



OPEN

Comparative transcriptome analysis reveals key genes associated with pigmentation in radish (*Raphanus sativus* L.) skin and flesh

Jifang Zhang^{1,2,4}✉, Jian Zhao^{3,4}, Qunyun Tan^{1,2}, Xiaojun Qiu^{1,2} & Shiyong Mei^{1,2}✉

Radish (*Raphanus sativus*) is an important vegetable worldwide that exhibits different flesh and skin colors. The anthocyanins responsible for the red and purple coloring in radishes possess nutritional value and pharmaceutical potential. To explore the structural and regulatory networks related to anthocyanin biosynthesis and identify key genes, we performed comparative transcriptome analyses of the skin and flesh of six colored radish accessions. The transcript profiles showed that each accession had a species-specific transcript profile. For radish pigmentation accumulation, the expression levels of anthocyanin biosynthetic genes (*RsTT4*, *RSC4H*, *RsTT7*, *RSCCOAMT*, *RSDFR*, and *RsLDOX*) were significantly upregulated in the red- and purple-colored accessions, but were downregulated or absent in the white and black accessions. The correlation test, combined with metabolome (PCC > 0.95), revealed five structural genes (*RsTT4*, *RSDFR*, *RSCCOAMT*, *Rsf3H*, and *RsbG8L*) and three transcription factors (*RsTT8-1*, *RsTT8-2*, and *RsPAR1*) to be significantly correlated with flavonoids in the skin of the taproot. Four structural genes (*RsbG8L*, *RSDFR*, *RSCCOAMT*, and *RsLDOX*) and nine transcription factors (*RsTT8-1*, *RsTT8-2*, *RsmYB24L*, *RsbHLH57*, *RsPAR2L*, *RsbHLH113L*, *RsOGR3L*, *RsmYB24*, and *RsmYB34L*) were found to be significantly correlated with metabolites in the flesh of the taproot. This study provides a foundation for future studies on the gene functions and genetic diversity of radish pigmentation and should aid in the cultivation of new valuable radish varieties.

Flavonoids are plant polyphenolic secondary metabolites that are widespread in the plant kingdom. Flavonoids can be classified into many subgroups based on their chemical structures and modifications, including flavanones, flavones, isoflavonoids, flavanols, anthocyanins, and flavonols¹. Anthocyanins are the major pigment metabolites of flavonoid compounds. They are abundant in the flowers, fruits, seeds, and leaves of many plant species² and play important roles in plant protection and reproduction^{3,4}. They can help plants attract pollinators and seed dispersers by stimulating red, purple, and blue pigments in plant tissues³. In recent years, there has been increased interest in anthocyanins because of their wide use as natural commercial food pigments and their potential health benefits⁵⁻⁷.

Flavonoid biosynthesis and regulatory pathways have been extensively described. Enzymes, such as CHS, F3H, F3'H, F3'5'H, and LDOX/ANS, are involved in flavonoid biosynthesis and may act as metabolic hubs that influence the overall efficiency, specificity, and regulation of these pathways⁸⁻¹⁰. DNA-binding R2R3 MYB transcription factors (TFs), MYC-like basic helix-loop-helix (bHLH), and WD40-repeat proteins interact in the regulation of flavonoid biosynthesis¹¹⁻¹⁹. Plant hormones²⁰⁻²² and environmental factors²³⁻²⁶ also play important roles in the regulation of anthocyanin biosynthesis. Nevertheless, the mechanisms that regulate anthocyanin metabolism in different plant species and tissues require further investigation²⁷.

Radishes (*Raphanus sativus*) are common vegetables worldwide that possess a wide variety of colors, shapes, and sizes. The pigments in radish taproots have health-promoting properties. Based on anthocyanin production,

¹Institute of Bast Fiber Crops, Chinese Academy of Agricultural Science, Changsha, China. ²Center for Southern Economic Crops, Chinese Academy of Agricultural Science, Changsha, China. ³Novogene Bioinformatics Institute, Beijing, China. ⁴These authors contributed equally: Jifang Zhang and Jian Zhao. ✉email: smilehome@163.com; hbvegt@163.com

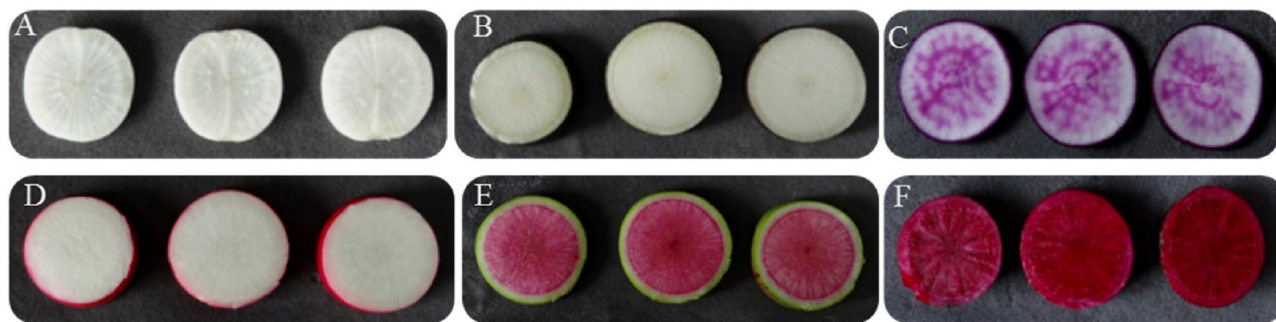


Figure 1. Taproot colors of the six radish accessions: (A) Shizhuangbai (SZB) has white taproot skin (WS) and white flesh (WF); (B) black radish has black taproot skin (BS) and white flesh (WF); (C) Zijinling (ZJL) has purple taproot skin (ZJLS) and purple flesh (ZJLF); (D) Manshenhong (MSH) has light red taproot skin (PS) and white flesh (WF); (E) Xinlimei has green taproot skin (XLMS) and red flesh (XLMF); (F) Touxinhong (TXH) has dark red taproot skin (RS) and dark red flesh (RF).

radishes can be categorized into those that have anthocyanins in their skin and flesh, and those that have anthocyanins only in their skin. Several varieties with different skin and/or flesh taproot colors (e.g., white, green, red, purple, and black) are widely cultivated and consumed. However, the mechanisms that regulate anthocyanin catabolism in radish have not received much research attention.

Fortunately, correlation and clustering analyses based on transcript and metabolite information have been used to further build networks between genes and metabolites in many plants, including *Arabidopsis*^{28,29}, tomato (*Solanum lycopersicum*)^{30–32}, potato (*Solanum tuberosum*)^{33–34}, grape (*Vitis vinifera*)²⁷, and *Actinidia arguta*³⁵. The integration of “omics” datasets (i.e., transcriptomes and metabolomes) is considered a useful approach for discovering and identifying potential genes that regulate the determination of radish pigmentation.

In this study, RNA sequencing (RNA-seq) analysis using the Illumina platform was used to study the genetic control of the differential accumulation of anthocyanins. By combining our previous metabolite datasets³⁶ in a bioinformatics analysis, the major metabolic pathways related to radish taproot pigmentation were deduced, and candidate genes targeting the regulation of radish pigmentation were examined.

Results

Anthocyanin contents of radish taproot skin and flesh. Six radish accessions (Fig. 1) with various skin and flesh colors were used for the anthocyanin content assay. The skin of the radish taproot (Fig. 2), ZJL, which is purple in color, exhibited the highest anthocyanin content, followed by the dark red TXH and the light red MSH. The skin of SZB, black radish, and XLM exhibited trace or undetectable contents. In terms of flesh color (Fig. 2), TXH, which has dark red flesh, showed the highest anthocyanin content, followed by the red-fleshed XLM and the purple-fleshed ZJL. The white flesh of SZB had no detectable anthocyanin content. The purple skin of ZJL had a higher anthocyanin content than the dark red flesh of TXH.

