Transcriptomic analysis reveals biosynthetic genes and transcription factors related to leaf anthocyanin biosynthesis in *Aglaonema commutatum*

Ji Li
South China Botanical Garden
Kunlin Wu
South China Botanical Garden
Lin Li
South China Botanical Garden
Guohua Ma
South China Botanical Garden
Lin Fang
South China Botanical Garden
Songjun Zeng (zengsongjun@scib.ac.cn)
South China Botanical Garden

Research Article

Keywords: Aglaonema commutatum 'Red Valentine', Transcriptomics, Anthocyanin biosynthesis, R2R3-MYB, Transcriptional regulation

Posted Date: July 22nd, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1834389/v1

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at BMC Genomics on January 17th, 2023. See the published version at https://doi.org/10.1186/s12864-022-09107-1.

Abstract

Backgrounds: *Aglaonema commutatum* 'Red Valentine', as an excellent foliage ornamental plant, is widely used in interior and exterior decoration because of its easy cultivation and management. However, reducing proportions of red color foliage during large-scale production of *A. commutatum* seedlings is of frequent occurence, which has considerable implications on the ornamental and market value of the plant, and the molecular mechanisms of this phenomenon remain unclear.

Results: To explore the molecular basis leading to the variation in leaf color in *A. commutatum* 'Red Valentine', transcriptome sequencing were performed based on the Illumina platform in two different varieties at three leaf developmental stages. We annotated 63621 unigenes and 14186 differentially expressed genes by pairwise comparison. Further, twenty-six anthocyanin biosynthetic structural genes were identified, the PKM values were significantly higher in 'Red Valentine' than those in the green mutant at the three stages examined, which is consistent with the high anthocyanin content of'Red Valentine' leaves. Positive transcription factors that may involved in regulating anthocyanin biosynthesis were detected by BLAST tool and correlation analysis and a downregulation of these TFs may downregulate the expression of anthocyanin genes. Full-length cDNA of anthocyanin biosynthesis and regulation genes was obtained and phylogenetic trees were performed to ensure the accuracy of the analysis.

Conclusion: Our study provide an insight into the further understanding of leaf variation and facilitate the breeding of ornamental cultivars with high-anthocyanins levels in *A. commutatum* 'Red Valentine'.

Background

The flower, fruit, and leaf colors of agricultural crops are important traits of commercial and ornamental value and are determined by natural plant pigments, including anthocyanins, betalains, chlorophylls, and carotenoids. Among them, anthocyanins are the major pigments that confer red, violet, or blue color [1]. Anthocyanins play vital roles in plant, seed dispersal, and stress responses [2, 3]. As a antioxidant, anthocyanins protect humans against cardiovascular diseases, cancers, and several chronic diseases, and they have been widely used in the food, medicine, and cosmetic industries [4-6].

Studies have shown that anthocyanin biosynthesis pathway is a ubiquitous branch of the flavonoid biosynthesis pathway in plants. Genes regulating anthocyanin biosynthesis and transport have been identified from numerous plant species [7-9]. The anthocyanin biosynthesis pathway starts from the phenylpropanoid pathway, and three major genes: *PAL*, *C4H* and *4CL* are included. The flavonoid biosynthetic pathways two classes of genes: early biosynthetic genes (*EBGs*) and late biosynthetic genes (*LBGs*) [10]. Chalcone synthase (CHS), chalcone isomerase (CHI), flavanone 3-hydroxylase (F3H), flavonoid 3-hydroxylase (F3'H), and flavonol synthase (FLS) participate in early biosynthesis. These genes regulate the biosynthesis of flavonols and other flavonoid compounds, meanwhile the DFR (dihydroflavonol-4-reductase), anthocyanidin synthase (ANS), and UDP-glucose flavonoid 3-O-glucosyltransferase (UFGT) regulate the biosynthesis of anthocyanins. Subsequently, water-soluble anthocyanins are transported into the vacuole for stable storage [8].

Transcriptional regulation is important to regulate the expression profiles of anthocyanin biosynthesis genes in several plants, which is basically regulated by three kinds of transcription factors (TFs), including the R2R3-MYB gene family, bHLH TFs, and WD40 repeat proteins [11]. The EBGs are regulated by R2R3-MYB genes, whereas the

LBGs are activated by a transcriptional activation complex consisting of MYB, bHLH and WD40 proteins (MBW) [10, 12]. In *Arabidopsis thaliana*, PAP1 is a key R2R3-MYB TF that regulates anthocyanins, and its overexpression can significantly induce anthocyanin accumulation, resulting in dark purple leaves [7]. Other R2R3-MYB TFs (PAP2, MYB113, and MYB114), bHLH gene family members (*TT8*, *GL3*, and *EGL3*) and the WD40 gene (*TTG1*), have been identified as key transcription factor involved in anthocyanin biosynthesis [13, 14]. Important transcription factors that regulate anthocyanin structural gene have also been found in some ornamental plants, including in *Gerbera hybrida* (GMYB10 and GMYC1; [15]), *Chrysanthemum morifolium* (CmMYB6 and CmbHLH; [16, 17]), ornamental kale (BoTTG1 and BoTT8; [14, 18]), *lilium sp*. (LhMYB6, LhMYB12, and LhbHLH1; [19, 20]), and *Anthurium andraeanum* (AaMYB2 and AabHLH1; [21, 22]).

Aglaonema is a collective name for a group of perennial flowering plants belonging to the Araceae family. It is native to tropical Asian countries, including India, Thailand, Vietnam, the Philippines, Malaysia, and Indonesia, *Aglaonema commutatum* 'Red Valentine' is widely used in interior decoration and as courtyard ornamental plants, because of its easy cultivation and management. Its leaves have a variety of colors and beautiful stripes, which make them increasingly popular in the market as excellent foliage and fresh-cut ornamental plants [23]. Presently, large-scale production of *A. commutatum* seedlings is mainly performed through tissue culture. However, during the tissue culture process, there are sometimes cases of reduced proportions of red color foliage, and in some cases, these foliage change from red to green. This phenomenon has considerable implications on the ornamental and market value of the plant. In recent years, most studies on *A. commutatum* have been mainly focused on tissue culture [24-27], stress resistance, breeding [27], and interspecies genetic relationships, at the phenotypic and biochemical level. However, relatively fewer studies have examined the mechanism underlying leaf color variations. Moreover, little is known about the mechanisms regulating anthocyanin biosynthesis in *A. commutatum*.

Here, the effect of developmental stages on the leaf color of two varieties of *A. commutatum* ('Red Valentine' and green mutant) was examined at the biochemical and genomic levels. The total anthocyanin and chlorophyll contents and anthocyanin monomer compositions of the leaves of the two varieties were examined at three developmental stages (S1–S3). Additionally, TFs and structural genes regulating anthocyanin biosynthesis were identified using RNA-Seq technology.

Materials

Plant materials

In this study, *A. commutatum* 'Red Valentine' and its green mutant (1-yr tissue culture seedlings, it happens in the process of mature stems culture) were selected as experimental materials(Fig. 1). The plants were grown in a greenhouse at the South China Botanical Garden, Chinese Academy of Sciences, Guangzhou, China. The use of this plant material was licensed and approved by the Guangzhou Flowers Research Centre. The different developmental stages of leaves were defined: Stage 1 (S1), white curly leaf with small amount of pigment (7 d); Stage 2 (S2), uncurled leaves exhibiting light pink coloration (28 d); Stage 3 (S3): the matured leaves, with characteristic dark coloration and oily surface (35 d). All samples were stored at -80 °C for further analysis in May 2019.

Anthocyanin and chlorophyll extraction and observation

Total anthocyanin content in *A. commutatum* was determined using the pH difference method [28] with modifications. Fresh leaf (1 g) of *A. commutatum* was extracted with 15ML methanol (containing 0.05% HCl) in darkness for 12 h at 0 °C, after that the ultrasonic agitation was used for 0.5h to improve the extraction efficiency.. Absorbance at 510 nm and 700 nm was measured in triplicate with a microplate reader (Tecan Infinity, Switzerland). For the measurement of anthocyanin monomers, total anthocyanin extracted from leaf samples was hydrolyzed using 3 mol/L muriatic acid for 1h at 97 °C in darkness, after which the solution was dried using a vacuum thickener (Concentrator plus, Eppendorf) to yield anthocyanin monomer powder. The powder was redissolved in acidic methanol (0.05% HCl) for HPLC analysis, using cyanidin chloride, peonidin chloride, delphinidin chloride, pelargonidin chloride, petunidin chloride, and malvidin chloride (Phyproof, Germany) as standards. The anthocyanin content was quantified by linear regression and expressed in mg/g of fresh weight (FW) of sample. The total chlorophyll content of the leaves was determined following the procedures of Lichtenthaler [29]. Approximately 1 g of leaf sample was extracted with 1 ml of methanol, and the absorbance was measured at 652 nm, 665 nm, and 470 nm with the microplate reader (Tecan Infinity, Switzerland).

