

RESEARCH

Open Access



Metabolome and transcriptome analysis of anthocyanin biosynthesis reveal key metabolites and candidate genes in red-stemmed alfalfa (*Medicago sativa*)

Yaqian Zong¹, Zhili Zhao¹, Kai Zhou¹, Xinhui Duan¹, Bo Han¹, Chenggang He^{1*}, Heping Huang^{1*} and Hua Jiang^{1*}

Abstract

Background Alfalfa (*Medicago sativa* L.) serves as a vital high-quality forage resource, especially in tropical and subtropical regions where there is a deficiency of protein-rich feed. The red pigmentation of stem of space mutated alfalfa was mainly caused by anthocyanin accumulation. However, investigations into the mechanisms governing anthocyanin biosynthesis in alfalfa stems have been scarce.

Result In this study, we conducted combined transcriptome and metabolome analyses on two types of alfalfa stems: space mutation red-stemmed alfalfa and non-space mutation green-stemmed alfalfa (control). Profiling of the anthocyanin metabolome unveiled 45 metabolites linked to anthocyanin biosynthesis, with cyanidin-3-O-glucoside, pelargonidin-3-O-arabinoside, delphinidin-3-O-(6-O-acetyl)-glucoside, and kaempferol-3-O-rutinoside identified as the primary anthocyanins of red-stemmed alfalfa. Transcriptome analysis revealed 72 differentially expressed genes related to anthocyanin biosynthesis pathways, of which 54 genes were highly expressed in red stems, including 12 *PALs* (phenylalanine ammonia-lyase), 22 *4CLs* (4-coumaroyl: CoA-ligase), eight *CHSs* (chalcone synthase), three *F3Hs* (flavanone 3-hydroxylase), two *ANRs* (anthocyanidin reductase), three *DFRs* (dihydroflavonol-4-reductase), three *ANSs* (anthocyanidin synthase), and one *FLS* (flavonol synthase) gene. These genes are likely pivotal for anthocyanin biosynthesis in red-stemmed. Co-expression analysis of differentially expressed genes and relative contents of differentially expressed anthocyanin showed that each anthocyanin was closely related to multiple genes, and anthocyanin accumulation process was regulated by multiple genes. The expressions of these genes were significantly positively correlated with the relative contents of cyanidin-3-O-glucoside, pelargonin-3-O-arabinoside, and kaempferol-3-O-rutin.

Conclusion Overall, the expression patterns of *PAL*, *4CL*, *CHS*, *F3H*, *ANR*, *DFR*, *ANS*, and *FLS* structural genes in anthocyanin biosynthesis pathway were closely related to the composition and content of anthocyanins. Different

*Correspondence:
Chenggang He
chengganghe63@163.com
Heping Huang
1329873649@qq.com
Hua Jiang
jianghua15@163.com

Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

anthocyanins' accumulation patterns may result in the different stem colors of alfalfa. These findings provide comprehensive insights into the molecular mechanisms for anthocyanin biosynthesis in red-stemmed alfalfa.

Keywords Space mutation, Anthocyanins, *Medicago sativa*. Cv. Deqin, Metabolome, Transcriptome

Background

Alfalfa, known as the “king of forages” is a highly valuable forage plant distributed in arid and semi-arid areas and cultivated worldwide as a superior-quality forage crop [1, 2]. As a perennial legume, alfalfa holds significant economic value in the livestock industry because of its high yield, nutritional value, and palatability [3]. Therefore, the cultivation of new alfalfa varieties has garnered extensive attention. With recent advancements in space technology, space breeding has emerged as a notable approach. Space breeding, also referred to as space mutagenesis breeding, is a novel method primarily conducted in space. Consequently, our team sent ‘Deqin’ alfalfa seeds into space aboard a new-generation manned spacecraft on May 5, 2020. Alfalfa seeds were planted in pots after returning to the ground, and the characteristic alfalfa resources with stable characters and red stems epidermal layers were obtained by observing their characters. Numerous scientific studies have indicated that changes in plant coloration are attributed to alterations in pigment content, including chlorophyll, lutein, carotene, and anthocyanin. Among these plant pigments, anthocyanins have garnered significant attention due to their role in controlling plant color changes. Hence, the identification of anthocyanin synthesis genes and effective metabolic components in red-stemmed alfalfa offers great promise for the further development and utilization of this plant.

Anthocyanins, a type of flavonoid widely found in plant stems, leaves, flowers, fruits, and seeds exhibit a broad spectrum of biological functions [4]. These functions include imparting rich colors to plant organs and attracting animals and insects, thereby facilitating pollination and seed dispersal [5, 6], as well as protecting plants from drought, cold conditions, ultraviolet radiation, and harmful microorganisms [7, 8]. Additionally, anthocyanins find wide applications in the medical industry for preventing cardiovascular and cerebrovascular diseases, combating chronic conditions such as cancer, scavenging pathogenic free radicals in the human body, and inhibiting the oxidation of low-density lipids [9]. Anthocyanins are frequently glycosylated with glucose, rhamnose, galactose, xylose, arabinose, and other glycosides to synthesize various aglycone forms of anthocyanin [10].

The basic structure of anthocyanins is C6-C3-C6, and the different types of anthocyanins are formed by varying positions of methylation and hydroxylation modifications. This enables plants to exhibit a diverse array of colors, including orange, red, purple, and blue. Anthocyanin glycosides in higher plants mainly include cyanidin,

delphinidin, malvidin, peonidin, petunidin, and pelargonidin [11]. The molecular pathways of anthocyanin biosynthesis have been intensively investigated in many plants, including pepper (*Capsicum annuum* L.) [12], petunia (*Calibrachoa*) [13], and tobacco (*Nicotiana tabacum*) [14]. The direct precursor of anthocyanin biosynthesis is phenylalanine, which undergoes three main stages of synthesis. The first stage involves the formation of 4-coumaroyl-CoA under the catalysis of PAL (phenylalanine ammonia-lyase), C4H (cinnamic acid 4-hydroxylase), and 4CL (4-coumaroyl: CoA-ligase). The second stage involves the conversion of 4-coumaroyl-CoA into chalcone by CHS (chalcone synthase), followed by the formation of dihydroflavonols by CHI (chalcone isomerase) and F3H (flavonoid 3-hydroxylase). The third stage is the synthesis of anthocyanins, which is mediated by DFR (dihydroflavonol-4-reductase) and ANS (anthocyanin synthase), leading to the production of various anthocyanins [15, 16]. These structural genes are closely associated with anthocyanin synthesis and accumulation. For example, differences in the color of ramie (*Boehmeria nivea* (L.) Gaud.) leaves were caused by differences in the upregulated expression of flavonoid synthesis genes, and the upregulated expression of these genes was positively correlated with total anthocyanin accumulation in ramie leaves [17]. Similarly, the accumulation of cyanidin-3-O-(6"-O-malonyl) glucoside was correlated with the reddening of *Camellia japonica* petals, and candidate genes identified as regulating anthocyanin accumulation in camellia petals included 37 structural genes and 18 differentially expressed transcription genes [18]. Additionally, a previous study used transcriptomics and metabolomics to analyze the molecular mechanism of anthocyanin formation in purple and Haifa white clover (*Trifolium repens*), revealing that DFR, LDOX/ANS (leucoanthocyanidin oxygenase/anthocyanidin synthase), ANR (anthocyanidin reductase), and UFGT (UDPglucose: flavonoid-3-O-glucosyltransferase) may be key determinants of differences in anthocyanin and proanthocyanidin biosynthesis [19].

Transcriptomics and metabolomics are powerful tools for uncovering differences in gene expression and metabolite levels across samples with varying phenotypes. Although biosynthesis of anthocyanins has been extensively studied in many crops and cash crops using these approaches, research on the transcriptional regulation and metabolic pathways of anthocyanins in red-stemmed alfalfa mutants by space mutagenesis remains limited. Thus, this study employs a combined metabolomics and

transcriptomics approach to assess and analyze the relationships among gene expression, metabolite accumulation, and color formation in alfalfa stems. By elucidating the molecular mechanisms of anthocyanin biosynthesis in alfalfa stems, aims to provide theoretical support for breeding new varieties of alfalfa.

Materials and methods

Plant material

Seeds of the alfalfa cultivar 'Deqin' were sent into space aboard a new recoverable manned spacecraft for three days (May 5–8, 2020). These space-exposed seeds were planted alongside the control at the Kunming Institute of Botany, Chinese Academy of Sciences (28°08' N 102°44' E, 1950 m above sea level). During phenotypic observation, red-stemmed alfalfa mutant (R) was found and used as experimental materials. Green-stemmed alfalfa without space mutagenesis was selected as a control (G). In October 2022, cuttings of healthy stems of R and G were propagated in 10 × 10 × 10 cm containers filled with a soil medium composed of humus soil and laterite at a ratio of 2:1, with 30 poles of cuttings for each material (10 cuttings per replicate × 3 replicates). The plants were cultivated outside the Herbolgy greenhouse in Yunnan Agricultural University (25°07' N 102°45' E, 1950 m above sea level) under the same management practices. At the branching stage, plant samples were collected, and each duplicate entire stem sample was mixed. Three biological samples of each stem color were selected for transcriptome sequencing, and six biological replicates were collected for metabolite extraction. Samples were harvested at the same time of day on February 20, 2023. All samples were frozen with liquid nitrogen and stored at –80 °C for transcriptome and metabolome analysis.