RNA sequencing and assembly. The six radish accessions (Fig. 1) with various skin and flesh colors were passed through RNA-seq and analysis, generating a total of 1,695,680,738 raw reads (Table 1). After adaptor sequence trimming and filtering out of low-quality reads, 1,671,174,278 (98.55%) clean reads were obtained with an average GC content of 46.74% and a Q20 average base quality score of 97.79%. A total of 250.71 Gb of high-quality reads were obtained for the radish samples. Of these clean reads, 68.62–87.7% were mapped to the reference genome WK10039, and 62.3–80.24% were uniquely mapped to WK10039. Furthermore, 31.15–40.09% of the clean reads were mapped to the positive strands, and 31.14–40.15% were mapped to the negative strands (Table S2).

Differential gene expression levels. Principal component analysis (PCA) was used to assess inter-group differences and sample duplication within groups. In the PCA plot, 10 samples from the six accessions clustered into six independent groups, suggesting that each group had a relatively distinct transcript profile (Fig. 3A). Group 1 included white-skinned and white-fleshed taproots of the SZB accession (WST and WFT), and group 2 included green-skinned and red-fleshed taproots of the XLM accession (XLMST and XLMFT). The dark red-skinned and dark-red-fleshed taproots of the TXH accessions (RST and RFT) clustered in group 3, and the purple-skinned and purple-fleshed taproots of the ZJL accessions (ZJLST and ZJLFT) clustered in group 6. Accessions with the same colored skin and flesh clustered together.

To understand the molecular basis and predict the candidate genes responsible for color polymorphisms in radish, differentially expressed genes (DEGs) were identified by pairwise comparisons of the expression levels between colored and white samples. The transcriptome analysis revealed that 19,576 (10,721 upregulated and 8855 downregulated), 22,319 (12,143 upregulated and 10,176 downregulated), 15,776 (8576 upregulated and 7200 downregulated), 17,474 (9358 upregulated and 8116 downregulated), and 25,105 (14,562 upregulated and 10,543 downregulated) DEGs had at least a two-fold change in the skin of TXH, black radish, XLM, MSH, and ZJL, respectively, compared to SZB (Fig. 3B). There were 15,787 (8452 upregulated and 7335 downregulated), 16,350 (8579 upregulated and 7771 downregulated), and 19,817 (11,174 upregulated and 8643 downregulated)

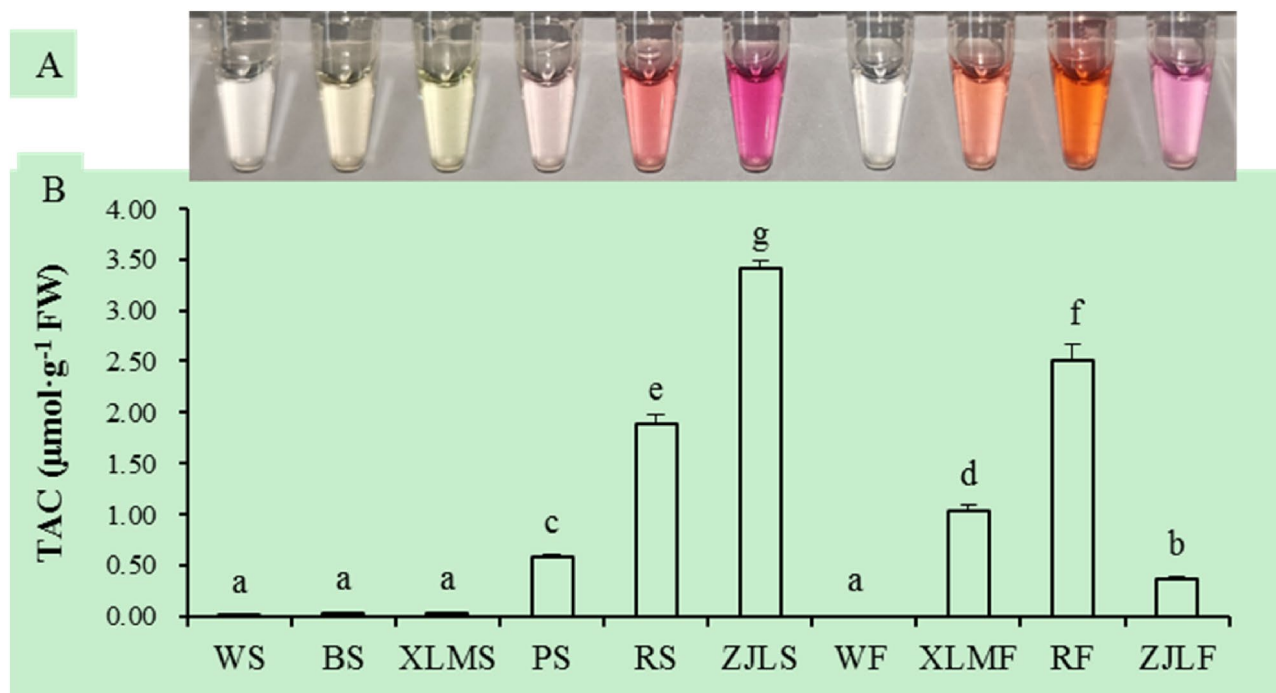


Figure 2. Pigment contents in the skin and flesh of radish taproots. (A) The color of the supernatant of the sample extracts. (B) Total anthocyanin content of the skin and flesh of radish taproots.

Sample	Raw reads	Clean reads	Clean bases (G)	Q20avg (%)	GC_pct (%)
WST	176,281,664	173,933,900	26.1	97.79	46.74
WFT	177,066,062	174,510,406	26.18		
RST	168,928,392	166,285,974	24.95		
RFT	168,355,314	165,648,598	24.85		
BST	170,937,186	168,309,646	25.25		
XL MST	163,260,394	160,683,094	24.1		
XL MFT	163,327,024	160,571,396	24.09		
PST	154,019,884	151,985,018	22.8		
ZJLST	167,115,082	165,168,530	24.77		
ZJLFT	186,389,736	184,077,716	27.62		
Total	1,695,680,738	1,671,174,278	250.71	-	-

Table 1. Details of the raw data and clean data of all radish samples of various colors. WST transcript data of SZB with white taproot skin, WFT transcript data of SZB with white flesh, RST transcript data of TXH with dark red taproot skin, RFT transcript data of TXH with dark red flesh, BST transcript data of black radish, XL MST transcript data of Xinlimei with green taproot skin, XL MFT transcript data of Xinlimei with red taproot flesh, PST transcript data of MSH with light-red taproot skin, ZJLST transcript data of ZJL taproot skin, ZJLFT transcript data of ZJL taproot flesh.

transcripts with at least a two-fold change in the flesh of TXH, XLM, and ZJL, respectively, compared to SZB (Fig. 3B). Additionally, the Venn diagram analysis showed that 6024 and 5333 DEGs were considerably differentially expressed in the skin and flesh in the pairwise comparisons, respectively (Fig. 3C,D).

A Kyoto Encyclopedia of Genes and Genomes (KEGG)³⁷ enrichment analysis was conducted to identify pathway-related DEGs in the colored skin and flesh of radish. A total of 111 pathways were enriched (Supplementary Table S3). The “biosynthesis of amino acids”, “glyoxylate and dicarboxylate metabolism”, “flavonoid biosynthesis”, “photosynthesis-antenna proteins”, “glycine, serine, and threonine metabolism”, and “carbon metabolism” pathways were significantly enriched in RS vs. WS (Fig. S1A). The “plant-pathogen interaction” pathway was significantly enriched in the skin of MSH (Fig. S1B). The “cyanoamino acid metabolism”, “glyoxylate and dicarboxylate metabolism”, “carbon fixation in photosynthetic organisms”, and “carbon metabolism” pathways were significantly enriched in the skin of XLM (Fig. S1D). DEGs related to the “ribosome” and “ribosome biogenesis in eukaryotes” pathways were significantly enriched in the flesh of XLM (Fig. S1G).

