OPEN ACCESS Freely available online

Journal of Glycobiology

ISSN: 2168-958X

Research Article

Transcriptomic and Proteomic Analyses of Gene Expression Differences and Functional Verification of the Anthocyanidins Synthesis Pathway in *Rhododendron*

Qinghao Wang[1*], Baoxin Jiang[1], Yonghong Jia[1], Huixia Shou[2], Zhihui Chen[3], Yuhao Cheng[1], Yueyan Wu[1]

[1]Department of Biology and Environment, Zhejiang Wanli University, Ningbo, China; [2]Department of Plant Biology, Institute of Plant Biology, Zhejiang University Academy of Sciences, Hangzhou, China; [3]Department of Biology, University of Dundee; Dundee, United Kingdom

ABSTRACT

Rhododendron (*Rhododendron simsii* Planch) is one of the most diverse genera of woody plants in the world. It is rich in germplasm resources, but the molecular regulation of *Rhododendron* color formation is poorly understood. In this study, two Belgian *Rhododendron* varieties with red and white flowers were subjected to RNA-sequencing and Protein sequencing analyses. Integrative analysis of transcriptome and proteome data was used to identify anthocyanidins synthesis genes and proteins specifically expressed in flowers of different colours. The key differences between red and white flowers across flower development stages were analysed using the KEGG (Kyoto Encyclopedia of Genes and Genomes) database, while the expression levels of differentially expressed genes in the anthocyanidins/flavonoid biosynthesis pathways were compared using RNA-seq and qRT-PCR (Real-Time Quantitative Reverse Transcription PCR) data. The key pathways affecting flower colour in *Rhododendron* were identified by correlation analysis of the transcriptome data. We identified 6 anthocyanidins biosynthesis-related genes. The differences in the gene sequences of *RsCHS*, *RsCHI*, *RsF3H*, *RsFLS*, *RsDFR*, and *RsANS* and gene expression might be related to the accumulation of anthocyanidins. *RsDFR* gene overexpression caused *Arabidopsis* leaves to turn red. These results provide valuable information on the molecular mechanism underlying *Rhododendron* flower colour formation.

Keywords: *Rhododendron*; Anthocyanidins; Transcriptome; Proteome; Overexpression

INTRODUCTION

Flower color is one of the most important characteristics of ornamental flowers. Plant breeders have invested great effort in producing varieties with novel flower colours. In recent decades, researchers have been attempting to change the color of flowers through genetic transformation [1].

For example, white flowers can be produced by antisense or RNAi (RNA interference) targeting of structural genes. Antisense expression of the Chalcone Synthase (*CHS*) gene can lead to the cultivation of white *Petunia hybrida*, tobacco and *Chrysanthemum* flowers. Moreover, some gene-deficient flower colour mutants are particularly valuable for producing ideal flower colours in genetic engineering experiments [2]. Anthocyanidins are important metabolites affecting petal color. Due to the presence of anthocyanidins, plants produce flowers of different colours, but these compounds also have pharmacological uses in preventing human disease [3,4]. The general biosynthetic pathway for anthocyanidins has been determined. Although most of the synthetic processes in plants are basically the same, there are some differences in the types of anthocyanidins produced by different plants. Cyanidin, delphinidin and pelargonidin, synthesized in the anthocyanidins synthesis pathway, confer different colours to plants in their natural state. However, not all plants can produce all anthocyanidins. *Chrysanthemum morifolium* and *Rosa rugosa* do not exhibit blue or purple flowers in nature

Correspondence to: Qinghao Wang, Department of Biology and Environment, Zhejiang Wanli University, Ningbo, China, E-mail: 502458114@qq.com

Received: 17-Oct-2023, Manuscript No. JGB-23-27603; **Editor assigned:** 20-Oct-2023, PreQC No. JGB-23-27603 (PQ); **Reviewed:** 03-Nov-2023, QC No. JGB-23-27603; **Revised:** 10-Nov-2023, Manuscript No. JGB-23-27603 (R); **Published:** 17-Nov-2023, DOI: 10.35841/2168-958X.23.12.266.

Citation: Wang Q, Jiang B, Jia Y, Shou H, Chen Z, Cheng Y, et al. (2023) Transcriptomic and Proteomic Analyses of Gene Expression Differences and Functional Verification of the Anthocyanidins Synthesis Pathway in *Rhododendron*. J Glycobiol. 12:266.

because they lack *F3'5'H* in their anthocyanidins synthetic pathways and cannot synthesize delphinidin, possibly having lost the ability to do so during the process of evolution [5,6]. However, some plants of the family Compositae regained *F3'5'H* function from *F3H* during convergent evolution. Transgenic blue or purple carnations, Dianthus caryophyllus, and roses have since been developed through heterologous expression of the *F3'5'H* gene [7,8]. *P. hybrida* lacks the *DFR* gene in its synthetic pathway; *DFR* has strict specificity for its substrate and cannot use DHK. Therefore, the species lacks pelargonidin, which is why there are no orange or brick red *Petunias* in nature. Finally, pink colour is not exhibited by anthocyanidins-deficient *Torenia fournieri* and *Gentiana scabra*.

Some newly discovered structural genes, such as *PlACLB2*, are associated with anthocyanidins. For example, herbaceous peony *PlACLB2* positively regulates red petal formation by promoting anthocyanidins accumulation [9]. In addition to the genes for the abovementioned compounds, there are several transcription factors, including MYB, bHLH and WD40 [10], associated with the biosynthetic pathway. Furthermore, silencing *PsMYB30* in *Paeonia suffruticosa* petals reduced blotch size, faded blotch colour, and decreased the anthocyanidins and *Pn3G5G* levels. Overexpressing *PsMYB30* increased the anthocyanidins content in tobacco petals [11]. *P. suffruticosa PsMYB44* negatively regulates the biosynthesis of anthocyanidins by directly binding to the PsDFR promoter and subsequently inhibiting blotch formation [12]. These factors cause differences in the levels of gene expression in the anthocyanidins synthesis pathway, thus affecting flower colour.

Rhododendron is one of the most diverse genera of woody plants, with more than 1000 species worldwide. The genus is distributed from the Arctic mountains to the tropics. China is one of the countries of origin for *Rhododendron* germplasms and has rich germplasm collections. By 1996, more than 1025 *Rhododendron* varieties had been collected worldwide, and >90% of these were distributed in Asia, many in the Himalayas and southwest China. *Rhododendron* plants produce colourful flowers, which are often used in urban greening projects, landscaped gardens and potted landscapes. The colour of *Rhododendron* flowers ranges from white to yellow, red and purple. The relationship between *Rhododendron* flower colour and anthocyanidins species and content has been well studied. According to HPLC analysis, red group samples had two to four major anthocyanidins, and those of the purple group had two to six major anthocyanidins. In contrast, no anthocyanidins were detected in petals of the white group, although anthocyanidins were detected [13]. Most of the genes were expressed at higher levels in flowers than in other organs [14]. The flower colours of both tree peony cultivars gradually deepened with flower development. The total anthocyanidins content in 'Caihui' increased during flower development, and the same trend was observed for the anthoxanthins and flavonoids of these two cultivars, but the levels of these two categories of flavonoids in 'Caihui' were always higher than those in 'Xueta' [15]. However, the regulation of the flavonoid biosynthesis pathway and the molecular basis of different petal colours in *Rhododendron* remain unclear.

Rhododendron 'Belgium', also known as *Rhododendron hybridum* Ker Gawl, is a variety obtained by repeated hybridization of *Rhododendron simsii* Planch, *Rhododendron indicum* (L.) sweet and *Rhododendron mucronatum* (Blume) G. Don. It produces flowers of various colours that bloom during all four seasons. It is often used as an indoor bonsai shrub and is an important material for *Rhododendron* flower colour breeding. Currently, flower colour breeding by molecular means is widely reported in *Lilium brownii* var. viridulum, *P. suffruticosa*, *R. rugosa* Thunb, *Rosa chinensis* Jacq and other flowers [16-18], whereas cross-breeding strategies are mostly adopted for *Rhododendron*. In this study, we combined transcriptome sequencing (RNA-seq), proteome sequencing and High-Performance Liquid Chromatography (HPLC) to analyse the differences in the anthocyanidins content of white- or red-flowered *Rhododendron* 'Belgium'. We aimed to study the correlation between flower colour and the expression of candidate genes for the anthocyanidins biosynthesis pathway. The key genes and proteins affecting *Rhododendron* flower colour were found through correlation analysis of related proteins to provide a reference for *Rhododendron* flower colour breeding.