Freehand sectionwas was performed to get a anatomical observation of the pigment tissue localization of the two varieties at the three developmental stages using a digital microscope(DVM6; Leica, USA)

RNA extraction , cDNA library construction and sequencing

Total RNA was isolated from the leaves using an RNA kit (Polysaccharides & Polyphenolics-rich) (Hua Yueyang, Beijing, China) with RNase-free DNase I (TaKara) to remove the genomic DNA contamination. The integrity and purity of the total RNA quality were determined using a 2100 Bioanalyzer (Agilent Technologies, Inc., Santa Clara, CA, USA) and quantified using the ND-2000 (NanoDrop Thermo Scientific, Wilmington DE, USA). Only high-quality RNA samples were applied to construct the sequencing library. RNA purification, reverse transcription, library construction, and sequencing were performed by Shanghai Majorbio Bio-pharm Biotechnology Co., Ltd. (Shanghai, China) (Supplementary Fig. S1). Independent triplicate whole-leaf materials, from A. commutatum 'Red Valentine' R and its green mutant G at S1, S2, and S3 were used to construct 18 cDNA libraries, and then named as R1_1, R1_2, R1_3, R2_1, R2_2, R2_3, R3_1, R3_2, R3_3, G1_1, G1_2, G1_3, G2_1, G2_2, G2_3, G3_1, G3_2, and G3_3. Libraries were prepared using Illumina TruSeqTM RNA sample preparation Kit (San Diego, CA) and sequenced on a flow cell using the Illumina HiSeq™ 6000 sequencing platform (Illumina, San Diego, CA).

De nova assembly and annotation

High quality read from the 18 libraries were assembled de novo using Trinity (http://trinityrnaseq.sourceforge.net/). To annotate the unigenes, the transcripts were aligned to public databases, including the NCBI protein non-redundant (Nr) database (http://www.ncbi.nlm.nih.gov), Clusters of Orthologous Groups (COG) database (http://www.ncbi.nlm.nih. gov/COG), and the Kyoto Encyclopedia of Genes and Genomes (KEGG) database using BLASTX. To determine the processes the unigenes are involved in, gene ontology (GO) function annotation was performed using BLAST2GO (http://www.blast2go.com/b2ghome) software. Metabolic pathway analysis was performed using the KEGG (http://www.genome.jp/kegg) database [30].

Identification of differentially expressed genes and functional enrichment analysis

To identify differentially expressed genes (DEGs) between samples from the two varieties, the expression level of each transcript was calculated according to the transcripts per million reads (TPM) method. RSEM (http://deweylab.biostat.wisc.edu/rsem/, [31]) was used to quantify the gene abundance. Differential expression

analysis was performed using the DESeq2 [32] with a Q value ≤ 0.05. Unigenes with fold change > 2 or < -2 and Q value ≤ 0.05 were considered significantly differentially expressed. The differentially expressed genes (DEGs) were used for GO function enrichment analysis and KEGG pathway analysis using Goatools (https://github.com/tanghaibao/ Goatools) and KOBAS (http://kobas.cbi.pku.edu.cn/home.do, [33]).

Quantitative RT-PCR analysis

Total RNA (the same sample using in RNA-Seq) was reverse-transcribed to first-strand cDNA for qRT-PCR using TransScript® One-Step gDNA Removal cDNA Synthesis SuperMix (Transgen, Beijing, China). Diluted cDNA was amplified with specific primers (listed in Additional file) and reacted with PerfectStart Green qPCR SuperMix (Transgen, Beijing, China) on a LightCycler 480 II (Roche) under the following conditions: 45 cycles at 94 °C for 5 s, 57 °C for 15 s, and 72 °C for 10 s. The expression levels of the genes were normalized using *A. commutatum* 'Red Valentine' actin as the internal control and calculated using the $2^{-\Delta\Delta CT}$ method. Three biological replicates were used for the analysis.

Results

Leaf color chages during different leaf development stages

Except for the variation in leaf color of the two varieties of *A. commutatum* examined in this study, there were no observable differences in the leaf shape and plant size of the two varieties (Fig. 1A). There were some changes in the appearance of the leaves during S1 to S2, which included the uncurling of the leaves and leaf color change from white to light pink and green. At maturity, the leaf color changed to deep red and dark green, and the leaf possessed a shiny surface, indicating wax coating. In the mutant, red patches were concentrated in the middle of the leaves (Fig. 1B). To obtain a detailed observation of pigment distribution inside leaf tissues, we performed anatomical observation of the lamina of the leaves of the two varieties at the three developmental stages. The results indicated the presence of a deep red pigmentation in the palisade tissues and the absence of pigments in the sponge tissues and epidermis cells. However, chlorophylls were present in both the sponge and palisade tissues. There were almost no pigments in the cells of the leaves during S1; however, there was a significant increase in pigments in the cells, including anthocyanin and chlorophyll from S2 to S3, which was consistent with the color change of the leaves during the growth stages (Fig. 1C).

Chlorophyll and anthocyanin contents of the leaves

To check whether the color change was determined by the total anthocyanin and chlorophyll content, we measured the leaf content of the two pigments at the three developmental stages. There was a significant increase in both the anthocyanin and chlorophyll contents of the leaves of both varieties of *A. commutatum* from S1 to S2. However, the anthocyanin content of the leaves of 'Red Valentine' was significantly higher than that of the green mutant in the three development stages. Leaves anthocyanin content of 'Red Valentine' was approximately 5.58-fold higher than that of the leaves of the green mutant at S3 (Fig. 2A). In contrast, the chlorophyll content of the leaves of the green mutant was significantly higher than that of 'Red Valentine' in the three developmental stages. The chlorophyll content in the green mutant leaves was approximately 5.57-fold higher than that of the leaves of 'Red Valentine' at S3 (Fig. 2B). Furthermore, we calculated the anthocyanin/chlorophyll ratio of the leaves at the three developmental stages and found that the ratio in the leaves of the mutant was not significantly affected by the developmental stages, with the highest value (0.01) observed at S3. However, there was a significant increase

in anthocyanin/chlorophyll ratio of the leaves of 'Red Valentine' from S2 (0.07) to S3 (0.39), indicating that the mutants did not accumulate anthocyanins as much as the 'Red Valentine' (Fig. 2C).

Furthermore, HPLC was used to determine the composition of anthocyanin monomers in the two varieties at S3. Three anthocyanin monomers, including cyanidin, paeonidin, and pelargonin, were identified in the leaves. Among them, cyanidin was the most abundant, accounting for 98.1% and 95.8% pigments in the leaves of 'Red Valentine' and the mutant, respectively, indicating that cyanidin was responsible for the red color of the leaves. Trace amounts of paeonidin and pelargonin were also detected (Fig. 2D). Our data imply that the composition and proportion of anthocyanin monomers did contribute to leaf color changes when the leaf is mature, because there are no significant difference between the ratio of the two monomers in two varieties.

Library Construction and de novo assembly

In this study, 18 total RNA samples of high quality were extracted from the leaves of the two varieties of *A. commutatum* at three developmental stages for transcriptome analysis using Illumina sequencing with the HiSeq Hiseq 6000 sequencer. The raw sequencing data were checked for quality and subjected to data filtering. The number of clean reads for each library ranged from 22791579–28919060, with a mapped ratio between 66.71% and 76.25%. The GC content ranged from 50.54%–54.95%, and the Q30 base percentage was ≥ 94.71%. A total of 63621 unigenes with N50 length of 1340 bp were generated (Table 1).

Functional annotation and classification

A total of 63621 assembled unigenes were BLAST-searched against the NCBI non-redundant protein (NR), Nt, SwissProt, KEGG, KOG, Pfam and GO databases using BLASTx, and 34554 unigenes were annotated (Fig. 3A). The unigenes were annotated and aligned in the NR database. Quercus suber had the highest matching degree (10.62%), followed by Elaeis guineensis (8.40%), and Phoenix dactylifera (7.57%). Approximately 42.29% of the unigene sequence did not match the protein sequences of other species (Fig. 3B). The result of the GO analysis showed that the genes were mainly involved in metabolic processes, including cell part and compound binding, which were the major metabolic processes. The KEGG analysis showed that the unigenes were involved in translation, carbohydrate metabolism, and folding, sorting, and degradation pathways (Supplementary Fig. S2).

Identification and KEGG classification of DEGs

The TPM values of the unigenes were used to identify DEGs in the leaves of 'Red Valentine' and the green mutant at the three developmental stages using the DEGseq R package. We identified 14186 DEGs (p-adjust≤0.05, FC≥2) by pairwise comparison (G1 vs. G2, G2 vs. G3, R1 vs. R2, and R2 vs. R3). A venn diagram of the DEGs showed that 843 DEGs were common to the four comparison groups (Fig. 4A). Among the four comparisons, the number of DEGs in S2 vs. S3 was higher than that in S1 vs. S2. The changes in gene expression is shown in Figure 4B. Additionally, The KEGG pathway enrichment analysis showed that the DEGs were significantly enriched in glyoxylate and dicarboxylate metabolism, photosynthesis, phenylpropanoid biosynthesis, fatty acid elongation, flavonoid biosynthesis, and amino sugar and nucleotide sugar metabolism (Fig. 4C).