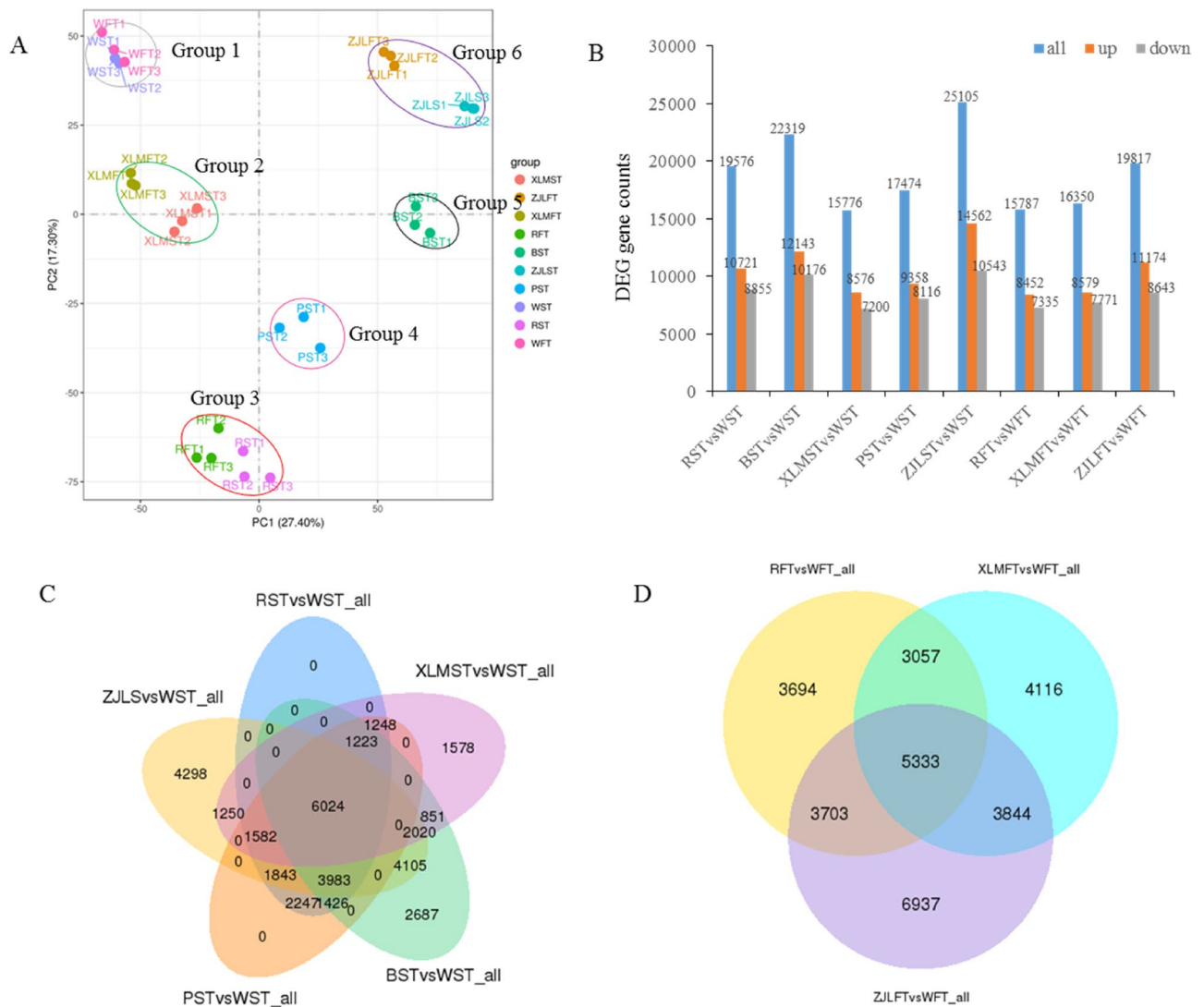


Figure 3. PCA and distribution of DEGs based on changes in the color of radish taproot skin and flesh. (A) PCA of the relative differences in the RNA-seq profiles among the different radish accessions. The sampling groups are color-coded as follows: pink-purple, SZB (WST and WFT); green-red XLM (XLMST and XLMFT); green-pink, TXH (RST and RFT); green, black radish (BST); blue-brown, ZJL (ZJLST and ZJLFT); blue, PST. (B) The numbers of up- and downregulated DEGs. (C) Venn diagram showing the number of DEGs revealed by pairwise comparisons in skin color. (D) Venn diagram showing the number of DEGs revealed by pairwise comparisons in flesh color. (A,C,D) were produced in R3.0.3 (<https://www.r-project.org/>).

Correlation analysis between transcripts and flavonoid derivatives revealed a differential network of flavonoid biosynthesis. To understand the network of flavonoids uncovered by the differential distribution of flavonoid derivatives between colored and white radishes, we conducted correlation tests between the quantitative changes in the metabolites and transcripts. A total of 11 structural genes were enriched in the red and purple radishes compared to the white and black radishes (Fig. 4A,B). Seven genes (*RsTT4*, 108843267; *RsC4H*, 108849675; *RsHCT*, 108812355; *RsTT7*, 108814778; *RsCCOAMT*, 108820088; *RsDFR*, 108826061; and *RsLDOX*, 108843686) were upregulated in the red and purple radishes, while two genes (*RsF3H* and *RsBAN*) did not change significantly. *RsFLS* (108839020) was downregulated in all of the samples. *RsTT5* (108821914) was upregulated in the skin of the five colored radishes, but was downregulated in the flesh of XLM, TXH, and ZJL.

The correlation analysis between transcripts and flavonoid derivatives showed that different interaction modes were responsible for the color diversity of the taproot skin and flesh. For the taproot skin, three genes involved in flavonoid biosynthesis, seven genes involved in phenylpropanoid biosynthesis, and 35 TFs were found to be strongly correlated (Pearson's correlation coefficient (PCC) > 0.95) with 48 metabolites (Fig. 5A, Table S4). A total of 11 flavonoids were found to be significantly correlated with *RsBGL* (108847438, encodes beta-glucosidase 8-like), 10 flavonoids were found to be significantly correlated with *RsDFR* (108826061, encodes dihydroflavonol-4-reductase), 10 metabolites were found to be significantly correlated with *RsCCOAMT* (108814129, encodes caffeoyl-CoA O-methyltransferase), six metabolites were found to be significantly correlated with *RsTT4*

