ACTA SCIENTIARUM POLONORUM
Acta Sci. Pol. Hortorum Cultus, 22(5) 2023, 3–17
https://czasopisma.up.lublin.pl/index.php/asphc ISSN 1644-0692 e-ISSN 2545-1405 https://doi.org/10.24326/asphc.2023.4831
ORIGINAL PAPER
Received: 6.07.2022
Accepted: 24.04.2023
Issue published: 30.10.2023

TRANSCRIPTOME ANALYSIS OF GENES INVOLVED IN FLOWER AND LEAF COLOR OF *Oncidium* BY RNA-SEQ

Ma-Yin Wang[1], Yu Ding[1], Ye Zhang[1], Lu Sun[1], Xi-Qiang Song[1], Dai-Cheng Hao[2], Wei-Shi Li[2], Min-Qiang Tang[1], Peng Ling✉[1], Shang-Qian Xie[1]

[1] Key Laboratory of Ministry of Education for Genetics and Germplasm Innovation of Tropical Special Trees and Ornamental Plants, College of Horticulture, Hainan University, Haikou 570228, China

[2] Hainan Boda Orchid Technology Co. Ltd, Haikou 570311, China

ABSTRACT

Oncidium, an important tropical orchid, has high ornamental value due to its specific color and occupies a significant market position for the worldwide flower. Transcriptome analysis of flower and leaf color formation provides new sources for producing novel *Oncidium hybridum* cultivars. We sequenced 12 samples of flowers (yellow and white) and leaves (striped and regular) of *O. hybridum* and assembled 381,136 and 453,566 unigene sequences from RNA-seq data, respectively. Among unigenes, 662 and 1,324 differentially expressed genes were identified in flower and leaf samples, respectively. Gene ontology and pathway enrichment showed that secondary metabolite biosynthetic pathways were responsible for flower and leaf color formation. It was determined that *UGT75C1*, *E2.4.1.115*, *CCD7*, *E2.1.1.76*, and *CCoAOMT* are involved in regulating flower color, and *UGT75C1*, *LHCB*, *UGT*, *RP-L18Ae*, and A*BCB1* play crucial roles in regulating leaf color. Kyoto Encyclopedia of Genes and Genomes analysis revealed that *UGT75C1* was significantly enriched in the anthocyanin biosynthetic pathway, showing effects on flower and leaf colors. This study was the first detailed analysis of the molecular mechanisms of *O. hybridum* flower and leaf colors, and the results advanced the understanding of the genetic basis of flower and leaf colors; they also provided additional support for improving commercial value and producing novel cultivars of *O. hybridum*.

Key words: *Oncidium hybridum*, RNA-seq, differential analysis, differentially expressed genes, key genes

INTRODUCTION

Oncidium is an important ornamental plant native to tropical and subtropical regions of Central and South America, with about 750 native species and more than 400 wild species worldwide. Most cultivars of *Oncidium* exhibit diverse flower colors, such as yellow and orange, and have a long flowering period that ensures high ornamental value. *Oncidium* is mainly sold as cut flowers and is an essential component in the international cut flower market [Yang et al. 2017]. To facilitate the color breeding of *Oncidium*, it is necessary to scientifically understand the formation and regulatory mechanism of flower color in *Oncidium*.

Previous studies have found that *Oncidium* flowers were pigmented by carotenoids and anthocyanins [Chiou and Yeh 2008]. Hieber identified five key pigment biosynthesis genes, including phytoene synthase, phytoene desaturase, carotenoid isomerase, the downstream 9-cis epoxycarotenoid dioxygenase, and dihydroflavonol 4-reductase, that generated a yellow pigment by mixing all-trans and 9-cis isomers of vi-

✉ 18389807612@163.com

olaxanthin [Hieber et al. 2006]. Simultaneously, four key genes, including chalcone synthase, chalcone isomerase, dihydroflavonol 4-reductase, and anthocyanidin synthase involved in the anthocyanin biosynthetic pathway were also identified in *Oncidium Gower Ramsey* [Chiou and Yeh 2008]. In addition to the study of flower color, leaf color has been increasingly discussed in recent years due to its unique value in genetics and breeding [Zhao et al. 2020]. Leaf color variation can be used as a marker to simplify the breeding of improved varieties and hybrid production. Some leaf color variations have excellent characteristics, which provide excellent germplasm resources for crop genetics and breeding. Moreover, leaf color variation can be used in studying photomorphogenesis, photosynthesis, hormone physiology, and disease resistance mechanisms [Zhang 2022]. Chlorophyll is the primary pigment in regular leaves, and the genes involved in chlorophyll biosynthesis and chloroplast development and division are critical factors in leaf color formation [Cai et al. 2023]. Chlorophyll-chlorophyllido hydrolases pheophorbide an oxygenase, red Chl catabolite reductase, and chlorophyll b reductase have been reported in the leaf color formation of *Cymbidium sinense* [Zhu et al. 2015]. Moreover, several genes or gene families, including the golden 2-like gene family, Arabidopsis pseudo-response regulator 2 (APRR2), and the accumulation and replication of the chloroplast gene family have been investigated [Fitter et al. 2010, Pan et al. 2013]. However, there are few studies on the genetic regulation of leaf color in *Oncidium hybridum*. Therefore, the formation and regulatory mechanisms of leaf color in *O. hybridum* warrant further study.

This study aimed to investigate the key genes involved in the regulation of flower as well as leaf color of the *O. hybridum* and how they affect its formation. In this context, we carried out transcriptome differential analysis on petal and leaf samples of 'Ningmenghuang' and 'Xuelian' and further detailed analysis of ten relevant differentially expressed genes.

MATERIAL AND METHODS

Sample collection and RNA extraction. The 12 experimental samples were collected from Hainan Boda Orchid Scientific Technology Company nursery (Hainan, China), including three samples each of yellow petal, white petal, standard leaf, and striped leaf samples of *O. hybridum* (Fig. 1). The six petal samples were in the full bloom stage of two cultivars (Ningmenghuang and Xuelian). The yellow samples were collected from one individual cultivar, Ningmenghuang, and three white samples were from one individual of Xuelian (Fig. 1A and B). The six leaf samples were from two individuals of cultivar Ningmenghuang, and the striped samples were from a natural variant of Ningmenghuang (Fig. 1C and D). The leaves were sampled at the flowering stage.

The ploidy of collected samples was detected in a preliminary experiment using the chromosome squashing technique before RNA extraction. Then, these samples were placed in liquid nitrogen and stored at –80°C. The samples were monitored for RNA degradation and contamination on 1% agarose gels. Next, RNA purity was checked using the NanoPhotometer spectrophotometer (Implen, Westlake Village, CA, USA), and RNA concentration was measured using Qubit RNA Assay Kit in Qubit 2.0 Fluorometer (Life Technologies, Carlsbad, CA, USA). Then, RNA integrity was assessed using the RNA Nano 6000 Assay Kit with the Bioanalyzer 2100 system (Agilent Technologies, Santa Clara, CA, USA).