Expression pattern of the genes involved in anthocyanin biosynthesis

The expression pattern of the unigenes was determined using transcript per million reads (PKM) values. We analyzed the expression profiles of genes coding enzymes related to anthocyanin biosynthesis in 'Red Valentine'

leaf and found that the phenylpropanoid pathway was regulated by three group of genes, which are five PAL unigenes, two C4H, and 4CL unigenes. Two CHS, one CHI, and one F3H unigenes were expressed in the early biosynthetic stage. Accordingly three DFR, one ANS, and five UGTs unigenes were expressed during the later biosynthetic stage. Consistent with the high anthocyanin content of 'Red Valentine' leaves, the PKM values of the anthocyanin structural unigenes were higher in the 'Red Valentine' leaves than those in the leaves of the green mutant at the three stages examined. From stage R1 to R2, there was a significant increase (10-fold) in the expression profiles of most anthocyanin related genes (22 of 26). Although the fold change was slightly lesser than that observed from stage R1 to R2, the expression profiles of the genes also exhibited an upward trend from stage R2 to R3. Contrarily, there was a significant decrease in the expression profiles of the genes in leaves of the green mutant leaves, with extremely low PKM values at the three stages examined (Fig. 5). The full-length cDNA of genes was obtained from RNA-Seq data and rapid amplification of cDNA ends (RACE) PCR, and phylogenetic trees were performed to ensure the accuracy of the analysis.

Identification of transcription factors (TFs) related to anthocyanin biosynthesis

It's reported that anthocyanin biosynthetic genes are mainly regulated by three different transcription factor families which are the MYB, bHLH, and WDR families. Hence, we examined the TFs regulating anthocyanin biosynthesis in the leaves of *A. commutatum*. We identified 1068 TFs belonging to 32 TF families by aligning to the PlantTFDB database using BLASTX (Fig 6). The top 10 annotated TFs were MYB_superfamily, AP2/ERF, C2C2, bZIP, bHLH, C3H, B3_superfamily, NAC, WRKY, LBD (AS2/LOB), each of which comprises over 48 unigenes in both varieties of *A. commutatum* 'Red Valentine'. To identify MYB and WDR TFs related to anthocyanin biosynthesis, we performed a functional search in the RNA-seq annotation file and obtained the following unigenes information: TRINITY_DN91028_c1_g1 (Unigene ID) was annotated to be a putative flavonoid or anthocyanin regulator (*Anthurium andraeanum*). This gene *AaMYB1* in combination with bHLH genes has been reported to be associated with anthocyanin biosynthesis through biolistic transformation in *Cymbidium* orchid [34]. TRINITY_DN62889_c2_g2 was annotated to have the highest homology with AaMYB2 (*Anthurium andraeanum*), which was proven to primarily contribute to the regulation of anthocyanin biosynthesis in the spathes and leaves of *A. andraeanum* [22]. The results of qRT-PCR confirmed the profile of these two unigenes in relation to anthocyanin biosynthesis in *A. commutatum*. TRINITY_DN84708_c1_g1 had a high similarity with TTG1, which was involved in regulating anthocyanin biosynthesis in many species according to report [35, 36].

Regarding bHLH TFs, we were unable to identify bHLH unigenes involved in anthocyanin biosynthesis in the two varieties of *A. commutatum* based on the RNA-seq annotation data. Therefore, we searched based on the expression profiles of anthocyanin biosynthesis genes. Overall, 'Red Valentine' 26 putative anthocyanin biosynthesis genes (Date A) (Supplementary Table S1) and all bHLH TFs (Date B) were first normalized to reduce the error caused by exact PKM means. A correlation analysis was performed between the mean value of Date A (which could represent the anthocyanin biosynthesis gene expression level as a whole) and Date B to determine the relationship between each bHLH TF and the expression profiles of the anthocyanin biosynthesis genes, and they were ranked in descending order (Table 2). The top 10 bHLH unigenes were aligned against NCBI database (https://blast.ncbi.nlm.nih.gov/Blast.cgi) using BLASTX to detailed information of the unigenes. We discovered that TRINITY_DN107490_c2_g1, RINITY_DN8 7484_c1_g1, and TRINITY_DN97087_c0_g10 had the highest correlation coefficients with the basic helix-loop-helix transcription factor 1 (*Anthurium andraeanum*) annotation. *AabHLH1* was reported to participate in the regulation of flavonoid metabolic pathway [21], suggesting that it may also be involved in regulating anthocyanin biosynthesis in 'Red Valentine'. We cloned this gene using RACE PCR

and found that the three unigenes belong to one gene, named *AcbHLH1*. Gene expression patterns of *AcbHLH1* were examined by qRT-PCR, which was consistent with the anthocyanin structural gene in the leaves of the two varieties of *A. commutatum* in this study. The CDS sequence of these TFs was cloned by sequence information and RACE PCR (Supplementary Table S2) for phylogenetic tree analysis (Supplementary Fig. S3) and further studies, name as *AcMYB1*, *AcMYB2*, *AcTTG1* and *AcbHLH1*, respectively .

qRT-PCR analysis of the anthocyanin related DEGs

To validate the RNA-seq data, qRT-PCR assays were performed for the DEGs and TFs involved in anthocyanin biosynthesis using gene-specific primers. Totally, 9 unigenes of anthocyanin biosynthesis and identified TFs in the two varieties of *A. commutatum* were selected for qRT-PCR. The results of the qRT-PCR analysis generally showed similar expression patterns with the PKM values, indicating that the expression data obtained by RNA-Seq were reliable (Fig. 7; Supplementary Table S3). Melting curves of selected gene was in Supplementary Fig. S4

Discussion

Transcriptome analysis of *A. commutatum*' Red Valentine' at the different leaves developmental stages

Several studies have been conducted on 'Red Valentine' [23-27]; however, these studies were mostly phenotypic and biochemical, and none were on molecular biology research. The red color of 'Red Valentine' leaves is of considerable ornamental value and a trait of interest by breeders. In the present study, we performed RNA-seq analysis of the leaves of the 'Red Valentine' and green mutant of *A. commutatum* at three developmental stages to identify key genes involved in leaf color variations and their expression patterns.

Variations in leaf anthocyanin and chlorophyll contents

Leaf color changes in ornamental plants are associated with the biosynthesis of pigments. Anthocyanin is a major pigment in plants and confers red, blue, and purple leaf colors, and chlorophyll is responsible for the green leaf colouration [1, 37, 38]. In the 'Red Valentine' variety of *A. commutatum*, anthocyanins accumulated mainly in the palisade tissues, while chlorophylls were present in both sponge and palisade tissues. However, there was a significant increase in both the anthocyanin and chlorophyll content of the leaves of the two varieties as they matured (S1 to S3). At S3, the anthocyanin content of the leaves of 'Red Valentine' was approximately 5.58-fold higher than that of the leaves of the green mutant. Contrarily, the chlorophyll content of the leaves of the mutant was approximately 5.57-fold higher than that of the leaves of 'Red Valentine' at S3, which was in accordance with the leaf color phenotype of the two varieties. Although cyanidin was the major anthocyanin in the leaves of the two varieties, the composition and proportion of anthocyanin monomers did not significantly affect the leaf color of the plants. In the present study, we first examined the phenotypic and biochemical basis for leaf color changes during three developmental stages, which was vital for further molecular studies.

Genes involved in anthocyanin and chlorophyll biosynthesis and degradation

RNA-Seq technology is a powerful tool to revealing the relationship between different biological processes and gene expression, as well as for identifying key genes under various conditions. Transcription sequencing technology has been widely applied in several studies on plants, especially in plants lacking genomic information [39]. In the present study, we identified 283130 unigenes in the leaves of *A. commutatum*. However, after filtering out low-expression unigenes, we identified 26 differentially expressed putative anthocyanin structural unnigenes,

including three phenylpropanoid pathway genes (*PAL*, *C4H*, and *4CL*), three early anthocyanin biosynthetic genes (*CHS*, *CHI*, and *F3H*), and three later biosynthetic genes (*DFR*, *ANS*, *UFGT*). The full-length cDNA of the genes were obtained and phylogenetic trees were constructed to get more functional information for future study. In accordance with the phenotype and anthocyanin content, the PKM values of most of the anthocyanin genes were significantly higher in the leaves of the 'Red Valentine' variety, compared with those of the leaves of the green mutant at the three developmental stages. Particularly, the PKM values of *CHS* (TRINITY_DN78614_c3_g2), *CHI* (TRINITY_ DN98854_c4_g1), *DFR* (TRINITY_DN75107_c2_g1, TRINITY_DN99635_c0_g4), and *UFGT* (TRINITY_ DN91923_c1_g1, TRINITY_DN90166_c0_g1) in 'Red Valentine' leaves were more than 5-fold higher than that in the leaves of the green mutant at S3.

Furthermore, we identified 31 unigenes involved in chlorophyll biosynthesis and 15 unigenes involved in chlorophyll degradation in the leaves of *A. commutatum*. The PKM values of the genes was provided in Supplementary Table S4 in order to elucidate the gene expression pattern during the three development stages. Among the unigenes, there was a 2-fold higher in the expression of 10 chlorophyll biosynthetic unigenes in the leaves of 'Red Valentine' compared with that in the leaves of the mutant at the same developmental stage, indicating that these genes play vital roles in chlorophyll biosynthesis and accumulation. However, there were no significant differences in the expression profiles of the other chlorophyll biosynthetic and degradation unigenes in the two varieties. Further research should be conducted to elucidate how chlorophyll biosynthetic and degradation genes influence leaf color changes.