MATERIALS AND METHODS

Materials and sampling

Experimental materials were provided by the Beilun *Rhododendron* Cooperation base at Zhejiang Wanli University. Five 5-year-old 'Belgian' *Rhododendron* shrubs with red and white flowers and similar growth potential were selected. From November 1st to December 30th, 2020, unopened, half-open and blooming red flowers at the flower bud stage (Rbud), initial opening stage (Rmiddle) and full opening stage (Rfull) and white petals of blooming flowers at the full opening stage (Wfull) were collected. The samples were stored at -80°C for later use. These four samples (Rbud, Rmiddle, Rfull and Wfull) were used for transcriptome sequencing. Two samples of the petals of red-coloured 'Belgium' (Rfull) and white-flower 'Belgium' (Wfull) in full bloom was used for proteome sequencing and anthocyanidins assays. Three biological replicates were used in the RNA-seq and proteome sequencing experiments for each sample (Figure 1A).

Detection and quantification of anthocyanidins

The frozen flower petals were ground into a fine powder and placed in a 1% solution of hydrochloric acid: Methanol. The samples were completely immersed in the solution and treated away from light. The samples were stored in a refrigerator at 4°C for 24 h and then filtered to obtain the filtrate. A 1% hydrochloric acid: Methanol solution was then introduced to the filter residue, and the samples were again stored away from light in a 4°C refrigerator for 24 h before being filtered. The two filtrates were then mixed to obtain the required pigment extract. A 1% hydrochloric acid: Methanol solution was used as the control. The samples were then scanned in a wavelength range of 300-700 nm to detect the characteristic absorption spectrum of anthocyanidins.

A total of 0.5 g of Rfull and Wfull samples was ground into a powder under liquid nitrogen. Ten millilitres of ethanol/

hydrochloric acid extraction buffer was added, and the sample was mixed with ultrasonication for 30 min. The samples were centrifuged at a low temperature for 5 min, the supernatant was removed, and the residue was extracted once again. Then, concentrated hydrochloric acid was added at 0.3× the final sample volume and the samples were placed in a water bath at 95°C for 40 min; after cooling, the samples were passed through a 0.22 μm filter membrane and stored at -20°C for later use. The standard samples are Cyanidin-3-O-glucoside chloride, Pelargonidin 3-glucoside chloride and Delphinidin 3-O-glucoside, all from Shanghai Yuanye Biotechnology Co. Ltd. Analysis was performed by HPLC-tandem mass spectrometry (LC-MS/MS) with an injection volume of 2 μl. The column temperature was set to 30°C; mobile phase A was a methanol solution containing 0.1% formic acid, and mobile phase B was an aqueous solution containing 0.1% formic acid. The gradient elution procedure was as follows: 0-2 min, A=20%; 2-14 min, an increased to 80%; 14-15 min, A=80%; 15.1 min, A decreased to 20%; and 15.1-20 min, A=20%. The MS conditions were as follows: Atomization temperature=400 °C; atomizer pressure=65 psi; curtain gas=15 psi; and spray voltage=4500 V. The results were analysed using waters synapt G2 MS software.

RNA extraction and RNA-seq analysis

The four samples (Rbud, Rmiddle, Rfull and Wfull) were used for transcriptome sequencing, which was repeated three times (Figure 1A). Samples were sent to BGI (Wuhan) for sequencing. The total RNA was treated with mRNA enrichment or rRNA removal. For mRNA enrichment, mRNA with a poly-A tail was enriched with magnetic beads harbouring oligo-DTs. For rRNA removal, rRNA was hybridized with a DNA probe. The DNA/RNA hybridization chain was selectively digested with RnaseH, and the DNA probe was digested with DnaseI to obtain the required RNA sample after purification. The obtained RNA was fragmented with fragmentation buffer, and random N6 primers were reverse transcribed. The cDNA was then synthesized to form double-stranded DNA. The end of the synthesized double-stranded DNA was flattened and phosphorylated at the 5'end; the 3'end formed a sticky end with a protruding "A" and was then connected at a bubble joint with a protruding "T" at the 3'end. The ligation products were amplified by PCR with specific primers. The PCR product was thermally denatured into a single strand, and then a bridge primer was used to cyclize the single-stranded DNA to obtain a single-stranded circular DNA library. Finally, the DNA was sequenced.

The raw data was filtered with SOAPnuke [19], Removing reads containing adapters (adapter contamination); Removing reads whose unknown base ('N' base) ratio is more than 5%; Removing reads whose low-quality base ratio (Base quality less than or equal to 15) is more than 20%, afterwards clean reads were obtained and stored in FASTQ format (https://github.com/BGI-flexlab/SOAPnuke).

Trinity (v2.0.6) was used to assemble the clean reads and BUSCO to assess the assembly quality (https://github.com/trinityrnaseq/trinityrnaseq/wiki) [20]. Clean data were mapped to the assembled unique gene by Bowtie2 (v2.2.5) (http://bowtie-bio.sourceforge.net/Bowtie2/index.shtml) and the expression level of genes was calculated by RSEM (v1.2.8) (http://deweylab.biostat.wisc.edu/rsem/rsem-calculate-expression.html) [21].

Annotating assembled Unigene with functional databases (Kyoto Encyclopedia of Genes and Genomes (KEGG), Gene Ontology (GO)), and predicting the transcription factors. Within-group differential gene analysis was performed using DESeq [22], under the conditions of Fold Change ≥ 2 and Adjusted P value ≤ 0.001. PossionDis was performed between-group differential gene analysis under the conditions of Fold Change ≥ 2 and FDR ≤ 0.001. Using the pheatmap function on the differential gene set to draw a heatmap of differential gene clusters. According to the GO and KEGG annotation results and official classifications, the differentially expressed genes were functionally classified, and the phyper (https://en.wikipedia.org/wiki/Hypergeometric_distribution) in R software was used for KEGG enrichment analysis, and the TermFinder package was used for GO Enrichment analysis (https://metacpan.org/pod/GO:TermFinder). With a Q value of ≤ 0.05 as the threshold, candidate genes that met this condition was defined as being significantly enriched.

Protein extraction, iTRAQ labelling and MS

Total protein was extracted from each sample as described previously [23]. The Shimadzu LC-20AB liquid phase system was used, and the column used for liquid phase separation of the sample was a 5 μm 4.6 × 250 mm Gemini C18 column. The dried peptide samples were reconstituted with mobile phase A (5% ACN, pH 9.8) and injected, eluting at a flow rate of 1 mL/min using the following gradient: 5% mobile phase B (95% ACN, pH 9.8) for 10 min; 5-35% mobile phase B for 40 min; 35-95% mobile phase B for 1 min; mobile phase B for 3 min; and 5% mobile phase B for 10 min. The elution peak was monitored at a wavelength of 214 nm, and one component was collected per minute. The samples were combined according to the chromatographic elution peak map to obtain 20 fractions, which were then freeze-dried. The dried peptide samples were reconstituted with mobile phase A (2% ACN, 0.1% FA) and centrifuged at 20,000 × g for 10 min, and the supernatant was taken for injection. Separation was performed using a Thermo UltiMate 3000 ultra-HPLC system. The sample was first enriched in a trap column and desalted and then entered a self-packed C18 column (75 μm internal diameter, 3 μm column size, 25 cm column length) and separated at a flow rate of 300 nL/min using the following elution gradient: 0~5 min, 5% mobile phase B (98% ACN, 0.1% formic acid); 5~45 min, mobile phase B linearly increased from 5–25%; 45~50 min, mobile phase B increased from 25–35%; 50~52 min, mobile phase B increased from 35–80%; 52~54 min, 80% mobile phase B; 54~60 min, 5% mobile phase B. The HPLC instrument was connected directly to the mass spectrometer. The peptides separated by liquid chromatography were ionized by a Nano Electrospray Ionization (ESI) source and then transferred to a tandem mass spectrometer Q-Exactive HF X (Thermo Fisher Scientific, San Jose, CA) for detection in data-dependent acquisition mode.

iTRAQ data analysis

Automated software called IQuant was applied for quantitatively analysing the labelled peptides with isobaric tags [24]. This software integrates Mascot Percolator, an effective machine learning method for rescoring database search results, to provide reliable significance measures. To assess the confidence of peptides, the PSMs were prefiltered at a PSM-level FDR of 1%. Then, based on the "simple principle" (the parsimony principle), the identified peptide sequences were assembled into a set of confident proteins. To control the rate of false positives at the protein level, a protein FDR of 1%, which was based on the picked protein FDR strategy [25], was estimated after protein inference (protein-level FDR ≤ 0.01). We defined a Differentially Expressed Protein (DEP) as being significantly regulated if the p value was less than 0.05. In the GO enrichment analysis, we used a hypergeometric test to identify the target GO terms.

