Integrated analysis of the metabolome and transcriptome in different developmental stages of *Cerasus humilis* peel coloration

Xiaolong Ji
Northeast Forestry University
Jing Ren
Northeast Agricultural University
Lijiao Sun
Northeast Forestry University
Shaoyu Lang
Northeast Forestry University
Liwei Zhu
Northeast Forestry University
Xing Shun Song (sfandi@163.com)
Northeast Forestry University https://orcid.org/0000-0001-8129-2612

Research article

Keywords: Cerasus humilis, anthocyanin, pigmentation, Transcriptome, Metabolites

Posted Date: August 12th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-41824/v1

Abstract

Background: *Cerasus humilis* is a unique dwarf shrub and fruit color is an important trait in the species. In this study, we evaluated the transcriptomic and metabolomic profiles of the plant at different developmental stages to elucidate the mechanism underlying color formation.

Results: In a metabolomics analysis, 16 anthocyanin components were identified at different developmental stages, and high levels of cyanidin *O*-syringic acid and pelargonidin 3-*O*-beta-d-glucoside (callistephin chloride) were correlated with the reddening of the fruit peel. Additionally, transcriptome analysis showed that most anthocyanin biosynthetic genes and two MYB transcription factors were significantly up-regulated. qRT-PCR results for these differentially expressed genes were generally consistent with the high-throughput sequencing. A co-expression analysis revealed that *ANS* and *UFGT* play keys role in pigmentation.

Conclusions: These findings provide insight into the mechanisms underlying anthocyanin accumulation and coloration during fruit peel development, providing a basis for the breeding of anthocyanin-rich *C. humilis* cultivars.

Background

Cerasus humilis (Bge.) Sok., which belongs to the genus *Cerasus* in the family Rosaceae, is a small dwarf shrub that originated from China. It is highly stress-resistant, especially to drought and low temperatures [1]. The fruits of *C. humilis* have a bright color and show a clustered arrangement, they are known as “calcium fruits” owing to their relatively high calcium contents [2]. The fruits are rich in flavonoids, including anthocyanins, proanthocyanidins, organic acids, and polyphenols [3–5]. Additionally, *C. humilis* seeds (semen pruni) have been used as a traditional Chinese medicine for heat-clearing and water-disinhibiting [6]. Based on its ecological and economic benefits, *C. humilis* has become an emerging multipurpose fruit tree in the last 10 years, with a total estimated planting area of approximately 20,000 ha [4, 5].

Anthocyanins belong to the flavonoid class of phenolic compounds [7]. They are a group of water-soluble pigments in plants that contribute to fruit development, tolerance to environmental stresses, and human health [4, 8, 9]. Genes involved in anthocyanin biosynthesis have been well-characterized in many plants, including *Arabidopsis thaliana* [10], apple (*Malus pumila*) [11], tomato (*Solanum lycopersicum*) [12], and pear (*Pyrus pyrifolia*) [13]. Briefly, anthocyanins are synthesized from phenylalanine via the phenylpropanoid and flavonoid pathways in the cytoplasm and transported to vacuoles [9]. Many functional genes, including genes encoding phenylalanine ammonia-lyase (*PAL*) [14], cinnamic acid 4-hydroxylase (*C4H*) [15], 4-coumarate-CoA ligase (*4CL*) [16], chalcone synthase (*CHS*) [17], chalcone isomerase (*CHI*) [18], flavanone 3-hydroxylase (*F3H*) [19], dihydroflavonol 4-reductase (*DFR*) [20], anthocyanidin synthase (*ANS*; also called leucoanthocyanidin dioxygenase [*LDOX*]) [21], and UDP-glucose: flavonoid-3-*O*-glycosyltranferase (*UFGT*) [22], play important roles in the production of

anthocyanins. Furthermore, MYB transcription factors (TFs) are key factors in regulatory networks controlling growth and development, metabolism, and responses to biotic and abiotic stresses [23]. For example, anthocyanin accumulation is regulated by the MYB transcription factor [7, 24–26]. Anthocyanin biosynthesis, controlled by a diverse array of exogenous and endogenous factors, is the most well-studied secondary metabolic pathway in plants. Gene expression controlling the accumulation of metabolites under different stages of development is a direct determinant of fruit color changes. However, the molecular and metabolic pathways involved in *C. humilis* fruit coloration are still elusive. Therefore, the genetic basis of anthocyanin biosynthesis and related metabolic pathways requires further exploration. Knowledge of the genetic basis of characters related to fruit coloration will facilitate the manipulation of this trait to obtain more attractive and healthier fruits for consumers [27].

