Comparative transcriptome analyses reveal genes related to pigmentation in the petals of a flower color variation cultivar of *Rhododendron obtusum*

Xiaobo Sun
Institute of Leisure Agriculture, Jiangsu Academy of Agricultural Sciences & Jiangsu Key Laboratory for Horticultural Crop Genetic Improvement

Lisi He
Institute of Leisure Agriculture, Jiangsu Academy of Agricultural Sciences & Jiangsu Key Laboratory for Horticultural Crop Genetic Improvement

Zhenhao Guo
Institute of Leisure Agriculture, Jiangsu Academy of Agricultural Sciences & Jiangsu Key Laboratory for Horticultural Crop Genetic Improvement

Jiale Su
Institute of Leisure Agriculture, Jiangsu Academy of Agricultural Sciences & Jiangsu Key Laboratory for Horticultural Crop Genetic Improvement

Xiaoqing Liu
Institute of Leisure Agriculture, Jiangsu Academy of Agricultural Sciences & Jiangsu Key Laboratory for Horticultural Crop Genetic Improvement

Huimin Zhou
Institute of Leisure Agriculture, Jiangsu Academy of Agricultural Sciences & Jiangsu Key Laboratory for Horticultural Crop Genetic Improvement

Chang Li (changli529@foxmail.com)
Institute of Leisure Agriculture, Jiangsu Academy of Agricultural Sciences & Jiangsu Key Laboratory for Horticultural Crop Genetic Improvement

Haidong Gao
Genepioneer Biotechnologies Co. Ltd

Research Article

Keywords: Rhododendron, anthocyanin biosynthesis, transcriptome, flower coloration, R2R3-MYB

Posted Date: March 3rd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-263287/v1

Version of Record: A version of this preprint was published at Molecular Biology Reports on January 21st, 2022. See the published version at https://doi.org/10.1007/s11033-021-07070-w.

Abstract

Rhododendron is an important woody ornamental plant and breeding varieties with different colors is vital research goal. In this study, a flower color variation cultivar 'Yanzhi Mi' (pink petals) and the wild-type (WT) cultivar 'Dayuanyangjin' (white petals with pink stripes) were used as research objects, the pigment and transcriptome of their petals during different flower development (stage S1, S2, S3, S4 and S5) were analyzed and compared. The results showed that the derivatives of cyanidin, peonidin and pelargonidin may be responsible for the pink of mutant petals and S2 stage (buds showing color at the top but with the scales still present) was the key stage of flower color formation of mutant. In total, 412,910 transcripts and 2,780 differentially expressed genes (DEGs) were identified in pairwise comparisons of WT and mutant petals. GO and KEGG enrichment analyses of the DEGs showed that the 'DNA-binding transcription factor activity', 'Flavonoid biosynthesis' and 'Phenylpropanoid biosynthesis' were more active in mutant petals. Early anthocyanin pathway candidate DEGs (*CHS3-CHS6*, *CHI*, *F3H*s and *F3'H*) were strongly correlated and up-regulated expression in mutant petals than in WT petals at S2 stage. These genes may be the key structural genes for the pink coloration of mutant petals. In the petals of mutant, two *R2R3-MYB* unigene (TRINITY_DN59015_c3_g2 and TRINITY_DN49281_c1_g6) were identified as repressors involved in anthocyanin regulation and were significantly down-regulated at S2 stage. This study shed light on the biochemistry and genetic mechanisms underlying the flower coloration in two *Rhododendron obtusum* cultivars.

Introduction

Rhododendron is not only the largest genus of the family *Ericaceae*, but also important ornamental woody plants, contains approximately 1,000 species and thousands of commercial hybrids [1]. *Rhododendron* is a familiar ornamental shrub that is widespread around the world with beautiful flowers. Flower color, as an important ornamental trait, greatly affects the economic value of *Rhododendron*. There is a remarkably broad range of *Rhododendron* flower colors, including white, red, pink, purple, yellow, green, and blue and so on[2]. Previous studies have shown that many factors (cell shape, copigmentation, pH, environmental conditions, and so forth) affect petal colors, but in which, pigment composition is the most important [3]. The petals of *Rhododendron* flowers contain flavonols and anthocyanins as the major pigments [4]. The main flavonols of *Rhododendron* petals are quercetin, kaempferol, and myricetin, though many others exist [5-7]. The flavonol composition and content have some effect on flower color by copigmentation, but might not play a major role in coloration [8,9].The major anthocyanidins of *Rhododendron* petals are cyanidin, peonidin, delphinidin, petunidin, malvidin, and pelargonidin [4,10,11]. Anthocyanins exist in the form of glycosides, which consist of an aglycone backbone to which sugars are attached. Most of anthocyanin glycoside compounds are 3-monoglycosides or 3,5-diglycosides [12]. Anthocyanin constitution and content are fundamental elements that determine *Rhododendron* flower colors [4,9,13]. Therefore, the study of anthocyanin biosynthesis in the petals of *Rhododendron* and the related regulatory mechanisms is important to elucidate the formation of different flower colors in *Rhododendron*.

The anthocyanin biosynthesis pathway is conserved in higher plants [14,15]. The biosynthesis of anthocyanin and the regulatory mechanisms of transcription factors have been carefully studied in most model plants [16]. The enzyme-coding structural genes, including phenylalanine-ammonia lyase (*PAL*), 4-coumaryl: CoA ligase (*4CL*), chalcone synthase (*CHS*), chalcone isomerase (*CHI*), flavonoid 3-hydroxylase (*F3H*), flavonoid 3'-hydroxylase (*F3'H*), flavonoid 3'5'-hydroxylase (*F3'5'H*), dihydroflavonol 4-reductase (*DFR*), anthyocyanidin synthase (*ANS*), and UDP-glucose:flavonoid 3-O-glucosyltransferase (*UFGT*) are involved in anthocyanins biosynthesis [17,18]. The expression of these structural genes is regulated at transcriptional level by a complex of transcription factors, *R2R3-MYB*, basic helix–loop–helix (*bHLH*), and WD40 proteins (*WD40*) (named *MBW* complex), which can bind to the promoters of these structural genes [19,20,21]. Among these transcription factors, the *R2R3-MYB* transcription factors play a major role in determining the spatio-temporal expression of anthocyanin biosynthetic genes [22,23].

So far, there have been a few reports on the molecular mechanisms of flower colors and color patterning in *Rhododendron*. *CHS* and *DFR* genes were cloned from *R. simsii* hybrids, and the expression of these two genes has been examined in eight azalea cultivars, and the results showed that there was no significant correlation between flower color phenotype and expression levels of the *CHS* and *DFR* genes [24,25]. The main enzymes of the anthocyanin biosynthesis pathway in azalea includes *CHS*, *CHI*, *F3H, F3′H*, *F3′5′H*, *DFR*, *ANS*, and flavonol synthase (*FLS*) [19]. The genes for these biosynthetic enzymes were isolated from evergreen azalea *R. × pulchrum* Sweet 'Oomurasaki' and the expression analysis showed that only *F3′5′H* gene expression was strongly correlated to the accumulation of total anthocyanin in petals, but expression of other genes are unclear for anthocyanin synthesis [26]. A comparative study between the purple-flowered 'Oomurasaki' and its red-flowered mutant showed that 'Oomurasaki' contained anthocyanidins from both the cyanidin and delphinidin series, whereas the red flower mutant contained only pigments from the cyanidin series. The transcript levels of *F3′5′H* in the mutant were 0.14-fold those in 'Oomurasaki', suggesting that low transcript levels of *F3′5′H* in the red-flowered mutant resulted in no accumulation of delphinidin-derived anthocyanin [10]. Despite these reports, it is still largely unknown what factors regulate flower pigmentation in *Rhododendron*.

Su (2012) bred a flower color variation cultivar 'Yanzhi Mi' (pink petals) which was from the bud spot of the *Rhododendron obtusum* 'Dayuanyangjin' (white petals with pink stripes) [27]. Compared with 'Dayuanyangjin', the mutant has only difference in flower color, and other genetic background is the same with 'Dayuanyangjin', which provides an ideal experimental material for studying the mechanism of *Rhododendron* petal coloration. In the current study, we qualitatively and quantitatively characterized and compared pigmentation in the petals of the mutant 'Yanzhi Mi' and its wild-type 'Dayuanyangjin' at five stages of flower development. Then, we used the Illumina sequencing platform to conduct the transcriptome sequencing analysis of mixed RNA separately extracted from petals of 'Yanzhi Mi' and 'Dayuanyangjin' at the same five stages of flower development. After analyzing the data, we identified some key candidate genes related to anthocyanin biosynthesis in *R. obtusum*. These transcriptome sequences may provide a valuable genomic resource to our further understanding of the molecular mechanisms of color formation in *R. obtusum*.

Results

Qualitative and quantitative pigments analyses

The total anthocyanin contents in ‘Yanzhi Mi’ and ‘Dayuanyangjin’ petals were similar at S1 stage. At other four floral development stages, the total anthocyanin accumulated significantly higher levels in ‘Yanzhi Mi’ petals, especially at S2 stage; the anthocyanin contents of ‘Yanzhi Mi’ petals were 35.9 times higher than those of ‘Dayuanyangjin’ petals at S2 stage (Fig. 1B). In ‘Yanzhi Mi’ petals, the anthocyanin contents were highest at S2 stage, decreased rapidly at S3 stage, and remained unchanged at S4 and S5 stages. In contrast, the anthocyanin contents of ‘Dayuanyangjin’ petals were highest at S1 stage, decreased rapidly at S2 stage, and remained unchanged until S5 stages. The flavonoid contents in ‘Dayuanyangjin’ petals were higher compared to in ‘Yanzhi Mi’ petals at S1 stage, and were similar at other four floral development stages (Fig. 1B). The results indicated that anthocyanins might play vital roles in pink flower color formation of ‘Yanzhi Mi’. Therefore, we went further to identify the anthocyanin composition of petals of ‘Yanzhi Mi’ and ‘Dayuanyangjin’. In total, seven anthocyanins were identified in the petals of the two cultivars, which included cyanidin-3-glucoside-5-xyloside, cyanidin-3-rutinoside, cyanidin-3-glucoside, cyanidin-3-xyloside, peonidin -3-glucoside-5-xyloside, peonidin 3, 5-diglucoside and Pelargonidin-3-diphrhamnoside. Among seven anthocyanins, Pelargonidin-3-diphrhamnoside, cyanidin-3-glucoside-5-xyloside and cyanidin-3-xyloside are the three most abundant anthocyanins of the petals of ‘Yanzhi Mi’ at S2 stage (Fig. 1C and Table S2). These results suggested that the derivatives of cyanidin, peonidin and pelargonidin might be responsible for the pink of ‘Yanzhi Mi’ petals.

Transcriptome sequencing, de novo assembly and annotation

The color variation of the petals of ‘Yanzhi Mi’ may be related with the accumulation of anthocyanins, which prompts us to further investigate the transcrptional variation of the candidate genes in the anthocyanin biosynthesis pathway. The transcriptomes of the petals of ‘Yanzhi Mi’ and ‘Dayuanyangjin’ at five floral development stages were obtained by Illumina technologe. Following data cleaning and quality checking, 722,321,099 reads with Q30 values ≥90.00% were obtained from the 30 libraries. Among the clean reads, >96% had quality scores at Q20 level (Table S3). The sequencing raw data were deposited in the NCBI database and can be accessed in the Short Reads Archive (SRP304986).