Correlation analysis of the proteome and transcriptome

Based on gene expression information, omics sequencing was conducted at the mRNA and protein levels, and the two sets of omics data were integrated to investigate the gene expression process based on data classification and correlation [26].

qRT–PCR

Total RNA of experimental materials was extracted using the RNAprep Pure Plant Plus Kit for total RNA extraction from polysaccharide- and polyphenol-rich plant samples, and fluorescent cDNA was synthesized according to the instructions for the NovoScript Plus All-in-One 1st Strand cDNA Synthesis Supermix (gDNA purge) reverse transcription kit. qRT-PCR analysis was performed according to the instructions for the Novostart SYBR qPCR Supermix Plus (Shanghai nearshore Technology Co., Ltd.) kit. The reaction system included 35 μL of 2 × Novostart SYBR qPCR Supermix Plus, 1.4 μL each of the forward and reverse primers, 1.4 μL of fluorescent cDNA, and 30.8 μL of ddH_2O, in a total volume of 70 μL. The reaction procedure was as follows: Predenaturation at 95°C for 5 min, followed by 30 cycles of 95°C for 1 min, 95°C for 20 s, and 60°C for 1 min. The relative expression level of the genes was determined using the $2^{-\Delta\Delta CT}$ method using 18S rRNA as the reference gene. The primers used for qRT–PCR are listed in the Supplementary material (Supplementary Table 1).

Overexpression of the *DFR* gene

The linearized vector pcambia1302 and sgRNA *DFR* reading frame were connected by DNA recombinase to construct the recombinant vector pCambia *DFR*. The recombinant product was transformed into *E. coli* (*Escherichia coli)* DH5 α using the freeze-thaw method. The competent cells, screened by kanamycin, produced single white colonies. The instructions of the plasmid extraction kit (Endo-free Plasmid Mini Kit I) were followed to extract plasmids. The *Agrobacterium tumefaciens* solution stored at -20°C.

Then, 300 μL of the first activated bacterial liquid was added to 25 mL of YB medium until the bacteria were shaken again. Specifically, they were centrifuged at 4°C and 5000 rpm for 10 min. The supernatant was discarded, the bacteria were collected and resuspended in 25 mL of YB medium, 25 μL of 0.1 mol/L Acetylsyringone (AS) was added to activate A. *tumefaciens*, and cultivation was continued for 5 hours. The bacteria were collected and centrifuged at 4°C and 5000 rpm for 10 min. Azalea culture medium WPM was used to adjust the OD value, with an OD 600 of 0.6-0.8. The inflorescence infection method was applied to infect *Arabidopsis* with the prepared Agrobacterium infection liquid. Arabidopsis seeds were collected, and positive seeds were screened on Rif and hygromycin resistance screening plates. After sowing, Rs-DFR-overexpressing plants (T1), empty control AT-N plants and wild-type AT-WT plants were obtained.

Statistical analysis

LC–MS was used to compare the anthocyanidins content between the two cultivars. Pearson correlation coefficients were calculated between the abundance of different proteins from proteomic profiling and between relative expression values from qRT–PCR and RNA-seq across stages. The data are expressed as the mean ± SD from three independent biological replicates. Significance was determined *via* one-way Analysis of Variance (ANOVA), and for differences between groups, the Least Significant Difference (LSD) t test was employed ($p<0.05$).

RESULTS

Composition and contents of anthocyanidins in different *Rhododendron* cultivars

The pigment absorption spectra of the red and white *Rhododendron* flower petals were obtained by scanning the petal extract with UV-visible spectrophotometry at 300~700 nm. The pigment absorption spectra of the two differently coloured *Rhododendron* 'Belgium' flowers showed characteristic absorption peaks at 345 nm, and the pigment absorption spectra of the red flowers also showed characteristic absorption peaks at 525 nm, indicating the existence of anthocyanidins and flavonoids in the red pigments. The single absorption peak of the white flower extract indicated that there were only flavonoids in the white flower pigment. The differences in the peak values and the number of absorption peaks indicated that the pigment types and structures contained in the two *Rhododendron* samples differed (Figure 1B).

The anthocyanidins contained in the petals were qualitatively and quantitatively examined by LC-MS. At the full flowering stage, the levels of anthocyanidins in the red flowers were higher than those in the white flowers. The cyanidin and delphinidin levels in the red flowers were 3.5 and 2.18 times higher than those in the white flowers, respectively. There was little difference in the levels of pelargonidin (Figure 1C and Supplementary Figure 1).

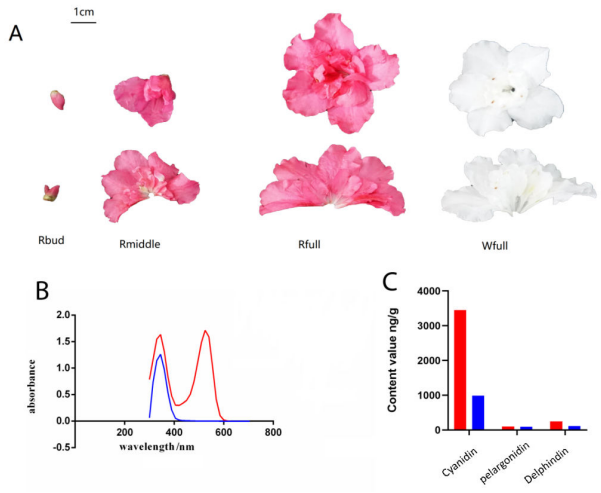

Figure. 1: Difference in anthocyanidins content in flower petals and anthocyanidins wavelength of Rhododendron. (A) Red Rhododendron and white Rhododendron; (B) Anthocyanidins wavelength detection; (C) Differences in anthocyanidins content between Rfull and Wfull. **Note:** ( ): Absorbance/white; ( ): Absorbance/Red.

De novo transcriptome assembly and quality assessment

We used Illumina paired-end sequencing to sequence a total of four RNA-seq libraries from three stages of red flower blooming and white flowers in full bloom. After quality control, approximately 1.2 billion (m) clean reads (approximately 180 GBP) remained, and the number of reads between libraries was very consistent (Supplementary Table 2).The overlapping group N50 values for Corset analysis across the two varieties were similar, ranging from 1443-1740 BP (Supplementary Table 3).Based on 140 conserved BUSCO embryonic plant homologous genes, the transcriptome integrity assessment found more than 140 genes in Belgian *Rhododendron* (Supplementary Table 4).Therefore, our results show that the transcriptome data were well assembled and relatively complete.

DEG analysis of RNA-seq data

The DEGseq method is based on the Poisson distribution [27]. A volcano plot was created showing the overall distribution of different genes. The statistical results showed that the total number of *DEGs* reached 18,901, of which 10,980 and 7921 were upregulated and downregulated, respectively. CL27, Contig1, CL6158, Contig6 are significant downregulation whereas CL4037, Contig11 CL10396, Contig9 are significantly upregulated (Figure 2A). The KEGG enrichment analysis showed that among the genes with pathway annotations, phenylpropane biosynthesis and flavonoid biosynthesis were important pathways for anthocyanidins synthesis. The first 20 KEGG pathways obtained are shown in Figure 2B. Through a heatmap hierarchical cluster analysis, it was easy to distinguish the *DEGs* in red and white flower samples, indicating that there were significant differences in gene expression levels between the red and white flower petals at the full flowering stage (Supplementary Figure 2).To further study the biochemical pathways in red and white flower petals, we identified all the *DEGs*. The most abundant enriched GO term was integral component of membrane. The largest proportion compared to the total number of genes was observed for monovalent inorganic cation transport (Figure 2C). Through statistical analysis of the transcriptome data from the red and white flower samples, genes with differential expression of the anthocyanidins synthesis pathway in red flowers and white flowers were found to include *CHS, CHI, F3H, DFR, 3GT, FLS, ANS, F3'H*, and *F3'5'H* (Figure. 2D). Transcriptome pathway analysis showed that the expression of the *RsCHS2* gene in red flowers was significantly higher than that in white flowers, the expression of *RsCHI3*, *RsCHI4*, *RsCHI5* and *RsCHI7* in red flowers was higher than that in white flowers, the expression of family genes in white flowers was significantly higher than that in red flowers (*RsCHI2*), the expression of *F3H* family genes in white flowers was higher than that in red flowers (*RsF3H2*), and the expression of other genes was the same between the two types of flowers. There was no significant difference in *F3'H* and *F3'5'H* between white flowers and red flowers. The expression of 3GT in red flowers was higher than that in white flowers, wherein the level in *Rs3GT3* was significantly higher than that in white flowers. The difference in the expression of the *ANS* gene between white flowers and red flowers was not obvious (Table 1).