Transcriptomics and metabolomics can provide insight into transcriptional regulation and metabolic flow. Bai et al. [28] performed a spatiotemporal transcriptomic analysis of the red Chinese sand pear peels, revealing that a light-responsive pathway functions in anthocyanin accumulation with different temporal expression patterns, they suggested that *PpMYB10* as well as other light-responsive transcription factors are involved in the regulation of anthocyanin biosynthesis. Using a combination of transcriptomic and metabolite profiling, Zhuang et al. [29] found that purple turnip essentially diverts dihydroflavonols to the biosynthesis of anthocyanins over flavonols by strongly down-regulating one flavonol synthase gene, while strikingly up-regulating dihydroflavonol 4-reductase, anthocyanidin synthase, and UDP-glucose: flavonoid-3-*O*-glucosyltransferase compared to levels in green turnip. In the same way, Meng et al. [30] demonstrated the molecular mechanism underlying the WP1-mediated regulation of floral carotenoid pigmentation by network analyses of metabolomics and transcriptomics data. Thus, multi-omics joint analysis is a powerful means to elucidate the complex process of plant development [7, 31].

In the present study, based on UPLC-MS/MS, metabolite changes in the process of peel color change in *C. humilis* were qualitatively and quantitatively analyzed, and the unique pattern of anthocyanin-related metabolites during peel color change was established for the first time. Additionally, a correlation analysis of the transcriptomic profiles and anthocyanin contents was performed to elucidate the mechanism underlying variation in fruit peel pigmentation in *C. humilis*, followed by further analyses of the regulatory mechanism based on gene expression assays. Our transcriptome analysis and metabolic profiling of four stages of fruit development, with a focus on anthocyanin biosynthetic pathways, will facilitate breeding and the functional characterization of genes of interest in *C. humilis* as well as several other species in *Cerasus* Mill.

Results

Principal component analysis (PCA) reveals differences in anthocyanin metabolite profiles

Traditionally and practically, the fruiting of *C. humilis* is classified into four developmental stages. Metabolite content data processing was performed to improve normalization using R (www.r-project.org/)

to compare the accumulation of metabolites in different samples by HCA. Based on relative differences in accumulation patterns among different stages, the HCA of anthocyanins resulted in four main clusters in *C. humilis* (Fig. 1a). Anthocyanins belonging to clusters 1 and 2 accumulated at the highest levels in S4, while anthocyanins within clusters 3 and 4 did not differ significantly among stages.

Metabolite profiles of the fruit peel were then subjected to a principal component analysis (PCA) [32]. The score plots for the PCA (Fig. 1b) showed obvious separation among fruit peel development stages. The first two principal components (PC1 and PC2) for the flowers explained 62.42% and 15.02% of the total variance, respectively. The fruit peel contained 16 anthocyanins based on subsequent HPLC-MS and MRM (multiple reaction monitoring) analyses (Additional file 1: Table S1), with 11 differential metabolites (fold change ≥ 2 or fold change ≤ 0.5), including peonidin *O*-hexoside, cyanidin 3-*O*-malonylhexoside, cyanidin *O*-syringic acid, pelargonidin *O*-acetylhexoside, cyanidin 3-*O*-glucoside (kuromanin), procyanidin A1, cyanidin 3,5-*O*-diglucoside (cyanin), pelargonidin 3-*O*-beta-d-glucoside (callistephin chloride), cyanidin 3-*O*-galactoside, and peonidin 3-*O*-glucoside chloride. Apart from procyanidins, the predominant anthocyanins were cyanidin and its glycoside-derived compounds. Among the anthocyanin compounds discovered in fruit peel, two were found at particularly high concentrations in S4, suggesting that these two compounds [cyanidin *O*-syringic acid and pelargonidin 3-*O*-beta-d-glucoside (callistephin chloride)] contribute substantially the mechanism underlying fruit peel coloration.