Construction of cDNA library and sequencing. In total, 1 μg RNA per sample was used as the input material to prepare samples, and mRNA was purified from the total RNA using poly-T oligo-attached magnetic beads (AMPure XP, Beckman Coulter, Indianapolis, IN, USA). Fragmentation was carried out using divalent cations under an elevated temperature in NEBNext First-Strand Synthesis Reaction Buffer. First-strand cDNA was synthesized using a random hexamer primer and M-MuLV Reverse Transcriptase (RNase H). Second-strand cDNA synthesis was subsequently performed using DNA polymerase I and RNase H (5×; NEBNext, New England Biolabs, Ipswich, MA, USA). The remaining overhangs were converted into blunt ends via exonuclease/polymerase activities. After the adenylation of 3′ ends of DNA fragments, a NEBNext Adaptor with a hairpin loop structure was ligated to prepare for hybridization. Then, 3 μL USER Enzyme (USER, New England Biolabs, Ipswich, MA, USA) was used with size-selected, adaptor-ligated cDNA at 37°C for 15 min followed by 95°C for 5 min before PCR. Then, PCR was performed

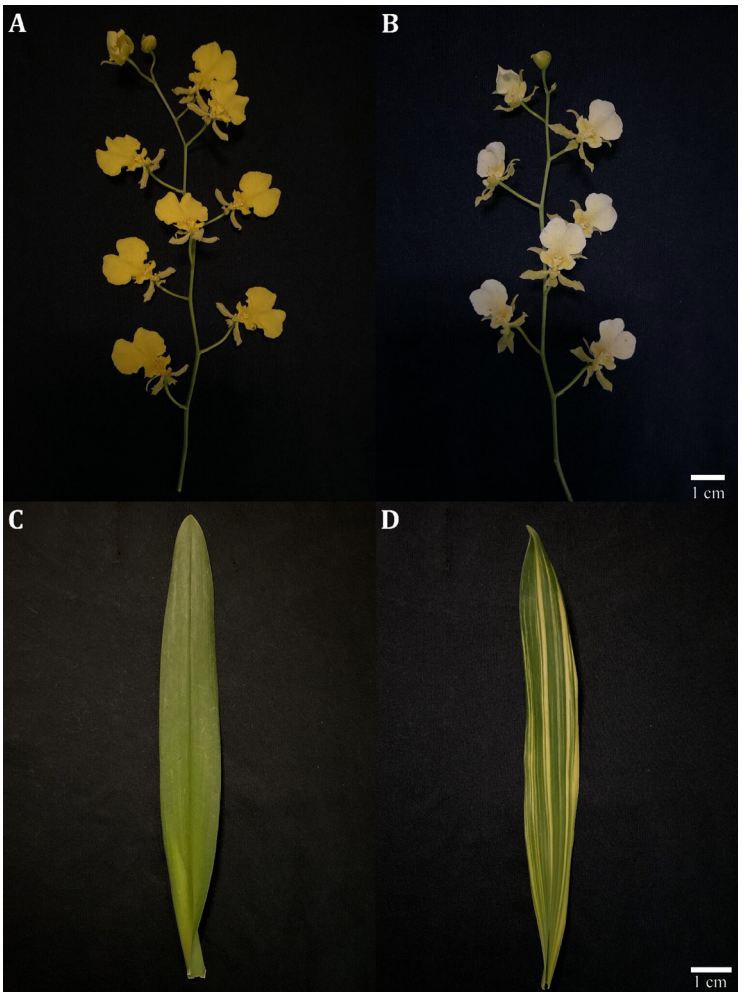

Fig. 1. The orchid materials used in this study. A – yellow petals of Ningmenghuang, B – white petals of Xuelian, C – typical leaf of Ningmenghuang, D – striped leaf natural variant Ningmenghuang

with Phusion High-Fidelity DNA polymerase (Phusion, New England Biolabs, Ipswich, MA, USA), universal PCR primers, and Index (X) Primer. Finally, PCR products were purified (AMPure XP system, Beckman Coulter, Indianapolis, IN, USA), and library quality was assessed on the Agilent Bioanalyzer 2100 system (Agilent 2100, Agilent Technologies, CA, USA). Sequencing was carried out using an Illumina HiSeq 2500 instrument (Illumina, San Diego, CA, USA).

De novo assembly and annotation. The quality of raw RNA-seq data was controlled by fastQC v0.11.7, and reads with joints and low quality (Q Phred ≤ 20 bases accounting for more than 50% of the whole read) were removed [Andrews 2014]. Then, reads were *de novo* assembled into contigs using Trinity v2.8.5 [Grabherr et al. 2013], and the *de novo* assembled unigenes from contigs were annotated by eggNOG v5.0 [Huerta-Cepas et al. 2019]. The protein sequences were aligned based on the diamond algorithm pattern to the Nr, Swiss-prot, Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG), and Cluster of Orthologous Groups of Proteins (COG) databases. According to protein sequence similarity, the COG database categorizes protein sequences into different

classes. Each class was assigned a COG number, representing a homologous protein, and all homologous proteins were subdivided into 25 major classes [Galperin et al. 2020].

Differential expression analysis. The Salmon algorithm was used to quantify the abundance of transcripts against unigenes [Patro et al. 2017]. Then, we used the DESeq2 package of R to normalize the abundance of gene expression and analyze DEGs [Love et al. 2014]. The threshold value for DEGs was $P < 0.05$, and DEGs with $|\log_2 \text{foldchange}| > 1$ were considered as up-regulated or down-regulated genes.

Enrichment analysis of DEGs. The DEGs were compared with the analysis results of eggNOG v2.0.0 software, and annotation information was obtained from the GO database. GO enrichment analysis of DEGs was implemented by the cluster Profiler package of R, and GO terms with corrected P-values of less than 0.05 were considered significantly enriched DEGs. Furthermore, we used KOBAS v2.0 to analyze the statistical enrichment of DEGs in KEGG pathways [Chen et al. 2011].

RESULTS

Unigene assembly and functional annotation. Six samples of yellow flowers (NF1, NF2, NF3) and white flowers (WF1, WF2, WF3) and six samples of normal leaves (NL1, NL2, NL3) and striped leaves (SL1, SL2, SL3) were collected, and the chromosome squashing technology showed that the chromosome number of collected samples was about 2n = 120 for both Ningmenghuang and Xuelian cultivars (Additional file 1: Fig. S1). Then, the extracted RNA from these samples was sequenced on an Illumina Hiseq platform. 52,095,898; 64,069,126; 47,136,772; 52,270,916; 55,795,220; and 54,045,378 raw reads were obtained from NF1, NF2, NF3, WF1, WF2 and WF3, respectively. The NL1, NL2, NL3, SL1, SL2 and SL3 samples generated 64,962,964; 52,742,122; 62,900,848; 62,097,324; 61,934,142 and 63,935,170 raw reads from RNA-seq datasets, respectively. All read rates and base rates of 12 samples were greater than 96%, and the size of clean reads of each sample was more than 6.82 Gb. The total sizes of clean reads of flower and leaf samples were 47.78 Gb and 54.5 Gb, respectively (Tab. 1). Moreover, all Q30 content (bases in clean reads with a base quality value of greater than 30, that is, have sequencing error rates less than 0.1%) was greater than 91%, and the GC content of each sample was around 45% (Tab. 1). The above results indicated that the quality of clean data was sufficient for subsequent analysis.