TFs related to anthocyanin biosynthesis

Previous studies have revealed that anthocyanin biosynthesis is transcriptionally modulated by TFs, including R2R3-MYB, bHLH, and WD40. Among them, MYB TFs can singly or with MBW (MYB, bHLH, and WD40) complexes to control anthocyanin synthesis and accumulation [10, 40]. In the present study, two MYB TFs, AcMYB1 and AcMYB2, were identified as candidate anthocyanin biosynthesis regulators. AcMYB1 and AcMYB2 are homologs of AaMYB1 and AaMYB2 in *Anthurium andraeanum*, which were reported to be involved in anthocyanin regulation. Phylogenetic analysis showed that the two TFs belonged to the same anthocyanin subgroup. An overexpression of AcMYB2 resulted in red leaf phenotype of transgenic plants, indicating its anthocyanin regulatory activity in *A. commutatum* 'Red Valentine' (Supplementary Figure S5). Transposon insertion and methylation of the promoter on R2R3-MYB TFs regulating anthocyanin biosynthesis have been reported in *Phalaenopsis* orchids [41], apple [42, 43], and pear [44] , which would directly control the MYB TF expression, consequently affecting anthocyanin accumulation. The candidate unigenes exhibited significantly higher expression profiles in the 'Red Valentine' variety than in the green mutant. Therefore, we speculated that the transposon insertion or methylation of the AcMYB2 promoter may occur during leaf development, consequently causing lower or no expression of AcMYB2 and leading to the green mutants. Further studies on how these key TFs regulate leaf coloring should be conducted.

Conclusion

Overall, the present study showed that high chlorophyll and low anthocyanin contents were responsible for the green color of the leaves of *A. commutatum* green mutant. Furthermore, twenty-six anthocyanin biosynthetic structural genes and four key regulatory transcription factors: AcMYB1, AcMYB2, AsTTG1, and AcbHLH1 were identified. A downregulation of these TFs may downregulate the expression of anthocyanin biosynthetic genes. Further molecular studies are needed for a detailed understanding of the role of each TF and

structural gene in anthocyanin biosynthesis in ‘Red Valentine’. The findings of the study provide an insight into anthocyanin candidate genes that can facilitate the breeding and development of ornamental cultivars, and contributes to the understanding of leaf color variation in *A. commutatum*.

Declarations

Ethics approval and consent to participate

This manuscript does not involve animal or human testing.

Consent for publication

Not Applicable.

Availability of data and materials

All data generated or analyzed during this study are included in this published article and the supplementary information files. The sequence data was deposited in the NCBI database under the BioProject PRJNA793608.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was funded by the Guangdong Key Technology Research and Development Program (2018B020202003). the Guangdong Modern Agricultural Industry Technology System Program (2018LM2176, 2021KJ121).

Authors’ contributions

LJ, LF, and SZ conceived the research project and designed research. LJ, LF performed research, LJ, LF LL, GM analyzed the data. LJ, LF and SZ wrote the manuscript.

Acknowledgements

We would like to express our gratitude to all of our colleagues in our laboratory for their insightful conversations and technical assistance.

References

1. Tanaka Y, Ohmiya A: **Seeing is believing: engineering anthocyanin and carotenoid biosynthetic pathways**. *Curr Opin Biotech* 2008, **19**(2):190-197.
2. Huang XW, Zou XB, Zhao JW, Shi JY, Zhang XL, Holmes M: **Measurement of total anthocyanins content in flowering tea using near infrared spectroscopy combined with ant colony optimization models**. *Food Chem* 2014, **164**:536-543.
3. Tanaka Y, Sasaki N, Ohmiya A: **Biosynthesis of plant pigments: anthocyanins, betalains and carotenoids**. *Plant J* 2008, **54**(4):733-749.

4. Ha US, Bae WJ, Kim SJ, Yoon BI, Hong SH, Lee JY, Hwang TK, Hwang SY, Wang ZP, Kim SW: **Anthocyanin Induces Apoptosis of DU-145 Cells In Vitro and Inhibits Xenograft Growth of Prostate Cancer**. *Yonsei Med J* 2015, **56**(1):16-23.
5. Hou DX: **Potential mechanisms of cancer chemoprevention by anthocyanins**. *Curr Mol Med* 2003, **3**(2):149-159.
6. Martin C, Butelli E, Petroni K, Tonelli C: **How Can Research on Plants Contribute to Promoting Human Health?** *Plant Cell* 2011, **23**(5):1685-1699.
7. Borevitz JO, Xia YJ, Blount J, Dixon RA, Lamb C: **Activation tagging identifies a conserved MYB regulator of phenylpropanoid biosynthesis**. *Plant Cell* 2000, **12**(12):2383-2393.
8. Zhao J, Dixon RA: **The 'ins' and 'outs' of flavonoid transport**. *Trends Plant Sci* 2010, **15**(2):72-80.
9. Zhao L, Gao LP, Wang HX, Chen XT, Wang YS, Yang H, Wei CL, Wan XC, Xia T: **The R2R3-MYB, bHLH, WD40, and related transcription factors in flavonoid biosynthesis**. *Funct Integr Genomic* 2013, **13**(1):75-98.
10. Petroni K, Tonelli C: **Recent advances on the regulation of anthocyanin synthesis in reproductive organs**. *Plant Sci* 2011, **181**(3):219-229.
11. Ramsay NA, Glover BJ: **MYB-bHLH-WD40 protein complex and the evolution of cellular diversity**. *Trends Plant Sci* 2005, **10**(2):63-70.
12. Xu WJ, Grain D, Le Gourrierec J, Harscoet E, Berger A, Jauvion V, Scagnelli A, Berger N, Bidzinski P, Kelemen Z *et al*: **Regulation of flavonoid biosynthesis involves an unexpected complex transcriptional regulation of TT8 expression, in Arabidopsis**. *New Phytol* 2013, **198**(1):59-70.
13. Payne CT, Zhang F, Lloyd AM: **GL3 encodes a bHLH protein that regulates trichome development in arabidopsis through interaction with GL1 and TTG1**. *Genetics* 2000, **156**(3):1349-1362.
14. Zhang B, Hu ZL, Zhang YJ, Li YL, Zhou S, Chen GP: **A putative functional MYB transcription factor induced by low temperature regulates anthocyanin biosynthesis in purple kale (Brassica Oleracea var. acephala f. tricolor)**. *Plant Cell Rep* 2012, **31**(2):281-289.
15. Elomaa P, Uimari A, Mehto M, Albert VA, Laitinen RAE, Teeri TH: **Activation of anthocyanin biosynthesis in Gerbera hybrida (Asteraceae) suggests conserved protein-protein and protein-promoter interactions between the anciently diverged monocots and eudicots**. *Plant Physiol* 2003, **133**(4):1831-1842.
16. Liu XF, Xiang LL, Yin XR, Grierson D, Li F, Chen KS: **The identification of a MYB transcription factor controlling anthocyanin biosynthesis regulation in Chrysanthemum flowers**. *Sci Hortic-Amsterdam* 2015, **194**:278-285.
17. Xiang LL, Liu XF, Li X, Yin XR, Grierson D, Li F, Chen KS: **A Novel bHLH Transcription Factor Involved in Regulating Anthocyanin Biosynthesis in Chrysanthemums (Chrysanthemum morifolium Ramat.)**. *Plos One* 2015, **10**(11).
18. Guo N, Han S, Zong M, Wang GX, Zheng SN, Liu F: **Identification and differential expression analysis of anthocyanin biosynthetic genes in leaf color variants of ornamental kale**. *Bmc Genomics* 2019, **20**.
19. Nakatsuka A, Yamagishi M, Nakano M, Tasaki K, Kobayashi N: **Light-induced expression of basic helix-loop-helix genes involved in anthocyanin biosynthesis in flowers and leaves of Asiatic hybrid lily**. *Sci Hortic-Amsterdam* 2009, **121**(1):84-91.
20. Yamagishi M, Shimoyamada Y, Nakatsuka T, Masuda K: **Two R2R3-MYB Genes, Homologs of Petunia AN2, Regulate Anthocyanin Biosyntheses in Flower Tepals, Tepal Spots and Leaves of Asiatic Hybrid Lily**. *Plant Cell Physiol* 2010, **51**(3):463-474.