The clean reads were 216.70 Gb (total length), which was equivalent to ~310 fold coverage of the genome of *R. obtusum* (about 0.7 Gb). All clean reads were merged and de novo assembled using the Trinity platform software, resulting in the generation of 412,910 transcripts with an average length of 1276.10 bp and a N50 length of 1881 bp (Table 1). Most of the reads could be mapped back to the assembled transcripts and the total length of all transcripts was approximately 527 Mb. These transcripts were further subjected to cluster and assembly analyses and resulted in 137,018 unigenes with an average length of 1069.22 bp and an N50 value of 1549 bp, of which 34.19% unigenes (46,853) were more than 1kbin length (Table 1).

Table 1. Length distribution of assemble transcripts and unigenes

Nucleotides length (bp)	Transcripts	Unigenes
300-500	104,752 (25.37%)	45,881 (33.49%)
500-1000	122,396 (29.64%)	44,284 (32.32%)
1000-2000	106,483(25.79%)	28,573 (20.85%)
>2000	79,279 (19.20%)	18,280 (13.34%)
Total Number	412,910	137,018
Total Length	526,912,718	146,502,783
N50 Length	1,881	1,549
Mean Lenth	1276.10	1069.22

Sequence similarity search was carried out against protein sequences available in various databases using the BLASTX algorithm with an E value threshold of 1e-10. Among the 137,018 unigenes, 77,578 (56.6%), 45,902 (33.5%), 42,496 (31.0%), 28,005 (20.4%), 18,018 (13.2%), 49,672 (36.3%) and 44,816 (32.7%) were blasted into the NCBI and database of Non-Redundant Protein Sequences (NR), Gene Ontology (GO), Eukaryotic Ortholog Groups (KOG), Cluster of Orthologous Groups of Proteins (COG), Kyoto Encyclopedia of Genes and Genomes (KEGG), Swiss-Prot protein (SwissProt), Protein Family (Pfam), respectively (Table S4).

DEG identification and GO and KEGG enrichment

The criterion false discovery rate (FDR) value ≤ 0.05 and $|\log2 (\text{fold change})| \geq 1$ was used as the standard for pairwise differential expression analysis of the five sample groups (Y-S1vs D-S1, Y-S2 vs D-S2, Y-S3 vs D-S3, Y-S4 vs D-S4 and Y-S5 vs D-S5) (Fig. 2). A total of 2780 differentially expressed genes (DEGs) were identified in ‘Yanzhi Mi’ vs ‘Dayuanyangjin’ comparison at the five floral development stages. No DEG was found at S1 stage. 2546 DEGs were identified with 1327 upregulated and 1219 downregulated genes in Y-S2 vs D-S2 comparison. The largest number of upregulated DEGs and downregulated DEGs were both observed in this comparison. A sharp decline in the total number of DEGs between ‘Yanzhi Mi’ vs ‘Dayuanyangjin’ at S3 stage was observed with a total of 213 upregulated and 33 downregulated genes. The total number of DEGs in Y-S4 vs D-S4 (44 DEGs with 31 upregulated and 13 downregulated) was similar to Y-S5 vs D-S5 (46 DEGs with 27 upregulated and 19 downregulated), and there are very few DEGs in both of them (Fig. 2). Overall, the number of upregulated genes in ‘Yanzhi Mi’ petals was greater than the number of downregulated genes during floral development. These results indicate that the S2 stage is the key stage to determine the flower color difference between ‘Yanzhi Mi’ and ‘Dayuanyangjin’.

To illustrate the main biological functions of DEGs, we conducted Gene Ontology (GO) enrichment and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses. A GO enrichment analysis provides a description of gene products in terms of their associated Biological Process (BP), Cellular Component (CC), and Molecular Function (MF) [28].

Totally, All DEGs were assigned into 53 GO categories and 1,607 DGEs in Y-S2 vs D-S2 were also categorized into 53 GO categories. 42 significant enriched GO categories (p-value ≤ 0.05) were identified in Y-S2 vs D-S2 and the most highly enriched terms were “cell” including 1169 DEGs (Fig. 3A). Significantly up-regulated DEGs were observed in the six GO categories (‘DNA-binding transcription factor activity; GO:0003700’, ‘sequence-specific DNA binding; GO:0043565’, ‘DNA replication initiation; GO:0006270’, ‘flavonoid biosynthetic process; GO:0009813’, ‘MCM complex; GO:0042555’and ‘naringenin-chalcone synthase activity;GO:0016210) at S2 stage (Fig. 3B). Of these, four GO categories (DNA-binding transcription factor activity, sequence-specific DNA binding, flavonoid biosynthetic process and naringenin-chalcone synthase activity) were found to be related to the color of flower. GO categories (pectin catabolic process; GO:0045490, apoplast; GO:0048046, chitinase activity; GO:0004568 and pectatelyase activity; GO:0030570) were found to be more rich among down-regulated DEGs (Fig. 3B).

All DEGs were annotated into 94 metabolic pathways using the KEGG database. Compared with ‘Dayuanyangjin’, 427 unigenes were differentially expressed in ‘Yanzhi Mi’ petals at S2 stage, including 220 up-regulated unigenes and 207 down-regulated unigenes (Table S5). These DEGs were categorized into 92 pathways and KEGG pathway analysis indicated that ‘Flavonoid biosynthesis’ and ‘Phenylpropanoid biosynthesis’ were the two most enriched pathways, which included 20 and 34 unigenes, respectively (Fig. 4).

According to the results of pigment determination and GO and KEGG annotation of DEGs, a total of 62 unigenes were involved in the anthocyanin biosynthesis pathway (Table 2 and Table S6).

Table 2. Candidate anthocyanin structural genes and transcription factor genes identified among DEGs in the petals.

Gene	Protein	No. All[a]	No. Up[b]	No. Down[c]
PAL	Phenylalanine ammonia-lyase	1	1	0
4CL	4-coumarate--CoA ligase	3	2	1
CHS	Chalcone synthase	6	6	0
CHI	Chalcone isomerase	1	1	0
F3H	Flavanone 3-hydroxylase	2	2	0
F3'H	Flavonoid 3'-hydroxylase	1	1	0
F3'5'H	Flavonoid 3'5'-hydroxylase	1	1	0
FLS	Flavone synthase	3	2	1
DFR	Dihydroflflavonol 4-reductase	2	2	0
ANS	Anthocyanidin synthase	1	1	0
FLS	flavonol synthase	3	2	1
3GT	Anthocyanidin 3-O-glucosyltransferase	2	1	1
5,3GT	Anthocyanidin 5,3-O-glucosyltransferase	1	0	1
R2R3-MYB	*R2R3-MYB* transcription factors	26	14	12
bHLH	Basic helix–loop–helix transcription factors	9	8	1

[a] The total number of unigens analyzed; [b] The number of unigens with expression significantly upegulated in ‘Yanzhi Mi’ petals compared with ‘Dayuanyangjin’ petals; [c] The number of unigens with expression significantly downregulated in ‘Yanzhi Mi’ petal compared with ‘Dayuanyangjin’ petals.

Expression Patterns of Genes Involved in Anthocyanin Biosynthesis Pathway

We identified 16 DEGs from 27 candidate anthocyanin structural genes, which encoded seven putative enzymes involved in anthocyanin biosynthesis. The synthetic pathway and expression profiles of these gens were shown in Fig. 5A. These genes included *CHS* (6), *CHI* (1), *F3H* (2), *F3’H* (1), *DFR* (2), *ANS* (1), and *GT* (3), and exhibited significant differential expression in the petals of ‘Yanzhi Mi’ vs ‘Dayuanyangjin’. We found that except for *GT*s, all of genes encoding six enzymes showed similar expression levels in the petals of ‘Yanzhi Mi’ and ‘Dayuanyangjin’ at S1 stage, and three genes encoding *GT*s all showed lower expression levels in ‘Yanzhi Mi’ petals than in ‘Dayuanyangjin’ petals. At S2 stage, the expression levels of thirteen gene (except for *GT*s) were higher in ‘Yanzhi Mi’ petals than in ‘Dayuanyangjin’ petals with the maximum multiple 2.82 (*DFR1*) and the minimum multiple 1.75 (*F3H2*) ; three *GT*s showed lower expression levels in the petals of ‘Yanzhi Mi’. At S3 stage, thirteen gene showed higher transcript levels in ‘Yanzhi Mi’petals than in ‘Dayuanyangjin’ petals with the maximum multiple 2.9 (*3GT1*) and the minimum multiple 1.19 (*CHI*); *F3H1* and *F3H2* showed similar expression levels in the petals of ‘Yanzhi Mi’ and ‘Dayuanyangjin’, but *5,3GT* showed lower expression levels in ‘Yanzhi Mi’ petals. However, only two genes *CHS5* (TRINITY_DN49577_c0_g1) and *F3’H* (TRINITY_DN68439_c2_g2) showed

higher expression levels at S4 stage in ‘Yanzhi Mi’ petals compared to in ‘Dayuanyangjin’ petals. At S5 stage, the expression of seven genes was similar in the petals of ‘Yanzhi Mi’ and ‘Dayuanyangjin’, eight gene higher in ‘Yanzhi Mi’ cpetals, and *5,3GT* (TRINITY_DN64021_c3_g2) was not expressed in ‘Yanzhi Mi’ and ‘Dayuanyangjin’ petals.

In ‘Yanzhi Mi’, many genes of the anthocyanin synthetic pathway showed the maximum expression at S2 stage and then decreased at each subsequent developmental stage. For example, *CHS3, CHS4, CHS5, CHS6, CHI, F3H1*, and *F3’H* were expressed more at S2 stage than at other floral development stages, of which *F3’H* (TRINITY_DN68439_c2_g2) was 1.14-, 4.22-, 5.20- and 5.83-fold upregulated at S2 stage, respectively, compared to its expression at S1, S3, S4 and S5 stages. However, in ‘Dayuanyanjin’ flowers, the expression levels of ten genes (*CHSs, CHI, F3H*s and *F3’H*) were highest at S1 stage. For example, *F3’H* (TRINITY_DN68439_c2_g2) was 2.21-, 6.13-, 6.24- and 5.93-fold upregulated at S1 stage, respectively, compared to its expression at S2, S3, S4 and S5 stages. The *DFR* genes (TRINITY_DN67151_c2_g1, TRINITY_DN67151_c2_g3) , *3GT1* (TRINITY_DN52552_c0_g1,) and *3GT2* (TRINITY_DN62746_c0_g1) in ‘Dayuanyanjin’ petals had different expression pattern and were strongly expressed at S4 stage, S5 stage and S4 stage, respectively, which were the same as in ‘Yanzhi Mi’ petals.

Correlations of the expression profiles of the above-mentioned 16 genes were detected by Pillet’s method [29], and positive correlations between these genes (Pearson’s $r > 0.65$, $P < 0.05$) were identified (Fig. 5B,5C). Compared with the gene expression correlation map of ‘Dayuanyangjin’ petals, we found eight early anthocyanin pathway genes (including *CHS3, CHS4, CHS5, CHS6, CHI, F3H1, F3H2* and *F3’H*) were strongly related and shared more than two positive correlations with other genes in ‘Yanzhi Mi’ petals. Interestingly, a late anthocyanin pathway gene, *3GT1*, showed different regulation processes and share negative correlation with six early anthocyanin pathway genes (*CHS3-CHS6, F3H1* and *F3H2*). This correlation pattern clearly suggests that these anthocyanin pathway genes might be co-regulated by the same transcription factors.