Anthocyanin content analysis

Weigh about 1 g of fresh stem tissue was ground into a powder with liquid nitrogen, approximately 100 mg of which was then extracted in 10 mL of 1% (v/v) HCl–ethanol in the dark at 4 °C for 24 h. The mixture was vortexed and centrifuged at 12 000 × g for 10 min. The absorbance of the supernatant was measured at 530, 620, and 650 nm using an ultraviolet-visible spectrophotometer (Shanghai Yuan Instrument Co., Ltd., Shanghai, China).

Metabolite extraction and analysis

One 50 mg powder sample was weighed and extracted with 500 µL of extraction solution (70% methanol aqueous solution, containing 0.1% hydrochloric) overnight at 4 °C. Then, the extract was vortexed for 5 min, subjected to ultrasound for 5 min, and centrifuged at 12 000 × g at 4 °C for 3 min. The residue was then re-extracted once by repeating the above steps under identical conditions. The supernatants were collected and filtered through a

0.22 µm membrane filter (Agilent SB-C18) before LC-MS/MS analysis. The analytical conditions were as follows: high-performance liquid chromatograph (UPLC): column, ACQUITY BEH C18 (1.7 µm, 2.1 × 100 mm) with a flow rate of 0.4 mL/min; mobile phases, 0.1% formic acid aqueous solution (A) and methanol (0.1% formic acid) (B); flow rate, 0.35 mL/min; temperature, 40 °C; injection volume, 2 µL and gradient program, based on the method of Shao et al. [20].

A triple quadrupole linear ion trap mass spectrometer (QTRAP) was operated in both positive and negative ion modes. The optimized parameters were as follows: ion source temperature 500 °C; ion spray voltage, 5.5 kV (positive), –4.5 kV (negative); and ion source gas 1, ion source gas 2, and curtain gas set at 50, 50, and 35 psi, respectively. The collision gas and ion spray voltages were set to medium and 4.5 kV, respectively. Data acquisition and analysis were performed using the Analyst Software (Sciex, Foster City, CA, USA).

Triple quadrupole (QQQ) and linear ion trap (LIT) scans were acquired on a triple quadrupole-linear ion trap mass spectrometer (QTRAP) LC-MS/MS system, equipped with an ESI Turbo Ion-Spray interface, operating in positive and negative ion modes, and controlled by Analyst 1.6.3 software (Sciex). The ESI source operation parameters were as follows: source temperature 500 °C; ion spray voltage (IS) 5.5 kV (positive), –4.5 kV (negative); ion source gas I (GSI), gas II (GSII), curtain gas (CUR) was set at 55, 60, and 25.0 psi, respectively; the collision gas (CAD) was set to high. Instrument tuning and mass calibration were performed with 10 and 100 µmol/L polypropylene glycol solutions in QQQ and LIT modes, respectively. A specific set of MRM transitions was monitored for each period according to the metabolites eluted within this period.

Identification and quantitative analysis of metabolites

Based on the Biomarker database (Biomarker Technologies Co., Ltd., Beijing, China), the data detected by mass spectrometry was analyzed qualitatively. Multiple-reaction monitoring (MRM) was used for the quantitative detection of metabolites. Principal component analysis (PCA) was performed on 12 samples to evaluate differences in metabolite composition among them. A correlation heatmap analysis was performed using the BMKCloud (www.biocloud.net). Differentially accumulated metabolites (DAMs) were identified based on a combination of variable importance in projection (VIP), fold change (FC), and P value. The screening criteria were set as FC > 1 or < 0.5, and VIP > 1.

RNA extraction, library construction, and sequencing

Total RNA was isolated from alfalfa stems using the TRIzol reagent (Tiangen, Beijing, China) following the

manufacturer's protocol. After extraction, the RNA concentration and purity were examined using a NanoDrop 2000 (Thermo Fisher Scientific, Wilmington, DE, USA). The RNA integrity was assessed using an Agilent Bioanalyzer 2100 system (Agilent Technologies, Palo Alto, CA, USA). Sequencing was performed using an Illumina HiSeq 2500 platform (Novogene, China). After removing reads containing adapters and low-quality sequences, the resulting high-quality clean data were mapped to the alfalfa reference genome (https://figshare.com/articles/dataset/genome_fasta_sequence_and_annotation_files/12327602).

Functional annotation of differentially expressed genes (DEGs)

DESeq2 and P values were used to evaluate the differential gene expression between red and green stems. The difference in gene abundance between the samples was calculated based on the ratio of the two FPKM (Fragments per kilobase of exon per million fragments mapped) values. The false discovery rate (FDR) control method was used to identify the threshold of p values in multiple tests to calculate the significance of differences. DEGs were screened using the thresholds of $|\log_2(\text{fold-change})| > 1$ and $\text{FDR} < 0.05$. Enrichment analyses of the DEG sets were conducted using Gene Ontology (GO) annotation terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) biochemical pathway information with the 'cluster Profiler' package in R, correcting for gene length bias. GO terms with corrected p values < 0.05 were considered significantly enriched.

Verification of candidate genes by using quantitative real-time polymerase chain reaction (qRT-PCR)

To verify the RNA sequencing results, six unigenes were selected for validation via qRT-PCR. Primers were designed according to the DEG sequences using the online tool Primer 5 and synthesized using the SYBR GREEN dye method. The specific genes and primers used are listed in Table S1. The qRT-PCR procedure was conducted using SynScript[®] III RT SuperMix for qPCR (LightCycler[®] 480 II cycler, Roche, Carlsbad, CA, USA). The thermal profile for PCR amplification was as follows: 95 °C for 5 min, followed by 40 cycles at 95 °C for 10 s and another 40 cycles at 60 °C for 30 s. The following dissolution curves were generated: 95 °C for 15 s, 60 °C for 60 s, 95 °C for 15 s with continuous detection. All qRT-PCR analyses were performed using three technical and three biological replicates. The reference gene (β -actin gene) was used as an internal expression control. The expression levels of different genes relative to the control were calculated according to the $2^{-\Delta\Delta\text{CT}}$ method.

Integrated transcriptome and metabolome analysis

Pearson's correlation coefficient (PCC), a statistical method used to measure the linear relationship between two variables with values between -1 and 1 , was used to calculate the correlation coefficient between the transcriptome and metabolome data, and further integration was used to analyze the relationships between the gene transcript and metabolite content. Associations between $\text{PCCs} > 0.8$ and $P < 0.05$ were selected, and a network of genes and metabolites was constructed using Cytoscape software (Cytoscape Consortium, San Diego, CA, USA).

Data analysis

Data were analyzed using the Social Sciences (SPSS Version 20.0, SPSS Inc., Chicago, USA). An independent sample t-test was used to analyze the differences in total anthocyanin content among the differently colored stems, which revealed significant differences at the 5% level. Data are presented as mean \pm standard deviation (SD).

Results

Alfalfa with different stem colors and total anthocyanin content

The alfalfa stems were categorized into green and red colors (Fig. 1a). Apart from the stem color, there were no significant differences observed in leaf or stem morphology between the green 'G' and the red mutant 'R'. To elucidate the disparities in anthocyanin accumulation in alfalfa stems, the anthocyanin content was analyzed for both stem types. The results revealed that the total anthocyanin content in red-stemmed alfalfa was 15.03 mg/g fresh weight (FW), which was 2.89 times higher than that in green-stemmed alfalfa, consistent with the stem color phenotype (Fig. 1b).

Metabolic differences among the alfalfa stems

To compare the anthocyanin metabolite content disparities between green- and red-stemmed alfalfas, UPLC-MS/MS analysis was employed. The metabolite profiles of the green and red stems underwent PCA. The PCA score plots (Fig. 2) clearly demonstrate a distinct separation between green- and red-stemmed materials. The contribution rates of PC1 and PC2 were 81.58% and 6.64%, respectively. A high PC1 value indicates significant genetic variance between the two samples.

Analysis of differential metabolite components in G and R

The anthocyanin compound content of G and R was utilized to construct a hierarchical heatmap (Fig. 3). Interestingly, a distinct separation was observed between the G and R samples, indicating different anthocyanin profiles in the differently colored stems. A total of 45 anthocyanin metabolites were detected in the alfalfa stem