Unigenes of clean reads were assembled by Trinity for flower and leaf samples. Among them, the total numbers of Unigene sequences in flower and leaf samples were 381,136 and 453,566, and the total lengths of the sequences were 200,010,879 base pairs (bp) and

Table 1. Statistical information of sequencing data of 12 samples

Sample ID	Raw reads no.	Clean reads no.	Read rate (%)	Total size of raw reads (Gb)	Total size of clean reads (Gb)	Base rate (%)	Q20 (%)	Q30 (%)	GC (%)
NF1	52,095,898	50,562,576	97.06	7.81	7.58	97.06	96.42	91.46	47.63
NF2	64,069,126	62,793,018	98.01	9.67	9.48	98.04	97.92	94.69	44.93
NF3	47,136,772	45,471,290	96.47	7.07	6.82	96.46	96.50	91.70	45.23
WF1	52,270,916	51,170,980	97.90	7.84	7.68	97.96	97.20	92.99	44.87
WF2	55,795,220	54,621,856	97.90	8.43	8.25	97.86	97.90	94.69	44.83
WF3	54,045,378	52,769,038	97.64	8.16	7.97	97.67	97.78	94.41	45.02
NL1	64,962,964	63,870,426	98.32	9.81	9.64	98.27	97.62	93.95	48.88
NL2	52,742,122	51,967,582	98.53	7.91	7.80	98.61	97.96	94.83	45.55
NL3	62,900,848	61,699,886	98.09	9.50	9.32	98.11	97.85	94.57	44.98
SL1	62,097,324	60,735,798	97.81	9.38	9.17	97.76	97.77	94.39	45.18
SL2	61,934,142	60,590,976	97.83	9.35	9.15	97.86	97.89	94.62	45.26
SL3	63,935,170	62,447,904	97.67	9.65	9.43	97.72	97.85	94.53	45.51

Table 2. Statistical results of unigene sequences

Assembled transcripts	Flower	Leaf
Total sequence number	381,136	453,566
Total sequence bases (bp)	200,010,879	226,354,851
GC (%)	37.61	39.36
Median contig length (bp)	349	334
Max. length (bp)	28,028	28,066
Min. length (bp)	201	201
Mean length (bp)	524	499
N50 length (bp)	603	561

226,354,851 bp, respectively (Tab. 2). All unigenes were longer than 200 bp, and the length was mainly distributed between 200–500 bp (Additional file 1: Fig. S2). The assembled unigenes were annotated by Nr, Swissprot, GO, KEGG, and COG databases with e-value = 1e – 5. The number of functional annotation sequences obtained from COG, GO, KEGG, and Swiss-prot databases were 38,090 (9.99%), 19,250 (5.05%), 15,030 (3.94%), and 54,291 (14.24%) in the flower samples, and 54,202 (11.95%), 30,840 (6.80%), 26,358 (5.81%), and 72,148 (15.91%) in the leaf samples, respectively (Additional file 1: Tab. S1).

COG database distribution. To further annotate homologous proteins of unigenes, we analyzed 38,090 and 54,202 unigenes from flower and leaf samples, respectively, in the COG database. Among them, about 71% of unigene sequences were grouped into three categories, including Information Storage and Processing (9,130), Cellular Processes and Signaling (9,704), and Metabolism (9,552), which accounted for 23.97%, 25.48%, and 25.08% of all sequences in flowers, respectively (Tab. 3 and Additional file 1: Fig. S3). For leaf samples, about 74% of unigene sequences were grouped into three categories, namely, Information Storage and Processing (12,332), Cellular Processes and Signaling (14,728), and Metabolism (15,274), which accounted for 22.75%, 27.17%, and 28.18% of all sequences, respectively (Tab. 3 and Additional file 1: Fig. S3).

Identification of DEGs. The results showed that there were 662 and 1,324 DEGs in flower and leaf samples, respectively (Fig. 2, Additional file 2: Tab. S2 and S3). In the flower dataset, 321 genes were up-regulated and 341 were down-regulated (Fig. 2A and C, Additional file 2: Tab. S2); upon annotating these genes to Swiss-prot, GO, KEGG, COG, and Nr databases, the resulting numbers of annotated genes were 385 (58.16%), 252 (38.07%), 178 (26.89%), 445 (67.22%), and 274 (41.39%), respectively. Notably, 126 DEGs were annotated in all these databases (Additional file 1: Fig. S4A and Tab. S4). Among 1,324 DEGs from the leaf samples, 307 genes were up-regulated and 1,017 were down-regulated (Fig. 2B and D, Additional file 2: Tab. S3), and there were 719 (54.31%), 430 (32.48%), 291 (21.98%), 779 (58.84%), and 649 (49.02%) annotated genes in Swiss-prot, GO, KEGG, COG and Nr databases, respectively. Furthermore, there were 125 DEGs annotated in all these databases (Additional file 1: Fig. S4B and Tab. S4).

GO and KEGG pathway enrichment analysis of DEGs. To further annotate the DEG function, GO analysis was carried out. In flower samples, DEGs were involved in 829 biological processes, 328 cell components, and 142 molecular functions (Fig. 3A and Additional file 2: Tab. S5). The biological processes were significantly enriched in cellular processes, metabolic processes, and single-organism processes. The enrichment of cell components included cell parts, organelles, and other complex cell components. Molecular functions with significant enrichment were mainly involved in catalytic activity and binding (Fig. 3A). In leaf samples, DEGs were involved in 1,645 biological processes, 1,031 cell components, and 332 molecular functions (Fig. 3B and Additional file 2: Tab. S6). The biological processes were significantly enriched in cellular processes, metabolic processes, single-organism processes, and responses to stimuli. The enrichment of cell components included cells, or-

Table 3. Distribution of COG categories

Information storage and processing	Flower	Leaf
Transcription (K)	2,473	2,851
Replication, recombination, and repair (L)	3,738	3,920
Translation, ribosomal structure, and biogenesis (J)	1,431	3,149
RNA processing and modification (A)	1,076	1,835
Chromatin structure and dynamics (B)	412	577
Total	9,130 (23.97%)	12,332 (22.75%)
Cellular processes and signaling		
Posttranslational modification, protein turnover, chaperones (O)	2,651	4,587
Signal transduction mechanisms (T)	3,960	5,041
Intracellular trafficking, secretion, and vesicular transport (U)	1,126	2,085
Cytoskeleton (Z)	556	1,108
Cell wall/membrane/envelope biogenesis (M)	428	625
Cell cycle control, cell division, chromosome partitioning (D)	476	686
Defense mechanisms (V)	241	291
Extracellular structures (W)	152	129
Nuclear structure (Y)	109	165
Cell motility (N)	5	11
Total	9,704 (25.48%)	14,728 (27.17%)
Metabolism		
Carbohydrate transport and metabolism (G)	1,803	3,054
Energy production and conversion (C)	3,424	4,680
Secondary metabolite biosynthesis, transport, and catabolism (Q)	1,214	1,821
Amino acid transport and metabolism (E)	833	1,661
Lipid transport and metabolism (I)	859	1,625
Inorganic ion transport and metabolism (P)	859	1,343
Coenzyme transport and metabolism (H)	324	645
Nucleotide transport and metabolism (F)	236	445
Total	9,552 (25.08%)	15,274 (28.18%)
Poorly characterized		
Function unknown (S)	11,017	13,967
Total	11,017 (28.92%)	13,967 (25.77%)