The gene expression of transcription factor related to flower color formation

The expression of most of the structural genes in the anthocyanin biosynthetic pathway was coordinately regulated by a ternary complex (*R2R3-MYB*, *bHLH*, and *WDR* transcription factors) [30]. The above-mentioned analysis showed that S2 stage is the key stage to determine the flower color difference between ‘Yanzhi Mi’ and ‘Dayuanyangjin’. Therefore, we selected the transcription factors related to flower color formation and differentially expressed in two cultivars at S2 stage for further analysis. We identified 35 DEGs from the DEG data encoding candidate transcription factors, including 26 unigenes encoding *R2R3-MYB*s and 9 encoding *bHLH*s (Table S6).

26 predicted *MYB* genes were divided into 2 clusters based on their expression profiles in ‘Yanzhi Mi’ and ‘Dayuanyanjin’ (Fig. 6A). The expression levels of *MYB*s in cluster 1 were more abundant in ‘Dayuanyanjin’ petals than in ‘Yanzhi Mi’ petals. By contrast, the transcription levels of *MYB*s in cluster 2 were significantly higher in ‘Yanzhi Mi’ petals compared to those in ‘Dayuanyanjin’ petals. 3 unigenes were differentially expressed with absolute log2 (FC values) >2, and they all were down-regulated genes

(TRINITY_DN62378_c2_g1, TRINITY_DN49281_c1_g6 and TRINITY_DN59015_c3_g2). Notably, TRINITY_DN59015_c3_g2 and TRINITY_DN49281_c1_g6 was found to be closely related to *VviR2R3-MYB* in grape and *EOBI* in petunia (Fig. S1) [31,32];

The expression profiles of 9 predicted *bHLH* genes were divided into 2 groups (Fig. 6B). The candidate genes in Cluster 1 showed higher expression levels in 'Dayuanyanjin' petals than in 'Yanzhi Mi' petals. In contrast, the candidate genes in Cluster 2 exhibited high expression levels in 'Yanzhi Mi' petals than in 'Dayuanyanjin' petals. 2 unigenes were differentially expressed with absolute log2 (FC values) >1.5, including one down-regulated genes (TRINITY_DN49581_c1_g2) and one upregulated genes (TRINITY_DN50930_c0_g3). TRINITY_DN46669_c0_g1 in the cluster 2 was found to be closely related to VcobHLH040 (Fig. S2) [33], the gene was expressed significantly higher in 'Yanzhi Mi'petals.

qPCR analysis validated the differentially expressed genes of the flavonoid pathway

To confirm the transcripts obtained in the sequencing analysis, the reliability of the comparative transcriptional data was further verified. A total of 10 structural genes transcripts and transcription factor transcripts that might be involved in anthocyanin biosynthesis were randomly selected for qPCR test. The relative expression levels of these transcripts at five floral development stages of 'Yanzhi Mi', and TPM (transcripts per million clean tags) values of these transcripts obtained from the sequencing data from the 15 samples of 'Yanzhi Mi' are shown in Fig. 7. These results showed that the expression patterns obtained by qPCR are consistent with the digital expression data, and indicated that the transcriptomic data used in this study for anthocyanin synthesis-related gene analysis were reliable and highly reproducible.

Discussion

Flower color is one of the most important ornamental traits of rhododendrons, and the study of the mechanism underlying the formation of different flower colors is of great significance. The major anthocyanidins in rhododendron petals are cyanidin, peonidin, delphinidin, petunidin, malvidin, and pelargonidin [4,10,11]. Cyanidin and peonidin are responsible for red flowers, and delphinidin, petunidin, and malvidin can result in blue flowers [34]. Seven anthocyanins were identified in petals of 30 *Rhododendron* species, and the red-flowered species mainly contained 2 cyanidin monoglycosides [9]. A total of 5 anthocyanin compounds were identified in 10 Rhododendron species and cyanidin derivatives were the major anthocyanins in most of the red-purple group and in *R. triflorum* (red flower) [35]. The flowers of *R.schlippenbachii Maxim* contained 2 anthocyanins, cyanidin-3,5-diglucoside and cyanidin-3-sambubioside, and among 3 different colored flowers of *R. schlippenbachii*, the red flowers exhibited higher amount of total anthocyanins than the violet flowers, and no anthocyanins were detected in the white flowers [13]. Previous studies did not clarify the presence of pink flowers, but some researchers considered pink to be a gradation in pigment [36]. De Keyser et al. also confirmed in rhododendron that pink can be seen as (carmine) red at a lower intensity level by means of image analysis [25]. In this study, we compared the pigmentation of the petals of the mutant 'Yanzhi Mi' and its wild-type 'Dayuanyangjin'

during floral development. According to HPLC-Q-TOF-MS analysis, ‘Yanzhi Mi’ and ‘Dayuanyangjin’ contained similar contents of flavonoids, and the same compositions but significantly different contents of anthocyanins. The pink petals of ‘Yanzhi Mi’ contained higher content of anthocyanins during floral development, which are mainly composed of seven derivatives of cyanidin, pelargonidin and peonidin. The results are consistent with previous reports that the presence of cyanidin and peonidin leads to the formation of red petals [34]. The white petals of ‘Dayuanyangjin’ mainly contained flavonoids, and the trace levels of anthocyanins in the white petals may be due to the pink stripes in them.

The biosynthesis of anthocyanin is crucial to expand a wide range of flower colors. Previous studies have shown that *C4H, CHS, F3H, F3H, F3′5′H, DFR* and *ANS* are the key enzymes involved in the biosynthesis of anthocyanin for the determination of different flower colors in plants [37]. Flower color regulation was thought to achieve mainly via a coordinated transcriptional control of structural genes [38, 39]. In the study, we identified 16 DEGs from the transcriptomes of the petals of ‘Yanzhi Mi’ vs ‘Dayuanyangjin’, which encoded seven putative enzymes involved in anthocyanin biosynthesis. By comparing their gene expression profiles during floral development, we found that except for *GT*s, all of genes encoding six enzymes showed similar expression levels in the petals of ‘Yanzhi Mi’ and ‘Dayuanyangjin’ at S1 stage. However, most of the genes showed significantly higher transcript levels in ‘Yanzhi Mi’ petals than in ‘Dayuanyangjin’ petals at S2 and at S3 stages, especially at S2 stage. The comparative analysis of transcriptome of petals of ‘Yanzhi Mi’ vs ‘Dayuanyangjin’ showed that from S1 to S5 stages, the DEGs were mainly concentrated at S2 stage, which accounted for 91.6% of the total DEGs. Based on the results that no DEGs were found and anthocyanin contents were similar at S1 stage of petals of ‘Yanzhi Mi’ vs ‘Dayuanyangjin’, we can conclude that S2 stage was the key stage of flower color formation of ‘Yanzhi Mi’.

In *R.×pulchrum* flowers buds at S1 stage, five genes (*CHS, CHI, F3H, DFR*, and *ANS*) were expressed, but there was no anthocyanin to be detected. Of these genes, the expression of *CHS, F3H*, and *ANS* was the highest at S1 stage and decrease during flower development; *CHI* and *DFR* transcripts expressed with the highest levels at S2 stage [26]. However, anthocyanin content was highest at candle stage (buds showing color at the top but without any scales left) [24,26]. Anthocyanin synthesis and gene expression in *R.×pulchrum* petals were not coincident. In our study, only two *CHS* candidate genes (*CHS1* and *CHS2*) accumulated the highest transcripts at S1 stage of ‘Yanzhi Mi’ flower, and other early pathway candidate genes (*CHS3-6, CHI, F3H*s, and *F3′H*) in ‘Yanzhi Mi’ petals were strongly expressed at S2 stage and then declined in the next stages of floral development. The maximum expression of late pathway candidate genes (*DFR*s, *ANS* and *UFGT*s) were at S3, S4 and S5 stages, respectively (Fig. 5A). Lowest content of anthocyanin was detected at S1 stage of ‘Yanzhi Mi’ flowers, and highest content of anthocyanin was found at S2 stage of ‘Yanzhi Mi’ flowers (Fig. 1B). However, in ‘Dayuanyanjin’, the expression levels of all early pathway candidate genes (*CHS1-6, CHI, F3H*s, and *F3′H*) were highest at S1 stage, and decreased at each subsequent developmental stage. The late pathway candidate genes (*DFR*s, *ANS* and *UFGT*s) were expressed highly at S2, S4 and S5 stages, respectively (Fig. 5A). The anthocyanin content of ‘Dayuanyangjin’ petals at S1 stage, which was highest among five floral development stages, were similar to that of ‘Yanzhi Mi’ petals at S1 stage. Combined with the result above mentioned, we can suggest that

in ‘Yanzhi Mi’ and ‘Dayuanyangjin’, anthocyanin synthesis and gene expression in petals were coincident at the key stage of flower color formation, S2 stage, which is different from that in *R.× pulchrum* flowers. The flower color formation of ‘Yanzhi Mi’ and ‘Dayuanyangjin’ began to diverge at S2 stage. The anthocyanin content and gene expression levels of ‘Yanzhi Mi’ petals at S1 stage were similar to those of ‘Dayuanyangjin’ petals at S1 stage. However, at S2 stage, the transcript levels of some early pathway candidate genes increased in ‘Yanzhi Mi’ and the anthocyanin content also increased with the increase of the gene expression; in ‘Dayuanyangjin’the expression levels of the early pathway candidate genes decreased shapely at S2 stage, and the anthocyanin content also decreased with the decrease of the gene expression. The combination of the early pathway genes (*CHS, CHI, F3H*, and *F3’H*) is strongly correlated to differentiating between pink flowers of ‘Yanzhi Mi’ and white flowers of ‘Dayuanyangjin’. Therefore, we speculated that *CHS, CHI, F3H* and *F3’H* were the most likely candidate responsible for flower coloration in ‘Yanzhi Mi’.

R2R3-MYB, bHLH and *WD40* proteins are the main transcription factors responsible for regulating the expression of structural genes in the anthocyanin biosynthesis pathway [30]. Often, these transcription factors form a complex (the *MYB-bHLH-WDR* or *MBW* complex) that can coordinately activate or repress the expression of a set of target genes in the anthocyanin biosynthesis pathway to modulate pigment production [21,40]. In *Phalaenopsis spp.*, *PeMYB11* has been shown to activate the expression of the anthocyanin biosynthetic genes *PeF3H5*, *PeDFR1* and *PeANS3* [41]. In *Dendrobium hybrids orchid*, *DhMYB2* has been reported to interact with *DhbHLH1* to regulate the expression of *DhDFR* and *DhANS* [42]. In addition, *MiMYB1* has been shown to interact with *MibHLH2* and *MiWDR1* to activate the transcription of the anthocyanin biosynthetic genes *MiF3H*, *MiDFR*, and *MiANS* in *Matthiola incana* flowers [43]. Moreover, some *R2R3-MYB* repressors have also been identified in plants in recent years, including *PtrMYB182* and *PtrMYB57* from poplar, *PpMYB17-20* from peach, *MdHB1*, *MdMYB16*, and *MdMYB15L* from apple, *NtMYB2* from Chinese narcissus, and *CmMYB#7* from chrysanthemum [44-51]. In apple, *MdHB1* have been supposed to constrain *MdMYB10*, *MdbHLH3*, and *MdTTG1* to the cytoplasm, and then represses the transcription of *MdDFR* and *MdUFGT* indirectly and resulted in reduced anthocyanin biosynthesis [47].