21. Li CH, Qiu J, Huang SR, Yin JM, Yang GS: **AaMYB3 interacts with AabHLH1 to regulate proanthocyanidin accumulation in Anthurium andraeanum (Hort.)-another strategy to modulate pigmentation**. *Hortic Res-England* 2019, **6**.
22. Li CH, Qiu J, Yang GS, Huang SR, Yin JM: **Isolation and characterization of a R2R3-MYB transcription factor gene related to anthocyanin biosynthesis in the spathes of Anthurium andraeanum (Hort.)**. *Plant Cell Rep* 2016, **35**(10):2151-2165.
23. Bao-Gui DU, Liang CH, Chen YX, Lan TWJgas: **Study on the factors influencing the bud induction of Aglaonemum cultured in vitro**. 2013.
24. Chen J, Henley RW, Henny RJ, Caldwell RD, Robinson CA: **A Simple Leaf-assay Method for Evaluating Aglaonema Sensitivity to Chilling Temperatures**. 1998.
25. Zhang JX, Wu KL, Li Q, Zeng SJ, Duan JJPPC: **In vitro Rapid Propagation of Aglaonema costatum N. E. Br**. 2009.
26. Zhou Z, Zhang G, Liu C, Li Z, University NJG, Biology A: **In Vitro Culture of Aglaonema commutatum Schott 'Red Valentine'**. 2018.
27. Budiarto KJfpuh: **CONVENTIONAL PROPAGATION OF SEVERAL AGLAONEMA ACCESSIONS USING SPLIT SINGLE-BUD STEM CUTTING**. 2011.
28. Wrolstad RE, Culbertson JD, Cornwell CJ, Mattick LR: **Detection of Adulteration in Blackberry Juice Concentrates and Wines**. *J Assoc Off Ana Chem* 1982, **65**(6):1417-1423.
29. Lichtenthaler HKJME: **Chlorophylls and carotenoids.Pigments of photosynthetic membranes**. 1987, **148**.
30. Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M: **KEGG: Kyoto Encyclopedia of Genes and Genomes**. *Nucleic Acids Res* 1999, **27**(1):29-34.
31. Li B, Dewey CN: **RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome**. *Bmc Bioinformatics* 2011, **12**.
32. Love MI, Huber W, Anders S: **Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2**. *Genome Biol* 2014, **15**(12).
33. Xie C, Mao XZ, Huang JJ, Ding Y, Wu JM, Dong S, Kong L, Gao G, Li CY, Wei LP: **KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases**. *Nucleic Acids Res* 2011, **39**:W316-W322.
34. Albert NW, Arathoon S, Collette VE, Schwinn KE, Jameson PE, Lewis DH, Zhang HB, Davies KM: **Activation of anthocyanin synthesis in Cymbidium orchids: variability between known regulators**. *Plant Cell Tiss Org* 2010, **100**(3):355-360.
35. deVetten N, Quattrocchio F, Mol J, Koes R: **The an11 locus controlling flower pigmentation in petunia encodes a novel WD-repeat protein conserved in yeast, plants, and animals**. *Gene Dev* 1997, **11**(11):1422-1434.
36. Walker AR, Davison PA, Bolognesi-Winfield AC, James CM, Srinivasan N, Blundell TL, Esch JJ, Marks MD, Gray JC: **The TRANSPARENT TESTA GLABRA1 locus, which regulates trichome differentiation and anthocyanin biosynthesis in Arabidopsis, encodes a WD40 repeat protein**. *Plant Cell* 1999, **11**(7):1337-1349.
37. Iwashina T: **Contribution to Flower Colors of Flavonoids Including Anthocyanins: A Review**. *Nat Prod Commun* 2015, **10**(3):529-544.
38. Lightbourn GJ, Griesbach RJ, Novotny JA, Clevidence BA, Rao DD, Stommel JR: **Effects of anthocyanin and carotenoid combinations on foliage and immature fruit color of Capsicum annuum L.** *J Hered* 2008, **99**(2):105-111.

39. Vilperte V, Lucaciu CR, Halbwirth H, Boehm R, Rattei T, Debener T: **Hybrid de novo transcriptome assembly of poinsettia (Euphorbia pulcherrima Willd. Ex Klotsch) bracts**. *Bmc Genomics* 2019, **20**(1).
40. Gonzalez A, Zhao M, Leavitt JM, Lloyd AM: **Regulation of the anthocyanin biosynthetic pathway by the TTG1/bHLH/Myb transcriptional complex in Arabidopsis seedlings**. *Plant J* 2008, **53**(5):814-827.
41. Hsu CC, Su CJ, Jeng MF, Chen WH, Chen HH: **A HORT1 Retrotransposon Insertion in the PeMYB11 Promoter Causes Harlequin/Black Flowers in Phalaenopsis Orchids**. *Plant Physiol* 2019, **180**(3):1535-1548.
42. Jiang SH, Wang N, Chen M, Zhang R, Sun QG, Xu HF, Zhang ZY, Wang YC, Sui XQ, Wang SF *et al*: **Methylation of MdMYB1 locus mediated by RdDM pathway regulates anthocyanin biosynthesis in apple**. *Plant Biotechnol J* 2020, **18**(8):1736-1748.
43. Ma CQ, Jing CJ, Chang B, Ya JY, Liang BW, Liu L, Yang YZ, Zhao ZY: **The effect of promoter methylation on MdMYBl expression determines the level of anthocyanin accumulation in skins of two non-red apple cultivars**. *Bmc Plant Biol* 2018, **18**.
44. Wang ZG, Meng D, Wang AD, Li TL, Jiang SL, Cong PH, Li TZ: **The Methylation of the PcMYB10 Promoter Is Associated with Green-Skinned Sport in Max Red Bartlett Pear**. *Plant Physiol* 2013, **162**(2):885-896.

Tables

Table 1 A summary of the sequencing data

Samples	Clean reads	Mapped reads	Mapped ratio %	Q30(≥%)	GC content(%)
G1_A	28919060	21553660	74.53	96.36	53.44
G1_B	26875623	20492131	76.25	96.37	53.33
G1_C	26320908	18925032	71.90	96.12	52.57
G2_A	22817364	16390186	71.83	95.62	53.34
G2_B	27104121	19606183	72.34	96.27	54.33
G2_C	23676524	16923445	71.48	95.86	53.4
G3_A	23320739	16022526	68.71	96.16	53.38
G3_B	25671029	17124158	66.71	95.97	52.65
G3_C	24603689	17458253	70.96	96.35	53.07
R1_A	23431089	17426287	74.37	96.24	54.63
R1_B	27039584	20294292	75.05	96.5	54.95
R1_C	23244555	16316706	70.20	95.67	53.9
R2_A	24510857	17071903	69.65	95.87	54.26
R2_B	24249840	17898965	73.81	96.39	54.05
R2_C	27739382	20209775	72.86	96.31	54.07
R3_A	22791579	15490014	67.96	94.71	50.54
R3_B	23917425	16698185	69.82	96.27	51.86
R3_C	23008761	16564656	71.99	96.49	53.21

Table 2 Correlation coefficient between the mean value of anthocyanin biosynthetic genes expression level and bHLH TFs, arranged in descending order of correlation coefficient

Unigene ID	Correlation coefficent	NR_description from Transcriptomic date	Description from NCBI by blastx	Percent identity of blastx	Protein sequence ID
TRINITY_DN107490_c2_g1	0.9984	basic helix-loop-helix protein A [*Jatropha curcas*]	basic helix-loop-helix transcription factor 1, partial [*Anthurium andraeanum*]	0.7339	AZS49191.1
TRINITY_DN85407_c0_g2	0.9347	hypothetical protein CCACVL1_14859 [*Corchorus capsularis*]	basic helix-loop-helix transcription factor 1, partial [*Anthurium andraeanum*]	0.8793	AZS49191.1
TRINITY_DN87484_c1_g1	0.8675	PREDICTED: transcription factor bHLH47-like [*Phoenix dactylifera*]	transcription factor BHLH062-like [Phoenix dactylifera]	0.6905	XP_008807714.1
TRINITY_DN97087_c0_g10	0.8166	None Pfam description bHLH-MYC and R2R3-MYB transcription factors N-terminal	basic helix-loop-helix transcription factor 1, partial [*Anthurium andraeanum*]	0.8462	AZS49191.1
TRINITY_DN89881_c0_g2	0.8144	PREDICTED: transcription factor HBI1 [*Nelumbo nucifera*]	PREDICTED: transcription factor HBI1 [*Nelumbo nucifera*]	0.4689	XP_010263637.1
TRINITY_DN87339_c0_g1	0.8140	PREDICTED: transcription factor FAMA [*Musa acuminata* subsp. malaccensis]	PREDICTED: transcription factor FAMA [*Musa acuminata* subsp. *malaccensis*]	0.6165	XP_018677740.1
TRINITY_DN67453_c0_g1	0.8058	PREDICTED: transcription factor UNE12-like isoform X1 [*Phoenix dactylifera*]	transcription factor UNE12 isoform X3 [*Elaeis guineensis*]	0.7469	XP_010936103.1
TRINITY_DN93501_c0_g1	0.7778	PREDICTED: transcription factor bHLH78-like isoform X1 [*Phoenix dactylifera*]	transcription factor bHLH62 [*Elaeis guineensis*]	0.4954	XP_029117287.1
TRINITY_DN100086_c1_g1	0.7504	PREDICTED:	transcription	0.6493	XP_008806486.1

		transcription factor bHLH13-like [*Phoenix dactylifera*]	factor bHLH13-like [*Phoenix dactylifera*]		
TRINITY_DN103801_c1_g1	0.7478	PREDICTED: transcription factor bHLH49 isoform X1 [Elaeis guineensis]	transcription factor bHLH49 isoform X1 [*Elaeis guineensis*]	0.5822	XP_010943352.1