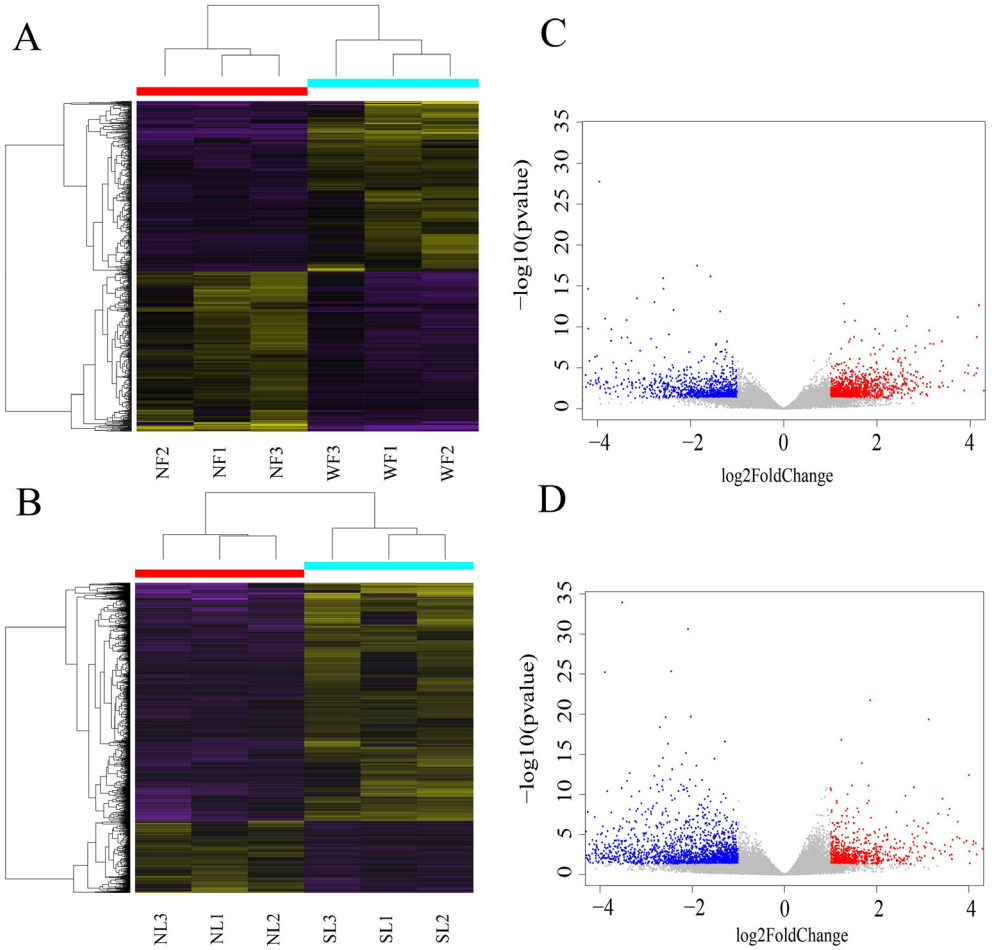

Fig. 2. Analysis of DEGs in flower and leaf samples. A, B – heatmaps of DEG expression in flower (A) and leaf (B) samples; C, D – volcano plots of DEG expression in yellow vs. white petals (C) and in striped vs. regular leaves (D). (NF – The samples of yellow flowers from Ningmenghuang; WF – The samples of white flowers from Xuelian; NL – The samples of typical leaves from Ningmenghuang; SL – The samples of striped leaves from Ningmenghuang)

ganelles, and membranes. The molecular functions showing significant enrichment were mainly involved in catalytic activity, binding, and nuclear acid binding transcription factor activity (Fig. 3B).

Furthermore, we used KOBAS v2.0 to analyze the KEGG pathway enrichment of DEGs (Fig. 4A and Additional file 2: Tab. S7). The results indicated that these DEGs were involved in 137 signaling pathways in flower samples, including biosynthesis of anthocyanins, carotenoids, flavones and flavonols, phenylpropanoids, flavonoids, and betalains; metabolism of xenobiotics by cytochrome P450; melanogenesis; and many other essential signaling pathways. Meanwhile, there were 143 signaling pathways in leaf samples, including photosynthesis, photosynthesis-antenna proteins, porphyrin and chlorophyll metabolism, ribosomes, biosynthesis of amino acids, metabolism of xenobiotics by cytochrome P450, and other signaling pathways (Fig. 4B and Additional file 2: Tab. S8).

Among flower samples, there were 5 DEGs (*UGT75C1*, *E2.4.1.115*, *CCD7*, *E2.1.1.76*, and *CCoAOMT*) enriched in the anthocyanin, carotenoid,