In the present study, 26 *R2R3-MYB*s and 9 *bHLH*s were identified and their expression profiles at S2 stage were analyzed. The *R2R3-MYB*s were used to construct a phylogenetic tree with those anthocyanin-related *R2R3-MYB*s in other plants. 2 *R2R3-MYB* transcription factors (TRINITY_DN59015_c3_g2 and TRINITY_DN49281_c1_g6), down-regulated significantly in ‘Yanzhi Mi’ petals at S2 stage, was found to be closely related to the anthocyanin regulated gene *VviR2R3-MYB* in grape and *EOBI* in petunia [31,32]. In addition, a *bHLH* unigene (TRINITY_DN46669_c0_g1) was up-regulated in ‘Yanzhi Mi’ petals at S2 stage and closely related to the known plant *bHLHs* involved in anthocyanin biosynthesis regulation, *VcobHLH040* [33]. Further study on these candidate transcription factors are needed to confirm our hypothesis.

Conclusions

In conclusion, a combination of analytical chemistry and transcriptome analysis were performed to uncover the molecular basis underlying the pink pigmentation in the flowers of mutant ‘Yanzhi Mi’ and white pigmentation in the flowers of wide type ‘Dayuanyangjin’. Seven derivatives of cyanidin, pelargonidin and peonidin were deemed to be the main contributors to pink color formation. We identified the potential candidate genes encoding key enzymes in the anthocyanin biosynthetic pathway, such as *CHS, CHI, F3H,* and *F3’H*, which are significantly differently expressed in pink flower of ‘Yanzhi Mi’ and white flowers of ‘Dayuanyangjin’ at the key stage of flower color formation, S2 stage. These genes were the most likely candidate responsible for flower coloration in ‘Yanzhi Mi’. In particular, 2 *R2R3-MYB* repressors that might be involved in pink flower coloration regulation in ‘Yanzhi Mi’ was also identified. The most likely cause of the color variation in the flowers of ‘Yanzhi Mi’ was proposed and discussed.

Materials And Methods

Plant materials

R. obtusum ‘Dayuanyangjin’ (D) and ‘Yanzhi Mi’ (Y) were grown in a greenhouse at the Jiangsu Academy of Agricultural Sciences (Nanjing, China 118.881E, 32.039N). In March of 2019, the flowers at five developmental stages (closed buds (S1 stage), buds showing color at the top but with the scales still present (S2 stage), initial flowering stage (S3 stage), full flowering stage (S4 stage) and last flowering stage (S5 stage)) from ten plants of each cultivar were sampled (Fig. 1A). The petals were cut off immediately and frozen in liquid nitrogen then stored at - 80°C for pigment measurement and RNA-seq analyses. Three biological replicates were used for each sample.

Extraction and measurement of pigment

The anthocyanins and flavonoids were extracted and measured according to the previous study [52]. One gram of fresh petal of azalea was immersed in 5 ml extraction mixture of acetonitrile, water and formic acid (5:5:1, v:v:v) at 4℃ for 24 h, and then extracted by ultrasonic at ambient temperature for 30min. The samples were then filtrated and the residue was extracted with 5ml extraction solution for 30min, and then filtered. After two times of filtration, the two filtrates were combined and concentrated to 5ml by nitrogen blowing. The concentrated extract was passed through SPEC18 Sep-pak (GRAE company, USA), eluent to be tested. All samples analyzed on an Agilent 1260 ultra-high performance liquid chromatography (UHPLC) system coupled with a 6530 quadrupole-time of flight (Q-TOF) mass spectrometer (HPLC-Q-TOF-MS) operating in positive ion mode. Chromatographic separation was performed on an Agilent Poroshell 120 SB-Aq column (4.6×100 mm, 2.7 μm). Separation was carried out following a 50 min multistep linear gradient using a mobile phase consisting of (A) 1% v aqueous formic acid in water and (B) acetonitrile. From 0 to 20 min 5% B increase to 25% B, from 20 to 40 min 25% B increase to 100% B, then 100% B was maintained for 10 min, not including post-time. Flow rate was 0.3 $mL \cdot min^{-1}$; column temperature was 35 °C; injection volume was 20 μL; detection was performed at 530 nm and 254 nm. The source temperature was set at 350 °C, Capillary voltage at +4.0 kV, N2 drying gas

flow at 10 mL•min^{-1}, Nebulizer pressure at 50 psi, Fragmentor voltage at 175 V. Collision energy was set at 25 V for MS/MS analysis.

The compounds in the sample extracts were identified by comparison with the retention times of standards. Characteristics of the UV–Vis spectra of peaks and the mass spectrometric information were analyzed using MassHunter B0.05.0 Workstation (Agilent, USA). The relative content of anthocyanins and flavonoids was calculated from the peak areas of the samples based on the intensity of the corresponding standard compounds, including cyanidin, pelargonidin, peonidin, cyanidin-3-rutinoside, peonidin 3,5-diglucoside, cyanidin-3-glucoside, kaempferol, quercetin, quercetin-3-glucoside, Isorhamnetin and proanthocyanidin. For compounds lacking corresponding standards, quantification was performed using similar compounds. The mean values and SDs (standard deviations) were calculated for three biological replicates.

RNA extraction, cDNA library construction and RNA-seq

Total RNA was isolated from 1mg of mixed petals power from ten plants using the TRIzol kit according to the manufacturer's instructions (Invitrogen, Carlsbad, CA, USA). RNA concentration and the purity were assessed using Qubit RNA Assay Kit in Qubit 2.0 Fluorometer (Life Technologies, CA, USA) and a NanoDrop2000 spectrophotometer (IMPLEN, CA, USA) separately. RNA integrity was examined using t Agilent 2100 Bioanalyzer (Santa Clara, CA, USA) with an RIN (RNA integrity number)> 8.0. Three biological replicates were included for each flower development stage of each variety. Library construction and RNA-seq analyses were performed by Genepioneer Biotechnology Corporation (Nanjing, China) using Illumina HiSeq^{TM} 2500 platform. A total of 30 RNA sequencing libraries were generated (Table S3). All sequencing data were deposited at NCBI Sequence Read Archive under accession number SRP304986.

Transcriptome assembly and gene annotation

After removing the poor-quality reads (adaptor reads, ambiguous nucleotides and low-quality reads) [53], the clean reads were assembed using Trinity software as previously described for de novo transcriptome assembly without a reference genome [54]. The assembled unigenes were searched against public databases, including the NCBI, NR, COG , SwissProt, Pfam, GO, KOG and KEGG database.

Differentially expressed gene analysis

Gene expression levels of unigenes were estimated using FPKM method in RSEM software [55]. Clean data from each sample were respectively mapped back onto the assembled transcripts, and the normalized expression values RPKM (reads Per Kilobase per Million mapped reads) of each unigene in the 30 libraries were used as the value of gene expression levels [56]. Differential expression analysis was carried out using the DESeq R package [57]. Threshold criteria based on the FDR statistical method was used to determine significant differences in gene expression and the normalized expression values RPKM was compared using a threshold value of p-value $\leq$ 0.01and |log2 (fold change)| $\geq$ 1 based on the FDR

value ≤0.05. DEGs (FDR value ≤0.05 and |log2 (fold change)| ≥ 1) were determined and then analyzed through GO and KEGG pathway enrichment analysis using Goseq and KOBAS (2.0), respectively [58,59].

Phylogenetic Analysis

A phylogenetic tree of selected *R2R3-MYB* and *bHLH* transcription factors was constructed using MEGA 5.1 software by the neighbor-joining method with bootstrap analysis of 500 replicates [60].

Validation of GEGs by quantitative real-time PCR

qPCR analysis was performed using a SYBR® Premix Ex TaqTM Kit (TaKaRa, Dalian, China) and the ABI 7500 Real-Time PCR system (Applied Biosystems, CA, USA). The RNA samples used for qRT-PCR assays were the same as the sequencing samples. Total RNA (1 μg) was used to synthesize cDNA with the PrimeScript RT Reagent Kit (TaKaRa, Dalian, China) following the recommended procedures. All qRT-PCR reactions were incubated in a total volume of 20 μL comtaining 10 μL SYBR Master Mix, 2 μL cDNA template, 0.4 μL of 10 μM forward primer and reverse primer and 7.2 μL dd H2O. The cycling conditions were one cycle of 95 °C for 30 s, followed by 45 cycles of 95 °C for 5 s, 60 °C for 15 s, and 72 °C for 30 s. Housekeeping gene *Actin* was used as the control gene [61]. All primers are listed in Table S1.

The specificity of each primer pair was checked by agarose gel electrophoresis and melting curve analysis. The relative expression level of genes was calculated using the $2^{-\Delta\Delta CT}$ formula [62]. Each sample was analysed in both biological and technical triplicate.

Statistical analysis

Statistical analysis of the results of anthocyanins and flavonoids contents, were conducted by the one-way ANOVA LSD test using the IBM SPSS Statistics (version 19) statistical software. P values < 0.05 were considered significant. Correlation tests were performed by means of Pearson correlation coefficient (Pearson's r) with a two-tailed test. Correlations were considered positive when Pearson's r > 0.65 and P < 0.05.

Abbreviations

WT, wild-type; **DEG**, differentially expressed gene; **PAL**, phenylalanine-ammonia lyase; **4CL**, 4-coumaryl: CoA ligase ; **CHS**, chalcone synthase ; **CHI**, chalcone isomerase ; **F3H**, flavonoid 3-hydroxylase ; **F3'H**, flavonoid 3'-hydroxylase ; **F3'5'H**, flavonoid 3'5'-hydroxylase; **DFR**, dihydroflavonol 4-reductase; **ANS**, anthyocyanidin synthase; **UFGT**, UDP-glucose:flavonoid 3-O-glucosyltransferase; **bHLH**, basic helix–loop–helix; **FLS**, flavonol synthase; **NR**, Non-Redundant Protein Sequences; **GO**, Gene Ontology; **KOG**, Eukaryotic Ortholog Groups; **COG**, Cluster of Orthologous Groups of Proteins;**KEGG**, Kyoto Encyclopedia of Genes and Genomes; **SwissProt**, Swiss-Prot protein; Pfam, Protein Family; **FDR**, false discovery rate; **BP**, Biological Process; **CC**, Cellular Component; **MF**, Molecular Function; **3GT**, Anthocyanidin 3-O-glucosyltransferase; **5,3GT**, Anthocyanidin 5,3-O-glucosyltransferase; **FPKM**, fragments per kilobase of

transcript per million mapped reads; **TPM**, transcripts per million clean tags; **S.E.**, standard errors; **UHPLC**, ultra-high performance liquid chromatography; **Q-TOF**, quadrupole-time of flight; **SD**, standard deviation; **RPKM**,reads per kilobase per million mapped reads; **qPCR**, quantitative real-time polymerase chain reaction

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

Transcriptome raw sequence data in this study are available in SRA at NCBI with the BioSample accession number of SAMN17817099.

Competing interests

All authors have read and approved the manuscript. It is not being submitted to any other journal. The authors have declared that no competing interests exist.

Funding

This research was funded by the Natural Science Foundation of Jiangsu Province (Grant no. BK20190271), the Jiangsu Provincial Agricultural Science and Technology Innovation Project (CX (20)2030) and Special fund for forestry development of Jiangsu Province (sczh [2020] No.26). All the funding bodies didn't participate in the design of the study and collection, analysis, and interpretation of data and writing the manuscript.