Overview of transcriptome sequencing

To assess the gene expression profile of the fruit peel at S1–S4, RNAs from 12 *C. humilis* fruit samples were used for RNA-seq with three replicates per stage. Before using Illumina platform to read the cDNA libraries, the low-quality reads and adapter sequences were deleted. The clean reads for each sample ranged from 40,704,832 to 47,000,394, and a total of 541,781,582 clean reads were obtained. The Q30 values > 92%, meeting the criterion for gene discovery. Based on filtered clean data, the full-length transcript sequence was assembled using Trinity, and the total number of assembled transcripts was 176,109, with a mean length of 1340.05 nt and N50 length of 2,227 nt. These were assembled into 73,035 unigenes with a mean length of 1,018 nt and N50 length of 1,917 nt (Additional file 2: Table S2).

Annotation and identification of unigenes

All assembled unigenes and ORFs were annotated using Trinotate against the public databases NT, NR, UniProt, RNAMMER, eggNOG, GO, and KEGG. Using the Trinotate annotation results, genes were subjected to a GO analysis. According to the GO database, 73,035 unigenes are divided into 60 functional groups,, in the three main functional categories (biological process, cellular component, and molecular function) (Additional file 6: Figure S1). In the biological process category, 42,236 (57.83%) unigenes were associated with metabolic processes.

The assembled unigenes were searched through the COG (Cluster of Orthologous Group, KOG) protein database to evaluate potential functions. By a KOG analysis, 25,127 unigenes were clustered into 25 functional categories (Additional file 7: Figure S2). Among them, general function prediction only (Class R, 5,338 unigenes) was most common, accounting for 21.24% of unigenes. Class R is a category for which only a general prediction regarding protein function is feasible (e.g., biochemical activity). Most of the COGs could be assigned to a well-defined functional category. However, the fact that the largest category remains unspecific in nature indicates insufficient information within the COG database. The next largest categories were signal transduction mechanisms (Class T, 3,071 unigenes, 12.22%), posttranslational modification, protein turnover, chaperones (Class O, 2,593 unigenes, 10.32%), and translation, ribosomal structure and biogenesis (Class J, 1,111 unigenes, 7.02%). The smallest group was cell motility (Class N, 39 unigenes, 0.02%). The COG and GO annotations showed that unigenes expressed in the *C. humilis* fruit peel encode various proteins related to metabolism.

To gain more insight into the transcriptomic differences, we compared DEGs in *C. humilis* fruit peel among different developmental stages. Using the DESeq2 package, normalized RPKM (Reads Per Kilobase Million Mapped Reads) values were obtained from RNA-seq data to quantify transcript expression. We identified 7,930 DEGs in comparisons among different fruit developmental stages. Among the three comparisons (S2_S1, S3_S2, and S4_S3), the largest number of DEGs was found between S3 and S2, with 4,333 DEGs, including 1,540 up-regulated and 2,793 down-regulated DEGs. In addition, there were 430 up-regulated DEGs in S2_S1, and 1,179 up-regulated DEGs in S4_S3 (Fig. 2). These results indicated that DEGs were most active during the second developmental stage of the fruit peel and were perhaps largely responsible for anthocyanin biosynthesis. Therefore, we speculate that the transition from S2 to S3 is a key stage in fruit peel development.

GO enrichment and KEGG pathway analyses of DEGs

We functionally categorized the DEGs based on GO terms. A total of 7,927 unigenes were annotated, including 429 up-regulated and 201 down-regulated unigenes in the S2_S1-enriched GO term libraries and 1,539 up-regulated and 2,794 down-regulated DEGs in the S3_S2-enriched GO term libraries. The subcategories with the highest degree of enrichment were all cell part, followed by binding and cellular process, as shown in a histogram in Additional file 8: Figure S3. Apart from this, 1,178 up-regulated and 1,786 down-regulated unigenes in the S4_S3 comparison were enriched for GO terms, particularly cell part, followed by cellular process, and binding.