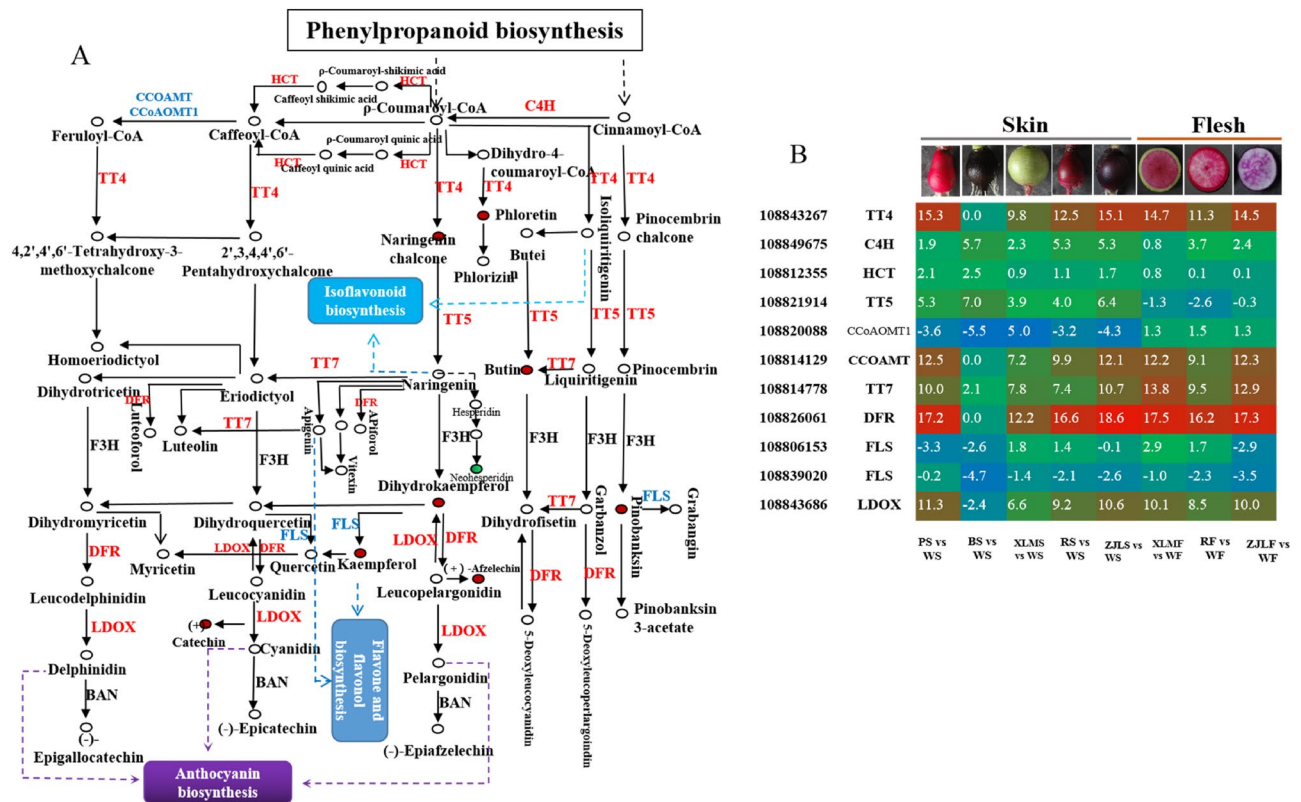


Figure 4. Biosynthetic pathway of flavonoids and the enriched or significantly expressed genes involved in the pathway. **(A)** The flavonoid biosynthesis pathway was constructed based on the KEGG pathway (ko00941)³⁷ and previous works. The genes shown in red were upregulated in samples, while genes shown in blue color were downregulated. Red circles indicate upregulated metabolites, and green circles indicate downregulated metabolites. **(B)** Expression patterns of the genes involved in flavonoid biosynthesis compared to WS. *TT4* chalcone synthase, *C4H* cinnamate-4-hydroxylase, *HCT* shikimate O-hydroxycinnamoyltransferase, *TT5* chalcone isomerase, *TT7* flavonoid 3'-monooxygenase, *CCoAOMT1* caffeoyl coenzyme A methyltransferase, *CCoAMT* caffeoyl-CoA O-methyltransferase, *DFR* dihydroflavonol-4-reductase, *FLS* flavonol synthase, *LDOX* anthocyanidin synthase, *F3H* 108852668, flavanone 3-hydroxylase, *BAN* anthocyanidin reductase.

(108843267, encodes chalcone synthase), and nine flavonoids were found to be significantly correlated with *RsF3H* (108852668, encodes flavanone 3-hydroxylase). Eleven flavonoids were found to be significantly correlated with *RsTT8-1* (108829780) and *RsTT8-2* (108838184) and nine metabolites were found to be significantly correlated with *RsPARI* (108837711), indicating that these five structural genes and three TFs play critical roles in flavonoid biosynthesis in radish skins.

For taproot flesh, four structural genes were found to be involved in flavonoid biosynthesis, 20 genes were found to be involved in phenylpropanoid biosynthesis, and 57 TFs were strongly correlated ($PCC > 0.95$) with 93 metabolites (Fig. 5B, Table S5). A total of 29 flavonoids were found to be significantly closely correlated with *RsBG8L*, 25 flavonoids were found to be significantly closely correlated with *RsDFR*, 20 metabolites were found to be significantly correlated with *RsLDOX* (108843686, encodes anthocyanidin synthase), 16 metabolites were found to be significantly correlated with *RsCCoAMT*. 29 metabolites were found to be significantly correlated with *RsbHHLH57* (10882212), 28 metabolites were found to be significantly correlated with *RsOGR3L* (1086444), and 26 metabolites were found to be significantly correlated with *RsbHHLH13L* (108810191), *RsMYB24* (108828405), *RsMYB34L* (108852392), and *RsPAR2L* (108860104). Twenty-four flavonoids were found to be significantly correlated with *RsTT8-2*, and 23 flavonoids were found to be significantly correlated with *RsTT8-1* and *RsbHHLH19* (108851776), suggesting that these four structural genes and nine TFs play important roles in flavonoid biosynthesis in radish flesh.

It has been reported that the overexpression of *MdMYB24L* results in higher anthocyanin content in transgenic apples than in the wild-type controls³⁸. *bHLH57* interacts with *rdo5* to control ABA biosynthesis and seed dormancy in *Arabidopsis*³⁹, and *AtbHLH19* plays roles in salt tolerance in *Arabidopsis*⁴⁰. *TT8* has been well characterized in *Arabidopsis*^{41,42}, *Brassica juncea*⁴³, and radishes⁴⁴. *AtTT8*, a positive regulator, controls its own expression in a feedback regulation involving *TG1* and homologous *MYB* and *bHLH* factors, allowing for the strong and cell-specific accumulation of flavonoids in *A. thaliana*^{41,42}. Natural mutations in two homoeologous *TT8* genes control the yellow seed coat trait in *B. juncea*⁴³. *RsTT8* interacts with *MYB1* to activate the promoter of *RsCHS* and *RsDFR* and so controls pigmentation accumulation in radishes⁴⁴.