Acknowledgements

Not applicable

Authors' information

Affiliations

Institute of Leisure Agriculture, Jiangsu Academy of Agricultural Sciences & Jiangsu Key Laboratory for Horticultural Crop Genetic Improvement, Nangjing 210014, China

Xiaobo Sun, Lisi He, Zhenhao Guo, Zheng Xiao, Jiale Su, Xiaoqing Liu, Huimin Zhou, Chang Li

Genepioneer Biotechnologies Co. Ltd, Nanjing, 210023, China

Haidong Gao

Contributions

JS, CL, LH and XL designed the experiments; XS, HZ, ZX handled experimental material; XS, ZX, HG and ZG were responsible for software and data processing; LH and XS writing original draft preparation; X.S. and Z.G. were responsible for review, editing and visualization. LH, HZ and CL were responsible for funding acquisition. XS and LH contributed equally to this work. All authors have read and agreed to the published version of the manuscript.

Corresponding author

Correspondence to Chang Li

References

1. Meijón M, Cañal MJ, Valledor L, Rodríguez R, Feito I. Epigenetic and physiological effects of gibberellin inhibitors and chemical pruners on the floral transition of azalea. Physiol. 2011; Plantarum 141(3): 276-288. https://doi.org/10.1111/j.1399-3054.2010.01430.x
2. Ureshino K., Nakayama M., Miyajima I. Contribution made by the carotenoid cleavage dioxygenase 4 gene to yellow colour fade in azalea petals. Euphytica. 2016;207:401-417. https://doi.org/10.1007/s10681-015-1557-2
3. Cooper-Driver GA. Contributions of Jeffrey Harborne and co-workers to the study of anthocyanins. Phytochemistry. 2001; 56:229-236. https://doi.org/10.1016/S0031-9422(00)00455-6
4. Mizuta D, Ban T, Miyajima I, Nakatsuka A, Kobayashi N. Comparison of flower color with anthocyanin composition patterns in evergreen azalea. Scientia Horticulturae. 2009;122(4):594-602. https://doi.org/10.1016/j.scienta.2009.06.027
5. Harborne JB. Flavonoid patterns and phytogeography: the genus Rhododendron section Vireya. Phytochemistry. 1986;25 (7):1641-1643. https://doi.org/10.1016/S0031-9422(00)81226-1
6. Chosson E, Chaboud A, Chulia AJ, Raynaud J. Dihydroflflavonol glycosides from Rhododendron ferrugineum. Phytochemistry. 1998;49 (5):1431-1433. https://doi.org/10.1016/S0031-9422(98)00080-6
7. Mok SY, Lee S. Identification of flavonoids and flavonoid rhamnosides from Rhododendron mucronulatum for. albiflorum and their inhibitory activities against aldose reductase. Food Chem. 2013;136 (2):969-974. https://doi.org/10.1016/j.foodchem.2012.08.091
8. Boulton R. The copigmentation of anthocyanins and its role in the color of red wine: a critical review. Am. J. Enol. Vitic. 2001;52 (2):67-87.

9. Du H, Lai L, Wang F, Sun W, Zhang L, Li X, et al. Characterisation of flower colouration in 30 Rhododendron species via anthocyanin and flavonol identification and quantitative traits. Plant Biol (Stuttg). 2018;20(1):121-129. https://doi.org/10.1111/plb.12649
10. Mizuta D, Nakatsuka A, Miyajima I, Ban T, Kobayashi N. Pigment composition patterns and expression analysis of flavonoid biosynthesis genes in the petals of evergreen azalea Oomurasaki and its red flower sport. Plant Breeding. 2010;129:558-562. https://doi.org/10.1111/j.1439-0523.2009.01714.x
11. Hang NTT, Miyajima I, Ureshino K, Kobayashi N, Kurashige Y, Matsui T, et al. Anthocyanins of wild Rhododendron simsii Planch. Flowers in Vietnam and Japan. J. Jpn. Soc. Hortic. Sci. 2011;80:206-213. https://doi.org/2503/jjshs1.80.206
12. Li CH, Wang LS, Shu QY, Xu YJ, Zhang J. Pigments composition of petals and floral color change during the blooming period in Rhododendron mucronulatum. Acta Hortic. Sin. 2008;35 (7):1023-1030. https://doi.org/10.16420/j.issn.0513-353x.2008.07.015
13. Park CH, Yeo HJ, Kim NS, Park YE, Park SY, Kim, JK, et al. Metabolomic Profiling of the White, Violet, and Red Flowers of Rhododendron schlippenbachii Maxim. Molecules. 2018;23(4):827. https://doi.org/10.3390/molecules23040827
14. Lloyd A, Brockman A, Aguirre L, Campbell A, Bean A. Advances in the MYB–bHLH–WD repeat (MBW) pigment regulatory model: addition of a WRKY factor and co-option of an anthocyanin MYB for betalain regulation. Plant Cell Physiol. 2017;58:1431-1441. https://doi.org/10.1093/pcp/pcx075
15. Naing AH, Kim CK. Roles of R2R3-MYB transcription factors in transcriptional regulation of anthocyanin biosynthesis in horticultural plants. Plant Mol. Biol. 2018;98:1-18. https://doi.org/10.1007/s11103-018-0771-4
16. Petroni K, Tonelli C. Recent advances on the regulation of anthocyanin synthesis in reproductive organs. Plant Sci. 2011;181 (3):219-229. https://doi.org/10.1016/j.plantsci.2011.05.009
17. Koes R, Verweij W, Quattrocchio F. Flavonoids: A colorful model for the regulation and evolution of biochemical pathways. Trends Plant Sci. 2005;10:236-242. https://doi.org/10.1016/j.tplants.2005.03.002
18. Saito K, Yonekura-Sakakibara K, Nakabayashi R, Higashi Y, Yamazaki M. The flavonoid biosynthetic pathway in Arabidopsis: structural and genetic diversity. Plant Physiol. Biochem. 2013; 72:21-34. https://doi.org/10.1016/j.plaphy.2013.02.001
19. Tanaka Y, Sasaki N, Ohmiya A. Biosynthesis of plant pigments: anthocyanins, betalains and carotenoids. Plant J. 2008;54 (4):733-749. https://doi.org/10.1111/j.1365-313X.2008.03447.x
20. Ferreyra MLF, Rius SP, Casati P. Flavonoids: Biosynthesis, biological functions, and biotechnological applications. Front. Plant Sci. 2012;3(222):222. https://doi.org/10.3389/fpls.2012.00222
21. Xu W, Dubos C, Lepiniec L. Transcriptional control of flavonoid biosynthesis by MYB-bHLH-WDR complexes. Trends Plant Sci. 2015;20:176-185. https://doi.org/10.1016/j.tplants.2014.12.001

22. Lai B, Li XJ, Hu B, Qin YH, Huang XM. LcMYB1 is a key determinant of diffffferential anthocyanin accumulation among genotypes, tissues, developmental phases and ABA and light stimuli in Litchi chinensis. PLoS One. 2014;9:e86293. https://doi.org/10.1371/journal.pone.0086293
23. Jin W, Wang H, Li M, Wang J, Yang Y. The R2R3 MYB transcription factor PavMYB10.1 involves in anthocyanin biosynthesis and determines fruit skin colour in sweet cherry (Prunus avium L.). Plant Biotechnol. J. 2016;14:2120-2133. https://doi.org/10.1111/pbi.12568
24. Schepper D, Debergh P, Bocktaele V, Loose D. Molecular characterisation of flower colour genes in azalea sports (Rhododendron simsii hybrids). Acta Hortic. 2001;552:143-150. https://doi.org/10.17660/ActaHortic.2001.552.15.
25. De KE, Lootens P, Van BE, De RJ. Image analysis for QTL mapping of flower colour and leaf characteristics in pot azalea (Rhododendron simsii hybrids). Euphytica. 2013;189:445-460. https://doi.org/10.1007/s10681-012-0809-7
26. Nakatsuka A, Mizuta D, Kii Y, Miyajima I, Kobayashi Isolation and expression analysis of flavonoid biosynthesis genes in evergreen azalea. Scientia Horticulturae. 2008;118(4):314-320. https://doi.org/10.1016/j.scienta.2008.06.016
27. Su JL, Liu XQ, Li C, He LS, Chen L. A New Rhododendron Cultivar'Yanzhi Mi'. Acta Horticulae Sinica. 2012;39(12):2555-2556. https://doi.org/10.16420/j.issn.0513-353x.2012.12.036
28. Berardini TZ, Mundodi S, Reiser L, Huala E, Garcia-Hernandez M, Zhang PF, et al. Functional annotation of the Arabidopsis genome using controlled vocabularies. Plant Physiol. 2004;135:745-755. https://doi.org/10.1104/pp.104.040071
29. Pillet J, Yu HW, Chambers AH, Whitaker VM, Folta KM. Identification of candidate flavonoid pathway genes using transcriptome correlation network analysis in ripe strawberry (Fragaria×ananassa) fruits. Journal of Experimental Botany. 2015;66(15):4455-4467. https://doi.org/10.1093/jxb/erv205
30. Shin DH, Choi MG, Bang G, Cho M, Choi SB, Choi G, et al. HY5 regulates anthocyanin biosynthesis by inducing the transcriptional activation of the MYB75/PAP1 transcription factor in Arabidopsis. FEBS Lett. 2013;587:1543-1547. https://doi.org/10.1016/j.febslet.2013.03.037
31. Matus JT, Aquea F, Arce-Johnson P. Analysis of the grape MYB R2R3 subfamily reveals expanded wine quality-related clades and conserved gene structure organization across Vitis and Arabidopsis genomes. BMC Plant Biol. 2008;8:83. https://doi.org/10.1186/1471-2229-8-83
32. Ben SR, Moran F, Boaz A, Alon C, Michal MBZ, Tania M, et al. The R2R3-MYB–Like Regulatory Factor EOBI, Acting Downstream of EOBII, Regulates Scent Production by Activating ODO1 and Structural Scent-Related Genes in Petunia. The Plant Cell. 2012;24 (12):5089-5105. https://doi.org/10.1105/tpc.112.105247
33. Yang S, Hongdi L, Qiang Z, Hongjun Z, Zhidong Z, Yadong L, et al. High-throughput sequencing of highbush blueberry transcriptome and analysis of basic helix-loop-helix transcription factors. Journal of Integrative Agriculture. 2017;16(3):591-604. https://doi.org/10.1016/S2095-3119(16)61461-2
34. Kunishige M, Kobayashi Y. Chromatographic identifification of Japanese azalea species and their hybrids. In: Luteyn JL, O'Brien ME, editors. Contributions Toward a Classifification of Rhododendron.

New York: Botanical Garden; 1980. p. 277-287.