Figures

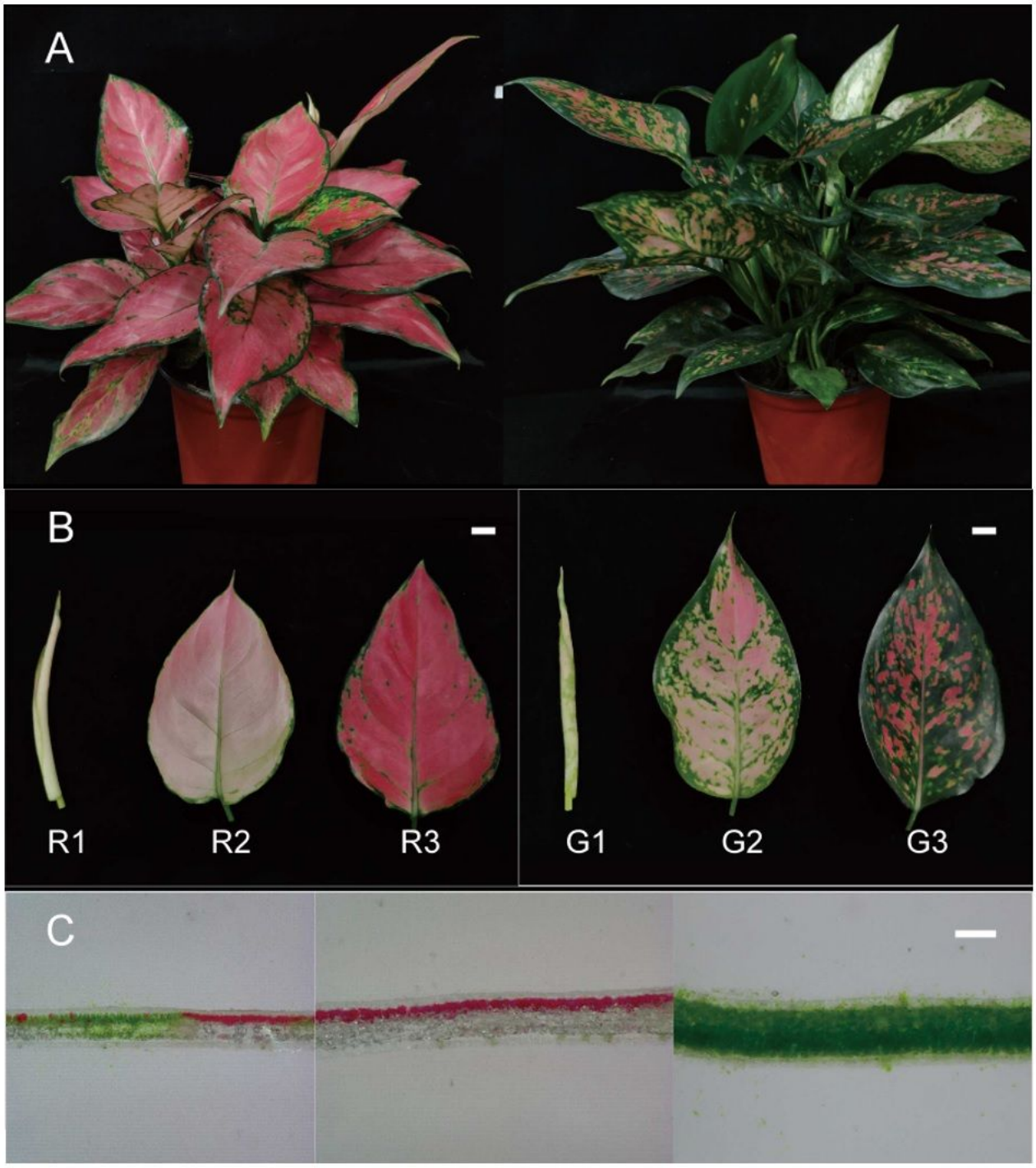

Figure 1

Morphological characteristics of leaves. (A) Pigmentation and phenotypes of *A. commutatum* 'Red Valentine' and the green mutant. (B) leaves phenotypes; Stage 1, white curly leaf with small amount of pigment (7 days); Stage 2, uncurled leaf exhibiting light pink coloration (28 days); Stage 3: The mature stage of the leaves exhibiting dark red coloration, with oily surface (35 days). R1, R2, R3: Stages 1, 2 and 3 of *A. commutatum* 'Red Valentine'; G1, G2, G3: Stage 1, 2 and 3 of the mutant. Scale bars are 1 cm. (C) anatomical observation in the lamina of *A. commutatum* 'Red Valentine' and the green mutant. Scale bars are 100 μm.

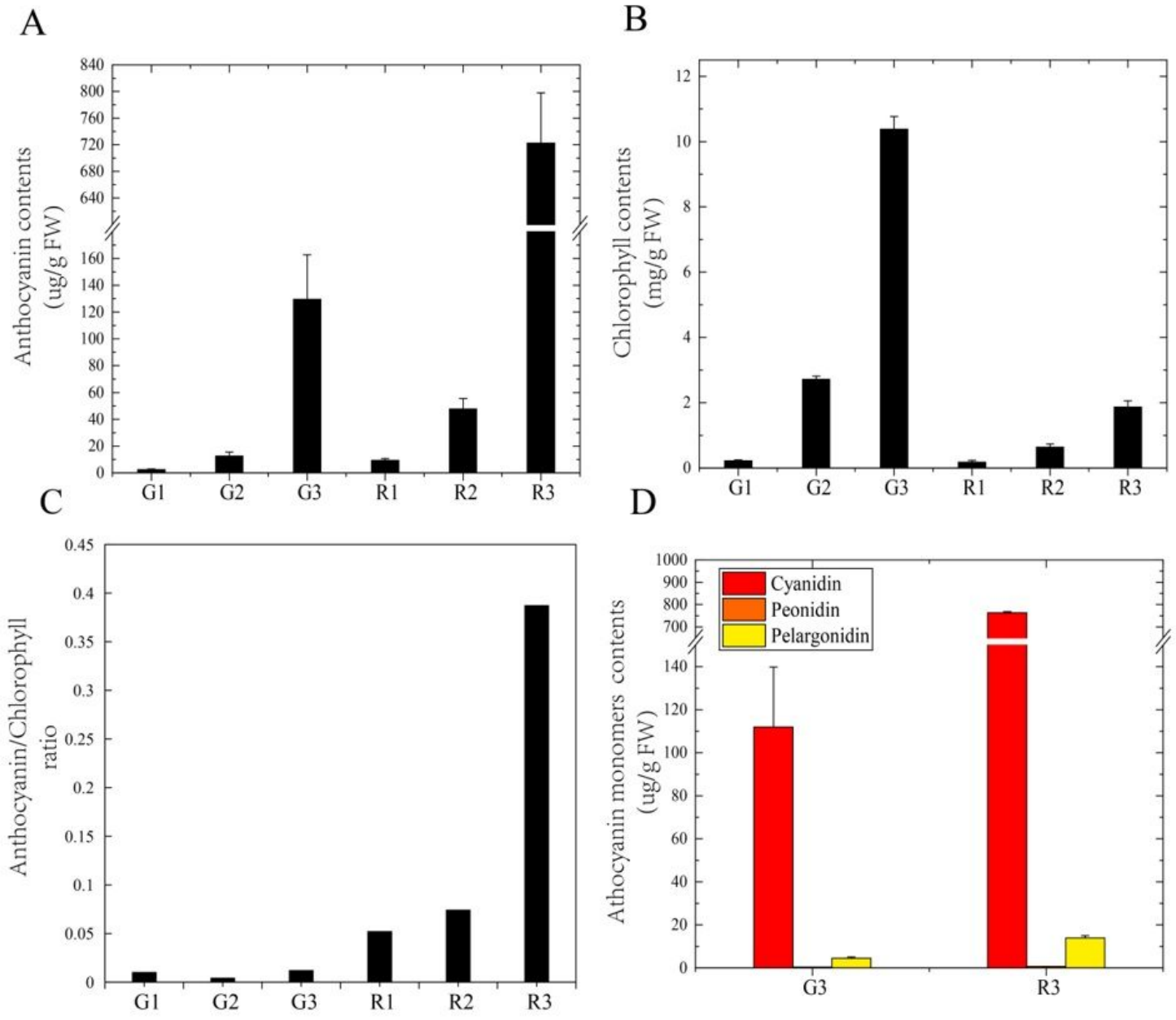

Figure 2

Pigment accumulation in the leaves of *A. commutatum* 'Red Valentine' and its green mutant at three stages. (A) Total anthocyanins,contents. (B) Total Chlorophyll contents. (C) Ratio of anthocyanin/chlorophyll. (D) Anthocyanin

monomers contents.