Gene	Wfull FPKM	Rfull FPKM	$\log_2$ (Rfull/Wfull)	Description
CHS	0		10.13887	Transferring acyl groups other than amino-acyl groups
CL11238.Contig1	0.066	7.596	10.413197	
Unigene3474		95.656		

CHI	0.196	1.116	2.4488925	Transferase activity, transferring acyl groups
CL11238.Contig1				
F3H	8.14	17.016	1.0319216	Naringenin 3-dioxygenase activity
CL7405.Contig3				
DFR	0.756	2.213	1.5661191	Dihydrokaempferol 4-reductase activity
Unigene22738				
3GT	0.46	11.85	4.5985508	Transferase activity, transferring hexosyl groups
CL10961.Contig11				
FLS	19.97	4.583	-2.1087368	Oxidoreductase activity, metal ion binding
CL11733.Contig3	21.14	5.083	-2.034999	
CL11733.Contig4				

Table 1: Differentially expressed anthocyanidins synthesis pathway genes.

To study the regulation of flower colour in different tissues during red flower development, we analysed the differential expression of genes between red flower petals at various stages (Figure 3A). The results showed that the numbers of upregulated genes in Rbud *vs.* Rmiddle, Rbud *vs.* Rfull and Rmiddle *vs.* Rfull were 5071, 13,830 and 20,372, respectively, and the numbers of downregulated loci were 17,122, 25,662 and 17,541, respectively (Figure 3B). The *DEGs* in the anthocyanidins synthesis pathway during red flower development were compared, and the results showed that *FLS*, *FLS*, *CHS*, and *CHS* were significantly upregulated and *F3H*, *DFR*, and *3GT* were downregulated (Figure 3B). The common DEGs in Rbud *vs.* Rmiddle, Rbud vs. Rfull and Rmiddle *vs.* Rfull were screened out. As a result, 102 Common DEGs were screened out (Figure 3C and Figure 3D). Perform KEGG analysis on these 102 common differentially expressed genes, KEGG function enrichment analysis can determine the metabolic and transduction pathways associated with DEGs, and we found that phenylpropanoid biosynthesis-related genes played a key role in flower development. In addition, high enrichment and richness of flavonoid biosynthesis genes was noted, and the genes related to flavonoid biosynthesis are likely to be key during flower development.

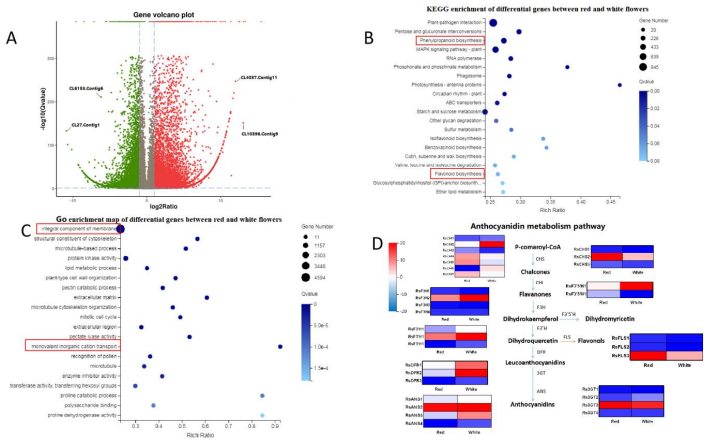

Figure 2: Anthocyanidins synthesis pathway and *DEGs* between red and white flowers; (A) Volcano plot of *DEGs* in Rfull and Wfull; (B) KEGG pathway enrichment of the *DEGs*; (C) GO enrichment of the *DEGs*; (D) Anthocyanidins synthesis pathway. **Note:** The x-axis represents the fold change after $\log_2$ conversion, and the y-axis represents the significance value after $\log_{10}$ conversion. (■) upregulated *DEGs*; (■) downregulated *DEGs*; (■) non-*DEGs*. (B and C) The x-axis is the enrichment ratio (the ratio of the number of genes annotated to an item in the selected gene set to the total number of genes annotated to the item in this species, and the calculation formula is rich ratio=term candidate gene num/term gene num), the y-axis is the GO term, the size of the bubble indicates the number of *DEGs* annotated to a GO term, and the (■) enrichment Q value; the darker the smaller the Q value. (D) The (■) higher the expression level, and the (■) the lower the expression level.

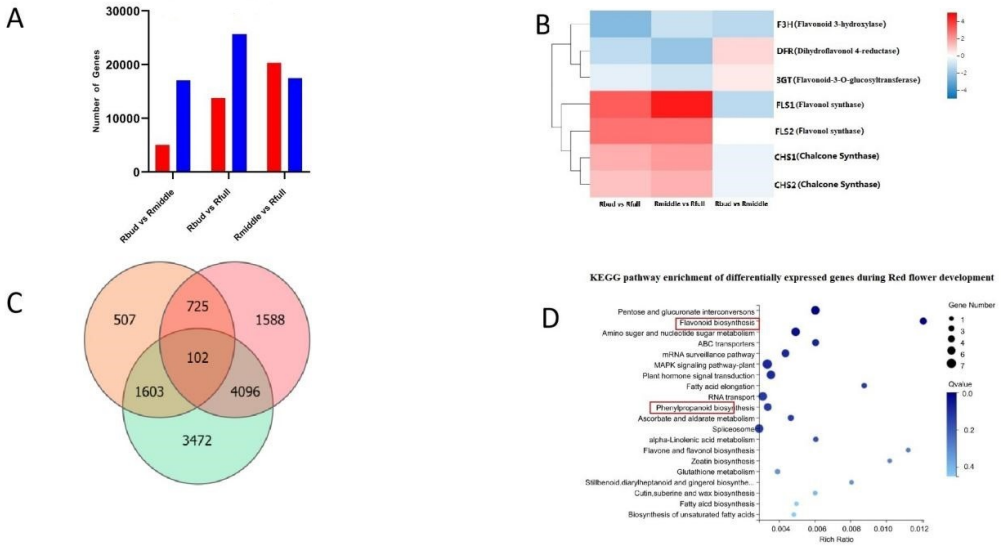

Figure 3: *DEGs* among different flower developmental stages.(A) Number of *DEGs*; (B) Clustering heatmap of anthocyanidins synthesis pathway-related genes in the azalea development stage; (C) Venn diagram; (D) KEGG pathway enrichment. **Note:** (▬) up regulation; (▬) down regulation. (▬) Rbud_3-*vs*-Rmiddle_3; (▬) Rbud_3-*vs*-Rfull_3; (▬) Rmiddle_3-*vs*-Rfull_3.

Protein data analysis

Under the "1% FDR" filtering standard, a total of 35,103 peptides and 8303 proteins were identified. The results showed that there were 1613 DEPs between red and white flowers, of which 1043 were upregulated and 570 were downregulated (Supplementary Figure3). The results showed that there were significant differences in protein levels between red and white flowers at the full flowering stage. According to the KEGG classification analyses of the DEPs, biosynthesis of secondary metabolites accounted for the largest proportion of up- and downregulated genes, indicating that secondary metabolites may be key factors affecting flower colour (Supplementary Figure 4). GO functional annotation was also carried out for the DEPs, which were divided into three categories: Biological process, cellular component and molecular function. Cellular and metabolic processes accounted for the highest proportion of biological process terms, and cell, cell part, organelle, membrane, membrane part, organelle part and polymer complex accounted for the highest proportion of cellular component terms. Catalytic activity and binding accounted for most of the molecular function annotations. The above enriched GO terms may represent the key proteins affecting flower colour in red and white flowers and should be investigated further (Supplementary Figure 5).

Transcriptome and proteome association analysis

We conducted a correlation analysis between the transcriptomic and quantitative proteomic data. Globally, the expression levels of all the transcripts and their corresponding quantified proteins in Wfull *vs*. Rfull showed a limited correlation (r=0.070) (Figure. 4A). However, a stronger correlation was observed between the *DEGs* and their corresponding DEPs (r=0.125) (Figure. 4B). The expression ratios of proteins and their corresponding mRNAs with the same or opposite trend (both>1 or both<1) were also plotted, and a higher positive or negative correlation was indicated (Figure 4C and 4D). Single-molecule long-read sequencing analysis, or iso-Seq, was performed as a reference. We quantified 28100 *DEGs* and 1613 DEPs in the transcriptome and proteome, respectively, and 409 members were related in the two omic data sets (Supplementary Figure 6). The molecular number and high correlation ratio of molecules in the omic data sets revealed that the sequencing data were of high quality. It can be seen from the box diagram that the gene expression levels of genes associated with red and white flower colour were higher than those of non-associated genes (Supplementary Figure 7).