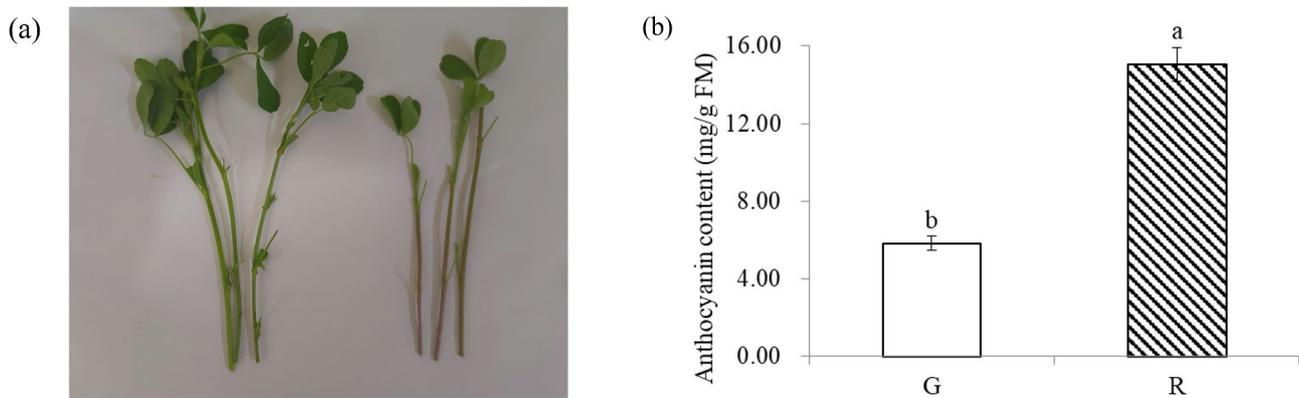


Fig. 1 Phenotypes and anthocyanins compounds of alfalfa materials. **(a)** Phenotypes of green- and red-stemmed alfalfa. **(b)** Anthocyanin content of green- and red-stemmed alfalfa. Values are the means \pm standard deviation (SD). G: green-stemmed alfalfa, R: red-stemmed alfalfa. Data is the means of three biological replicates. Different lowercase letters indicate significant differences at the 0.05 level

samples. Anthocyanin metabolites were categorized into eight groups: seven cyanidins, four delphinidins, four malvidins, nine pelargonidins, seven peonidins, four petunidins, five procyanidins, and five flavonoids. The levels of 33 metabolites differed between G and R (Table S2). Overall, the accumulation of anthocyanin metabolites was greater in the R sample than in the G sample, with significantly higher levels of cyanidin, delphinidin, pelargonidin, and peonidin in the R sample compared to the G sample. Specifically, cyanidin-3-O-glucoside, pelargonidin-3-O-arabinoside, delphinidin-3-O-(6-O-acetyl)-glucoside, and kaempferol-3-O-rutinoside in the R samples exhibited the highest levels. Among the 45 anthocyanin compounds detected, 12 were not detected in the green stems. These results suggest that the type and content of compounds anthocyanins contributing could be crucial to the color differences between G and R.

Transcriptome sequencing and identification of DEGs in alfalfa stems

To elucidate the molecular mechanisms of anthocyanin biosynthesis in red and green alfalfa stems, we performed transcriptome sequencing on the G and R samples. With three biological replicates for each sample, six cDNA libraries were constructed. The clean reads of the G and R samples were compared with those of the alfalfa reference genome. The number of clean reads in each library ranged from 20,339,182 to 27,075,917, and the comparison rate ranged from 90.27 to 91.37%. The percentage of Q20 bases in the obtained sequences was greater than 96.40%, and the percentage of Q30 bases was greater than 90.60%. The base error rate of each sample was relatively low, and the GC content of the six samples exceeded 41%, indicating a high GC content in the genes. The overall comparison efficiency exceeded 90% for similar species, confirming the correctness of the selected reference

genomes (Table 1). These results indicate that the constructed database meets the requirements for transcriptome analysis.

PCA was performed for transcription sample expression (Fig. 4a). The two samples were clearly distinguished by PCA, indicating that the stem transcripts of the two materials were different. Based on a $|\log_2(\text{FC})| \geq 1$ and $\text{FDR} < 0.05$, we identified 24,606 genes that were differentially expressed in G vs. R. Among them, 12,301 genes were upregulated, and 12,305 genes were downregulated (Fig. 4b and c).

GO enrichment and KEGG pathway analysis of DEGs

In this study, differentially expressed unigenes were functionally categorized using GO terms to better understand their roles in alfalfa stems of different colors. The GO analysis results revealed that these DEGs were classified into three major categories and 40 subcategories, among which three were cellular components, 23 were biological processes, and 14 were molecular functions. Among the 23 GO terms related to biological processes, DEGs were mainly annotated to cellular processes, metabolic processes, biological regulation, stimulus response, protein localization, and other functions. Notably, 9,494 genes were annotated to metabolic processes, including physiological processes related to photo responses, phenylpropane synthesis, flavonol synthesis, and pigment synthesis (Fig. 5). These results represent a general picture of DEG functions and provide information for understanding inter-sample differences at the gene function level.

To clarify the associated biological pathways, all DEGs were annotated using the KEGG database and functional enrichment analysis was performed. KEGG analysis was used to further investigate the transcriptome and the regulation of biological functions, secondary metabolite biosynthesis, and multigene interaction mechanisms in alfalfa at the molecular level. The KEGG pathway

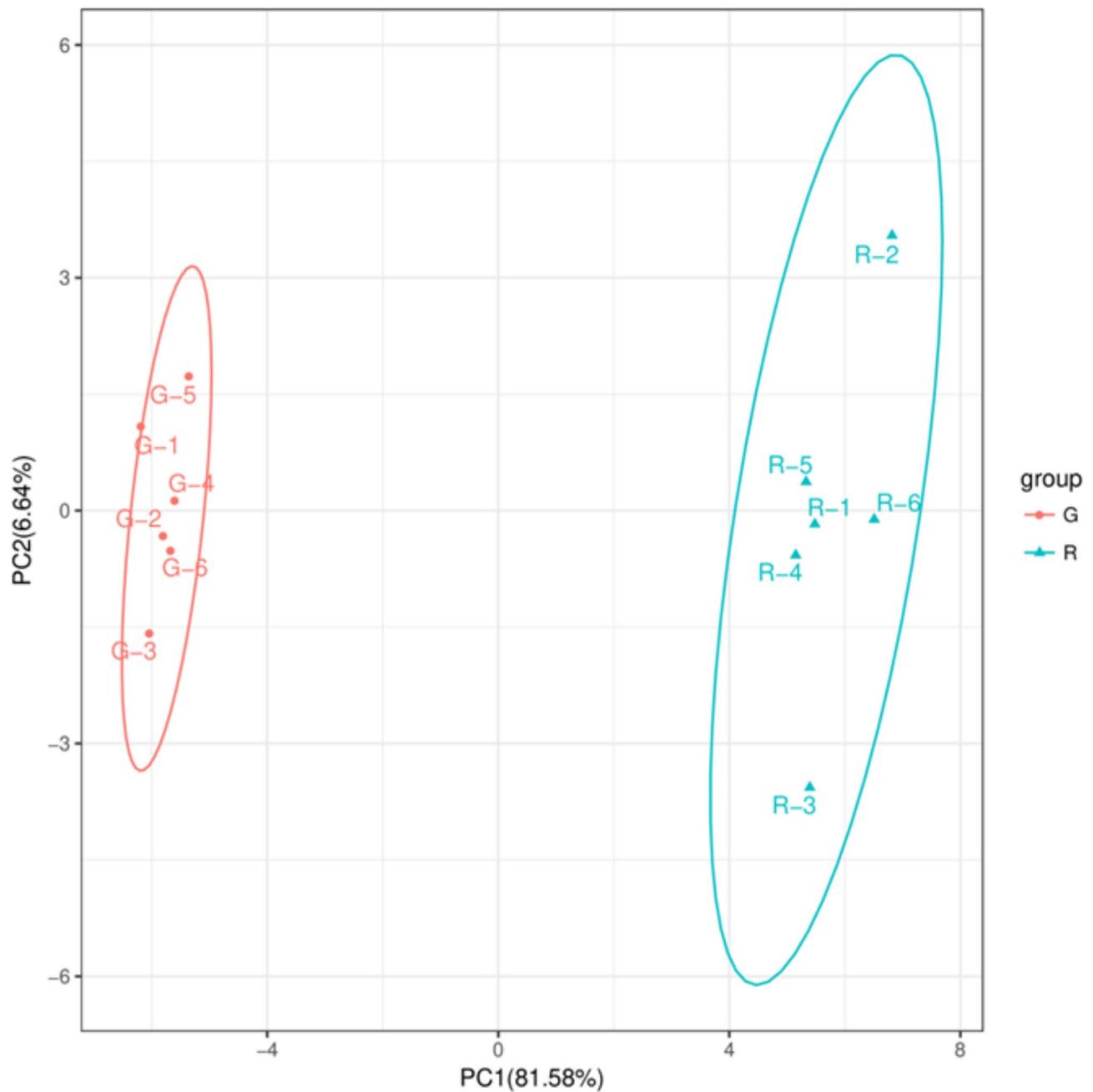


Fig. 2 The PCA score plot of alfalfa with green- and red-stem alfalfa. G: green-stemmed alfalfa, R: red-stemmed alfalfa

enrichment results revealed enrichment in 135 metabolic pathways, including plant-pathogen interaction, plant hormone signal transduction, starch and sucrose metabolism, ribosome, spliceosome, MAPK signaling pathway-plant, protein processing in endoplasmic reticulum, carbon metabolism, biosynthesis of amino acids, endocytosis, and phenylpropanoid biosynthesis (Fig. 6). Among these 135 pathways, phenylpropanoid biosynthesis (ko00940), phenylalanine metabolism (ko00360), flavonoid biosynthesis (ko00941), isoflavone biosynthesis (ko00943), and flavonol biosynthesis (ko00944) were

closely related to anthocyanin biosynthesis and significantly enriched in the stems.