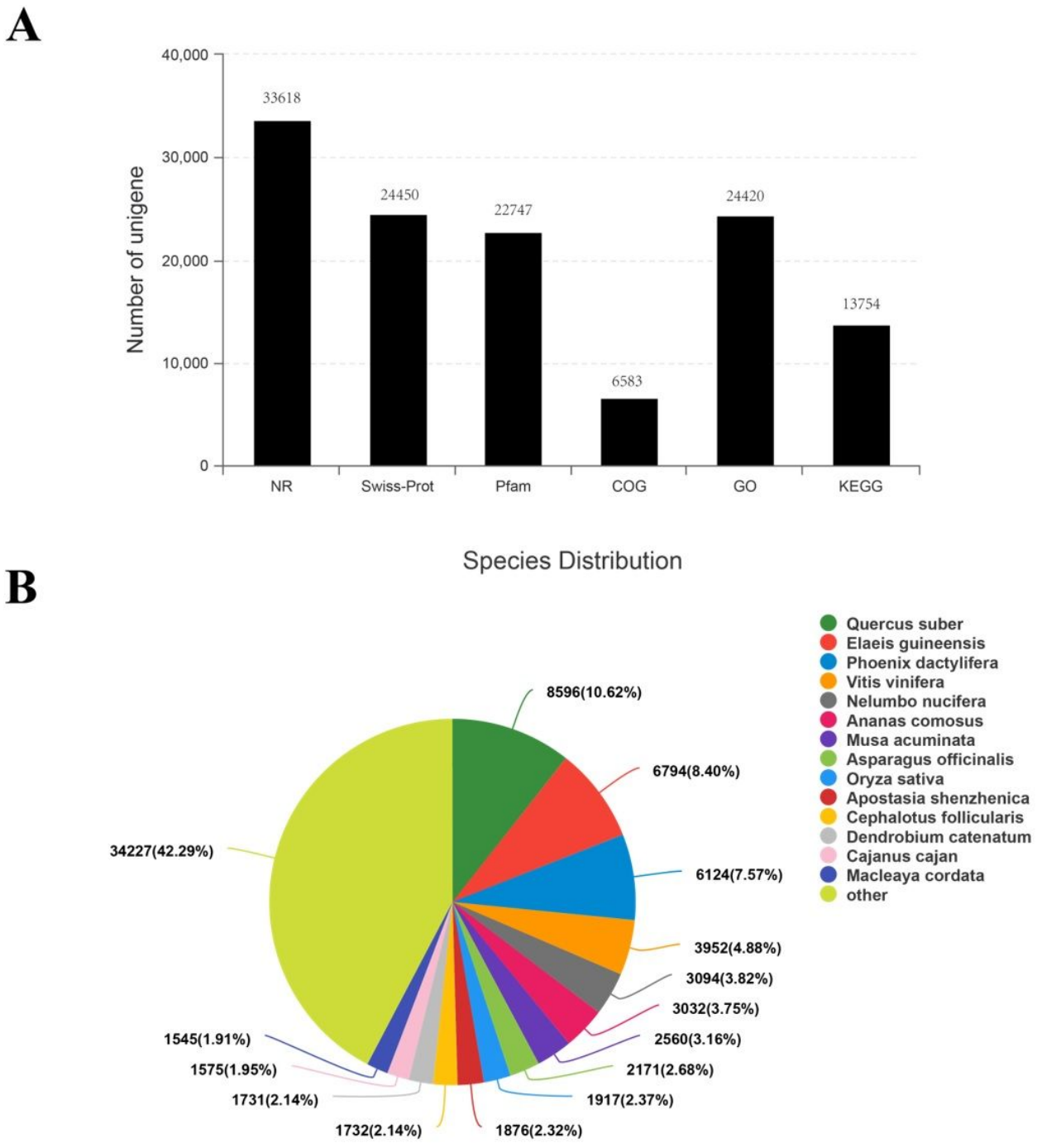

Figure 3

Functional annotation and classification. (A) Gene ontology (GO) classification of unigenes; (B) Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway classification; (C) Numbers of functional annotation of all unigenes in six databases. (D) Species distribution of unigenes matching the main species using BLASTx in the NR database.

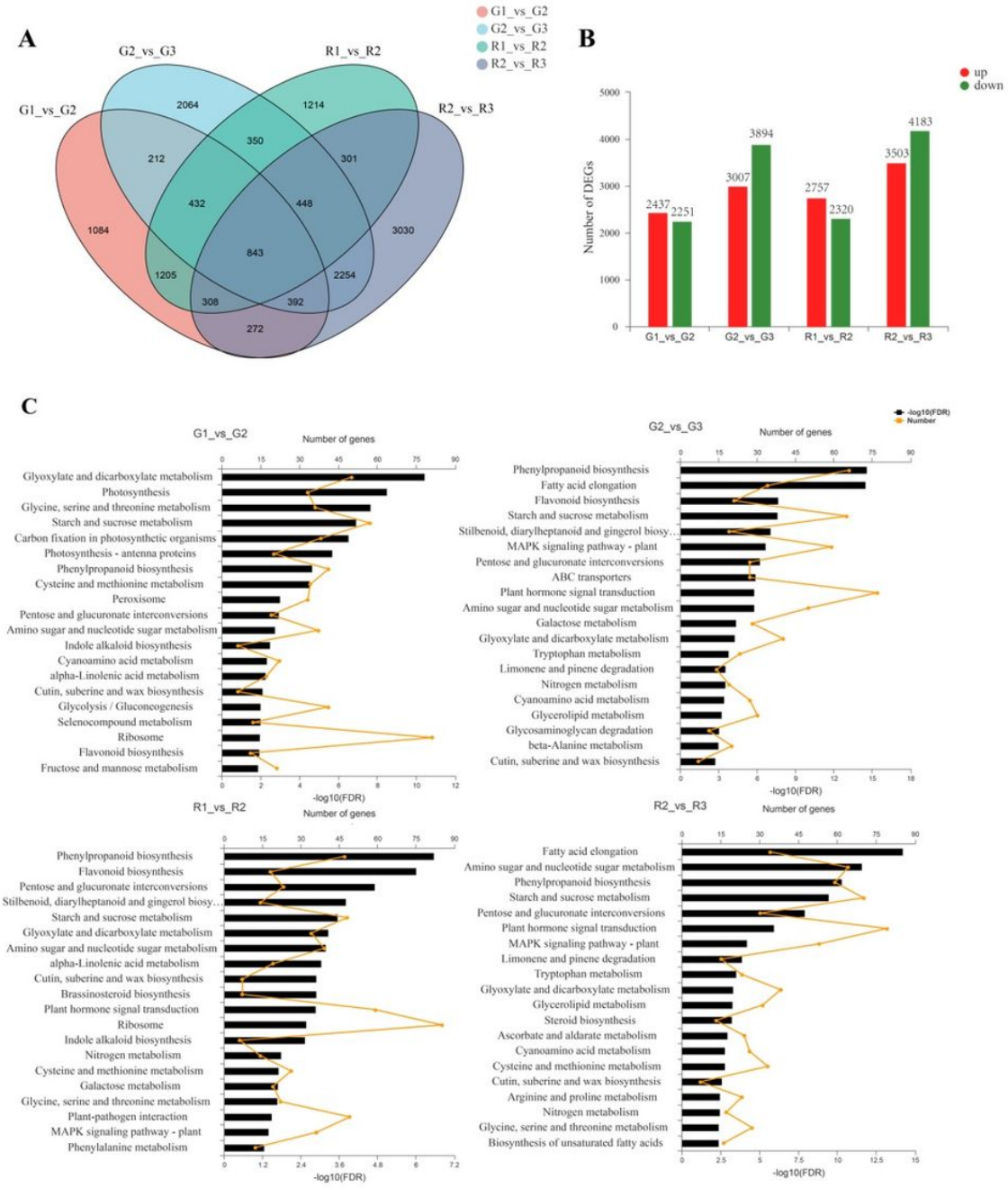

Figure 4

The identification of DEGs in *A. commutatum* 'Red Valentine'. (A) Venn diagrams of DEGs; (B) List of DEGs between pairs of treatments; (C) KEGG pathway enrichment of DEGs in G1 vs. G2, G2 vs. G3, R1 vs. R2, R2 vs. R3.

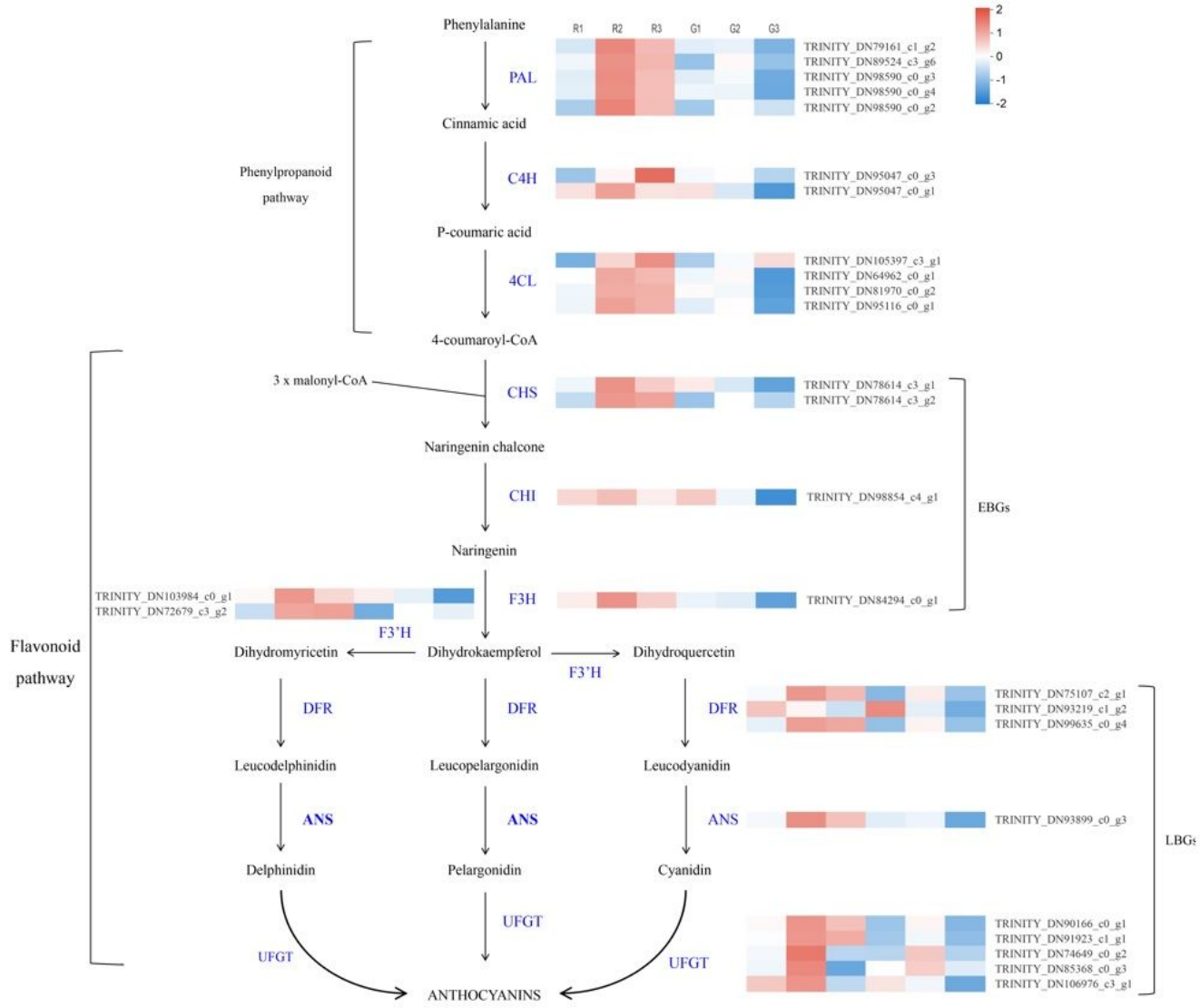

Figure 5

Heatmap of DEGs related to anthocyanin biosynthesis. The pathway can be divided into two sections: the phenylpropanoid and the flavonoid pathways and early biosynthetic genes (EBGs) and late biosynthetic genes (LBGs) are involved in the flavonoid pathway.