The KEGG pathways of the *cor-DEG-DEP* genes were identified using a p value of less than .05 as the cut-off, and the 409 *cor-DEG-DEP* genes were mapped to 128 KEGG pathways (Supplementary Table 5). The results showed that three KEGG pathways were highly enriched at both the mRNA and protein levels, including "RNA transport" (ko03013), "Metabolic pathways" (ko01100), and "Biosynthesis of secondary metabolites" (ko01110). Moreover, the KEGG pathways of secondary metabolism were significantly enriched in the *cor-DEG-DEP* genes, such as "Flavonoid biosynthesis" (ko00941), "Phenylpropanoid biosynthesis" (ko00940), and "Carotenoid biosynthesis" (ko00906), suggesting that these processes were highly active after light exposure (Supplementary Table 5).

To study the mechanisms underlying the production of red and white flower colour, we performed combinatorial analysis, including RNA-seq and label-free quantification proteomics based on MS at different time points. Except for *F3'5'H*, other proteins related to the anthocyanidins metabolism pathway were significantly different. *ANS*, *CHI* and *3GT* were the most significant, and the mean ratio (red/white) for these genes was 1.187, 0.975 and 0.968, respectively.

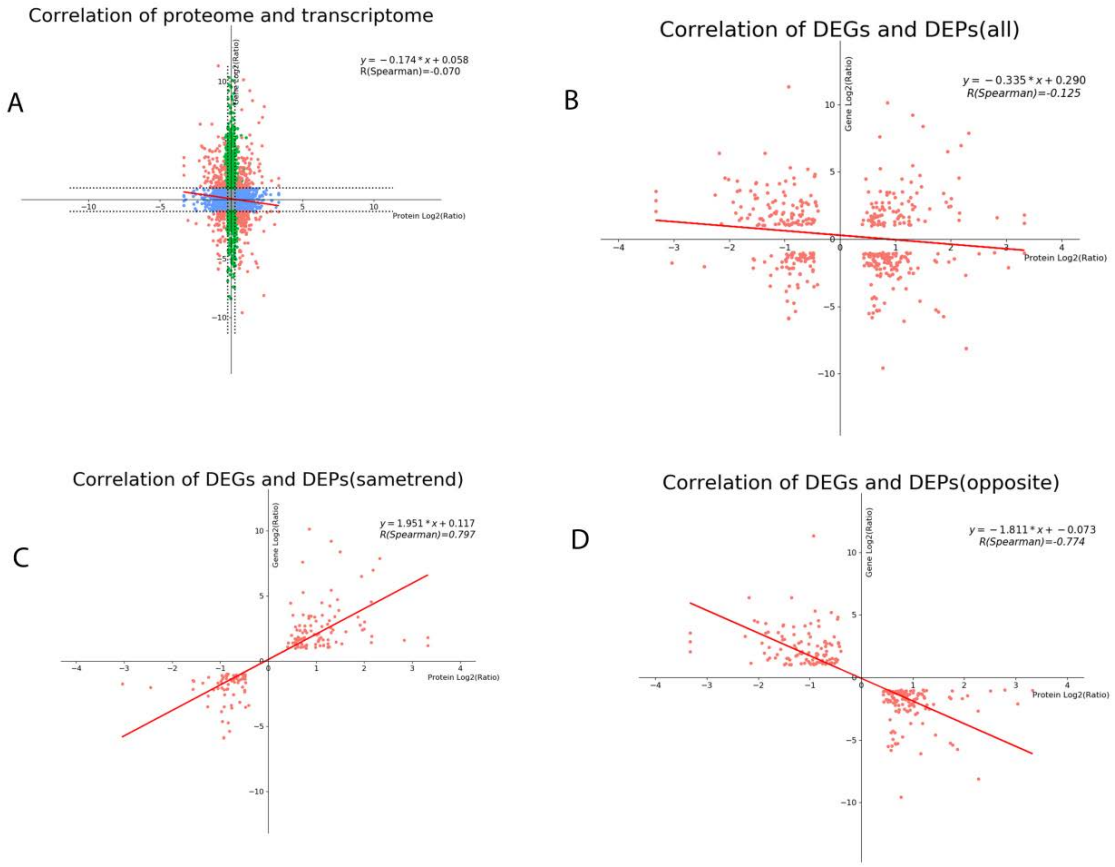

Figure 4: Correlations between protein and messenger Ribonucleic Acid (mRNA) expression. The x-axis represents the protein expression level, and the y-axis represents the gene expression level. (A) Scatterplots of the relationship between genes quantified in both the transcriptomic and proteomic data sets; (B) Scatterplots and correlation coefficients between Differentially Expressed Proteins (DEPs) and Differently Expressed Genes (*DEGs*). Scatterplots and correlation coefficients between proteins and mRNA expression ratios with the same (C) or opposite (D) changing tendency. **Note:** ( ) dots indicate no DEPs and *DEGs*; ( ) dots indicate DEPs but no *DEGs*; ( ) dots indicate *DEGs* but no DEPs; ( ) dots indicate DEPs and *DEGs*. All data were $\log_2$-transformed.

Real-time quantitative PCR analysis

To verify the RNA-seq results, red and white 'Belgium' petal samples were examined by qPCR. The key genes of the anthocyanidins synthesis pathway, *RsCHS*, *RsCHI*, *RsF3H*, *RsFLS*, *RsDFR*, *RsF3'H*, *RsF3'5'H* and *RsANS*, were selected. The results showed that the expression trends of Differentially Expressed Genes (*DEGs*) measured by qRT-PCR were consistent with the results of transcriptome sequencing, indicating that the transcriptome data were reliable [28]. The *CHS* gene was enriched in both red flowers and white flowers from the bud to full bloom stages. The *DFR* gene level in red flowers increased from the bud to full bloom stages; the level in white flowers increased from the bud to middle stages and declined from the middle to full bloom stages; *DFR* gene expression in red flowers was higher than that in white flowers; and FLS gene expression in red flowers and white flowers decreased from the bud to full bloom stages [29]. These findings demonstrate competitive common substrate relationships with *DFR* genes. FLS expression in white flowers was significantly higher than that in red flowers (Figure. 5). To verify the RNA-seq results, red and white 'Belgium' petal samples were examined by qPCR. The key genes of the anthocyanidins synthesis pathway, *RsCHS*, *RsCHI*, *RsF3H*, *RsFLS*, *RsDFR*, *RsF3'H*, *RsF3'5'H* and *RsANS*, were selected. The results showed that the expression trends of differentially expressed genes (*DEGs*) measured by qRT-PCR were consistent with the results of transcriptome sequencing, indicating that the transcriptome data were reliable. The *CHS* gene was enriched in both red flowers and white flowers from the bud to full bloom stages. The *DFR* gene level in red flowers increased from the bud to full bloom stages; the level in white flowers increased from the bud to middle stages and declined from the middle to full bloom stages; *DFR* gene expression in red flowers was higher than that in white flowers; and *FLS* gene expression in red flowers and white flowers decreased from the bud to full bloom stages. These findings demonstrate competitive common substrate relationships with *DFR* genes. FLS expression in white flowers was significantly higher than that in red flowers (Figure 5).