35. Liu L, Zhang LY, Wang SL, Niu XY. Analysis of anthocyanins and flavonols in petals of 10 Rhododendron species from the Sygera Mountains in Southeast Tibet. Plant Physiology and Biochemistry. 2016;104:250-256. https://doi.org/10.1016/j.plaphy.2016.03.036
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;29:223-227. https://doi.org/10.5511/plantbiotechnology.12.0120a
37. Winkel-Shirley B. Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. Plant Physiol. 2001;126:485-493. https://doi.org/10.1104/pp.126.2.485.
38. Mol J, Grotewold E, Koes R. How gens paint flowers and seeds. Trends Plant Sci. 1998; 6:212-217.
39. Schwinn K, Venail J, Shang YJ, Mackay S, Alm V, Butelli E, et al. A small family of MYB-regulatory genes controls floral pigmentation intensity and patterning in the genus Antirrhinum. Plant Cell. 2006;18:831-851. https://doi.org/10.1105/tpc.105.039255
40. Albert NW, Davies KM, Lewis DH, Zhang HB, Montefifiori M, Brendolise C, et al. A conserved network of transcriptional activators and repressors regulates anthocyanin pigmentation in eudicots. Plant Cell. 2014;26:962-980. https://doi.org/10.1105/tpc.113.122069
41. Hsu CC, Chen YY, Tsai WC, Chen WH, Chen HH. Three R2R3-MYB Transcription factors regulate distinct floral pigmentation patterning in Phalaenopsis spp. Plant Physiol. 2015;168:175-191 https://doi.org/10.1104/pp.114.254599
42. Li CH, Qiu J, Ding L, Huang MZ, Huang SR, Yang GS, et al. Anthocyanin biosynthesis regulation of DhMYB2 and DhbHLH1 in Dendrobium hybrids petals. Plant Physiol Biochem. 2017;112:335-345. https://doi.org/10.1016/j.plaphy.2017.01.019
43. Nuraini L, Ando Y, Kawai K, Tatsuzawa F, Tanaka K, Ochiai M, et al. Anthocyanin regulatory and structural genes associated with violet fower color of Matthiola incana. Planta. 2020;251(3):61. https://doi.org/10.1007/s00425-020-03351-z
44. Yoshida K, Ma D, Constabel CP. The MYB182 protein down-regulates proanthocyanidin and anthocyanin biosynthesis in poplar by repressing both structural and regulatory flflavonoid genes. Plant Physiol. 2015;167:693-710. https://doi.org/10.1104/pp.114.253674
45. Wan S, Li C, Ma X, Luo K. PtrMYb57 contributes to the negative regulation of anthocyanin and proanthocyanidin biosynthesis in poplar. Plant Cell Rep. 2017;36:1263-1276. https://doi.org/10.1007/s00299-017-2151-y
46. Zhou H, Peng Q, Zhao J, Owiti A, Ren F. Multiple R2R3-MYB transcription factors involved in the regulation of anthocyanin accumulation in peach flflower. Front. Plant Sci. 2016;7:1557. https://doi.org/10.3389/fpls.2016.01557
47. Jiang Y, Liu C, Yan D, Wen X, Liu Y, Wang H, et al. MdHB1 down-regulation activates anthocyanin biosynthesis in the white-fleshed apple cultivar 'Granny Smith'. J Exp Bot. 2017;68(5):1055-1069. https://doi.org/10.1093/jxb/erx029

48. Xu H, Wang N, Liu J, Qu C, Wang Y. The molecular mechanism underlying anthocyanin metabolism in apple using the MdMYB16 and MdBHLH33 genes. Plant Mol. Biol. 2017;94:149-165. https://doi.org/10.1007/s11103-017-0601-0
49. Xu H, Yang G, Zhang J, Wang Y, Zhang T. Overexpression of a repressor MdMYB15L negatively regulates anthocyanin and cold tolerance in red-fleshed callus. Biochem. Biophys. Res. Commun. 2018;500:405-410. https://doi.org/10.1016/j.bbrc.2018.04.088
50. Anwar M, Wang G, Wu J, Waheed S, Allan AC. Ectopic overexpression of a novel R2R3-MYB, NtMYB2 from Chinese narcissus represses anthocyanin biosynthesis in tobacco. Molecules. 2018;23: 781. https://doi.org/10.3390/molecules23040781
51. Xiang L, Liu X, Li H, Yin X, Grierson D, Li F, et al. CmMYB7, an R3MYB transcription factor, acts as a negative regulator of anthocyanin biosynthesis in chrysanthemum. J Exp Bot. 2019;70(12):3111-3123. https://doi.org/10.1007/s00299-019-02391-4
52. Monica GM, Luis E, Rodrı'guez-Saona. Electrospray and Tandem Mass Spectroscopy as Tools for Anthocyanin Characterization. J Agr Food Che. 1999;47:4657-4664. https://doi.org/1021/jf981242+
53. Liu SS, Chen J, Li SC, Zeng X, Meng ZX, Guo SX. Comparative Transcriptome analysis of genes involved in GA-GID1-DELLA regulatory module in symbiotic and asymbiotic seed germination of Anoectochilus roxburghii (Wall.) Lindl. (Orchidaceae). Int J Mol Sci. 2015;16:30190-30203. https://doi.org/10.3390/ijms161226224. PMID: 26694378, PMCID: PMC4691166.
54. Grabherr MG, Haas B, 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;29:644-652. https://doi.org/10.1038/nbt.1883
55. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinfor. 2011;12:323-323. https://doi.org/10.1186/1471-2105-12-323
56. Mortazavi A, Williams BA, Mccue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods. 2008;5(7):621–8. https://doi.org/1038/nmeth.1226
57. 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):136-8. https://doi.org/10.1093/bioinformatics/btp612
58. Young MD, Wakefield MJ, Smyth GK, Oshlack A. Gene ontology analysis for RNA-seq: Accounting for selection bias. Genome Biol. 2010;11:R14. https://doi.org/10.1186/gb-2010-11-2-r14
59. Xie C, Mao X, Huang J, Ding Y, Wu J, Dong S, et al. KOBAS 2.0: A web server for annotation and identification of enriched pathways and diseases. Nucleic Acids Res. 2011;39:316-322. https://doi.org/10.1093/nar/gkr483
60. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol. 2011;28:2731-2739. https://doi.org/1093/molbev/msr121
61. Xiao Z, Sun X, Liu X, Li C, He L, Chen S, et al. Selection of reliable reference genes for gene expression studies on Rhododendron molle G. Don. Front Plant Sci. 2016;7:1547. https://doi.org/

10.3389/fpls.2016.01547

62. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2-ΔΔCT method. Methods. 2001;25:402-408. https://doi.org/10.1006/meth.2001.1262

Figures

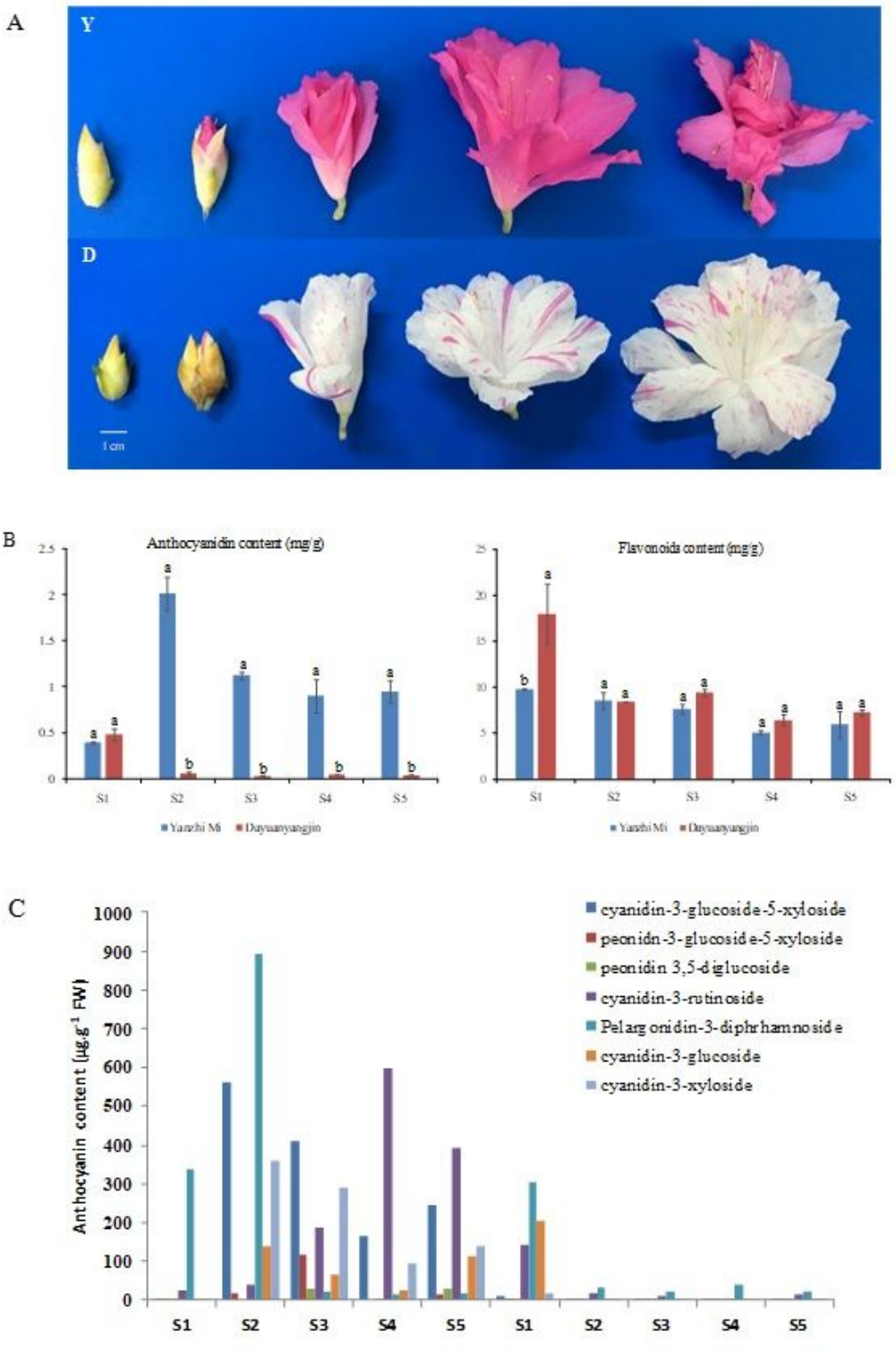

Figure 1

Phenotypic characteristic and pigment analysis during floral development stages of 2 R.obtusum cultivars. Y stands for ‘Yanzhi Mi’ and D stands for ‘Dayuanyangjin’, same as below. A: Phenotypic characteristic of flowers at different floral development stages. S1, S2, S3, S4 and S5 represent closed buds, buds showing color at the top but with the scales still present, initial flowering stage, full flowering stage and last flowering stage, respectively. Bar = 1 cm. B: The total contents of anthocyanin and flavonoids (mg/g) of petals at five floral development stages (S1-S5), C: Comparison the content of individual compound of anthocyanins at different floral development stages.