To classify the DEGs based on related pathways, a KEGG enrichment analysis was performed. In total, 125 biosynthetic and metabolic pathways were enriched in the pairwise comparisons of fruit peel developmental stages. The flavonoid biosynthetic pathway began to be enriched (ko00941, 10 genes) in S3_S2, and the anthocyanin biosynthetic pathway began to be enriched (ko00942, 2 genes) in S4_S3. We obtained 12 DEGs across the three comparisons related to pigment synthesis, suggesting that these genes are associated with fruit peel coloration (Additional file 9: Figure S4).

The changes of TFs in the process of peel coloring

Next, we focused on differentially expressed TFs based on our transcriptome data. The MYB and bHLH TF families were the most abundant TF families during the fruit coloring process, followed by the B3, AP2/ERF, and NAC families (Table 1), many of these were expressed before samples began to show color. Additionally, the TFs were mostly up-regulated before the color change and down-regulated after the color change. Notably, the MYB and bHLH TF families are important transcription factors for anthocyanin transcriptional regulation, and 68% of bHLHs and 70% of MYB genes were down-regulated. Interestingly, we found that 2 MYB genes (TRINITY_DN26515_c0_g1 and TRINITY_DN21536_c0_g1) were significantly up-regulated after the fruit peel color changes, and their expression levels were consistent with the pattern of fruit peel pigment accumulation, indicating that they might be key regulators of fruit peel color.

Table 1
Differentially expressed transcription factors (TFs) in fruit peel

DEG set	TF family	Number of DEGs	Up-regulated	Downregulated	Description
S2_S1	ARF/ERF	20	12	8	Ethylene-responsive transcription factor
	B3	12	7	5	B3 DNA-binding domain
	bHLH	22	11	11	Helix-loop-helix DNA-binding domain
	bzip	8	5	3	Basic region leucine zipper
	C2H2	8	5	3	Zinc fifinger C2H2 domain-containing protein
	C3H	9	7	2	Zinc fifinger CCCH domain-containing protein
	DBB	8	7	1	double B-box zinc finger protein
	G2-like	11	5	6	Golden2-Like protein
	GRAS	8	7	1	C-terminal GRAS domain
	HSF	22	18	4	Heat stress transcription factor
	MYB	28	20	8	MYB-related protein
	NAC	15	14	1	NAC domain-containing protein
	S1Fa-like	12	6	6	Leucine-rich repeat protein
	other TFs	104	56	48	
	total	287	180	107	
S3_S2	ARF/ERF	98	24	74	Ethylene-responsive transcription factor
	B3	121	27	94	B3 DNA-binding domain
	bHLH	152	53	99	Helix-loop-helix DNA-binding domain
	bzip	62	25	37	Basic region leucine zipper
	C2H2	76	27	49	Zinc fifinger C2H2 domain-containing protein

Notes: $q < 0.05$ and $|\log2(\text{ratio})| \geq 1$ were set as the threshold values for significantly different expression.

DEG set	TF family	Number of DEGs	Up-regulated	Downregulated	Description
	C3H	52	17	35	Zinc fifinger CCCH domain-containing protein
	DBB	17	3	14	double B-box zinc finger protein
	G2-like	67	22	45	Golden2-Like protein
	GRAS	47	14	33	C-terminal GRAS domain
	HSF	64	14	50	Heat stress transcription factor
	MYB	159	48	111	MYB-related protein
	NAC	92	25	67	NAC domain-containing protein
	S1Fa-like	32	7	25	Leucine-rich repeat protein
	other TFs	717	252	465	
	total	1756	558	1198	
S4_S3	ARF/ERF	78	25	53	Ethylene-responsive transcription factor
	B3	91	32	59	B3 DNA-binding domain
	bHLH	116	33	83	Helix-loop-helix DNA-binding domain
	bzip	55	17	38	Basic region leucine zipper
	C2H2	47	13	34	Zinc fifinger C2H2 domain-containing protein
	C3H	39	13	26	Zinc fifinger CCCH domain-containing protein
	DBB	6	3	3	double B-box zinc finger protein
	G2-like	40	17	23	Golden2-Like protein
	GRAS	30	15	15	C-terminal GRAS domain
	HSF	39	25	14	Heat stress transcription factor
	MYB	112	32	80	MYB-related protein
	NAC	74	32	42	NAC domain-containing protein