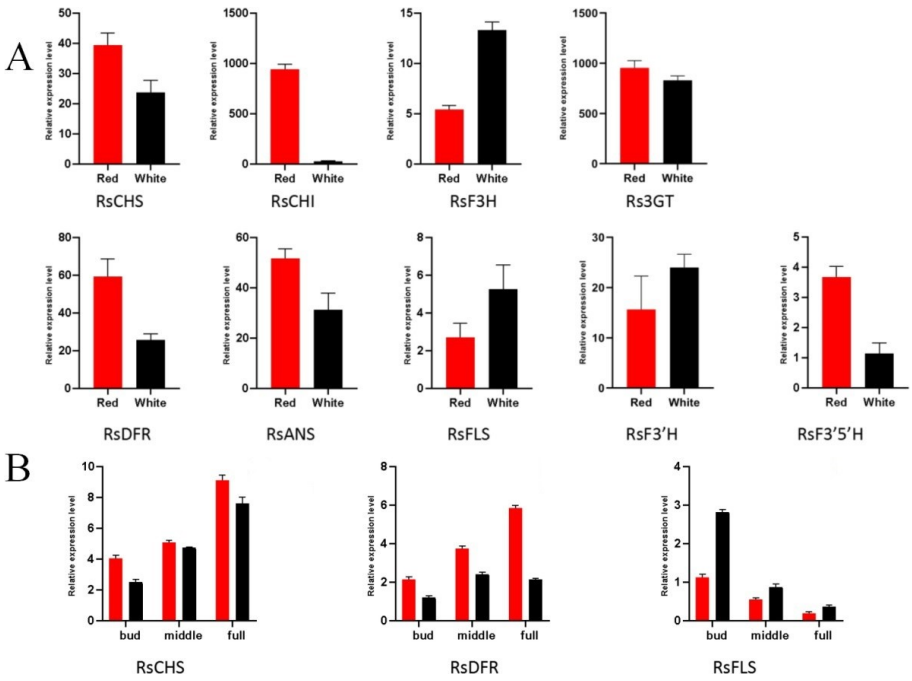

Figure 5: qRT-PCR analysis. Error bars show the mean ± SD of three independent biological experiments. (A) qRT-PCR of different genes of the anthocyanidins synthesis pathway in Rfull and Wfull; (B) Expression of *CHS*, *DFR*, and *FLS* in the bud, middle and full stages of red flowers and white flowers [30]. Note: (▬) indicates red colour; (▬) indicates white colour.

Comparison of overexpression phenotypes in *Arabidopsis thaliana*

To verify whether overexpression of *Rs-DFR* can change the colour of plants, this gene was overexpressed in *Arabidopsis*, and the overexpression plants and wild-type plants were compared. The *Rs-DFR* leaves were purplish red, and the AT-WT and AT-N leaves were green (Figure 6A). The expression level in *Rs-DFR* was 2 times that in AT-N and AT-WT, showing that the DFR gene plays an important role in anthocyanidins synthesis in plants (Figure 6B and Figure 6C). LC-MS analysis showed that the pelargonin content in *Rs-DFR* was much higher than that in AT-WT and AT-N (Supplementary Figure 8). Cyanidin was detected in Rs-DFR but not in AT-WT and AT-N (Figure 6D and Figure 6E).

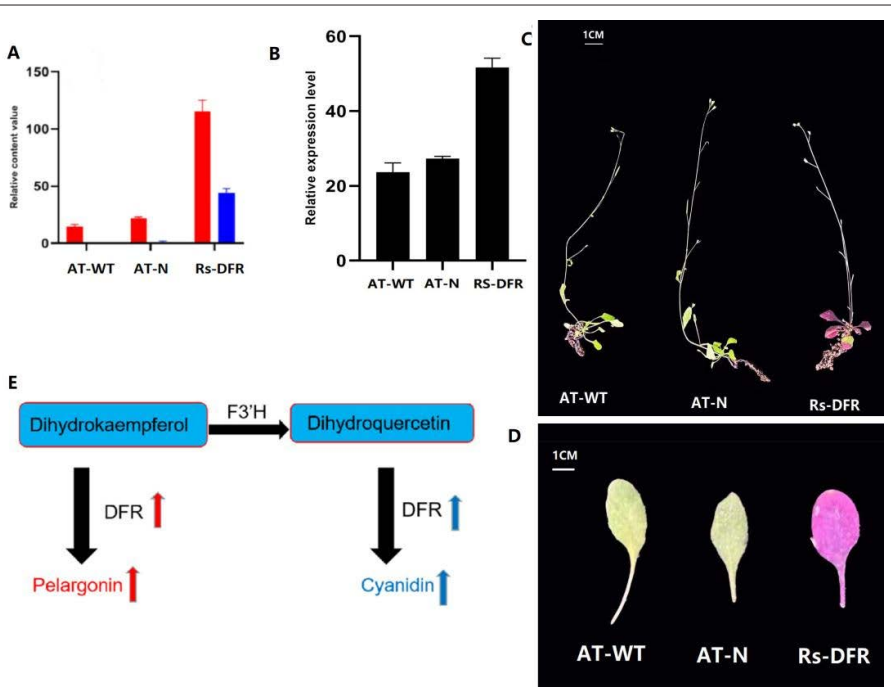

Figure 6: Overexpression results in Arabidopsis. (A) Analysis of the anthocyanidins content, (B) Analysis of *DFR* gene expression, (C) Comparison of Arabidopsis plants, (D) Comparison of Arabidopsis thaliana leaves, (E) DFR gene regulation mode diagram. **Note:** (▬) indicates pelargonidin; (▬) indicates cyanidin.

DISCUSSION

Rhododendron is a non-model plant genus, and there is little available genetic information about this genus [31]. The continuous development of new varieties makes it increasingly difficult to study the genetic diversity of their germplasm resources. In this study, scanning UV spectrophotometry indicated the presence of anthocyanidins and flavonoids in Belgian *Rhododendron*, and anthocyanidins with different structures occurred in flowers of different colours. Similar results have been found in other studies on anthocyanidins. For example, Li et al [32], analysed the anthocyanidins components of Alpine *Rhododendron* and found that it contains flavonoids, although their structures require further study [33]. Zheng and others found that in Qinling beauty Rhododendron with different degrees of red flower colour intensity, the factors determining colour development were mainly related to the types and distribution of anthocyanidins, and compound anthocyanidins formed by other flavonoids were also involved [34]. Plant species generally have six main anthocyanidins, namely, pelargonidin pigment, cornflower pigment, delphinium pigment, peony pigment, morning glory pigment and Mallow pigment. The anthocyanidins in the petals of two *Rhododendrons* with different flower colours were analysed by LC-MS, and the anthocyanidins content in the petals at different developmental stages was determined. The results showed that the anthocyanidins in the two differently coloured Rhododendron 'Belgium' flowers were basically the same, and the anthocyanidins content of the red flowers was approximately three times that of the white flowers. The cyanidin content was the highest in both red and white flowers, while the pelargonidin content was the lowest. The results showed that the colour difference between red and white flowers was mainly caused by cyanidin. The second most prevalent anthocyanidins was delphinidin, which is consistent with the result showing that *F3'5'H* in raspberry has no activity at the gene and protein levels, resulting in a lack of delphinidin [35].

The flavonoid pathway has the greatest impact on flower colour. In recent years, a variety of regulatory mechanisms have been reported through transcriptome sequencing. Following these studies, although there are still many unknown genes, many genes have been annotated and shown to participate in different regulatory mechanisms. In this study, we comprehensively analysed the flower colour formation mechanism of Belgian *Rhododendron*. During the development of Rhododendron 'Belgium', we found several key metabolic pathways for flower colour through enrichment analysis. The Spearman correlation coefficient was calculated through the correlation analysis of all the expression data correlated at the protein level and transcriptome level. Most previous studies have revealed that protein expression is moderately correlated. Although the correlation was not highly consistent, resulting from the different experimental processes and data types, the complex regulatory mechanisms of gene expression had an important effect (Figure. 4A), and differential expression of flavonoid pathway genes was found through analysis of *cor-DEG-DEPs* (Supplementary Table 6).

In addition, several *DEGs* required for anthocyanidins biosynthesis were detected by comparative transcriptome analysis. As mentioned above, 9 candidate genes involved in anthocyanidins biosynthesis were verified by qRT–PCR. Previous studies have shown that the depth of flower colour is related to the anthocyanidins content; the greater the anthocyanidins content is, the darker the colour. This is consistent with the results of this study, in which we found that different gene expression levels may account for the differences in the flower colours of *Rhododendron* 'Belgium' [36,37]. In the flavonoid/anthocyanidins pathway, the largest gene expression difference was observed in *CHI*. The expression of *CHI* in red flowers was 30 times that in white flowers (Figure 5). *CHI* can convert chalcone to flavonoids, making it a key enzyme in the anthocyanidins synthesis pathway. These results are consistent with the high expression of *CHI* in dark *Rhododendron* and lettuce with red leaves [38,39]. *CHS* is the first rate-limiting enzyme and a key enzyme in the anthocyanidins biosynthesis pathway. In this study, the expression of *CHS* in red flowers was nearly twice that in white flowers (Figure 2D and 5). The expression level of the *CHS* gene is directly related to the response of the flavonoid synthesis pathway in plants. For example, *Petunia* "Red Star" is a star-patterned red and white variety in which the sequence-specific degradation of *CHS* RNA causes colour variation [40]. In *Brunfelsia acuminata* flowers, a striking colour change from dark purple to light purple and ultimately to pure white resulted from a decline in the anthocyanidins content of the petals and was preceded by a decrease in the expression of BaCHS, which was consistent with the results of this study [41]. ANS is the last enzyme in the anthocyanidins synthesis pathway. In this study, the expression of ANS in red flowers was higher than that in white flowers. In *Paeonia lactiflora*, the low expression of ANS leads to the yellowing of flower colour [42,43]. *FLS* and *DFR* compete for a common substrate [44]. Research has shown that the high expression of *FLS* causes red tobacco flowers to become white. Our transcriptome data showed that, in contrast, the expression of *FLS* in red flowers of *Rhododendron* was higher than that in white flowers. It is possible that the high expression of *FLS* leads to increased flavonol content. Copigmentation with flavonol affects anthocyanidins and may turn the colour to dark red. This is consistent with the results found in lotus [45]. However, qPCR results showed that FLS expression in red flowers was lower than that in white flowers at all stages of flowering, and the highest expression was observed in white flowers at the flower bud stage, indicating that the *FLS* gene plays an important role in regulating the formation of white flowers at the flower bud stage. Copigmentation with anthocyanidins and flavonols plays an important role in the formation of flower colour, and the expression level of *F3'5'H* is closely related to petal pigmentation and fusion. For example, in Chrysanthemum, *F3'H* and *F3'5'H* play a key role in the anthocyanidins synthesis pathway [46,47], but the expression difference in *Rhododendron* is not significant.