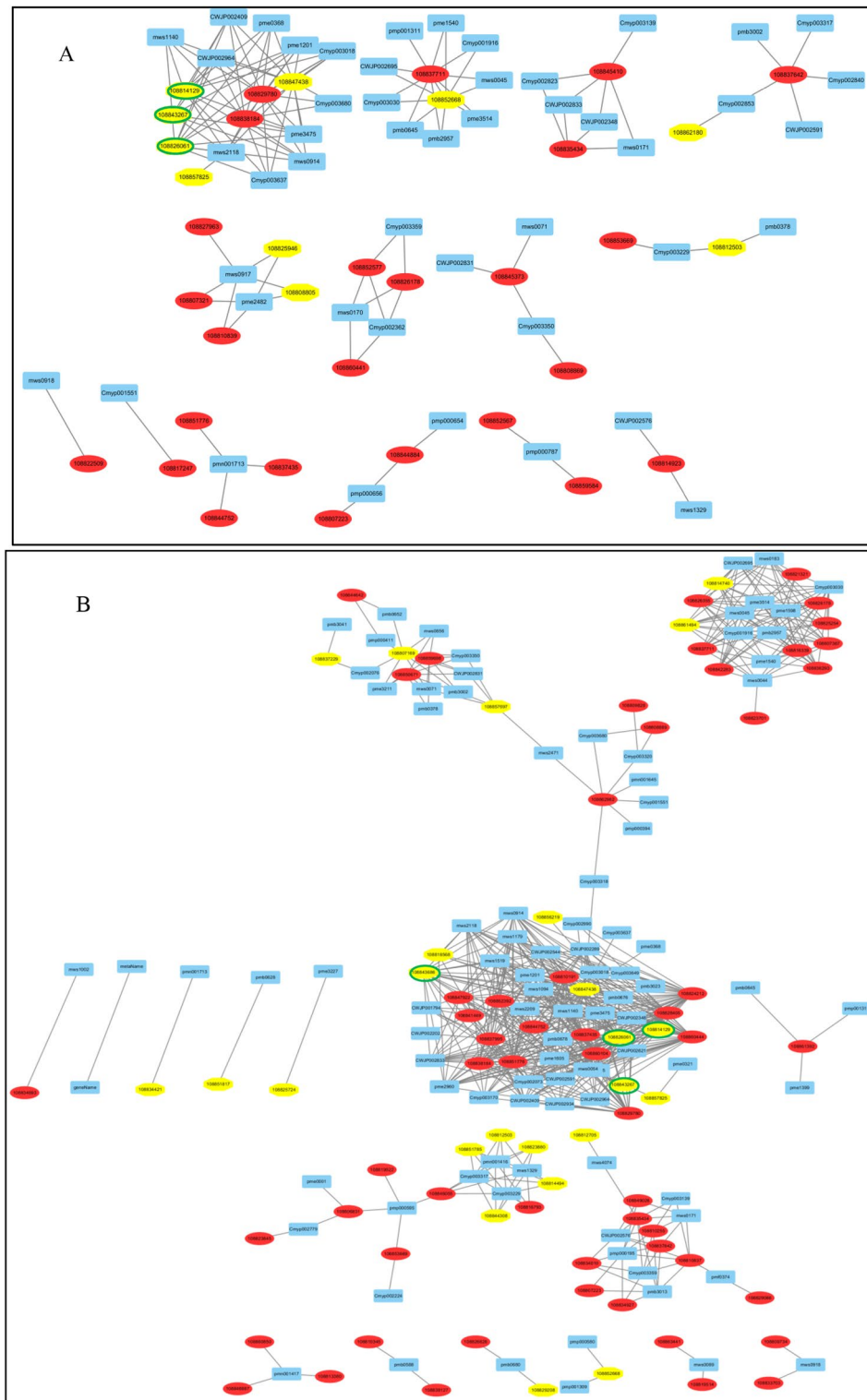


Figure 5. Connection network between the structural genes, TFs, and flavonoid-related metabolites. **(A)** Network between the genes and flavonoid metabolites in the taproot skin. **(B)** Network between the genes and flavonoid metabolites in the taproot flesh. Only genes with correlation coefficients > 0.95 are shown. Yellow circles with green frames represent the ID of structural genes involved in flavonoid biosynthesis; the other yellow circles represent the ID of structural genes involved in phenylpropanoid biosynthesis; the red circles indicate the TFs, and the blue boxes represent flavonoid-related metabolites. A and B were visualized using Cytoscape v2.8.2 (<https://cytoscape.org/>).

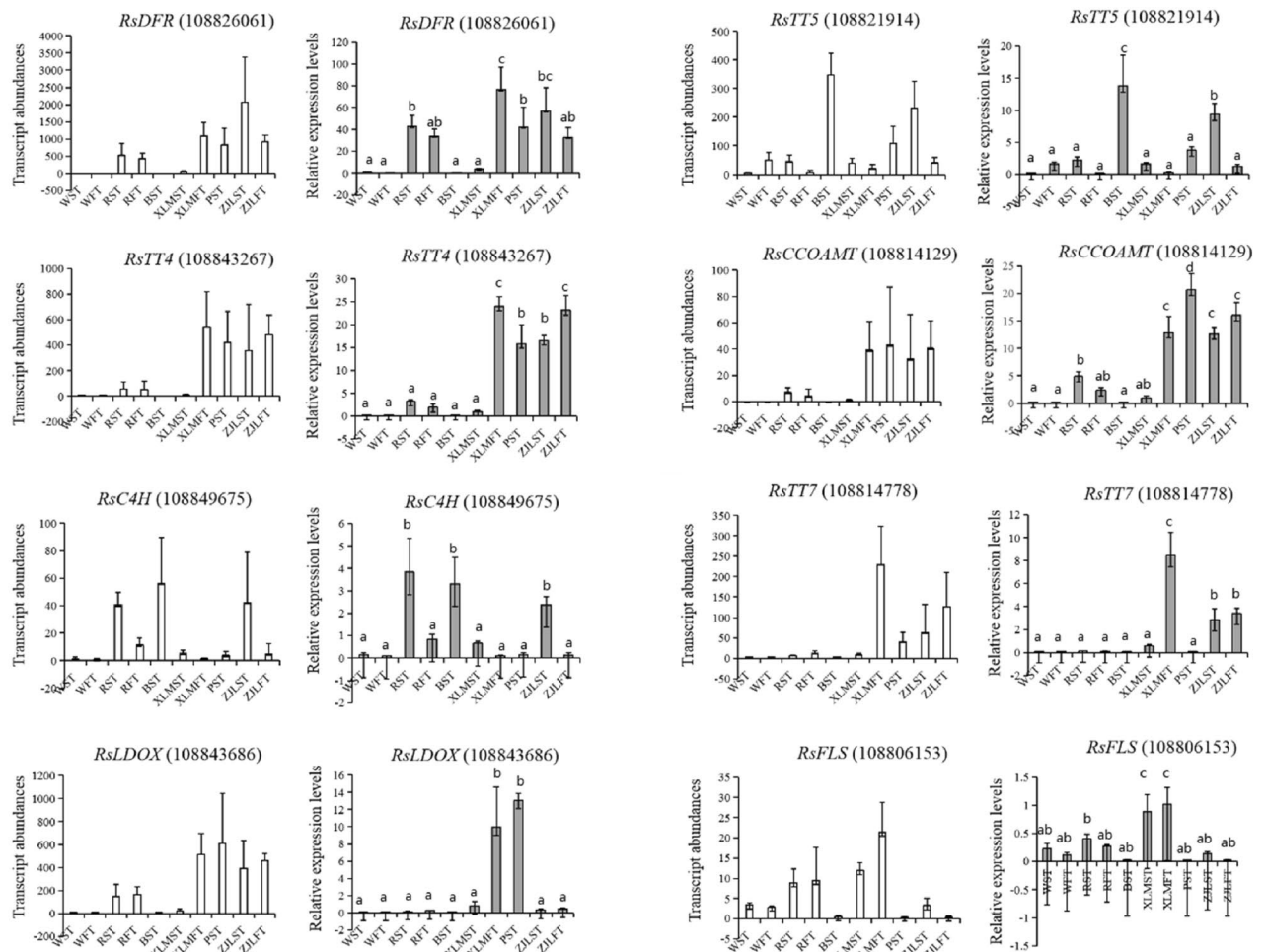


Figure 6. Validation of the transcriptomic data using RT-qPCR. Transcript abundances detected by transcriptome sequencing and expressed in FPKM are shown on the left. Relative expression levels detected by RT-qPCR and expressed in $2^{-\Delta\Delta CT}$ are shown on the right. Data are presented as the mean of three replicates with three biological repeats. The significant differences among samples were evaluated by Dunnett's test at $P < 0.05$.

Confirmation of the transcriptome data using real-time quantitative PCR (RT-qPCR). Eight genes involved in flavonoid metabolism were selected for RT-qPCR assays to validate the transcriptome datasets obtained from RNA-seq. The results of these assays were consistent with the transcriptome analysis (Fig. 6).