DEG expression of anthocyanin biosynthesis structural genes

To further identify the functional genes involved in regulating red stems formation, the functional genes involved in anthocyanin biosynthesis in green- and red-stemmed alfalfa were analyzed. A total of 72 candidate genes involved in the anthocyanin biosynthetic pathway were identified based on KEGG pathway and Swiss-Prot

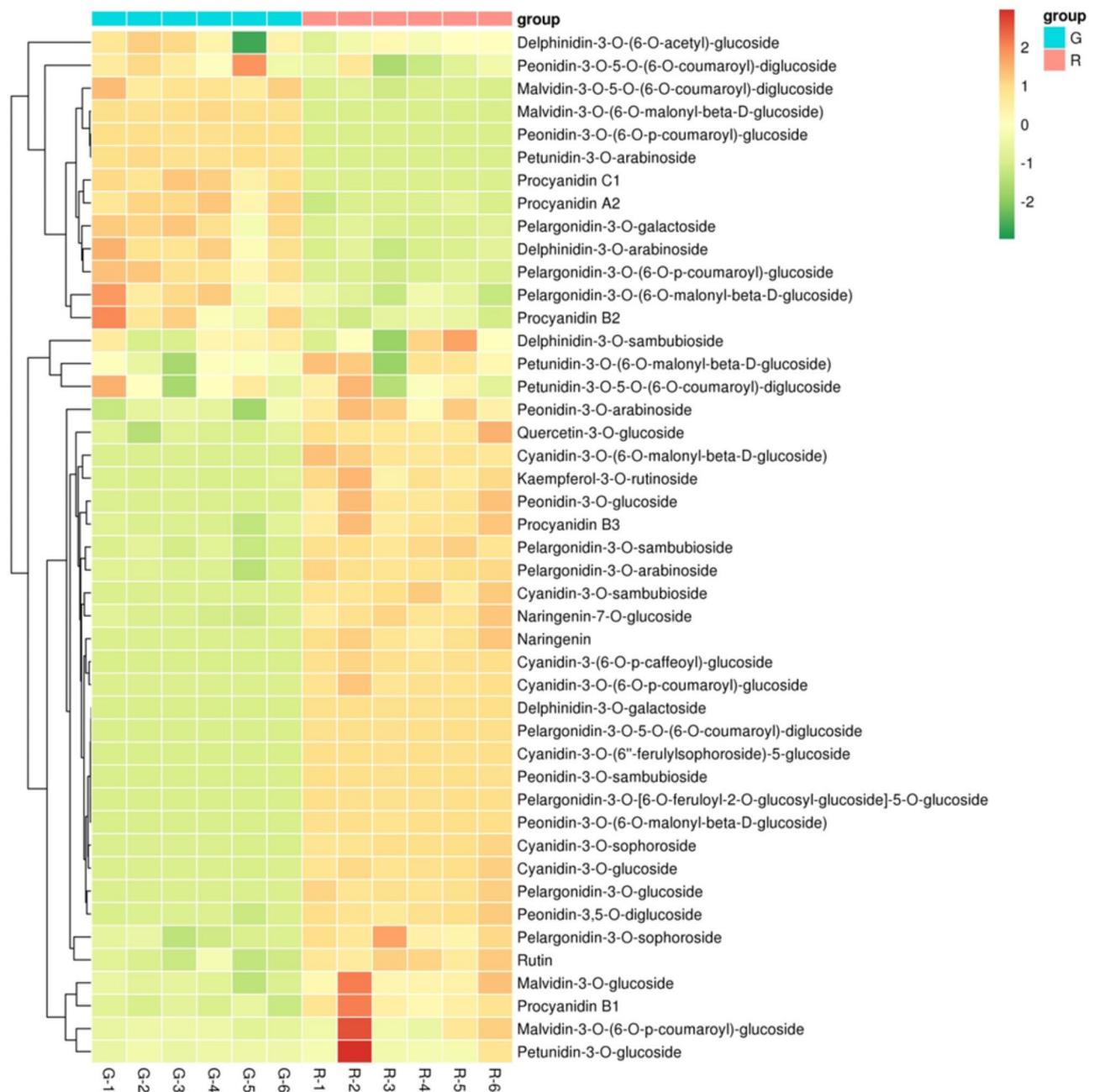


Fig. 3 Hierarchical clustering and correlation analysis of metabolites in the stems of green- and red-stemmed alfalfa. G: green-stemmed alfalfa, R: red-stemmed alfalfa

Table 1 Summary of transcriptomic data

Sample	Clean reads	Mapped reads	Q20 (%)	Q30 (%)	GC (%)
G1	22,016,159	90.27%	96.43	90.65	41.94
G2	23,123,183	91.37%	98.27	94.85	42.35
G3	27,075,917	90.56%	98.30	94.90	42.20
R1	20,681,044	90.31%	96.40	90.60	41.88
R2	24,376,752	90.82%	98.08	94.29	42.03
R3	20,339,182	91.13%	98.32	94.94	42.30

Note: G: green-stemmed alfalfa, R: red-stemmed alfalfa

annotations. Among these, 60 structural DEGs were related to upstream enzymatic reactions and the remaining 12 genes are downstream of the anthocyanin synthesis pathway. To compare the changes in gene expression levels between stems of different colors, six samples were stratified and clustered to observe the overall gene expression patterns. Notably, these 72 individual genes exhibited distinct expression patterns in the two stem colors: one group of genes was upregulated in the red stems, while another group was upregulated in the green

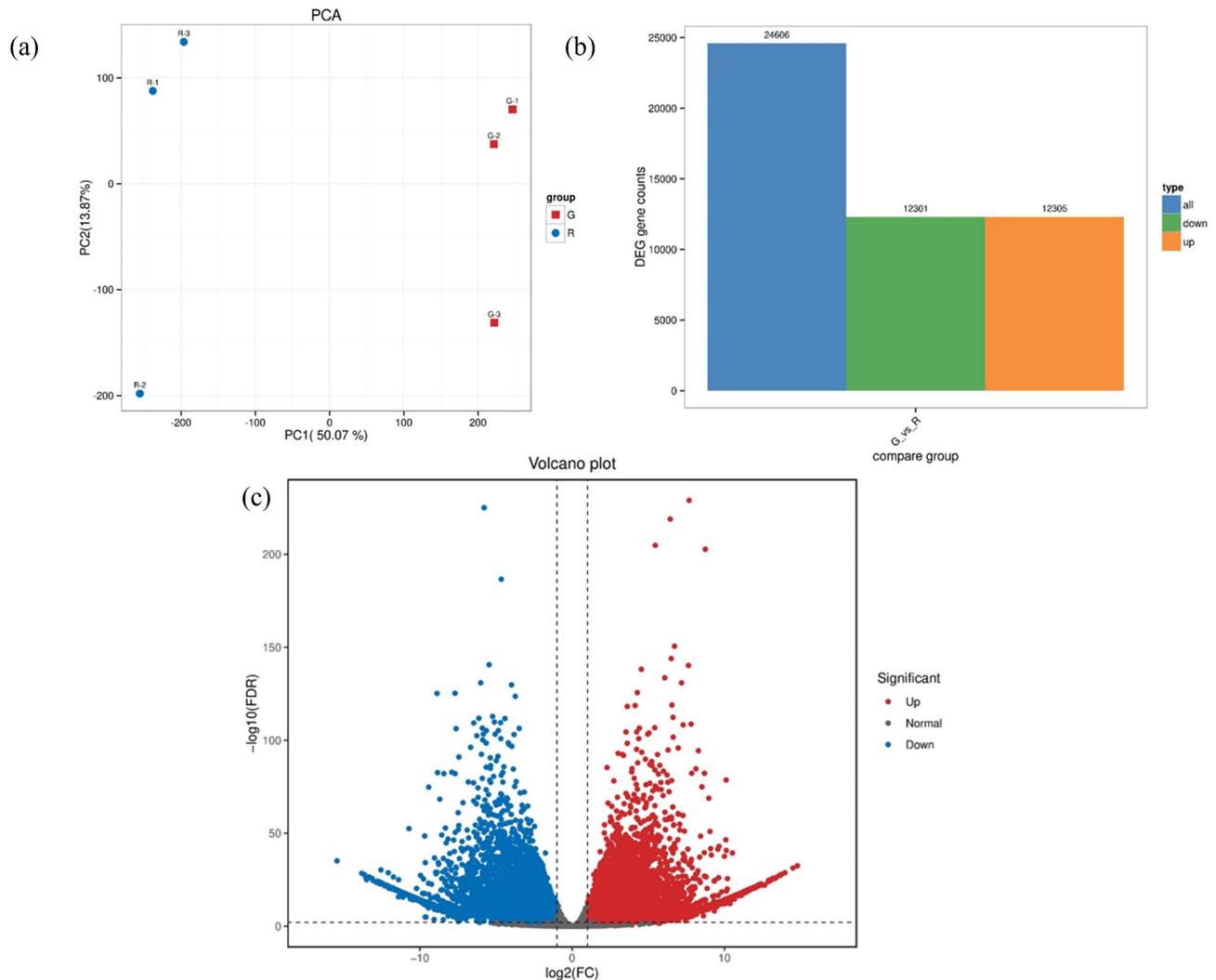


Fig. 4 (a) Principal component analysis results of the transcriptome. (b) Number of up- and down-regulated differentially expressed genes in green- and red-stemmed alfalfa. (c) Volcano plot of DEGs in green- and red-stemmed alfalfa. G: green-stemmed alfalfa, R: red-stemmed alfalfa

stems (Fig. 7). In these genes, 12 *PALs*, 22 *4CLs*, eight *CHSs*, three *F3Hs*, two *ANRs*, three *DFRs*, three *ANSs*, and one *FLS* gene showed higher expression levels in R, implying that these genes may be vital for the accumulation of anthocyanins (Table S3).