However, all the above candidate genes are classified as structural genes. There was no significant difference in the gene expression of the *MBW* (MYB-bHLH-WD40) protein complex. This may have been because the compared objects were petals in

full bloom [48]. Studies have shown that *MBW* transcription factors jointly regulate anthocyanidins accumulation at the transcriptional level, especially at the initial stage of flower colouration, but may not exist during the later stages of flower budding. In our study, the cluster heatmap showed that the expression of the MYB-bHLH transcription factor significantly differed during the development of red flowers, and the temporal and spatial changes were basically consistent, while there was relatively little difference between red flowers and white flowers (Supplementary Figure 9). In addition, the MBW complex can bind to the promoter of the anthocyanidins gene, and the MYB binding site of the promoter sequence plays an important role in anthocyanidins synthesis. Some studies have shown that the methylation level of the mdGST [49], and ogCHS [50], promoter regions is significantly correlated with the level of gene expression. However, there was little difference in GC content among the varieties, indicating that DNA methylation is unlikely to be the reason for the colour diversity of *Rhododendron*. Therefore, we believe that the regulation and transcription of anthocyanidins pathway genes may not be independent processes. Although they are expressed alone, they jointly promote the accumulation of anthocyanidins [51]. We have obtained preliminary data about the flavonoid/anthocyanidins pathway in *Rhododendron*, but more experiments are needed to further study the role of each genetic component.

A. thaliana transformation experiments showed that the high expression of the *DFR* gene could promote anthocyanidins synthesis. In this study, we overexpressed the *DFR* gene in *A. thaliana*, and the pelargonin level increased approximately 5-fold. Cyanidin was not detected in wild-type *Arabidopsis* but was detected in overexpressed *Arabidopsis*, indicating that the *DFR* gene plays a positive role in the regulation of cyanidin synthesis, and the leaves changed from green to purplish red, consistent with the results of previous studies [52]. The *DFR* gene of *Rhododendron* plays an important role in anthocyanidins synthesis, which lays an important foundation for flower colour regulation in *Rhododendron*. In the anthocyanidins synthesis pathway, dihydrokaempferol is converted to pelargonin via a reaction catalysed by *DFR*, *F3'H* catalyses the conversion of dihydrokaempferol to dihydroquercetin, and then the *DFR* gene synthesizes cyanide. In this study, we verified this result by overexpressing the *RsDFR* gene in Arabidopsis, which resulted in an increase in cyanidin levels (Figure 6D). Based on the current results, gene silencing vectors can be constructed to inhibit the expression of *RsDFR* in *Arabidopsis*, observe biological traits, and more specifically interpret the function of *DFR* genes.

CONCLUSION

In this study, we combined transcriptome and proteome analyses to identify *DEGs* and protein anthocyanidins biosynthetic pathways in *Rhododendron*. The *DEGs* of the anthocyanidins synthesis pathway in red flowers and white flowers were *CHS, CHI, DFR, F3H, FLS*, and *3GT*. Flavonoid biosynthesis plays an important role in flower development. This may be an important reason for changes in appearance. In addition, the difference in proteins between red flowers and white flowers is the result of differential expression of anthocyanidins biosynthetic pathway genes. In the future, we will focus on isolating some key transcription factors that regulate anthocyanidins synthesis and further elucidate the regulatory mechanism of the anthocyanidins biosynthesis pathway. These results provide insight into the factors that augment the effect of the anthocyanidins synthesis pathway on flower colour in *Rhododendron* and will, therefore, contribute to the molecular breeding of *Rhododendron* with ideal features and high quality.

REFERENCES

1. Li Y, Gao R, Zhang J, Wang Y, Kong P, Lu K, et al. The biochemical and molecular investigation of flower color and scent sheds lights on further genetic modification of ornamental traits in *Clivia miniata*. Hortic Res-Englang. 2022;9:114.
2. Sundaramoorthy J, Park GT, Jo H, Lee JD, Seo HS, Song JT. A novel pinkish-white flower color variant is caused by a new allele of flower color gene w1 in wild soybean (*Glycine soja*). Agrono. 2021;11(5). 1001.
3. Liu J, Wang Y, Zhang M, Wang Y, Deng X, Sun H, et al. Color fading in lotus (*Nelumbo nucifera*) petals is manipulated both by anthocyanidins biosynthesis reduction and active degradation. Plant Physiol Bioch. 2022;179:100-107.
4. Li S, He Y, Li L, Li D, Chen H. New insights on the regulation of anthocyanidins biosynthesis in purple Solanaceous fruit vegetables. Sci Hortic-Amsterdam. 2022, 297.
5. Noda N, Aida R, Kishimoto S, Ishiguro K, Fukuchi-Mizutani M, Tanaka Y, et al. Genetic engineering of novel bluer-colored *chrysanthemums* produced by accumulation of delphinidin-based anthocyanidins. Plant Cell Physiol. 2013;54(10):1648-1695.
6. Noda N. Recent advances in the research and development of blue flowers. Breeding Sci. 2018;68(1):79-87.
7. Morimoto H, Umakoshi K, Sugihara H, Narumi-Kawasaki T, Takamura T, Fukai S. Characteristics of delayed anthocyanidins accumulation of *Dianthus* "Colour Magician" after anthesis. Acta Horticulturae, 2019; 1263:207-214.
8. Sasaki N, Nakayama T. Achievements and perspectives in biochemistry concerning anthocyanin modification for blue flower coloration. Plant Cell Physiol. 2015;56(1):28-40.
9. Luan Y, Tang Y, Wang X, Xu C, Tao J, Zhao D. Tree peony R2R3-MYB transcription factor PsMYB30 promotes petal blotch formation by activating the transcription of the anthocyanidins synthase gene. Plant Cell Physiol. 2022;63(8):1101-1116.
10. Yin X, Zhang Y, Zhang L, Wang B, Zhao Y, Irfan M, et al. Regulation of MYB transcription factors of anthocyanidins synthesis in lily flowers. Front Plant Sci. 2021;12:761668.
11. Li X, Wang Y, Jin L, Chen Z, Jiang J, Jackson A. Development of fruit color in Rubus chingii Hu (Chinese raspberry): A story about novel offshoots of anthocyanidins and carotenoid biosynthesis. Plant Sci. 2021;311:110996.
12. Luan Y, Chen Z, Wang X, Zhang H, Tao J, Zhao D. Herbaceous peony PlACLB2 positively regulates red petal formation by promoting anthocyanidins accumulation. Fron Plant Sci. 2022;13:992529.
13. Mizuta D, Ban T, Miyajima I, Nakatsuka A, Kobayashi N. Comparison of flower color with anthocyanin composition patterns in evergreen azalea. Sci Hortic-Amsteranm. 2009;122(4):594-602.
14. Zhao D, Tao J, Han C, Ge J. Flower color diversity revealed by differential expression of flavonoid biosynthetic genes and flavonoid accumulation in herbaceous peony (*Paeonia lactiflora* Pall.). Mol Biol Rep. 2012;39(12):11263-11275.