Discussion

Taproot skin and flesh color determine the appearance, quality, and nutritional value of radishes. To understand the molecular basis of the differential distribution of flavonoid derivatives between colored and white radishes, core genes in the flavonoid pathway were studied in detail. The results showed that most of the UniGenes demonstrated significant changes in expression level, regardless of whether they were in the early stage (RsTT4, chalcone synthase; RsC4H, cinnamate-4-hydroxylase; RsTT7, flavonoid 3'-monooxygenase; RsCCOAMT, caffeoyl-CoA O-methyltransferase) or late stage (RsDFR, dihydroflavonol-4-reductase; RsLDOX, anthocyanidin synthase), and showed higher transcript abundances in red and purple radishes than in the white radishes (Figs. 4, 5). Additionally, core genes involved in upstream phenylpropanoid biosynthesis also affect the productivity of the following flavonoid pathway. These genes, absent or undetectable, are thought to be flux-limiting genes that lead to colorant elimination in white radishes.

It is already known that TT4 catalyzes the first reaction in anthocyanin biosynthesis and helps form the intermediate chalcone, which is the primary precursor of all flavonoid classes⁴⁵. Thus, if TT4 reactions are strongly constrained, then anthocyanin production and that of nearly all other flavonoids could be effectively eliminated⁴⁶. DFR reduces dihydroflavonols to colorless leucoanthocyanidins (apigenin-related glycosides, leucocyanidin, and leucopelargonidin) that are subsequently catalyzed by LDOX to form cyanidin- and pelargonidin-related glycosides, which are then catalyzed by ANS into colored anthocyanidins⁴⁷. Kaempferol-related glycosides are downstream of dihydrokaempferol, catalyzed by FLS, and fluxed into flavone and flavonol biosynthesis. DFR genes have been studied thoroughly in many plant species, including *Medicago truncatula*⁴⁸, *Camellia sinensis*

⁴⁹, *Ginkgo biloba*⁵⁰, *Populus trichocarpa*⁵¹, and *Malus crabapples*^{52,53}. Overexpression of *DFR* genes from *Rosa rugosa* and *Petunia hybrida* in white tobacco flowers promotes anthocyanin biosynthesis, corresponding to increased red pigmentation accumulation⁵⁴. The altered expression of *DFR* can affect the expression of other anthocyanin biosynthesis genes⁵⁴. The expression of *DFR* can also be positively and negatively affected by MYB TFs. For example, the apple TF MdMYB10 has been shown to regulate the red pigmentation in the fruit flesh, and MdMYB1 controls the color of fruit skin. They interact with bHLH genes to trans-activate the *DFR* promoter and increase red pigmentation⁵³. TT8 acts a positive regulator of anthocyanin biosynthesis in *Arabidopsis* siliques and *Raphanus sativus*^{41,42,44}. The expression of the AtMYB7 gene can repress the expression of *DFR* and UGT⁵⁵.

In our study, the expression levels of *RsDFR*, *RsTT4*, *RsCCOAMT*, and *RsLODX* were undetectable in white radishes, but were abundant in the red-color radishes (Fig. 6). The correlation analysis verified that most flavonoids were significantly correlated with the transcript expression of *RsBG8L*, *RsCCOAMT*, and *RsDFR* both in the skin and in the flesh of the radish taproot. These three genes may lose their biocatalyst function and become pseudogenes due to variations in gene structure or regulation by correlated TFs, which were significantly associated with kaempferol-related glycosides. The radish basic helix-loop-helix TFs, *RsTT8-1* and *RsTT8-2*, were both significantly associated with flavonoids in the skin and flesh of taproot, indicating that these two TFs play crucial roles in the accumulation of pigments in the taproots. It has been reported that *RsTT8* interacts with *RsMYB1* to activate the promoters of *RsCHS* and *RsDFR*⁴⁴, which confirmed the results of our correlation analysis. We here propose that the limitations of flux in upstream reactions and the multishunt process in downstream reactions could lead to the elimination of colored pigmentation in white radishes.

Materials and methods

Plant materials. Six advanced radish inbred lines were used in this study: Touxinhong (TXH), with dark red taproot skin and flesh; Xinlimei, with red taproot flesh and green skin; Shizhuangbai (SZB), with white taproot skin and flesh; black radish with black taproot skin and white flesh; Zijinling (ZJL), with purple taproot skin and flesh; and Manshenhong (MSH), with light red taproot skin and white flesh. All plants were grown in pots at the Changsha experimental station of the Institute of Best Fiber Crops, Chinese Academy of Agricultural Sciences, in April 2019. Our research complied with local guidelines and legislation, and the required permissions and licenses for the study were obtained. Developing taproots were harvested at 50 days after sowing. The skin and flesh of the taproots were separated using a blade on a clean table, sampled, and stored at -80°C for further analysis. Each sample included three biological repeats and each repeat included five individual plants.

Total anthocyanin content measurement. Total anthocyanins were extracted using a plant anthocyanin content assay kit (Solarbio corporation, Beijing, China) with modifications. The samples of radish taproot skin and flesh were powdered, and 0.1 g of powder was placed into 1 mL of extraction buffer and incubated for 30 min at 60°C . This was then centrifuged for 10 min at 12,000 rpm. The absorbance values (A530 and A700) of the supernatant reaction solution were measured using a microplate reader. All samples were evaluated in three independent biological replicates.

RNA isolation and transcriptome sequencing. Total RNA of each sample was isolated using a Quick RNA isolation kit (Biotek Corporation, Beijing, China), characterized on a 1% agarose gel, and examined with an RNA Nano 6000 assay kit on a Bioanalyzer 2100 system (Agilent Technologies, Santa Clara, CA, USA). Construction of the sequencing libraries and RNA-seq were performed by the Biomarker Biotechnology Corporation (Beijing, China). mRNA was purified from total RNA using poly-T oligo-attached magnetic beads. Fragmentation was conducted using divalent cations under elevated temperatures in NEB next first-strand synthesis reaction buffer (5 \times). The first-strand cDNA was synthesized using the random hexamer primer and M-MuLV Reverse Transcriptase (RNase H-). The second-strand cDNA synthesis was subsequently performed using DNA Polymerase I and RNase H. The remaining overhangs were converted into blunt ends via exonuclease/polymerase activities. After adenylation of the 3' ends of the DNA fragments, NEB next adaptors with a hairpin loop structure were ligated to prepare for hybridization. To select cDNA fragments (preferentially 250–300 bp in length), library fragments were purified with an AMPure XP system (Beckman Coulter, Beverly, MA, USA). Then, PCR amplification was performed to enrich the purified cDNA template. Finally, 30 libraries were sequenced using an Illumina NovaSeq platform and 150-bp paired-end reads were generated^{56,57}.

Transcript profiles and annotation. Reference genome and gene model annotation files were downloaded directly from the genome website (https://www.ncbi.nlm.nih.gov/assembly/GCF_000801105.1/#/st). An index of the reference genome was built, and paired-end clean reads were aligned to the reference genome using Hisat2 v2.0.5⁵⁸. The mapped reads of each sample were assembled using StringTie v1.3.3b⁵⁹ in a reference-based approach. FeatureCounts v1.5.0-p3⁶⁰ was used to count the read numbers mapped to each gene. Then, the Fragments Per Kilobase of transcript, per Million mapped reads (FPKM) of each gene was calculated based on the length of the gene, and the read count was mapped to this gene. The differential expression analysis of the two groups was performed using the DESeq2 R package v1.16.1⁶¹. The resulting *P*-values were adjusted using the Benjamini and Hochberg's approach for controlling the false discovery rate. Genes with an adjusted *P*-value < 0.05 found by DESeq2 were determined to be differentially expressed. Gene Ontology (GO) and KEGG pathway functional enrichment analyses of the DEGs were implemented based on *Arabidopsis* gene sets using the clusterProfiler R package. PCA using the software R3.0.3 (<https://www.r-project.org/>) was performed to study accession-specific expression patterns.