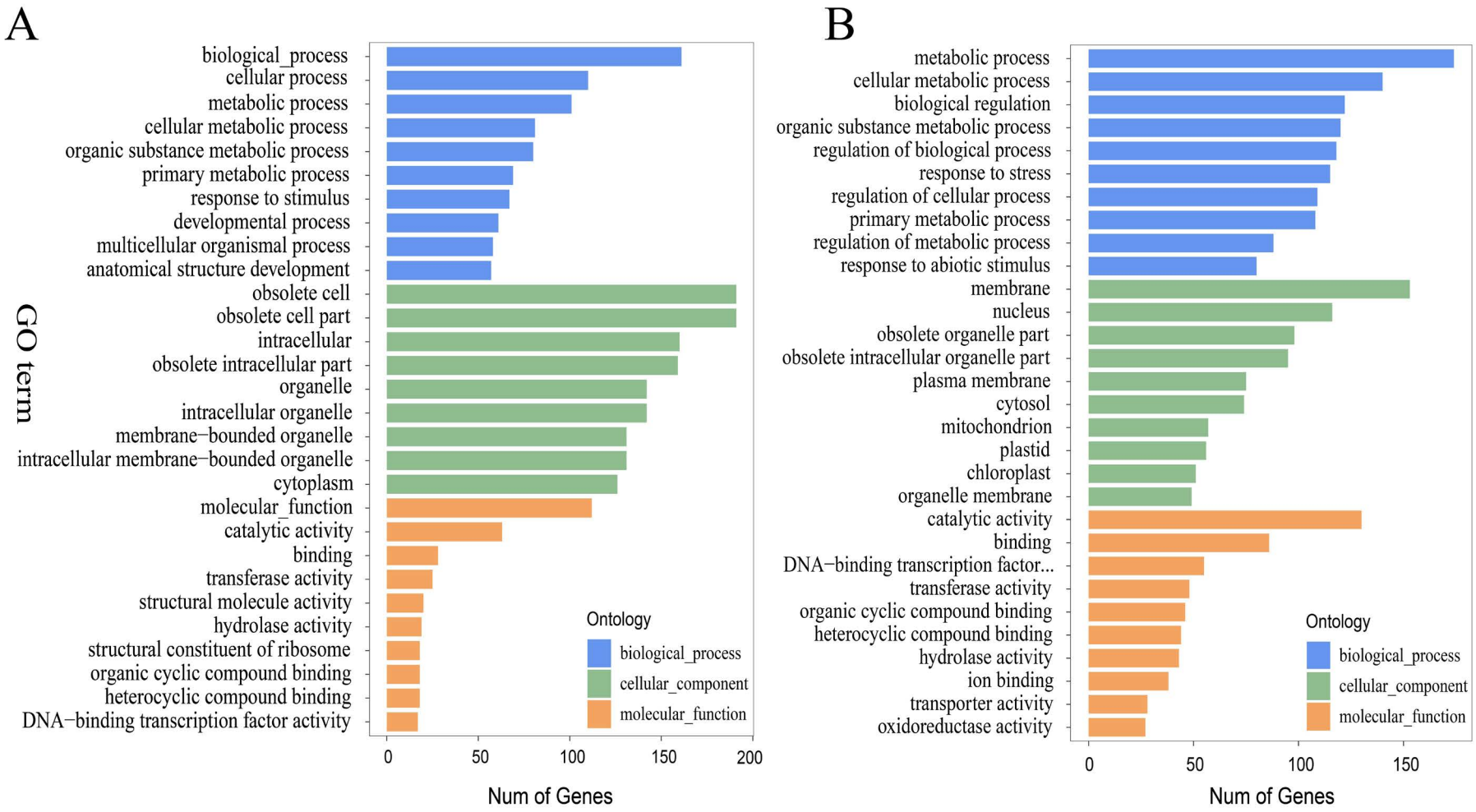

Fig. 3. GO classification enrichment pathways of DEGs. A – flower samples; B – leaf samples

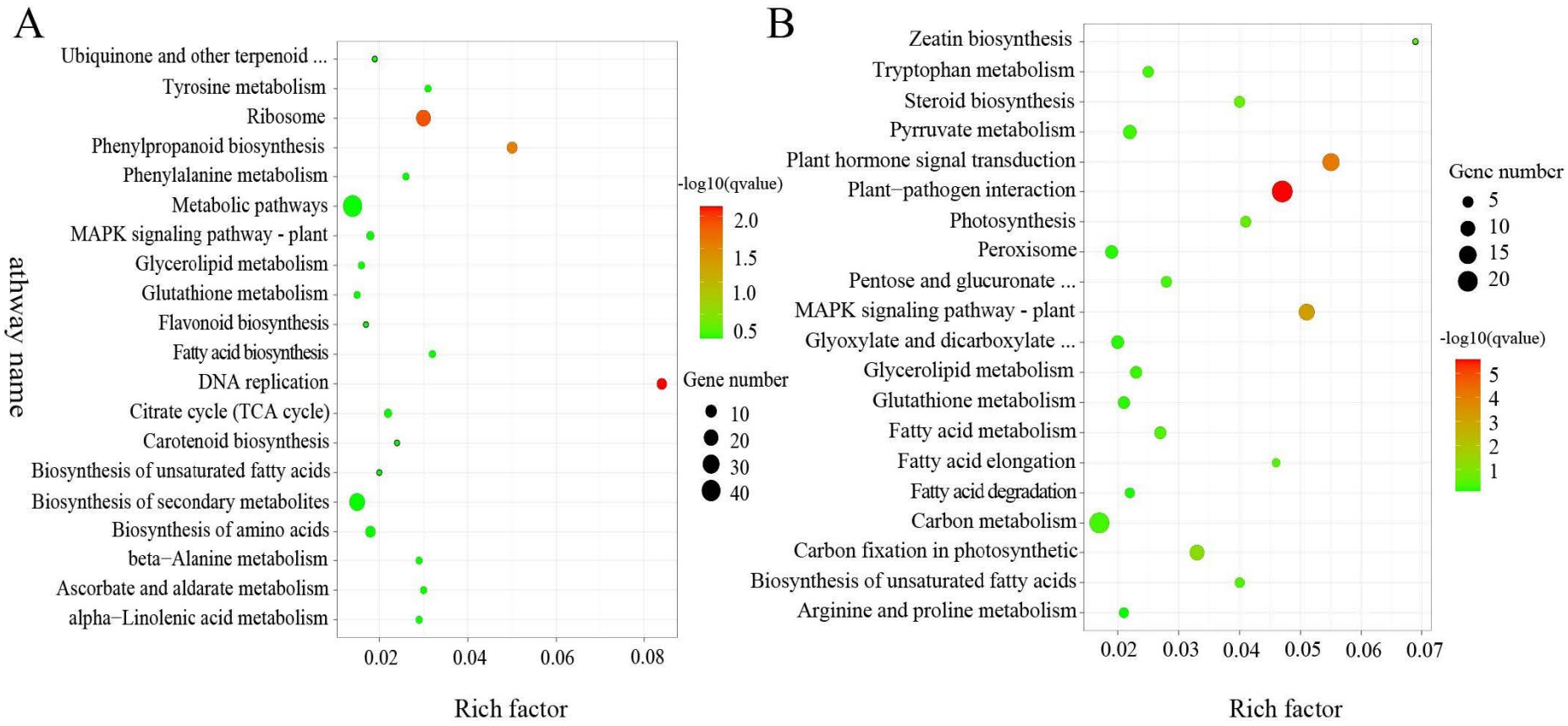

Fig. 4. The top 20 enriched KEGG pathways of DEGs. A – flower samples; B – leaf samples

Table 4. Annotation pathways of identified unigenes of *Oncidium hybridum*

Unigene ID	Gene_name	Definition	Pathway	Expression	Tissue
DN30777_c0_g1_i2 DN30120_c0_g1_i2 DN79020_c0_g1_i2	UGT75C1	anthocyanidin 3-O-glucoside 5-O-glucosyltransferase	anthocyanin biosynthesis	down	flower/leaf
DN6864_c0_g2_i6	E2.4.1.115	anthocyanidin 3-O-glucosyltransferase	anthocyanin biosynthesis	up	flower
DN351_c0_g1_i2	CCD7	9-cis-beta-carotene 9', 10'-cleaving dioxygenase	carotenoid biosynthesis	up	flower
DN44326_c0_g1_i6	E2.1.1.76	flavonol 3-O-methyltransferase	flavone and flavonol biosynthesis	up	flower
DN6254_c0_g1_i3 DN10386_c0_g1_i1	CCoAOMT	caffeoyl-CoA O-methyltransferase	flavonoid biosynthesis	up	flower
DN12198_c0_g1_i16 DN42588_c0_g2_i4 DN45643_c0_g1_i4	LHCB	light-harvesting complex II chlorophyll a/b-binding protein	photosynthesis – antenna proteins	down	leaf
DN4894_c1_g2_i1	UGT	glucuronosyltransferase	porphyrin and chlorophyll metabolism	down	leaf
DN47031_c0_g1_i1	RP-L18Ae	large subunit ribosomal protein L18Ae	ribosome	down	leaf
DN10768_c0_g1_i10	ABCB1	ATP-binding cassette, subfamily B (MDR/TAP), member 1	ABC transporters	down	leaf

flavone and flavonol, and flavonoid biosynthetic metabolic pathways in KEGG The expression of *E2.4.1.115*, *CCD7*, *E2.1.1.76*, and *CCoAOMT* were up-regulated, and that of *UGT75C1* was down-regulated (Tab. 4 and Additional file 2: Tab. S9). Meanwhile, we identified five down-regulated DEGs (*LHCB*, *UGT*, *RP-L18Ae*, *ABCB1*, and *UGT75C1*) enriched in photosynthesis-antenna proteins, porphyrin, and chlorophyll metabolism, ribosomes, ABC transporters, and anthocyanin biosynthetic metabolic pathways in leaf samples, which were the key pathways that regulated leaf color (Tab. 4 and Additional file 2: Tab. S9). Among these DEGs, the *UGT75C1* gene was identified in both the leaf and flower datasets.