Notes: $q < 0.05$ and $|\log2(\text{ratio})| \geq 1$ were set as the threshold values for significantly different expression.

DEG set	TF family	Number of DEGs	Up-regulated	Downregulated	Description
	S1Fa-like	43	4	39	Leucine-rich repeat protein
	other TFs	484	160	324	
	total	1254	421	833	

Notes: $q < 0.05$ and $|\log2(\text{ratio})| \geq 1$ were set as the threshold values for significantly different expression.

Analysis of unigenes related to anthocyanidin biosynthetic pathways in the fruit peel

Previous studies have suggested that the anthocyanin biosynthetic pathway is an important metabolic branch of the flavonoid pathway, responsible for the production of anthocyanins in different plant tissues. Thus, we evaluated the mechanisms underlying red pigmentation in fruit peel based on a comparative transcriptome analysis. The compositions of compounds in anthocyanin biosynthesis pathways differed significantly depending on the growth stage (Fig. 3). The *PAL* gene regulates the synthesis of cinnamic acid and consistently showed a gradual decrease over the four growth stages. Seven *4CL* genes exhibited significant differences in expression levels, including 2 up-regulated and 5 down-regulated genes. The *CHS* gene is the first key enzyme in the flavonoid pathway, which regulates the synthesis of chalcone; it consistently showed a gradual increase over the four growth stages. The expression levels of *CHI*, *F3H*, *DFR*, and *ANS* were significantly up-regulated from S1 to S4, and the expression levels were positively correlated with the accumulation of anthocyanin. UDP-glucose: flavonoid-3-*O*-glycosyltranferase (UFGT), which mainly catalyzes the transformation of unstable anthocyanin glycosides into stable anthocyanin. By using a transcriptomic analysis, two *UFGT* genes (TRINITY_DN23815_c1_g1 and TRINITY_DN19893_c1_g5) involved in the anthocyanin biosynthesis pathway were screened (Fig. 3). In order to test the accuracy of the transcriptome data, we selected 12 anthocyanin synthesis genes and 2 MYB genes for qRT-PCR verification. The results showed that the expression trend of qRT-PCR was consistent with the transcriptome data, indicating that the transcriptome data has high reliability (Fig. 4).

Integrated analysis of the transcriptome and metabolome

By transcriptome and metabolome analyses, we constructed a co-expression network to identify interactions (Fig. 5). We found that 11 anthocyanin biosynthesis DEGs and 11 differential metabolites exhibited interactions. Highly positive correlations were obtained among most differential metabolites, and only pmb0542 (cyanidin 3-*O*-malonylhexoside) was minimally relevant to all other differential

metabolites. The correlation analysis between DEGs and differential metabolites also showed that levels of 5 DEGs (TRINITY_DN15939_c0_g1, TRINITY_DN20784_c0_g2, TRINITY_DN21962_c0_g2, TRINITY_DN23815_c1_g1, and TRINITY_DN26276_c0_g1) were positively correlated with all differential metabolites, while the other 6 DEGs (TRINITY_DN17945_c0_g2, TRINITY_DN21141_c0_g1, TRINITY_DN25788_c2_g8, TRINITY_DN22885_c0_g4, TRINITY_DN25099_c1_g5, and TRINITY_DN24108_c0_g2) were negatively correlated with all differential metabolites (Fig. 5a, Additional file 4: Table S4). In addition, the combined metabolome and transcriptome analysis revealed a significant positive correlation between both *ANS* (TRINITY_DN15939_c0_g1) and *UFGT* (TRINITY_DN23815_c1_g1) (PCC > 0.82) and most metabolites (Fig. 5b). These data indicated that expression patterns of *ANS* and *UFGT* have decisive roles in color development, and their regulatory relationships need to be further studied.