Expression patterns of anthocyanin biosynthesis pathway genes

To further determine the relationship between key genes involved in anthocyanin accumulation in alfalfa stems and the type of anthocyanin compound. A pathway map showing the expression of structural genes and anthocyanin compounds in green and red stems was constructed. We identified 72 genes associated with anthocyanin biosynthesis among the DEGs (Table S3). Anthocyanins are flavonoids that share a common main chain with other flavonoids upstream and form various anthocyanins via branch synthesis reactions from dihydrokaempferol. As

depicted in Fig. 8, upstream of the elements of the synthesis pathway, most of the *PAL*, *4CL*, and *CHS* genes had relatively high expression levels in red stems and a small number of genes had relatively high expression levels in green stems, providing substrates for the synthesis of downstream anthocyanin substances. Downstream of the synthesis pathway, the *F3H*, *ANS*, and *DFR* genes, which are involved in the synthesis of colored substances, showed the highest expression levels in red stems. Metabolome analysis revealed that the levels of anthocyanin-related metabolites significantly increased. These three genes play crucial roles in the transformation of naringenin into delphinidin, cyanidin, and pelargonidin. Further, the expression of genes encoding anthocyanins is related to the type of anthocyanin.

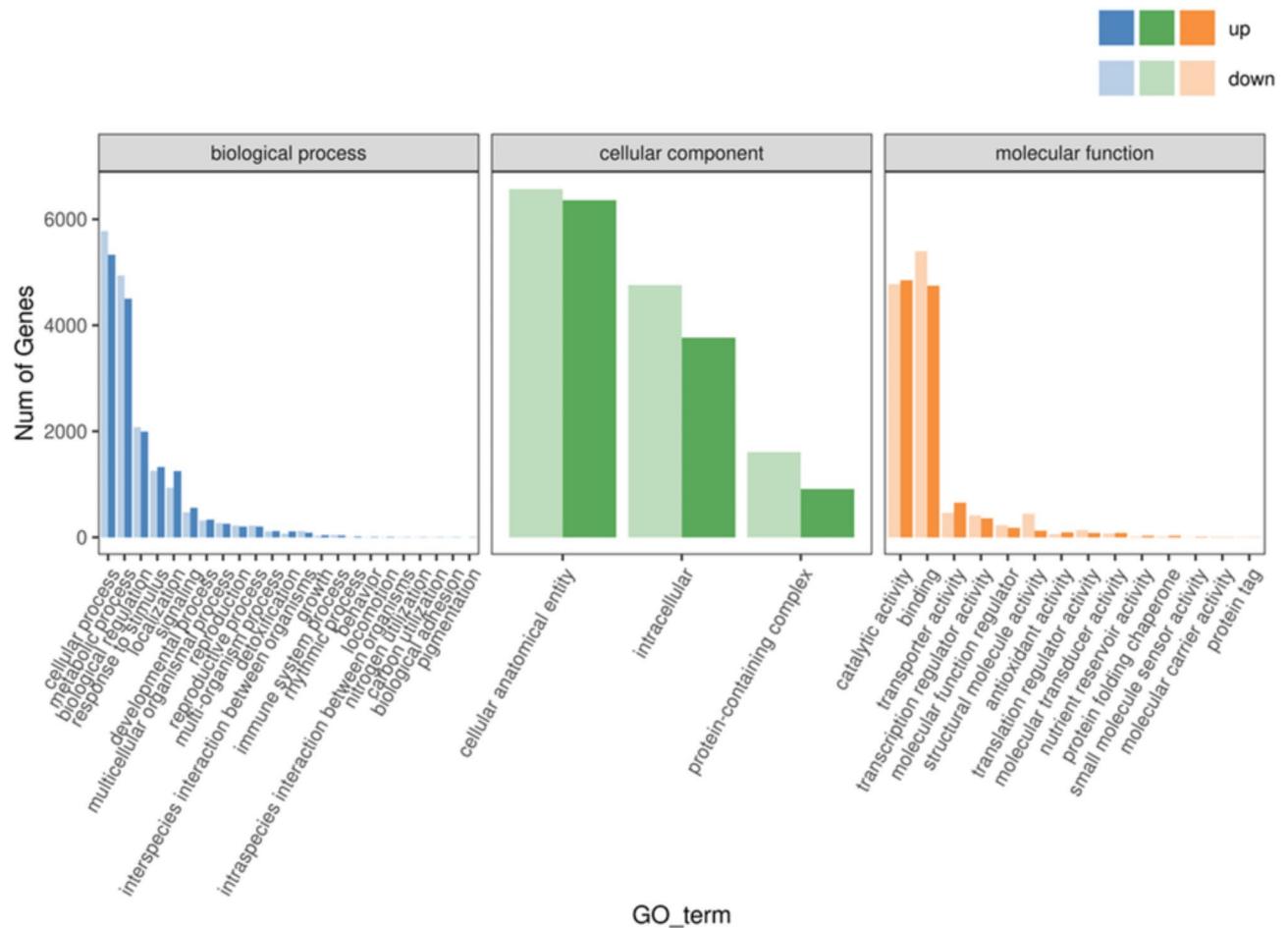


Fig. 5 Differentially expressed genes based on Gene Ontology enrichment in green and red-stemmed alfalfa

Validation of gene expression using qRT-PCR analysis

To validate the differences in the expression of genes associated with anthocyanin biosynthesis between the R and G varieties, six genes were selected from a pool of 72 candidate genes for qRT-PCR validation. The results presented in Fig. 9 reveal significant differences in gene expression between the two stem colors. These six genes were consistent with the DEG results from previously performed transcriptome analyses, further confirming the reliability of the sequencing data.

Correlation analysis of DEGs and DAMs

To comprehensively understand the regulatory network of anthocyanin synthesis in alfalfa stems, a correlation network analysis was conducted on 72 DEGs and 33 DAMs (Fig. 10). Correlation analysis showed that each anthocyanin was closely related to multiple genes, indicating that anthocyanin accumulation process was regulated by multiple genes, not just a single gene. The relative contents of cyanidin-3-O-glucoside, pelargonin-3-O-arabinoside and kaempferol-3-O-rutin were significantly positively correlated with the expressions of *4CL*, *PAL*,

CHS, *ANR*, *DFR*, *ANS*, *FLS*, and *F3H* ($PCC > 0.8$, $P < 0.05$) (Table S4). These genes are potential genes that regulate stem anthocyanin biosynthesis of alfalfa.

Discussion

Anthocyanin identification in alfalfa stems

Studies on color changes and anthocyanins in plants have primarily focused on the petals and pericarps, whereas few studies have focused on color changes in the stems of herbaceous plants. Anthocyanins play a vital role in the color formation of plant tissues [17], with plant color being largely determined by the type and content of anthocyanins and their derivatives [21, 22]. In this study, the anthocyanin content in red alfalfa stems was significantly greater than that in green alfalfa stems, indicating that anthocyanin accumulation is the primary cause of red color of alfalfa stems. This is in accordance with the findings of previous studies, which showed that most dark-colored plants contain anthocyanins as their main pigments [23]. Anthocyanins are important pigmentation compounds in plants, with colors ranging from red to purple. Brick red anthocyanins are mainly geranium

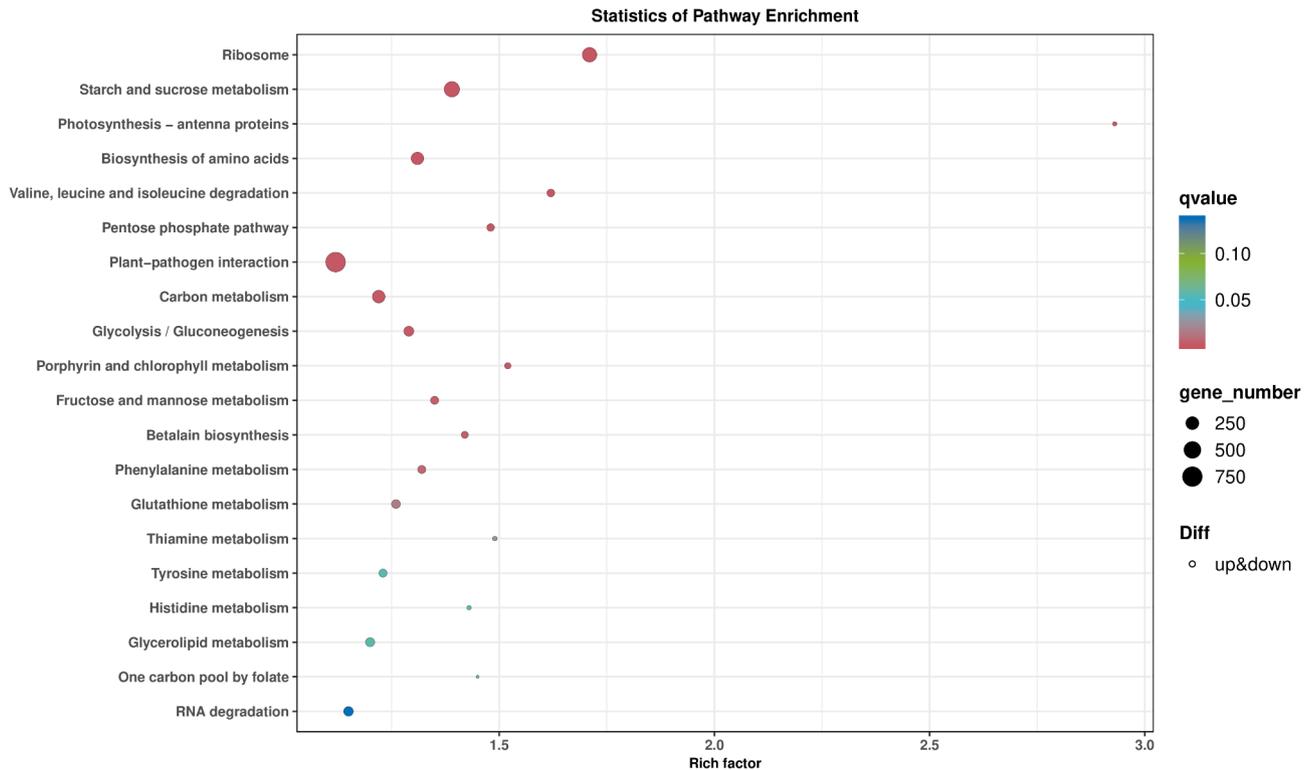


Fig. 6 The top 20 pathways of differentially expressed genes identified between green- and red-stemmed alfalfa