DISCUSSION

In view of the imperfect genomics information of *O. hybridum*, obtaining DEGs using transcriptome sequencing technology is an effective way to systematically mine the regulatory genes related to flower color and leaf color. In this study, we identified a total of 381,136 and 453,566 unigenes in flower and leaf samples, respectively, of which 662 and 1,324 were differentially expressed. Both GO enrichment and KEGG pathway analyses showed that DEGs were significantly enriched in the biosynthetic pathways of secondary metabolites.

Through the functional annotation results of NR, Swissprot, and COG, we obtained a more precise and accurate understanding of the classification and function of these DEGs. In combination with the results of GO and KEGG enrichment analyses, we finally screened the DEGs related to the regulation of color function and those that were enriched in related metabolic pathways (Tab. 4). Interestingly, the gene UGT75C1 was differentially expressed in both samples with enrichment in a key regulatory embolic pathway.

In addition, transcription factors play a crucial regulatory role in many biosynthetic pathways. MYB

transcription factor genes were identified in petal and leaf samples. These factors play an essential role in the synthesis of secondary metabolites in plants and mainly participate in the phenylpropanoid metabolic pathway [Niu et al. 2016]. The synthesis of anthocyanins and flavonoids in secondary products is closely related to MYB transcription factors; thus, they can regulate plant color and leaf color. Such factors may play a regulatory role in flower color and leaf color to a certain extent by affecting the synthesis of flavonoid biosynthetic pathway products.

DEGs related to flower color. In this study, we identified four transferase enzymes (UGT75C1, E2.4.1.115, E2.1.1.76, and CCoAOMT), which were enriched in anthocyanin, flavone, flavonol, and flavonoid biosynthetic pathways, these key pathways were reported to play crucial roles in regulating flower color [Patra et al. 2013, Yue et al. 2022]. It was found that flower colors were derived from three types of chemically distinct pigments: anthocyanins, carotenoids, and betalains. The major pigments were anthocyanins, which are water-soluble flavonoids accumulated in the vacuoles of petal epidermal cells or anthocyanoplasts [Zhang et al. 2023]. Most of the functional diversity displayed by flavonoids results from modification of the core flavone structure by transferase, which catalyzes the attachment of functional groups, such as methyl groups, sugars, prenyl groups, and sulfhydryl groups, to the hydroxyl groups on the flavone cores [Liao et al. 2023]. Previous studies also demonstrated that anthocyanidin 3-O-glucosyltransferase (E2.4.1.115) plays a vital role in modifying flower color as it catalyzes the transfer of the glycosyl moiety from UDP-glucose to the 3-hydroxyl group of the corresponding acceptor molecule, forming first stable product of the anthocyanin pathway through glycosylation [Wang et al. 2023]. In addition, the expression of E2.4.1.115 was up-regulated, which was consistent with the previously reported results for the Freesia hybrid [Sui et al. 2011]. Previous studies found that the gene anthocyanidin 3-O-glucoside 5-O-glucosyltransferase (UGT75C1) belonged to the 5GT gene subfamily, and its activity causes a change in anthocyanin synthesis [Zhao et al. 2022]. In the carotenoid biosynthetic pathway, 9-cis-beta-carotene 9',10'-cleaving dioxygenase (CCD7) had been shown to mediate strigolactone biosynthesis by catalyzing the oxidative cleavage of carotenoids [Yue et al. 2022]. CCD7 is necessary for AMF-induced C14 carotenoid biosynthesis and participates in nutrient absorption under AMF-plant environmental stress [Pan et al. 2016]. CCoAOMT converts caffeoyl CoA to feruloyl CoA via the phenylpropane biosynthetic pathway, is part of a multifunctional protein family that exhibits activity against flavonoid substrates (anthocyanin 3-O-glucoside chloride) and anthocyanins and can be methylated. The down-regulation of its expression would lead to the activation of anthocyanin biosynthesis [Giordano et al. 2016].

The phenylpropanoid metabolic pathway is upstream of the flavonoid biosynthetic pathway. The MYB transcription factor family mainly plays roles in the phenylpropanoid metabolic pathway, ultimately producing phenylalanine, which in turn produces various secondary metabolites under the action of phenylalanine ammonia-lyase, including anthocyanins and flavonoids, which are also substrates for plant pigment synthesis [Niu et al. 2016]. The obtained DEGs UGT75C1, UGT, E2.1.1.76, and CCoAOMT all encode enzymes that modify anthocyanins and flavonoids, and CCD7 is a carotene-modified enzyme. Differential expression of these transcription factors and enzyme genes ultimately influenced anthocyanin synthesis and played a regulatory role in flower color.

DEGs related to leaf color. We found 5 DEGs that were significantly enriched in regulatory metabolic pathways in leaf samples. Chlorophyll a/b-binding protein belongs to the light-harvesting pigment-protein complexes (LHC) encoded by the LHCB gene in the nucleus, which is responsible for capturing and transmitting light energy to the reaction center of the thylakoid membrane. It was reported that the LHC protein was significantly reduced or even absent in Arabidopsis, resulting in chlorophyll deficiency [Kim et al. 2009]. In the photosynthesis-antenna protein pathway, LHCB, which is related to chloroplast development and function, was down-regulated. The ribosome was composed of large and small subunits, and the main components were RNA and protein. In the RPS gene family of the large ribosomal subunit in the ribosomal pathway, RP-L18Ae was down-regulated. Moreover, there were six other DEGS (RP-L3e, RP-L4e, RP-L8e, RP-S20e, RP-S2e, and RP-SAe) that were enriched in the ribosome metabolic pathway without expression.

It had been reported that after the deletion of the RPS18 coding gene of the Tobacco chloroplast ribosomal small subunit, the leaves were deformed to varying degrees with an albino phenotype and eventually died [Rogalski et al. 2006]. The down-regulation of multiple LHCB family genes and the significant changes in the expression of multiple genes encoding proteins in the large-small ribosomal subunits indicated that the chloroplast structure of striped leaves was damaged, and thylakoid membrane development was defective. Moreover, in porphyrin and chlorophyll metabolic pathways, the down-regulated DEG UGT affected tetrapyrrole biosynthesis to a certain extent. The pyrrole biosynthetic pathway includes two branches of chlorophyll biosynthesis and heme, finally forming a photosensitive pigment chromophore after a series of reactions. The rate of chlorophyll synthesis was affected by the amount of heme in the cell. With an increase in cellular heme, feedback of the excess heme would occur to inhibit chlorophyll synthesis [Adjei et al. 2023]. The ABC transporter is a transmembrane transport protein present in most biofilms and is involved in the transport of substances between chloroplasts, mitochondria, peroxides, and vacuoles. The substrates for transportation mainly include sugars, amino acids, proteins, plant hormones, and secondary metabolites [Chang et al. 2023].