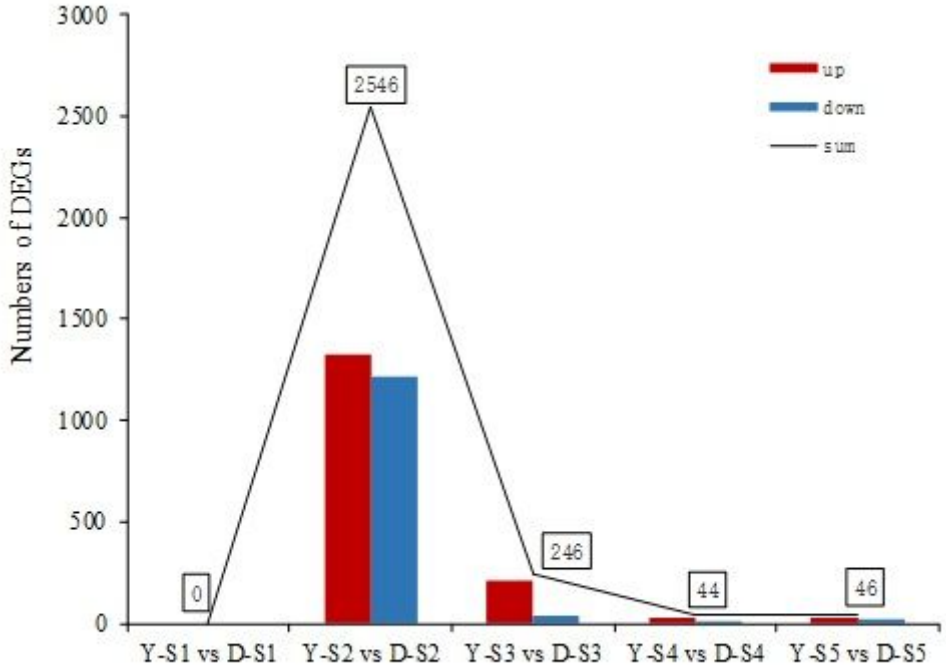

Figure 2

Pairwise comparison of DEGs in the petals of ‘Yanzhi Mi’ vs ‘Dayuanyangjin’. The five floral development stages S1-S5. S1, S2, S3, S4 and S5 represent closed buds, buds showing color at the top but with the scales still present, initial flowering stage, full flowering stage and last flowering stage, respectively.

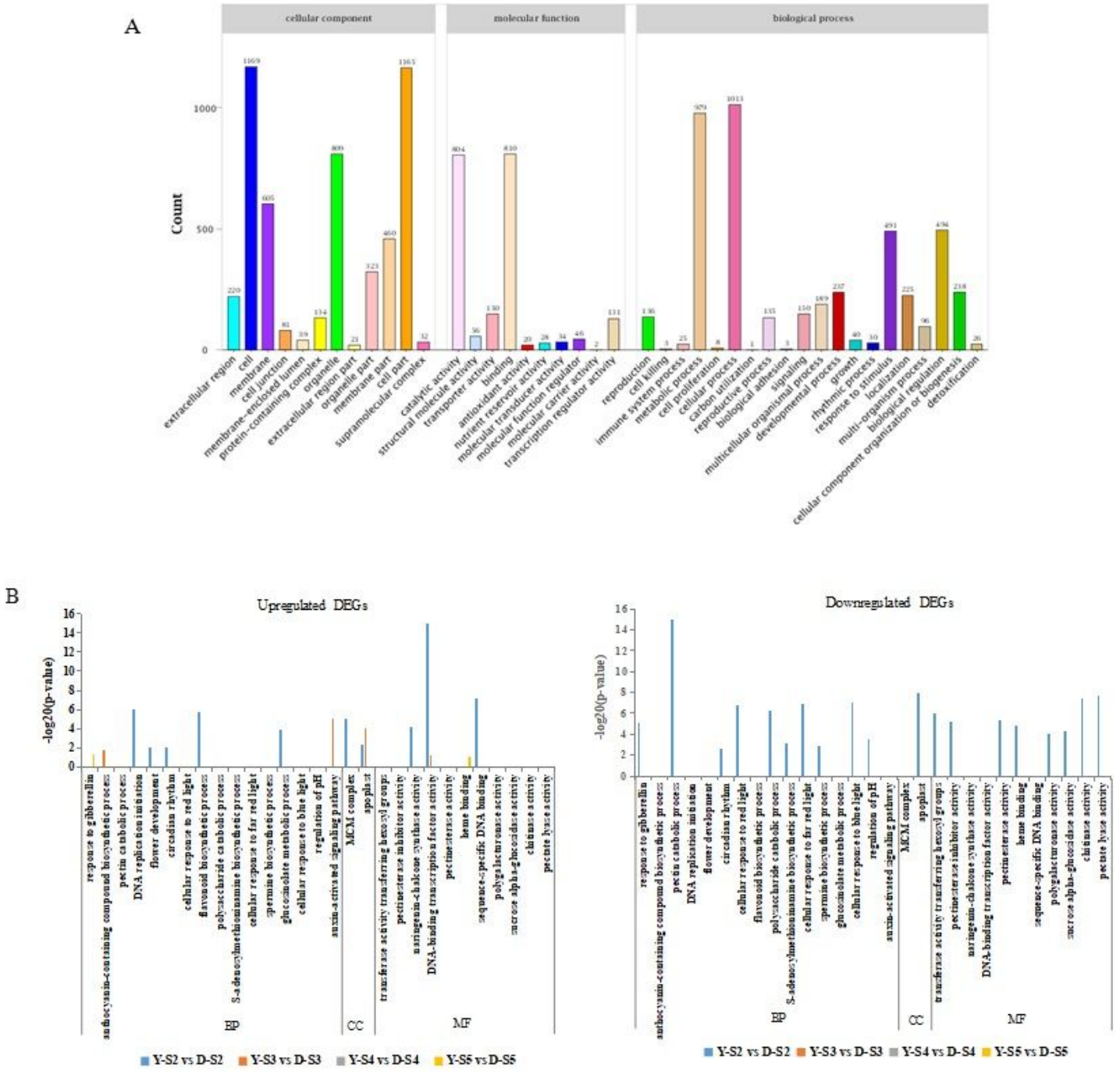

Figure 3

GO analysis of DEGs in the petals of ‘Yanzhi Mi’ vs ‘Dayuanyanjin’. A: The distribution of GO term categories of DEGs in the petals of ‘Yanzhi Mi’ vs ‘Dayuanyanjin’ at S2 stage. B: GO analysis of the upregulated and downregulated DEGs in the petals of ‘Yanzhi Mi’ vs ‘Dayuanyanjin’ at five floral development stages in three main classes (‘BiologicalProcess’,‘Cellular Component’, and ‘Molecular Function’).The figures show the distribution of GO terms exhibiting significant differences (p-value ≤ 0.05 and absolute log2 fold change ≥1).

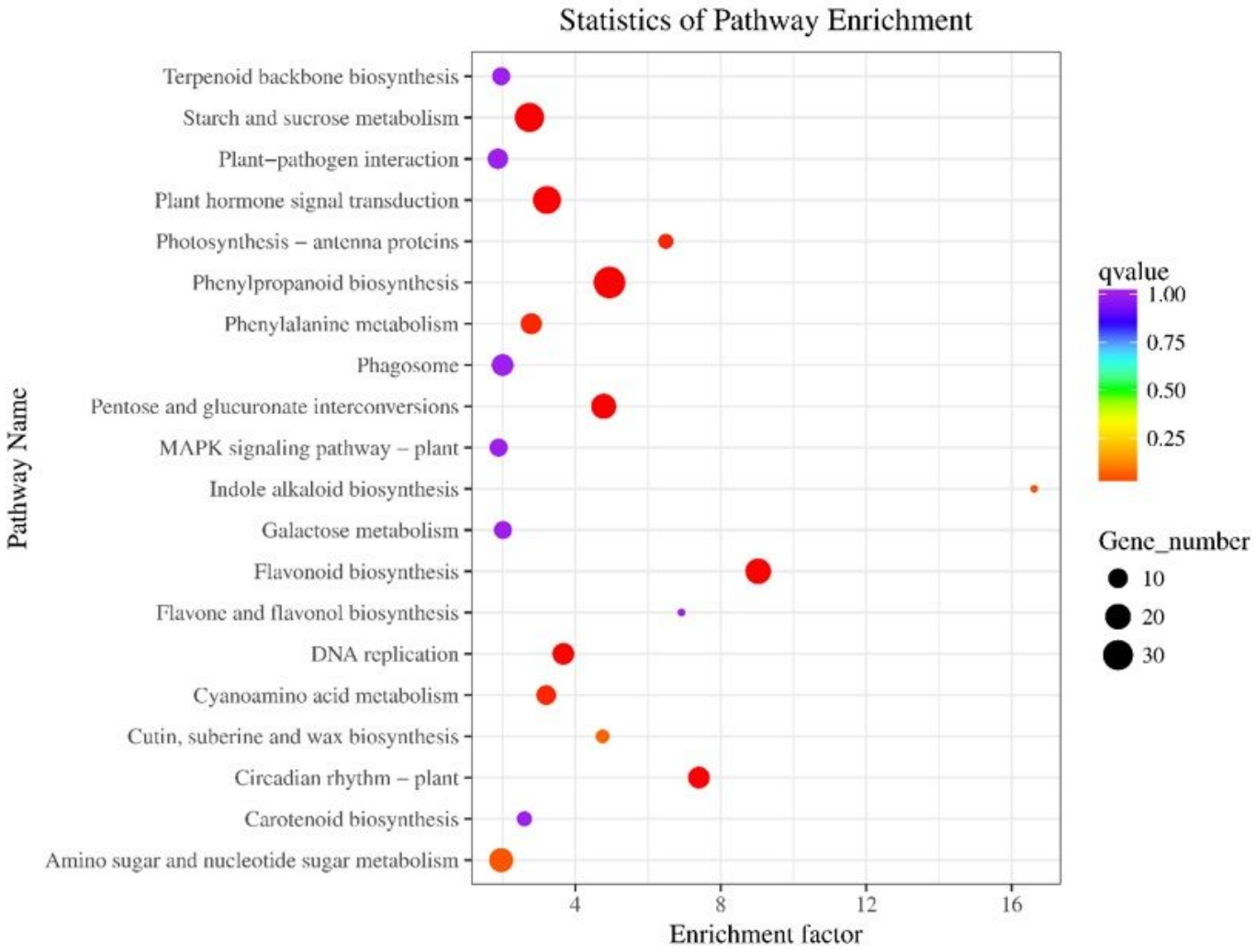

Figure 4

The top 20 pathway terms enriched in the KEGG database in Y-S2 vs D-S2 comparison. The circle size represents the quantity of DEGs. The rich factor is the ratio of DEG numbers annotated in a given pathway term to all gene numbers that were annotated in the pathway term, and greater rich factor values indicate greater intensiveness. The q-values coloring indicates the significance of the rich factor ranging from 0 to 1, and lower q-values indicate greater intensiveness.