pigments, whereas mallow pigments, petunia pigments, and delphinidin are associated with blue through violet coloration [24]. In recent years, various anthocyanins have been identified in different plant tissues. For example, high levels of cyanidin-3,5-O-diglucoside, cyanidin-3-O-galactoside, cyanidin-3-O-glucoside, and pelargonidin-3-O-glucoside are responsible for sepal reddening in *Heptacodium miconioides* [25]. Cyanidin-3-O-galactoside is the main anthocyanin in ‘Jinhongguan’ a new cultivar of *Actinidia arguta* that has a red peel and flesh after harvest [26]. Cyanidin-3-O-glucoside and cyanidin-3-O-rutinoside are the two major anthocyanins in the *Sophora japonica* ‘AM’ mutant and the key metabolites responsible for the red color of the wing and keel petals [23]. Petunidin 3-O-glucoside, cyanidin 3-O-galactoside, and cyanidin 3-O-glucoside were the main components of the new purple-leaf tea cultivar Zikui (*Camellia sinensis* cv. Zikui) [27]. Similarly, in the present study, anthocyanin-targeted metabolome analysis revealed that seven anthocyanin aglycones, including cyanidin, pelargonidin, malvidin, petunidin, delphinidin, proanthocyanidin, and peonidin, were detected in both G and R. However, the composition and accumulation of anthocyanins were significantly different, which further indicated that anthocyanin composition and accumulation are associated with the stem color of alfalfa. Cyanidin-3-O-glucoside, pelargonidin-3-O-arabinoside, delphinidin-3-O-(6-O-acetyl)-glucoside,

and kaempferol-3-O-rutinoside were main anthocyanin components in red stem. These findings further reveal the composition and form of pigments in red-stemmed alfalfa.

Structural gene expression analysis of stem color in alfalfa

The regulatory mechanism of plant color formation is closely related to the expression of structural genes in the anthocyanin biosynthesis pathway [28]. Many structural genes are involved in the biosynthesis of anthocyanins, including the early structural genes *PAL*, *CAH*, *4CL*, *CHS*, *CHI*, *F3H*, and *F3'H*, and the late structural genes *DFR*, *UFGT*, and *ANS*. Previous research has shown that most of these structural genes are highly expressed in the tissues of purple plants [22]. In this study, we identified eight structural gene families associated with stem anthocyanin biosynthesis pathways in alfalfa. These eight structural gene families comprise five early biosynthesis genes (*PAL*, *4CL*, *CHS*, *FLS*, and *F3H*) and three anthocyanin biosynthesis genes (*DFR*, *ANR*, and *ANS*), most of them are highly expressed in red-stemmed. Anthocyanin biosynthesis in higher plants requires the coordinated participation of various enzymes. *PAL* is the first key enzyme in the phenylpropane metabolic pathway and is the rate-limiting enzyme for anthocyanin synthesis. The *CHS* regulates continuous decarboxylation and condensation reactions using acyl-CoA as a precursor to form a polyketone intermediate, which provides a basic carbon

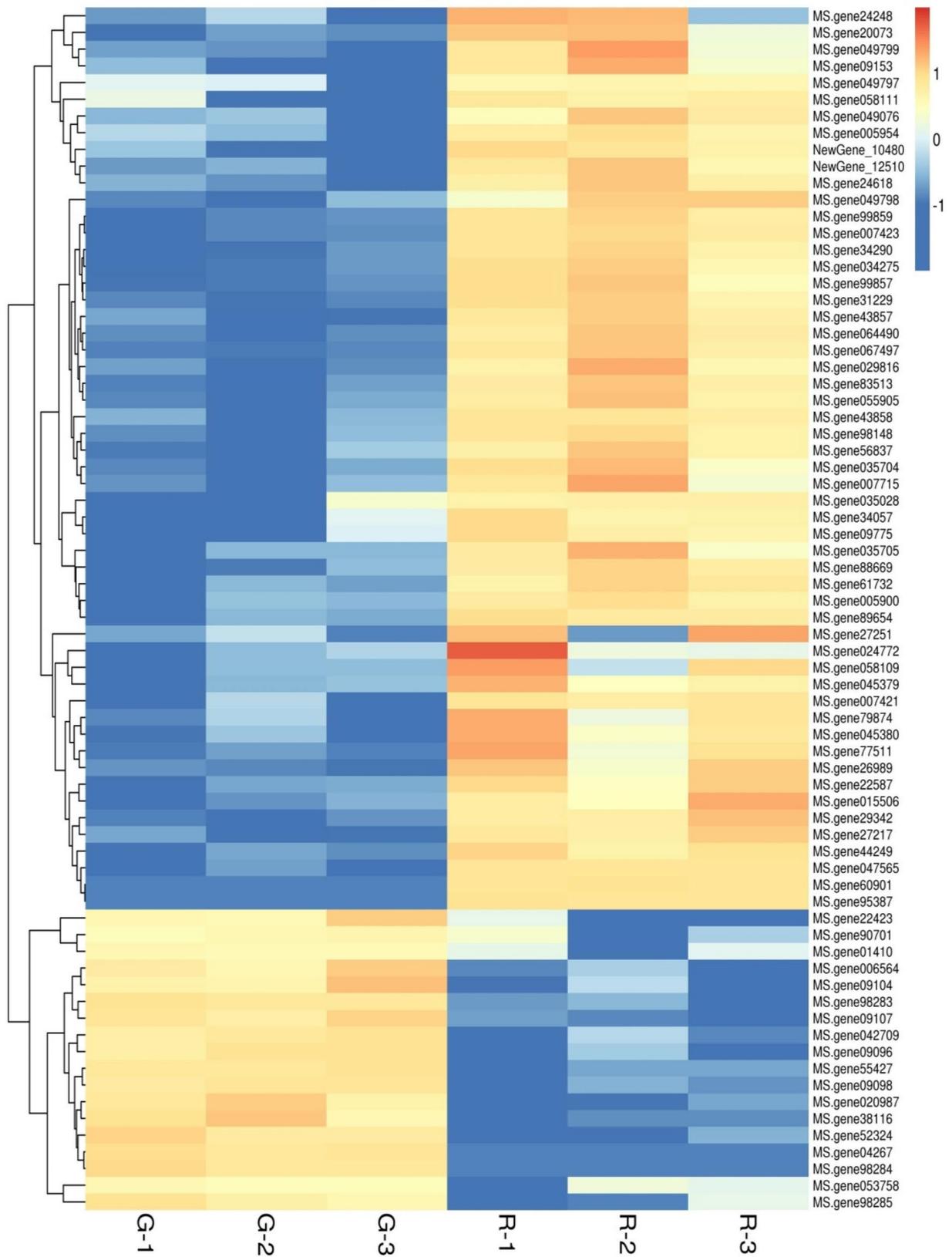


Fig. 7 Heatmap of the expression levels of anthocyanin biosynthesis pathway genes