Gene validation and expression analysis. To verify the reliability of the RNA-seq results, eight anthocyanin-related genes were subjected to RT-qPCR with specific primers identified using Primer Premier software (Premier, Canada) (Supplementary Table S1). cDNA synthesis and RT-qPCR were performed as previously described⁶². SYBR Green was used for the detection of PCR products. *Actin* was used as an internal control for the normalization of gene expression. Three independent biological replicates with three technical replicates each were analyzed by RT-qPCR to ensure reproducibility and reliability. Relative expression levels were estimated using the $2^{-\Delta\Delta CT}$ method⁶³.

Integrative analysis of the metabolome and transcriptome. Pearson's correlation coefficients were calculated for previous metabolome data³⁶ and transcriptome data integration by Metware Biotechnology (Wuhan, China). The mean of all biological replicates of each cultivar in the metabolome data and the mean value of expression of each transcript in the transcriptome data were calculated. The fold-change of each pigmented radish accession (TXH, XLM, and ZJL) was calculated for both the metabolome and transcriptome data, and then compared with the control accession (SZB with white skin and white flesh). Finally, the coefficients were calculated from the log2fold change of each metabolite and transcript using Excel (Microsoft Corporation, Redmond, WA, USA). Correlations corresponding to a coefficient with PCC > 0.95 were selected (Supplementary Table S4). Metabolome and transcriptome relationships were visualized using Cytoscape v2.8.2 (<https://cytoscape.org/>)⁶⁴.

Received: 9 November 2020; Accepted: 12 May 2021

Published online: 01 June 2021

References

1. Yao, L. H. *et al.* Flavonoids in food and their health benefits. *Plant Foods Hum. Nutr.* **59**, 113–122 (2004).
2. Owens, D. *et al.* Functional analysis of a predicted flavonol synthase gene family in Arabidopsis. *Plant Physiol.* **147**, 1046–1061 (2008).
3. Mazza, G. *Anthocyanins in Fruits, Vegetables, and Grains* (CRC Press, 2018).
4. Gould, K. S. Nature's swiss army knife: The diverse protective roles of anthocyanins in leaves. *J. Biomed. Biotechnol.* **2004**, 314–320 (2004).
5. He, J. & Giusti, M. M. Anthocyanins: Natural colorants with health-promoting properties. *Annu. Rev. Food Sci. Technol.* **1**, 163–187 (2010).
6. Wang, H. *et al.* Functional characterization of dihydroflavonol-4-reductase in anthocyanin biosynthesis of purple sweet potato underlies the direct evidence of anthocyanins function against abiotic stresses. *PLoS ONE* **8**, e78484 (2013).
7. Lee, S. L. *et al.* Purple sweet potato leaf extract induces apoptosis and reduces inflammatory adipokine expression in 3T3-L1 differentiated adipocytes. *Evid. Based Complement. Altern. Med.* **2015**, 1–9 (2015).
8. Harborne, J. B. & Williams, C. A. Advances in flavonoid research since 1992. *Phytochemistry* **55**, 481–504 (2000).
9. Tanaka, Y., Sasaki, N. & Ohmiya, A. Biosynthesis of plant pigments: Anthocyanins, betalains and carotenoids. *Plant J.* **54**, 733–749 (2008).
10. Hichri, I. *et al.* Recent advances in the transcriptional regulation of the flavonoid biosynthetic pathway. *J. Exp. Bot.* **62**, 2465–2483 (2011).
11. Allan, A. C., Hellens, R. P. & Laing, W. A. MYB transcription factors that colour our fruit. *Trends Plant Sci.* **13**(3), 99–102 (2008).
12. Koes, R., Verweij, W. & Quattrocchio, F. Flavonoids: A colorful model for the regulation and evolution of biochemical pathways. *Trends Plant Sci.* **10**, 236–242 (2005).
13. Matus, J. T., 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.* **8**, 83 (2008).
14. Hichri, I. *et al.* The basic helix-loop-helix transcription factor MYC1 is involved in the regulation of the flavonoid biosynthesis pathway in grapevine. *Mol. Plant.* **3**, 509–523 (2010).
15. Zheng, T. *et al.* Regulation of anthocyanin accumulation via MYB75/HAT1/TPL-mediated transcriptional repression. *PLoS Genet.* **15**, e1007993 (2019).
16. Jian, W. *et al.* SLMYB75, an MYB-type transcription factor, promotes anthocyanin accumulation and enhances volatile aroma production in tomato fruits. *Hortic. Res.* **6**, 1–15 (2019).
17. Colanero, S., Tagliani, A., Perata, P. & Gonzali, S. Alternative splicing in the anthocyanin fruit gene encoding an R2R3 MYB transcription factor affects anthocyanin biosynthesis in tomato fruits. *Plant Commun.* **1**, 100006 (2020).
18. Xu, Z. S., Yang, Q. Q., Feng, K., Yu, X. & Xiong, A. S. DcMYB113, a root-specific R2R3-MYB, conditions anthocyanin biosynthesis and modification in carrot. *Plant Biotechnol. J.* **18**, 1585 (2020).
19. Wang, Q., Wang, Y., Sun, H., Sun, L. & Zhang, L. Transposon-induced methylation of the RsMYB1 promoter disturbs anthocyanin accumulation in red-fleshed radish. *J. Exp. Bot.* **71**, 2537–2550 (2020).
20. Jia, H. *et al.* Abscisic acid plays an important role in the regulation of strawberry fruit ripening. *Plant Physiol.* **157**, 188–199 (2011).
21. Jeong, S. W. *et al.* Ethylene suppression of sugar-induced anthocyanin pigmentation in Arabidopsis. *Plant Physiol.* **154**, 1514–1531 (2010).
22. Das, P. K. *et al.* Cytokinins enhance sugar-induced anthocyanin biosynthesis in Arabidopsis. *Mol. Cells* **34**, 93–101 (2012).
23. Takos, A. M. *et al.* Light-induced expression of a MYB gene regulates anthocyanin biosynthesis in red apples. *Plant Physiol.* **142**, 1216–1232 (2006).
24. Huang, D. *et al.* Retrotransposon promoter of Ruby1 controls both light- and cold-induced accumulation of anthocyanins in blood orange. *Plant Cell Environ.* **42**, 3092–3104 (2019).
25. An, J. P. *et al.* Md BBX 22 regulates UV-B-induced anthocyanin biosynthesis through regulating the function of Md HY 5 and is targeted by Md BT 2 for 26S proteasome-mediated degradation. *Plant Biotechnol. J.* **17**, 2231–2233 (2019).
26. Bai, S. *et al.* BBX16, a B-box protein, positively regulates light-induced anthocyanin accumulation by activating MYB10 in red pear. *Plant Biotechnol. J.* **17**, 1985–1997 (2019).
27. Lou, Q. *et al.* Transcriptome sequencing and metabolite analysis reveals the role of delphinidin metabolism in flower colour in grape hyacinth. *J. Exp. Bot.* **65**, 3157–3164 (2014).
28. Hirai, M. *et al.* Omics-based identification of Arabidopsis Myb transcription factors regulating aliphatic glucosinolate biosynthesis. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 6478–6483 (2007).