Discussion

Fruit peel color is an important trait that effects the edible value of many fruits. Several recent reports have described key genes and enzymes for fruit peel color formation and anthocyanin accumulation in fruit trees, including pear and apple [13, 26]. However, the mechanism underlying *C. humilis* anthocyanin biosynthesis is not well understood. In addition, as far as we have known, the differences of metabolites of *C. humilis* peel color changes at different developmental stages have not been studied. We first completed a transcriptome analysis and metabolic profiling for four developmental stages of fruit development to focus on anthocyanin biosynthetic pathways.

Based on UPLC-MS/MS, changes in metabolites during peel color changes in *C. humilis* were qualitatively and quantitatively analyzed. The unique patterns of anthocyanin-related metabolite changes in the process of peel color change were preliminarily established for the first time. Using a targeted metabolomics approach, we detected and annotated 11 metabolites with significant differences among developmental stages of fruit peel, all of which were up-regulated. Notably, the levels of two anthocyanins, cyanidin *O*-syringic acid and pelargonidin 3-*O*-beta-d-glucoside (callistephin chloride), increased significantly during fruit peel development in *C. humilis*. Zhou et al. [33] found that cyanidin 3-*O*-glucoside and pelargonidin 3-*O*-beta-d-glucoside are unique anthocyanins in pink tea flowers. Accordingly, we suggested that the accumulation of these anthocyanins accounted for the red pigmentation of fruit peel during fruit development.

It has been deeply studied in antho biosynthesis patyway, and the structure genes who play important roles in this pathway has also been well know [33]. We found that most of upstream structural genes (i.e. *PAL*, *C4H*, and *4CL*) in the early committed steps were highly active in young fruits but were gradually silenced during fruit development (Fig. 4). In contrast, *UFGT* genes (TRINITY_DN23815_c1_g1 and TRINITY_DN19893_c1_g5), which are involved in the synthesis and accumulation of anthocyanins, were gradually activated during ripening process of *C. humilis*. *UFGT* played an important role in anthocyanin metabolism, which catalyzes the glycosylation of anthocyanins, and makes anthocyanins exist in plant vacuoles in the form of stable anthocyanins. Our results suggested that the accumulation of anthocyanin

metabolites is positively correlated with *UFGT* expression levels during the reddening of *C. humilis* fruit peel. The accumulated metabolites are converted by *UFGT* into the stable anthocyanins involved in fruit peel reddening with fruit ripening[31]. Similarly, Steyn et al. [34] reported that *UFGT* activity increases in pear fruit development. Therefore, we proposed that *UFGT* during fruit ripening could potentially promote *C. humilis* peel coloring.

Studies of *C. humilis* fruit color have revealed that in addition to DEGs, TFs play a key role in the regulation of color changes and the activity of pigment biosynthetic pathways [35]. MYB TFs contribute to the regulation of plant secondary metabolism, including anthocyanin biosynthesis [26, 36]. For example, *AcMYB110a* regulates anthocyanin production in petals by promoting the expression of *F3GT1* in *Actinidia* [37]. *McMYB10* plays an important role in ever-red leaf coloration by positively regulating *McF3'H* in crabapple [38]. Consistent with these results, we identified several candidate *MYB* genes (TRINITY_DN26515_c0_g1 and TRINITY_DN21536_c0_g1) with the potential to positively regulate the majority of flavonoid biosynthetic genes, their expression abundance was consistent with anthocyanin accumulation. Therefore, the results suggested that these 2 MYB genes might play an important role in the temporal and spatial regulation of anthocyanin synthesis.