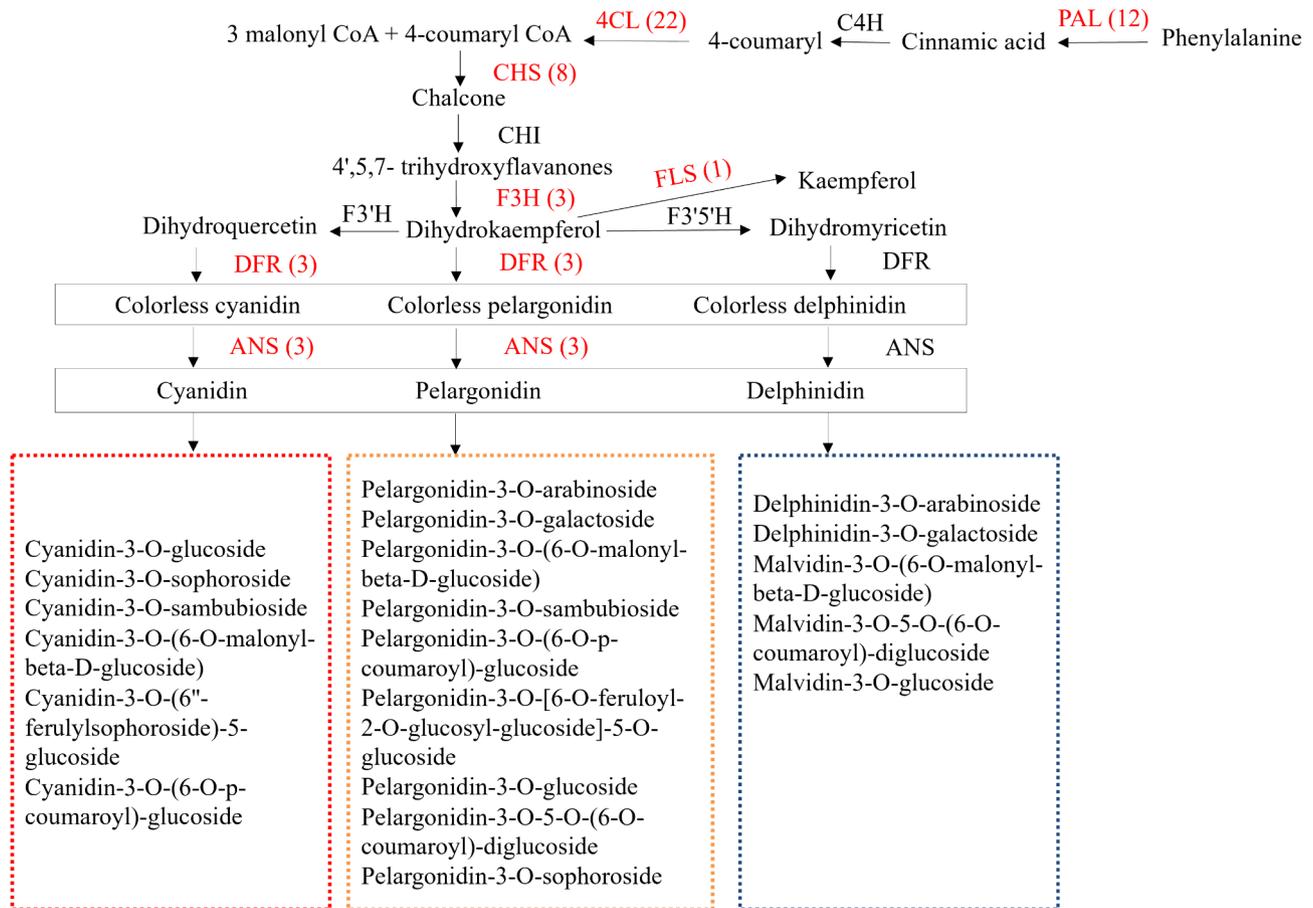


Fig. 8 Expression patterns of anthocyanin biosynthesis pathway genes. Red font indicates genes upregulated in red-stemmed alfalfa. PAL: phenylalanine ammonia-lyase; C4H: cinnamic acid 4-hydroxylase; 4CL: 4-coumarate: CoA- ligase; CHS: chalcone synthase; CHI: chalcone isomerase; F3'H: flavanoid 3'-hydroxylase; F3'5'H: flavanoid 3',5'-hydroxylase; F3H: flavanone 3-hydroxylase; DFR: dihydroflavonol 4-reductase; ANS: anthocyanidin synthase; FLS: flavonol synthase

skeleton for anthocyanin synthesis [23]. *4CL* is a key gene encoding a 4-coumarate: CoA ligase that affects the synthesis of 4-coumaroyl CoA, which is an important limiting intermediate, before branching to synthesize aromatic volatiles, including coumarin and lignin [29]. *F3H* catalyzes the conversion of flavanone to dihydroflavanol, and the upregulation of *F3H* facilitates the accumulation of anthocyanins [30]. In the present study, 12 *PALs*, 8 *CHSs*, 22 *4CLs*, and 3 *F3H* genes were highly expressed in red stems. These genes can promote the synthesis of upstream substrates can promote and provide sufficient precursors for downstream metabolism, thereby affecting anthocyanin synthesis and stem color. A similar finding was observed in the sepals of *H. miconioides*, where the expression of *HmPAL*, *Hm4CL*, *HmC4H*, *HmCHS*, *HmF3H*, and *HmANS* showed a strong positive correlation with the content of anthocyanins across the four sepal stages [25].

Plant anthocyanin synthesis is highly susceptible to interruptions at downstream positions in the biosynthetic pathway. This vulnerability may be attributed

to alterations in the activity or content of one or more enzymes involved in anthocyanin synthesis, which in turn affects the accumulation of these compounds in fruits. *DFR* is a key gene involved in the synthesis of anthocyanins from flavonols. Under the action of *DFR*, flavonols are converted into colorless anthocyanin precursors (e.g., colorless pelargonidin and colorless cyanidin). Many studies have demonstrated that *DFR* is a critical gene associated with anthocyanin biosynthesis in plant tissues and organs [31, 32]. In this study, three *DFR* genes (*MS.gene24248*, *MS.gene26989*, and *MS.gene27251*) exhibited significantly increased expression levels in red stems compared to those in green stems. It is closely associated with the accumulation of cyanidin-3-O-glucoside, pelargonidin-3-O-arabinoside, and kaempferol-3-O-rutinoside. In purple sweet potatoes, the expression of structural genes such as *DFR* is upregulated, leading to quantitative and qualitative variations in anthocyanin content [33]. High *ANS* expression promotes the conversion of colorless anthocyanins into anthocyanins. Anthocyanins are stabilized by glycosyltransferases through

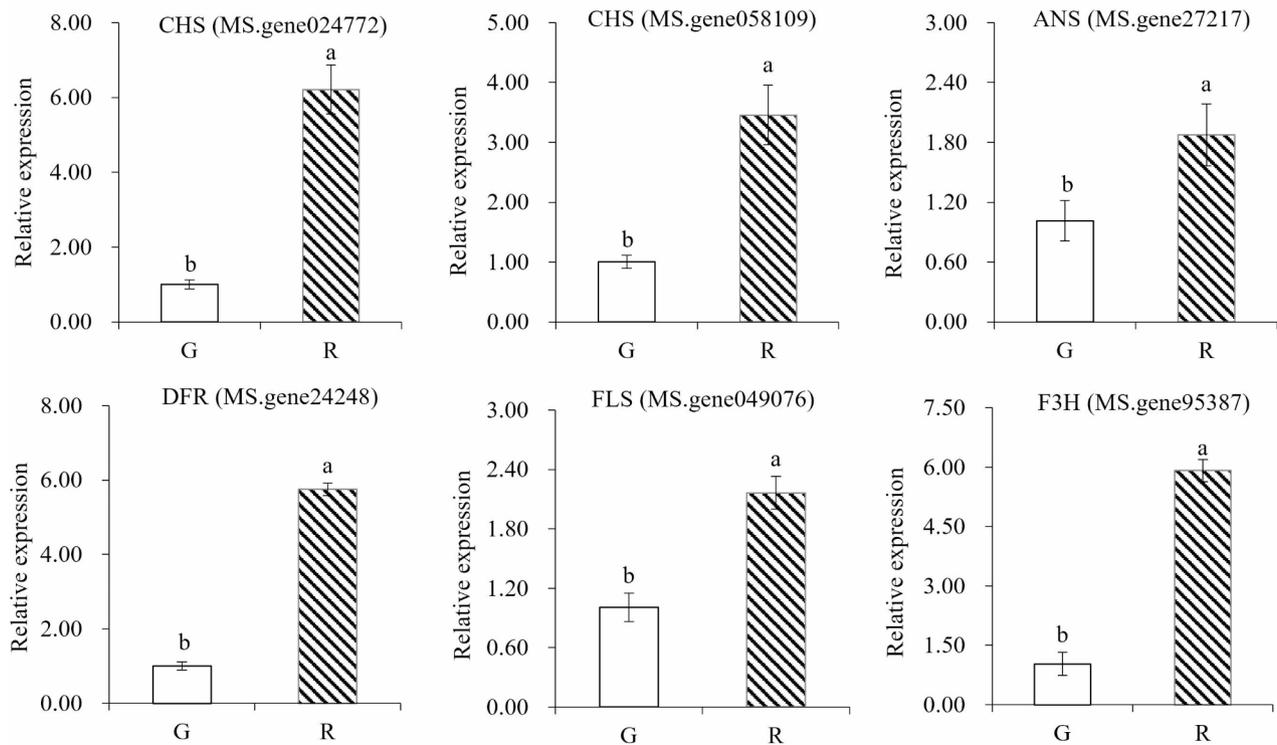


Fig. 9 Transcription levels of six anthocyanin biosynthesis genes as determined by qRT-PCR. Error bars represent the standard deviations of the mean in three replicates, and different lowercase letters indicate significant differences at 0.05

glycosylation, resulting in the formation of stable anthocyanins (such as cyanidin-3-

O-glucoside and cyanidin-3-O-rutinoside). This process has been confirmed as a key step in the regulation of anthocyanin types [33]. In this study, three *ANS* genes (MS.gene22587, MS.gene27217, and MS.gene015506) exhibited significantly increased expression levels in red stems compared to those in green stems. Overexpression of these *DFR* and *ANS* genes enhanced the accumulation of anthocyanins and resulted in the production of red-stemmed with expression patterns consistent with previous findings. For example, in the bicolor petal winter pot kalanchoe (*Kalanchoe blossfeldiana*), the biosynthetic genes *KbANS* and *KbDFR* are significantly upregulated in the red petal region [34].

The color change is influenced by the synthesis and transport of anthocyanins, as well as the expression levels of related transcription factors. In eukaryotes, RNA polymerase and certain transcription factors are often required to co-anchor to specific cis-acting elements in gene promoters to initiate gene transcription. However, no differentially expressed transcription factors associated with anthocyanin synthesis were detected in alfalfa stems. This further suggests that the structural genes may be the key genes responsible for anthocyanin accumulation in red stems. The up-regulated expression of key enzyme gene family members in the anthocyanin

synthesis pathway promoted the synthesis and accumulation of anthocyanin metabolites, ultimately contributing to the red coloration of stems.