29. Hirai, M. *et al.* Integration of transcriptomics and metabolomics for understanding of global responses to nutritional stresses in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 10205–10210 (2004).
30. Alba, R. *et al.* Transcriptome and selected metabolite analyses reveal multiple points of ethylene control during tomato fruit development. *Plant Cell* **17**, 2954–2965 (2005).
31. Mounet, F. *et al.* Gene and metabolite regulatory network analysis of early developing fruit tissues highlights new candidate genes for the control of tomato fruit composition and development. *Plant Physiol.* **149**, 1505–1528 (2009).
32. Zhu, G. *et al.* Rewiring of the fruit metabolome in tomato breeding. *Cell* **172**, 249–261 (2018).
33. Stushnoff, C. *et al.* Flavonoid profiling and transcriptome analysis reveals new gene-metabolite correlations in tubers of *Solanum tuberosum* L.. *J. Exp. Bot.* **61**, 1225–1238 (2010).
34. Cho, K. *et al.* Network analysis of the metabolome and transcriptome reveals novel regulation of potato pigmentation. *J. Exp. Bot.* **67**, 1519–1533 (2016).
35. Li, Y. *et al.* Combined analysis of the fruit metabolome and transcriptome reveals candidate genes involved in flavonoid biosynthesis in *Actinidia arguta*. *Int. J. Mol. Sci.* **19**, 1471 (2018).
36. Zhang, J., Qiu, X., Tan, Q., Xiao, Q. & Mei, S. A comparative metabolomics study of flavonoids in radish with different skin and flesh colors (*Raphanus sativus* L.). *J. Agric. Food Chem.* **68**, 14463–14470 (2020).
37. Kanehisa, M. & Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).
38. Wang, Y. *et al.* The R2R3-MYB transcription factor MdMYB24-like is involved in methyl jasmonate-induced anthocyanin biosynthesis in apple. *Plant Physiol. Biochem.* **139**, 273 (2019).
39. Fei, L. *et al.* Reversal of rdo5 1, a homolog of rice seed dormancy4, interacts with bhlh57 and controls aba biosynthesis and seed dormancy in *Arabidopsis*. *Plant Cell* **32**, 1933–1948 (2020).
40. Ogawa, T. *et al.* Participation of transcription factor *AtbHLH19* in salt tolerance in *Arabidopsis*. *Plant Cell Physiol.* **48**, 140–140 (2007).
41. Nesi, N. The *TT8* gene encodes a basic helix-loop-helix domain protein required for expression of DFR and BAN genes in *Arabidopsis* Siliques. *Plant Cell* **12**, 1863 (2000).
42. Baudry, A., Caboche, M. & Lepiniec, L. *TT8* controls its own expression in a feedback regulation involving TTG1 and homologous MYB and bHLH factors, allowing a strong and cell-specific accumulation of flavonoids in *Arabidopsis thaliana*. *Plant J.* **46**, 768–779 (2006).
43. Padmaja, L. K. *et al.* Natural mutations in two homoeologous *TT8* genes control yellow seed coat trait in allotetraploid *Brassica juncea* (AABB). *Theor. Appl. Genet.* **127**, 339–347 (2014).
44. Lim, S.-H. *et al.* A radish basic helix-loop-helix transcription factor, RsTT8 acts a positive regulator for anthocyanin biosynthesis. *Front. Plant Sci.* <https://doi.org/10.3389/fpls.2017.01917> (2017).
45. Koes, R. E., Spelt, C. E., van den Elzen, P. J. M. & Mol, J. N. M. Cloning and molecular characterization of the chalcone synthase multigene family of *Petunia hybrida*. *Gene* **81**, 245–257 (1989).
46. Clark, S. T. & Verwoerd, W. S. A systems approach to identifying correlated gene targets for the loss of colour pigmentation in plants. *BMC Bioinform.* **12**, 343 (2011).
47. Holton, T. A. & Cornish, E. C. Genetics and biochemistry of anthocyanin biosynthesis. *Plant Cell* **7**, 1071–1083 (1995).
48. Xie, D. Y., Jackson, L. A., Cooper, J. D., Ferreira, D. & Paiva, N. L. Molecular and biochemical analysis of two cDNA clones encoding dihydroflavonol-4-reductase from *Medicago truncatula*. *Plant Physiol.* **134**, 979–994 (2004).
49. Singh, K., Kumar, S., Yadav, S. K. & Ahuja, P. S. Characterization of dihydroflavonol 4-reductase cDNA in tea (*Camellia sinensis* (L.) O. Kuntze). *Plant Biotechnol. Rep.* **3**, 95–101 (2009).
50. Hua, C. *et al.* Molecular cloning and characterization of three genes encoding dihydroflavonol-4-reductase from *Ginkgo biloba* in anthocyanin biosynthetic pathway. *PLoS ONE* **8**, e72017 (2013).
51. Huang, Y. *et al.* Molecular cloning and characterization of two genes encoding dihydroflavonol-4-reductase from *Populus trichocarpa*. *PLoS ONE* **7**, e30364 (2012).
52. Tian, J. *et al.* The balance of expression of dihydroflavonol 4-reductase and flavonol synthase regulates flavonoid biosynthesis and red foliage coloration in crabapples. *Sci. Rep.* **5**, 12228 (2015).
53. Tian, J. *et al.* Characteristics of dihydroflavonol 4-reductase gene promoters from different leaf colored *Malus* crabapple cultivars. *Hortic. Res.* **4**, 17070 (2017).
54. Luo, P. *et al.* Disequilibrium of flavonol synthase and dihydroflavonol-4-reductase expression associated tightly to white vs red color flower formation in plants. *Front. Plant Sci.* **6**, 1257 (2015).
55. Fornalé, S. *et al.* *AtMYB7*, a new player in the regulation of UV-sunscreens in *Arabidopsis thaliana*. *Plant Cell Physiol.* **55**, 507–516 (2014).
56. Wang, Z., Gerstein, M. & Snyder, M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* **10**, 57–63 (2009).
57. Parkhomchuk, D. *et al.* Transcriptome analysis by strand-specific sequencing of complementary DNA. *Nucleic Acids Res.* **37**, e123 (2009).
58. Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* **5**, 621–628 (2008).
59. Pertea, M. *et al.* Stringtie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* **33**, 290–295 (2015).
60. Yang, L., Smyth, G. K. & Wei, S. FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
61. Love, M., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 1–21 (2014).
62. Zhang, J. *et al.* A naturally occurring variation in the BrMAM-3 gene is associated with aliphatic glucosinolate accumulation in *Brassica rapa* leaves. *Hortic. Res.* **5**, 1–11 (2018).
63. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using realtime quantitative PCR and the 2^{-ΔΔCT} method. *Methods* **25**, 402–408 (2001).
64. Shannon, P. *et al.* Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).

Acknowledgements

The authors thank Shuaifeng Geng and Metware Biotechnology for their technical assistance in this study. The image for the flavonoid biosynthetic pathway is based on ko00941 by KEGG (<https://www.kegg.jp/kegg/kegg1.html>). Thanks Kanehisa Laboratories for their copyright permission.

Author contributions

S.M. and J.Z. contributed to the conception and design of the study; all authors contributed to the acquisition, analysis data; J.Z., J.Z. and S.M. contributed greatly to the drafting and revision of the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding

This research was funded by the National Key Research and Development Program of China (Grant No.: 2017YFD0101806), Natural Science Foundation of Hunan Province (Grant No.: 2020JJ5642), Agricultural Talents Program of the Chinese Academy of Agricultural Sciences (Grant No.: CAASQNYC-KYYJ-38), and Central Public-interest Scientific Institution Basal Research Fund (No. 1610242021008).

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-90633-5>.

Correspondence and requests for materials should be addressed to J.Z. or S.M.

Reprints and permissions information is available at www.nature.com/reprints.

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



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, 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 changes were made. 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/4.0/>.

© The Author(s) 2021