Interestingly, UGT75C1 (anthocyanidin 3-O-glucoside 5-O-glucosyltransferase) was down-regulated in both flower and leaf samples, which plays an essential role in regulating flower and leaf color formation in the anthocyanin biosynthetic metabolic pathway (Additional file 1: Fig. S5). It affects anthocyanin synthesis through glycosyl transfer and regulates anthocyanin content in petals and leaves, leading to changes in flower color and leaf color. Anthocyanin is a principal pigment in flowers and leaves, which confers intense red-to-blue cyanic colors on petals and helps to attract pollinators. Furthermore, representative secondary products found in most plant species are formed through two stages of biosynthetic reactions. The early-stage reactions common to most anthocyanins are responsible for the formation of anthocyanidin 3-O-glucosides, which are the first stable colored metabolites in anthocyanin biosynthesis. The late stage involves further modification reactions, such as glycosylation, acylation, and methylation, which are concerned with fine adjustments of floral color in cultivars. Such biosynthesis involves glycosylation steps that are important for the stability of the pigment and its aqueous solubility in vacuoles [Liao et al. 2023]. UGT75C1 is a glycosyltransferase gene involved in flavonoid metabolism that plays a vital role in the flavonoid anabolic pathway. It belongs to the 5GT gene subfamily and is a functional UDP glucose: anthocyanin 5-O glucosyltransferase. Glycosylation at the 5-position is believed to allow for more stable complexes in the co-pigmentation of anthocyanins, resulting in flowers with a bright reddish-purple to dull violet color, which is pure red when 5-O-glycosylation does not take place. Thus, 5GT is one of the most important enzymes that modifies flower color, particularly for creating a purple hue. Anthocyanins of some plant cultivars, e.g., *Perilla frutescens* (L.) Britton and *Verbena* × *hybrida* Voss, are 5-glucosylated [Zhao et al. 2022]. The results indicated that differential expression of UGT5C1 in both petals and leaves contributed to the color change by influencing the glycosylation of anthocyanin.

Previous studies further confirmed the rationality and reliability of the genes identified in *O. hybridum* by RNA-seq in this study. In summary, this study provided a good reference and basis to verify the selection and identification of regulatory genes for flower and leaf colors and to elucidate regulatory pathways. Furthermore, the DEG UGT75C1 plays a vital role in regulating color formation in both plant petals and leaves. Our results revealed the key genes related to the regulation of flower color and leaf color. We speculated that variations in the expression of these genes affected pigment synthesis in the petals and leaves of *Oncidium* to a certain extent, causing the different colors in petals and yellow stripes in the leaves. These results would provide a useful basis for further research on the regulation of flower and leaf color in *O. hybridum*.

CONCLUSIONS

From the results of this study, UGT75C1, E2.4.1.115, CCD7, E2.1.1.76, and CCoAOMT are key genes that regulate secondary metabolite mod-

ifying enzymes in the carotenoid and anthocyanin formation pathways, which are responsible for methylation and glycosylation of carotenoids, flavonoids, and anthocyanins, resulting in differences in flower color phenotypes; LHCB, UGT, RP-L18Ae, ABCB1, and UGT75C1 are key genes involved in the regulation of chlorophyll and anthocyanin anabolism in leaves. UGT, RP-L18Ae, ABCB1, and UGT75C1 are involved in the regulation of chlorophyll and anthocyanin anabolism in leaves. Among them, LHCB, UGT, and UGT75C1 affect the enzyme activities in the chloroplast metabolic pathway and the glycosylation of metabolites in the anthocyanin synthesis pathway, which have an impact on the synthesis of chlorophyll and anthocyanin in leaves. The differential expression of ribosomal subunit genes, as well as protein transporter genes, has been shown to contribute to the alteration of leaf color. It is noteworthy that UGT75C1 was found in both groups of samples, and its glycosylation modification of anthocyanins led to an effect on the products of the anthocyanin synthesis pathway, affecting the content and enzymatic activity of anthocyanins in petals and leaves, and thus contributing to the formation of flower and leaf color differences in *O. hybridum*.

In addition, we identified differential expression of MYB family transcription factor genes in both groups of samples, and it was demonstrated that MYB transcription factors play an essential role in the synthesis of secondary metabolites in plants, mainly participating in the phenylpropanoid metabolic pathway and the secondary product synthesis pathway, and the synthesis of anthocyanins and flavonoids is closely related to MYB transcription factors. It was shown that MYB transcription factors affect the metabolites in the anthocyanin synthesis pathway, resulting in differences in flower color and leaf color phenotypes.

SOURCE OF FUNDING

This work was supported by grants from Hainan Provincial Natural Science Foundation of China (320RC500 and ZDKJ201815), the National Natural Science Foundation of China (grant number 32060149 and 31760316), Hainan University 2020 National and Provincial College Student Innovation and Entrepreneurship Training Program (SA2100001230), Priming Scientific Research Foundation of Hainan University (grant number KYQD (ZR) 1721).

SUPPLEMENTARY MATERIAL

The online version contains supplementary material (Additional file 1 and 2) available at https://doi.org/10.24326/asphc.2023.4831.

REFERENCES

Adjei, M.O., Luo, J., Li, X., Du, J., Luan, A., Li, S., Ma, J. (2023). Function of ALA content in porphyrin metabolism regulation of *Ananas comosus* var. *bracteatus*. Int. J. Mol. Sci., 24(6), 5274. https://doi .org/10.3390/ijms24065274

Andrews, S. (2014). FastQC a quality control tool for high throughput sequence data. Babraham Bioinforma. Available: https://www.bioinformatics.babraham.ac.uk/projects/fastqc [date of access: 6.07.2022].

Cai, W., Zhang, D., Zhang, X., Chen, Q., Liu, Y., Lin, L., Xiang, L., Yang, Y., Xu, L., Yu, X., Li, Y. (2023). Leaf color change and photosystem function evaluation under heat treatment revealed the stress resistance variation between *Loropetalum chinense* and *L. chinense* var. *rubrum*. Peer J., 11, e14834. https://doi.org/10.7717/peerj.14834

Chang, Y.L., Huang, L.M., Kuo, X.Z., Chen, Y.Y., Lin, S.T., Jeng, M.F., Yeh, H.H., Tsai, W.C., Chen, H.H. (2023). PbABCG1 and PbABCG2 transporters are required for the emission of floral monoterpenes in *Phalaenopsis bellina*. Plant J., 114(2), 279–292. https://doi.org/10.1111/tpj.16133

Chen, X., Mao, X., Huang, J., Yang, D., Wu, J., Dong, S., Lei, K., Ge, G., Li, C.Y., Wei, L. (2011). KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. Nucleic Acids Res., 39 (S2), W316–W322. https://doi.org/10.1093/nar/gkr483

Chiou, C.Y., Yeh, K.W. (2008). Differential expression of *MYB* gene (*OgMYB1*) determines color patterning in floral tissue of Oncidium Gower Ramsey. Plant Molecular Biol., 66(4), 379–388. https://doi.org/10.1007/s11103-007-9275-3

Fitter, D.W., Martin, D.J., Copley, M.J., Scotland, R.W., Langdale, J.A. (2010). GLK gene pairs regulate chlo-

roplast development in diverse plant species. Plant J., 31(6), 713–727. https://doi.org/10.1046/j.1365-313X.2002.01390.x