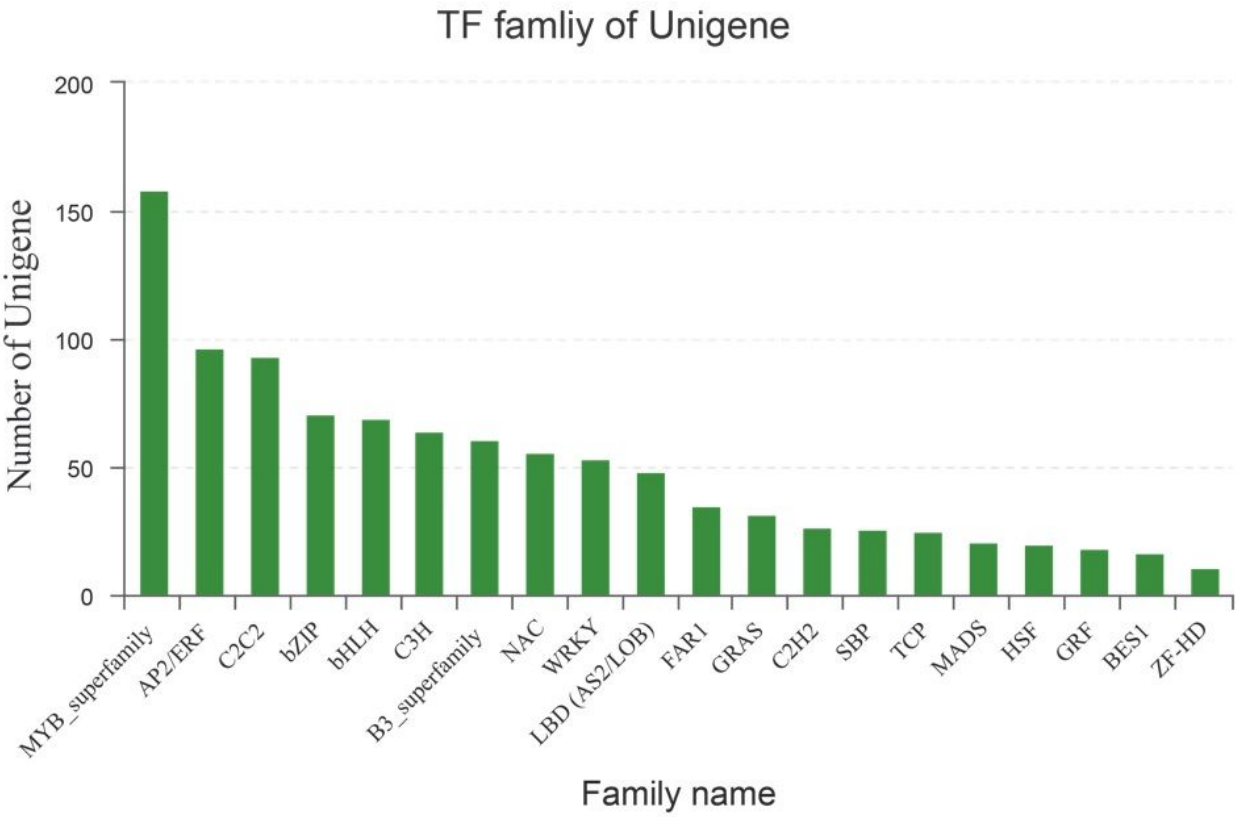

Figure 6

Top 20 transcription factors (TFs) in *A. commutatum* 'Red Valentine' and the green mutant leaves.

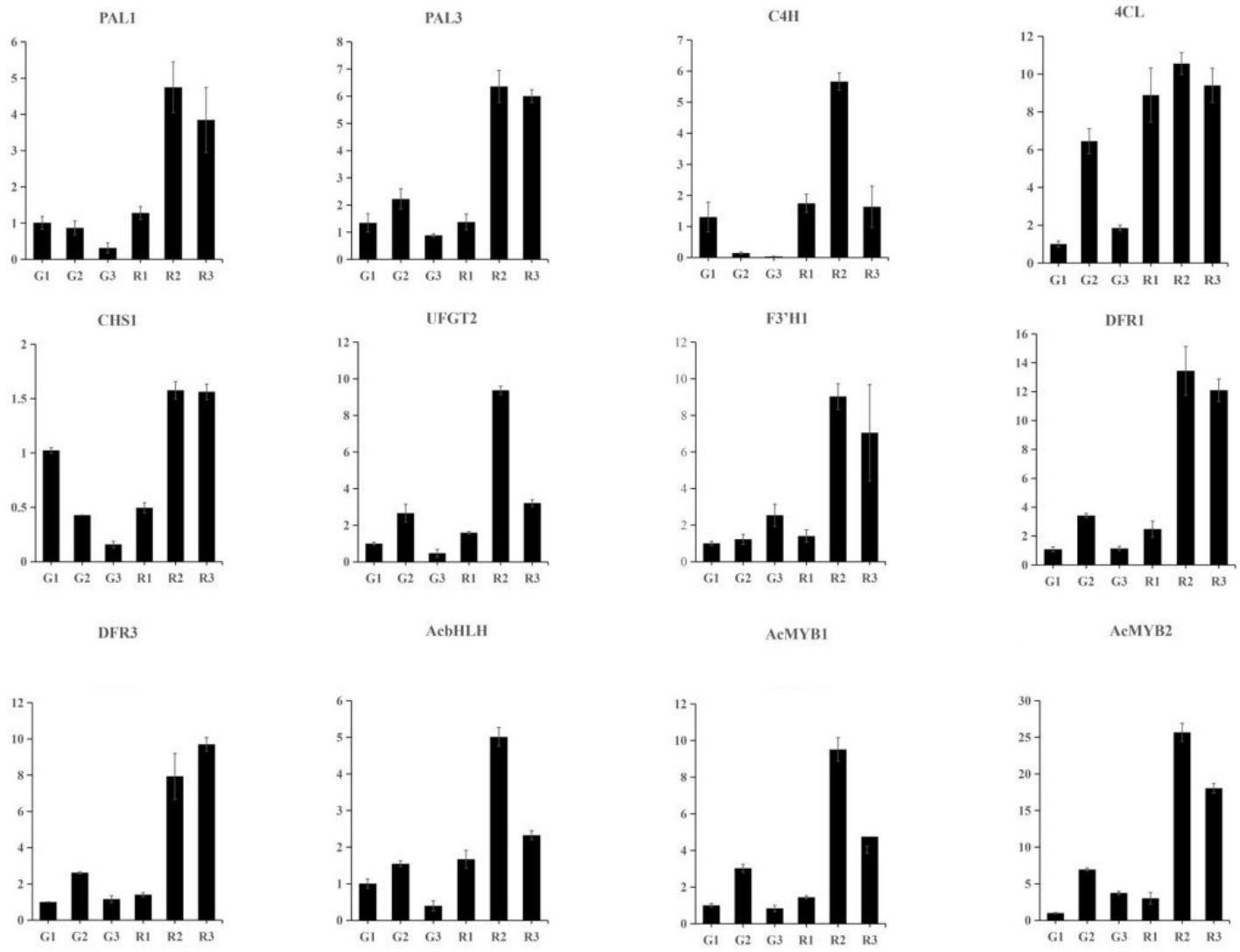

Figure 7

Comparative expression of candidate genes in the 'Red Valentine' variety.

Real-time qPCR of *PAL1* (TRINITY_DN98590_c0_g3); *PAL3* (TRINITY_DN89524_c3_g6); *C4H* (TRINITY_DN95047_c0_g1); *4CL* (TRINITY_DN95116_c0_g1); *CHS1* (TRINITY_DN78614_c3_g1); *UFGT2* (TRINITY_DN90166_c0_g1); *F3'H1* (TRINIY_DN103984_c0_g1); *DFR1* (TRINITY_DN75107_c2_g1); *DFR3* (TRINITY_DN99635_c0_g4); *AcbHLH*; *AcMYB1*; *AcMYB2* in *A. commutatum* 'Red Valentine' leaves. The columns represent means ± standard deviation (SD) from three independent biological replicates

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementarydata.pdf