in marigold (*Tagetes erecta* L.). Sci Rep 10(1):16835–16845. https://doi.org/10.1038/s41598-020-73859-7

Zhang YX, Zhang GQ, Zhang DY, Liu XD, Xu XY, Sun WH, Yu XE, Zhu X, Wang ZW, Zhao X, Zhong WY, Chen HF, Yin WL, Huang TB, Niu SC, Liu ZJ (2021) Chromosome-scale assembly of the *Dendrobium chrysotoxum* genome enhances the understanding of orchid evolution. Hortic Res 8(1):183. https://doi.org/10.1038/s41438-021-00621-z

Zhang DY, Zhao XW, Li YY, Ke SJ, Yin WL, Lan SR, Liu ZJ (2022a) Advances and prospects of orchid research and industrialization. Hortic Res 9:uhac220–uhac239. https://doi.org/10.1093/hr/uhac220

Zhang XY, Ruan RX, Zhao FK, Li CN, Fu QJ (2022b) Comprehensive transcriptome analysis discovers genes related to leaf color in *Cymbidium* 'Sakura' etiolated mutant. Acta Physiol Plant 45(1):12–21. https://doi.org/10.1007/s11738-022-03471-x

Zhang X, Wang J, Li PH, Sun C, Dong WX (2023) Integrative metabolome and transcriptome analyses reveals the black fruit coloring mechanism of Crataegus maximowiczii C. K. Schneid. Plant Physiol Biochem 194:111–121. https://doi.org/10.1016/j.plaphy.2022.11.008

Zhao XC, Zhang YR, Long T, Wang SC, Yang J (2022) Regulation mechanism of plant pigments biosynthesis: anthocyanins, carotenoids, and betalains. Metabolites 12(9):871–884. https://doi.org/10.3390/metabo12090871