15. Zhao D, Tang W, Hao Z, Tao J. Identification of flavonoids and expression of flavonoid biosynthetic genes in two coloured tree peony flowers. Biochem Bioph Res Co. 2015, 459, 450-456.
16. Yamagishi M. How genes paint lily flowers: Regulation of colouration and pigmentation patterning. Sci Hortic-Amsterdam. 2013;163:27-36.
17. Henz A, Debener T, Linde M. Identification of major stable QTLs for flower color in roses. Mol Breeding, 2015;35(10):1-2.
18. Huang Y, Xing X, Tang Y, Jin J, Ding L, Song A, et al. An ERF transcription factor and a FLK homologue jointly modulate photoperiodic flowering in *chrysanthemum*. Plant Cell Environ. 2022; 45(5):1442-1456.
19. Jiang S, Chen M, He N, Chen X, Wang N, Sun Q, et al. MdGSTF6, activated by MdMYB1, plays an essential role in anthocyanidins accumulation in apple. Hortic Res. 2019;6:40.
20. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Full-length transcriptome assembly from RNA-seq data without a reference genome. Nat Biotechnol. 2011;29(7):644-52.
21. Li B, Dewey CN. RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics.2011;12:323.
22. Wang L, Feng Z, Wang X, Wang X, Zhang X. DEGseq: An R package for identifying differentially expressed genes from RNA-seq data. Bioinformatics. 2010;26(1):136-138.
23. Yang LT, Qi YP, Lu YB, Guo P, Sang W, Feng H, et al. iTRAQ protein profile analysis of *Citrus sinensis* roots in response long-term boron-deficiency. J Proteomics. 2013;93:179-206.
24. Wen B, Zhou R, Feng Q, Wang Q, Wang J, Liu S. IQuant: An automated pipeline for quantitative proteomics based upon isobaric tags. Proteomics. 2014;14(20):2280-2285.
25. Savitski MM, Wilhelm M, Hahne H, Kuster B, Bantscheff M. A scalable approach for protein false discovery rate estimation in large proteomic data sets. Mol Cell Proteomics. 2015;14(9):2394-2404.
26. Lan P, Li W, Schmidt W. Complementary proteome and transcriptome profiling in phosphate-deficient *Arabidopsis* roots reveals multiple levels of gene regulation. Mol Cell Proteomics. 2012;11(11): 1156-1166.
27. Shi Q, Zhou L, Wang Y, Li K, Zheng B, Miao K. Transcriptomic analysis of *Paeonia delavayi* wild population flowers to identify differentially expressed genes involved in purple-red and yellow petal pigmentation. PLoS ONE. 2015;10(8):e0135038.
28. Cong L, Zhang F. Genome engineering using CRISPR-Cas9 system. Methods Mol Biol. 2015;1239:197–217.
29. Conrad AC, Mathabatha MF. Characterization and expression analyses of Chalcone Synthase (CHS) and Anthocyanidin Synthase (ANS) genes in *Clivia miniata*. Transcriptomics. 2016;4(2):136.
30. Zhang C, Wang W, Wang Y, Gao S, Du D, Fu J, et al. Anthocyanidins biosynthesis and accumulation in developing flowers of tree peony (*Paeonia suffruticosa*) 'Luoyang Hong'. Postharvest Biol Tec. 2014;97: 11-22.
31. Luan Y, Chen Z, Tang Y, Sun J, Meng J, Tao J, et al. Tree peony PsMYB44 negatively regulates petal blotch distribution by inhibiting dihydroflavonol-4-reductase gene expression. Ann Bot-London, 2023;131(2):323-334.
32. Huang G, Liao X, Han Q, Zhou Z, Liang K, Li G, et al. Integrated metabolome and transcriptome analyses reveal dissimilarities in the anthocyanidins synthesis pathway between different developmental leaf color transitions in *Hopea hainanensis* (Dipterocarpaceae). Front Plant Sci. 2022;13:830413-830413.
33. Janga MR, Campbell LM, Rathore KS. CRISPR/Cas9-mediated targeted mutagenesis in upland cotton (*Gossypium hirsutum* L.). Plant Mol Biol. 2017;94(4-5):349-360.
34. Wu Y, Liang D, Wang Y, Bai M, Tang W, Bao S, et al. Correction of a genetic disease in mouse *via* use of CRISPR-Cas9. Cell Stem Cell. 2013, 13, 659-62.
35. Yin X, Zhang Y, Zhang L, Wang B, Zhao Y, Irfan M, et al. Regulation of MYB transcription factors of anthocyanidins synthesis in lily flowers. Front Plant Sci. 2021;12:761668.
36. Sasaki N, Nishizaki Y, Uchida Y, Wakamatsu E, Umemoto N, Momose M, et al. Identification of the glutathione S-transferase gene responsible for flower color intensity in carnations. Plant Biotechnol. 2012; 29(3):223–227.
37. Miyagawa N, Miyahara T, Okamoto M, Hirose Y, Sakaguchi K, Hatano S, et al. Dihydroflavonol 4-reductase activity is associated with the intensity of flower colors in *Delphinium*. Plant Biotechnol. 2015; 32(3):249–255.
38. Zhang Y, Xu S, Cheng Y, Peng Z, Han J. Transcriptome profiling of anthocyanin-related genes reveals effects of light intensity on anthocyanin biosynthesis in red leaf lettuce. PeerJ. 2018;6:e4607.
39. Ye LJ, Möller M, Luo YH, Zou JY, Zheng W, Wang YH, et al. Differential expressions of anthocyanidins synthesis genes underlie flower color divergence in a sympatric *Rhododendron sanguineum* complex. BMC Plant Biol. 2021;21(1)204.
40. Wang Z, Yu Q, Shen W, El Mohtar CA, Zhao X, Gmitter FG. Functional study of CHS gene family members in citrus revealed a novel CHS gene affecting the production of flavonoids. BMC Plant Biol. 2018;18(1):189.
41. Li M, Cao YT, Ye SR, Irshad M, Pan TF, Qiu DL. Isolation of CHS gene from *Brunfelsia acuminata* flowers and its regulation in anthocyanidins biosysthesis. Molecules. 2016;22(1):44.
42. Zhao D, Jiang Y, Ning C, Meng J, Lin S, Ding W, et al. Transcriptome sequencing of a chimaera reveals coordinated expression of anthocyanidins biosynthetic genes mediating yellow formation in herbaceous peony (*Paeonia lactiflora* Pall.). BMC Genomics. 2014;15(1):689.
43. He H, Ke H, Keting H, Qiaoyan X, Silan D. Flower colour modification of chrysanthemum by suppression of *F3'H* and overexpression of the exogenous Senecio cruentus *F3'5'H* gene. PLoS ONE. 2013;8(11):e74395.
44. Wang X, Chen X, Luo S, Ma W, Li N, Zhang W, et al. Discovery of a *DFR* gene that controls anthocyanidins accumulation in the spiny Solanum group: Roles of a natural promoter variant and alternative splicing. Plant J. 2022;111(4):1096-1109.
45. Huyen DT, Ureshino K, Van, DT, Miyajima I. Co-pigmentation of anthocyanidins-flavonol in the blotch area of *Rhododendron simsii* planch. flowers. Hortic J. 2016;85(3):232–237.
46. Chen Y, Chen Y, Shi C, Huang Z, Zhang Y, Li S, et al. SOAPnuke: A mapreduce acceleration-supported software for integrated quality control and preprocessing of high-throughput sequencing data. Gigascience. 2017;7(1):gix120.
47. Zhang Y, Cheng Y, Xu S, Ma H, Han J, Zhang Y. Tree peony variegated flowers show a small insertion in the *F3'H* gene of the acyanic flower parts. BMC Plant Biol. 2020;20(1):211.
48. Yang FS, Nie S, Liu H, Shi TL, Tian XC, Zhou SS, et al. Chromosome-level genome assembly of a parent species of widely cultivated azaleas. Nat Commun. 2020;11(1)5269.
49. El-Sharkaw I, Liang D, Xu K. Transcriptome analysis of an apple (Malus × domestica) yellow fruit somatic mutation identifies a gene network module highly associated with anthocyanin and epigenetic regulation. J Exp Bot. 2015;66(22):7359-7376.
50. Liu XJ, Chuang YN, Chiou CY, Chin DC, Shen FQ, Yeh KW. Methylation effect on chalcone synthase gene expression determines anthocyanidins pigmentation in floral tissues of two *Oncidium orchid* cultivars. Planta. 2012;236(2):401–409.

51. Sapir Y, Gallagher MK, Senden E. What maintains flower colour variation within populations? Trends Ecol Evol. 2021;36(6):507-519.

52. Sun J, Wu Y, Shi M, Zhao D, Tao J. Isolation of *PlANS* and *PlDFR* genes from herbaceous peony (*Paeonia lactiflora* Pall.) and its functional characterization in Arabidopsis and tobacco. Plant Cell Tiss Org. 2020; 141(2):435–445.