Conclusions

In this study, transcriptome and metabolome analyses were integrated to identify key metabolites and candidate genes involved in anthocyanin biosynthesis in alfalfa stems. Determination of total anthocyanin content showed that anthocyanin accumulation may be the main cause of the red color change in alfalfa stems. Additionally, the anthocyanin-targeted metabolome revealed that the anthocyanin composition and content of anthocyanins among stems of different colors. Cyanidin-3-O-glucoside, pelargonidin-3-O-arabinoside, delphinidin-3-O-(6-O-acetyl)-glucoside, and kaempferol-3-O-rutinoside are the main anthocyanin components in red stem. Transcriptome analysis indicated that *PAL*, *4CL*, *CHS*, *F3H*, *ANR*, *DFR*, *ANS*, and *FLS* were likely pivotal for anthocyanin synthesis in red-stemmed alfalfa. These results provide valuable insights into the biosynthesis of anthocyanins in red-stemmed alfalfa.

Data availability

The reference genome data of alfalfa was obtained from (https://figshare.com/articles/dataset/genome_fasta_sequence_and_annotation_files/12327602). Data are provided within the manuscript or supplementary information files.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Faculty of Animal Science and Technology, Yunnan Agricultural University, Kunming 650201, China

Received: 15 October 2024 / Accepted: 25 March 2025

Published online: 31 March 2025

References

- McDonald I, Baral R, Min D. Effects of alfalfa and alfalfa-grass mixtures with nitrogen fertilization on dry matter yield and forage nutritive value. *J Anim Sci Technol.* 2021;63(2):305.
- Berti MT, Lukaszewsky J, Samarappuli DP. Intercropping alfalfa into silage maize can be more profitable than maize silage followed by spring-seeded alfalfa. *Agronomy.* 2021;11(6):1196.
- Duan HR, Wang LR, Cui GX, et al. Identification of the regulatory networks and hub genes controlling alfalfa floral pigmentation variation using RNA-sequencing analysis. *BMC Plant Biol.* 2020;20:1–17.
- Xue Q, Zhang X, Yang H, et al. Transcriptome and metabolome analysis unveil anthocyanin metabolism in Pink and red testa of peanut (*Arachis Hypogaea* L.). *Int J Genomics.* 2021;2021(1):5883901.
- Zhang Y, Hu Z, Zhu M, et al. Anthocyanin accumulation and molecular analysis of correlated genes in purple Kohlrabi (*Brassica Oleracea* Var. Gongylodes L.). *J Agric Food Chem.* 2015;63(16):4160–9.
- Castellarin SD, Pfeiffer A, Sivilotti P, et al. Transcriptional regulation of anthocyanin biosynthesis in ripening fruits of grapevine under seasonal water deficit. *Plant Cell Environ.* 2007;30(11):1381–99.
- Zhuang H, Lou Q, Liu H, et al. Differential regulation of anthocyanins in green and purple turnips revealed by combined de Novo transcriptome and metabolome analysis. *Int J Mol Sci.* 2019;20(18):4387.
- Xi D, Gao L, Li XF et al. Metabolome and transcriptome revealed genes associated with anthocyanin accumulation in purple Caitai (*Brassica campestris*. var. *tsai-tai* Hort.). *Scientia Horticulturae.* 2022;303: 111171.
- Lee DS, Jeon DS, Park SG, et al. Effect of cold storage on the contents of glucosinolates in Chinese cabbage (*Brassica Rapa* L. Ssp. *pekinensis*). *South Indian J Biol Sci.* 2015;1(1):38–42.
- Guo Y, Zhang H, Shao S, et al. Anthocyanin: A review of plant sources, extraction, stability, content determination and modifications. *Int J Food Sci Technol.* 2022;57(12):7573–91.
- Carreno-Quintero N, Bouwmeester HJ, Keurentjes JJB. Genetic analysis of metabolome–phenotype interactions: from model to crop species. *Trends Genet.* 2013;29(1):41–50.
- Shen Y, Mao L, Zhou Y, et al. Integrated transcriptome and metabolome analysis revealed the molecular mechanism of anthocyanin synthesis in purple leaf pepper (*Capsicum annuum* L.) under different light intensities. *Horticulturae.* 2023;9(7):814.
- Fu Z, Shang H, Jiang H, et al. Systematic identification of the light-quality responding anthocyanin synthesis-related transcripts in Petunia petals. *Hortic Plant J.* 2020;6(6):428–38.
- Jiao F, Zhao L, Wu X, et al. Metabolome and transcriptome analyses of the molecular mechanisms of flower color mutation in tobacco. *BMC Genomics.* 2020;21:1–10.
- Zhang J, Li S, An H, et al. Integrated transcriptome and metabolome analysis reveals the anthocyanin biosynthesis mechanisms in blueberry (*Vaccinium corymbosum* L.) leaves under different light qualities. *Front Plant Sci.* 2022;13:1073332.
- Jaakola L. New insights into the regulation of anthocyanin biosynthesis in fruits. *Trends Plant Sci.* 2013;18(9):477–83.
- Feng X, Gao G, Yu C, et al. Transcriptome and metabolome analysis reveals anthocyanin biosynthesis pathway associated with Ramie (*Boehmeria Nivea* (L.) Gaud.) leaf color formation. *BMC Genomics.* 2021;22:1–15.
- Fu M, Yang X, Zheng J, et al. Unraveling the regulatory mechanism of color diversity in camellia Japonica petals by integrative transcriptome and metabolome analysis. *Front Plant Sci.* 2021;12:685136.
- Ma S, Hu R, Ma J, et al. Integrative analysis of the metabolome and transcriptome provides insights into the mechanisms of anthocyanins and proanthocyanidins biosynthesis in trifolium repens. *Ind Crops Prod.* 2022;187:115529.
- Shao D, Liang Q, Wang X, et al. Comparative metabolome and transcriptome analysis of anthocyanin biosynthesis in white and Pink petals of cotton (*Gossypium hirsutum* L.). *Int J Mol Sci.* 2022;23(17):10137.
- Ilahy R, Tilili I, Siddiqui MW, et al. Inside and beyond color: comparative overview of functional quality of tomato and watermelon fruits. *Front Plant Sci.* 2019;10:769.
- Xiao L, Cao S, Shang X, et al. Metabolomic and transcriptomic profiling reveals distinct nutritional properties of cassavas with different flesh colors. *Food Chem.* 2021;2:100016.
- Guan L, Liu J, Wang R, et al. Metabolome and transcriptome analyses reveal flower color differentiation mechanisms in *Various sophora Japonica* L. Petal Types Biology. 2023;12(12):1466.
- Yan J, Qian L, Zhu W, et al. Integrated analysis of the transcriptome and metabolome of purple and green leaves of *Tetragium Hemsleyanum* reveals gene expression patterns involved in anthocyanin biosynthesis. *PLoS ONE.* 2020;15(3):e0230154.
- Li Y, Sun Z, Li J. Integrated transcriptomics and metabolomics analysis provide insight into anthocyanin biosynthesis for sepal color formation in *Heptacodium Miconioides*. *Front Plant Sci.* 2023;14:1044581.
- Ye L, Bai F, Zhang L, et al. Transcriptome and metabolome analyses of anthocyanin biosynthesis in post-harvest fruits of a full red-type Kiwifruit (*Actinidia arguta*) ‘‘hongguan’’. *Front Plant Sci.* 2023;14:1280970.
- Cai J, Lv L, Zeng X, et al. Integrative analysis of metabolomics and transcriptomics reveals molecular mechanisms of anthocyanin metabolism in the Zikui tea plant (*Camellia sinensis* Cv. Zikui). *Int J Mol Sci.* 2022;23(9):4780.
- Nakatsuka T, Suzuki T, Harada K, et al. Floral organ- and temperature-dependent regulation of anthocyanin biosynthesis in *Cymbidium* hybrid flowers. *Plant Sci.* 2019;287:110173.
- Nishiyama Y, Yun CS, Matsuda F, et al. Expression of bacterial tyrosine ammonia-lyase creates a novel p-coumaric acid pathway in the biosynthesis of phenylpropanoids in Arabidopsis. *Planta.* 2010;232:209–18.
- Xu J, Fan Y, Han X, et al. Integrated transcriptome and metabolomic analysis reveal the underlying mechanism of anthocyanin biosynthesis in Toona sinensis leaves. *Int J Mol Sci.* 2023;24(20):15459.
- Huang X, Liu L, Qiang X, et al. Integrated metabolomic and transcriptomic profiles provide insights into the mechanisms of anthocyanin and carotenoid biosynthesis in petals of *Medicago Sativa* Ssp. *Sativa* and *Medicago Sativa* Ssp. *Falcatia.* *Plants.* 2024;13(5):700.
- Katsumoto Y, Fukuchi-Mizutani M, Fukui Y, et al. Engineering of the Rose flavonoid biosynthetic pathway successfully generated blue-hued flowers accumulating Delphinidin. *Plant Cell Physiol.* 2007;48(11):1589–600.
- Xiao JP, Xu XY, Li MX, et al. Regulatory network characterization of anthocyanin metabolites in purple sweet potato via joint transcriptomics and metabolomics. *Front Plant Sci.* 2023;14:1030236.
- Feng G, Wang J, Pan Z, et al. Integrative metabolome and transcriptome analyses reveal the molecular mechanism of yellow-red bicolor formation in *Kalanchoe Blossfeldiana* petal. *Horticulturae.* 2023;9(7):844.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.