Galperin, M.Y., Wolf, Y. I., Makarova, K.S., Vera Alvarez, R., Landsman, D., Koonin, E.V. (2020). Cog database update: focus on microbial diversity, model organisms, and widespread pathogens. Nucleic Acids Res., 49(D1), D274–D281. https://doi.org/10.1093/nar/gkaa1018

Giordano D., Provenzano S., Ferrandino A., Vitali M., Pagliarani C., Roman F., Cardinale F., Castellarin S.D., Schubert A. (2016). Characterization of a multifunctional caffeoyl-CoA O-methyltransferase activated in grape berries upon drought stress. Plant Physiol. Bioch., 101, 23–32. https://doi.org/10.1016/j.plaphy.2016.01.015

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Amit, I. (2013). Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. Nat. Biotechnol., 29, 644–652. https://doi.org/10.1038/nbt.1883

Hieber, A.D., Mudalige-Jayawickrama, R.G., Kuehnle, A.R. (2006). Color genes in the orchid *Oncidium* Gower Ramsey: identification, expression, and potential genetic instability in an interspecific cross. Planta, 223(3), 521–531. https://doi.org/10.1007/s00425-005-0113-z

Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernandez-Plaza, A., Forslund, S.K., Cook, H., Mende, D.R., Letunic, I., Rattei, T., Jensen, L.J., von Mering, C., Bork, P. (2019). eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. Nucleic Acids Res., 47(D1), D309–D314. https://doi.org/10.1093/nar/gky1085

Kim, E.H., Li, X.P., Razeghifard, R., Anderson, J.M., Niyogi, K.K., Pogson, B.J., Chow, W.S. (2009). The multiple roles of light-harvesting chlorophyll *a/b*-protein complexes define structure and optimize function of *Arabidopsis* chloroplasts: a study using two chlorophyll *b*-less mutants. Biochim Biophys Acta., 1787(8), 973–984. https://doi.org/10.1016/j.bbabio.2009.04.009

Liao, L., Li, Y., Lan, X., Yang, Y., Wei, W., Ai, J., Feng, X., Chen, H., Tang, Y., Xi, L., Wang, Z. (2023). Integrative analysis of fruit quality and anthocyanin accumulation of plum cv. 'Cuihongli' (*Prunus salicina* Lindl.) and its bud mutation. Plants (Basel), 12(6), 1357. https://doi.org/10.3390/plants12061357

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

Niu, Y., Jiang, X.,Xu, X. (2016). Reaserch advances on transcription factor MYB gene family in plant. Mol. Breed., 14(8), 2050–2059. https://doi.org/10.1080/15592324.2019.1613131

Pan, X., Zheng, H., Zhao, J., Xu, Y.,Li, X. (2016). *ZmCCD7*/*ZpCCD7* encodes a carotenoid cleavage dioxygenase mediating shoot branching. Planta., 243(6), 1407–1418. https://doi.org/10.1007/s00425-016-2479-5

Pan, Y., Bradley, G., Pyke, K., Ball, G., Lu, C., Fray, R., Marshall, A., Jayasuta, S., Baxter, C., Wijk, R.V. (2013). Network inference analysis identifies an *APRR2*-like gene linked to pigment accumulation in tomato and pepper fruits. Plant Physiol., 161(3), 1476–1485. https://doi.org/10.1104/pp.112.212654

Patra, B., Schluttenhofer, C., Wu, Y., Pattanaik, S., Yuan, L. (2013). Transcriptional regulation of secondary metabolite biosynthesis in plants. Biochim. Biophys. Acta Gene Regul. Mech., 1829(11), 1236–1247. https://doi.org/10.1016/j.bbagrm.2013.09.006

Patro, R., Duggal, G., Love, M.I., Irizarry, R.A., Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. Nat. Methods, 14(4), 417–419. https://doi.org/10.1038/nmeth.4197

Rogalski, M., Ruf, S., Bock, R. (2006). Tobacco plastid ribosomal protein S18 is essential for cell survival. Nucleic Acids Res., 34(16), 4537–4545. https://doi.org/10.1093/nar/gkl634

Sui, X., Gao, X., Ao, M., Wang, Q., Yang, D., Wang, M., Fu, Y., Wang, L. (2011). cDNA cloning and characterization of UDP-glucose: anthocyanidin 3-O-glucosyltransferase in *Freesia hybrida*. Plant Cell Rep., 30(7), 1209–1218. https://doi.org/10.1007/s00299-011-1029-7

Wang, Q., Zhu, J., Li, B., Li, S., Yang, Y., Wang, Q., Xu, W., Wang, L. (2023). Functional identification of anthocyanin glucosyltransferase genes: a Ps3GT catalyzes pelargonidin to pelargonidin 3-O-glucoside painting the vivid red flower color of *Paeonia*. Planta., 257(4), 65. https://doi.org/10.1007/s00425-023-04095-2

Yang, L., Gan, C.Y., Wei, M.Q., Zeng, J., Shen, L.J., Su, D.H. (2017). Research advance on cultivation management technology of *Oncidium hybridum*. J. Green Sci. Technol., 2017(17), 114–118. https://doi.org/10.16663/j.cnki.lskj.2017.17.041

Yue, X.Q., Zhang, Y., Yang, C.K., Li, J.G., Rui, X., Ding, F., Hu, F.C., Wang, X.H., Ma, W.Q., Zhou, K.B. (2022). Genome-wide identification and expression analysis of carotenoid cleavage oxygenase genes in Litchi (*Litchi*

chinensis Sonn.). BMC Plant Biol., 22(1), 394. https://doi.org/10.1186/s12870-022-03772-w

Zhang, Q.S. (2022). Analysis of pigment content and photosynthetic characteristics in three leaf color mutants of *Pachira glabra*. J. Minnan Normal Univ. (Nat. Sci.), 35(3), 97–104. https://doi.org/10.16007/j.cnki.issn2095-7122.2022.03.015

Zhang, X., Zhang, L., Zhang, D., Su, D., Li, W., Wang, X., Chen, Q., Cai, W., Xu, L., Cao, F., Zhang, D., Yu, X., Li, Y. (2023). Comprehensive analysis of metabolome and transcriptome reveals the mechanism of color formation in different leave of *Loropetalum Chinense* var. *Rubrum*. BMC Plant Biol., 23(1), 133. https://doi.org/10.1186/s12870-023-04143-9

Zhao, M., Li, X., Zhang, X., Zhang, H.,Zhao, X. (2020). Mutation mechanism of leaf color in plants: a review. Forests, 11(8), 851. https://doi.org/10.3390/f11080851

Zhao, Y., Qi, X., Liu, Z., Zheng, W., Guan, J., Liu, Z., Ren, J., Feng, H., Zhang, Y. (2022). Transcriptome and metabolome profiling to explore the causes of purple leaves formation in non-heading Chinese cabbage (*Brassica rapa* L. ssp. *chinensis* Makino var. *mutliceps* Hort.). Foods, 11(12), 1787. https://doi.org/10.3390/foods11121787

Zhu, G., Yang, F., Shi, S., Li, D., Wang, Z., Liu, H., Huang, D., Wang, C. (2015). Transcriptome characterization of *Cymbidium sinense* 'Dharma' using 454 pyrosequencing and its application in the identification of genes associated with leaf color variation. PLoS One, 10(6), e0128592. https://doi.org/10.1371/journal.pone.0128592