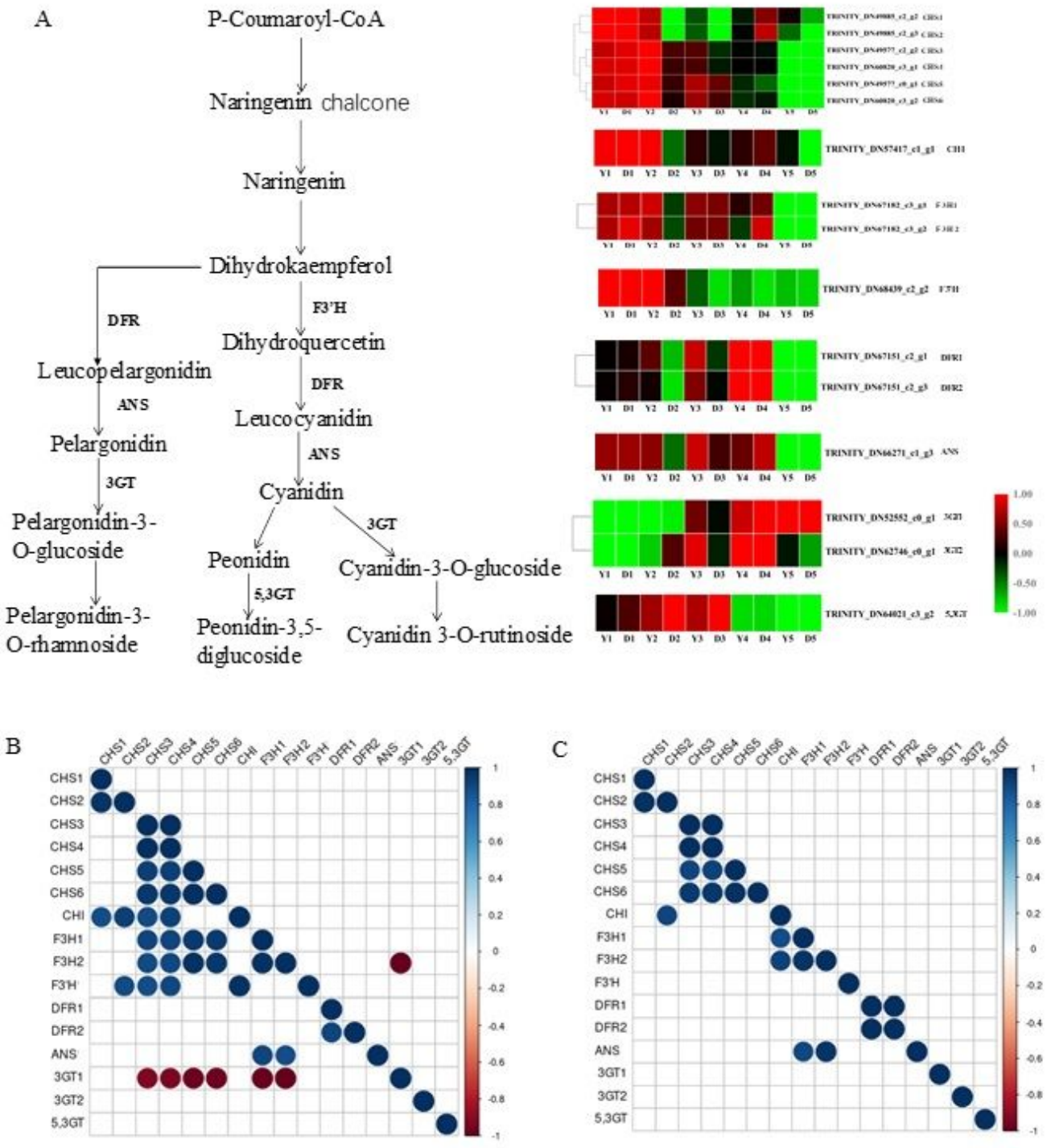

Figure 5

Expression profiles and expression correlation of the DEGs related to anthocyanin biosynthesis pathway. A: Heatmap of DEGs related to anthocyanin biosynthesis pathway. Expression profiles of these genes across five tissues (S1, closed buds; S2, buds showing color at the top but with the scales still present; S3, initial flowering stage; S4, full flowering stage; S5, last flowering stage) of 'Yanzhi Mi' (Y) and 'Dayuanyangjin' (D) are shown (FDR value ≤0.05, |log2 (fold change)| ≥ 1)). Green and red colors are

used to represent low-to- high expression levels; color scale represents log2-transformed FPKM (fragments per kilobase of transcript per million mapped reads) values. B: the gene expression correlation map of these genes in ‘Yanzhi Mi’. C: the gene expression correlation map of these genes in ‘Dayuanyangjin’. Correlation are considered positive when Pearson’s r >0.65 (P<0.05).

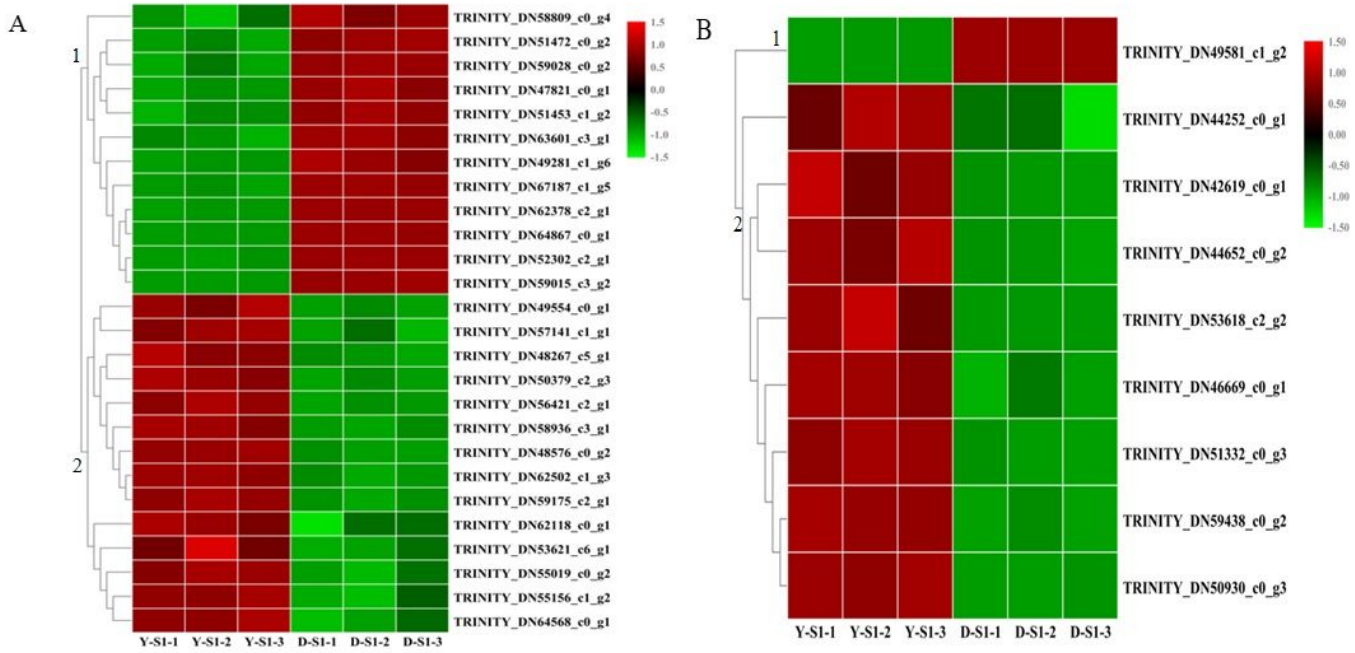

Figure 6

Heatmaps of unigenes encoding the R2R3-MYB and bHLH transcription factors at S2 stage. A and B stands for R2R3-MYB and bHLH respectively. The unigenes were hierarchically clustered and mapped using FPKM (fragments per kilobase of exon per million mapped reads) values. The numbers in the heat map represent different expression clusters. The color scale represents log2-transformed FPKM values. Red represents high expression, and green represents low expression. S2 represents S2 stage of flower development.

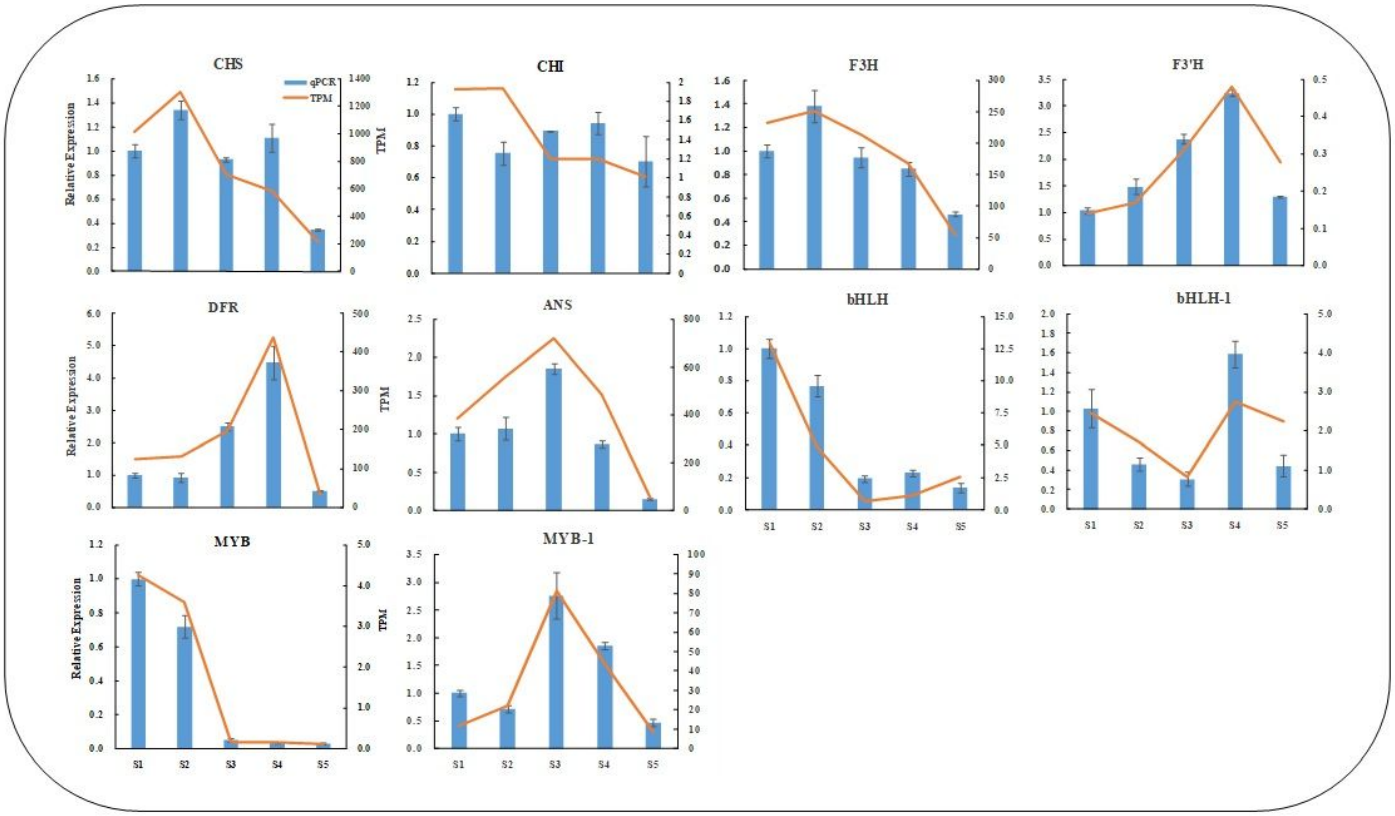

Figure 7

qPCR validations of 10 putative genes involved in the anthocyanin pathway of ‘Yanzhi Mi’ flower. The polylines represent TPM in DEG database, and the histograms represent the qPCR results. The expression at S1 was used as a calibration standard. Actin was used as an internal control and expression level present as means with standard errors (S.E.) of three replications. S1, S2, S3, S4 and S5 represent closed buds, buds showing color at the top but with the scales still present, initial flowering stage, full flowering stage and last flowering stage, respectively.

Supplementary Files

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

- Additionalfile1Fig.S1..pdf
- Additionalfile2Fig.S2..pdf
- Additionalfile3TableS1..xlsx
- Additionalfile4TableS2..xlsx
- Additionalfile5TableS3..xlsx
- Additionalfile6TableS4..xlsx

- Additionalfile7TableS5..xlsx
- Additionalfile8TableS6..xlsx