Metabolites can provide direct insight into biological processes and mechanisms. In this study, we detected 16 metabolites, including 11 differential metabolites, providing reference values for the isolation and identification of functional compounds in fruit peel. Our combined metabolome and transcriptome analysis uncovered a significant positive correlation between the synthesis of two genes (*ANS*, TRINITY_DN15939_c0_g1; *UFGT*, TRINITY_DN23815_c1_g1) in the anthocyanin pathway (ko00942) and differential metabolites (Fig. 5b), establishing an new interesting connection between relative secondary metabolites and genes. The high expression of *ANS* and *UFGT*, as major determinants of anthocyanin biosynthesis in fruit peel, might account for *C. humilis* fruit coloration.

Conclusions

Collectively, the mechanisms underlying anthocyanin accumulation in *C. humilis* fruit peel were analyzed by MRM, transcriptomics, and qRT-PCR. At least 16 anthocyanins exhibited differences among developmental stages, and cyanidin *O*-syringic acid and pelargonidin 3-*O*-beta-d-glucoside (callistephin chloride) were identified as the major anthocyanins. FurthermoreMoreover, The anthocyanin biosynthesis genes were identified and analyzed by transcriptome data and qRT-PCR. These results indicated that *ANS* and *UFGT* were the key genes involving in anthocyanin accumulation. Meanwhile, two MYB genes (TRINITY_DN26515_c0_g1 and TRINITY_DN21536_c0_g1) participate anthocyanin synthesis pathway and potentially regulated the expression of related genes in *C. humilis*. This study improves our understanding of anthocyanin accumulation and the molecular mechanism underlying anthocyanin biosynthesis in *C. humilis* fruit peel. However, the precise mechanism requires further studies.

Methods

Plant materials

C. humilis (cv. Zhangwu) was grown at the Xiangyang Farm of Northeast Agricultural University in Harbin, Heilongjiang Province, China. Fruits samples were collected from 30, 42, 57, and 70 days post-anthesis (DPA), designated young fruit (S1), green fruit (S2), slightly red fruit (S3), and red fruit (S4) (Fig. 6). The samples were cored, frozen immediately in liquid nitrogen, and stored at -80 °C until use.

Metabolome analysis

The sample preparation, analysis, metabolite qualitative and quantification were carried out by MWDB (Metware Biotechnology Co., Ltd., Wuhan, China) according to their standard procedures [39–42]. The instrument system of metabolite data acquisition was mainly ultra-high performance liquid chromatography (HPLC). The data of metabolite content were normalized by range method and R software was used (www.r-project.org/), and the accumulation patterns of metabolites among different samples were analyzed by hierarchical cluster analysis (HCA). The analysis of metabonomics data combines the methods of univariate and multivariate statistical analysis, and according to the characteristics of data, multi-perspective analysis is also carried out, and finally the differential metabolites will be accurately mined.

RNA isolation and Illumina sequencing

Total RNA was extracted from four developmental stages of the fruit peel using the Omega RNA Extraction Kit (Biel, Switzerland). The 1% agarose gel electroassay was used to detect degradation and impurities. RNA quantity and quality were determined using a Keio K5500 spectrophotometer (Kaio, Beijing, China), Agilent Bioanalyzer 2100 system, and Agilent RNA 6000 Nano Kit (Agilent Technologies, Palo Alto, CA, USA). A total of 4 g of total RNA per sample (12 samples = 4 stages × 3 biological replicates) was used as input material for RNA-seq library preparation (Annoroad Biotechnology, Zhejiang, China). Trinity was for de novo assembly. Raw data (in FASTq format) were processed using Perl scripts to measure Q30 values and GC contents. After removing low-quality reads, the remaining clean data were used for all downstream analyses.

Transcriptome data analysis

Differential expression analysis was performed using DESeq2 v1.4.5 according to the standard procedures of Annoroad Biotechnology, The corrected $p < 0.05$ and $|\log2Ratio| \geq 1$ was used to identify the differentially expressed genes (DEGs), Using gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses of differentially expressed genes.

Quantitative real-time PCR and expression validation

Based on the results of the transcriptomic analysis, 12 anthocyanin biosynthetic genes and 2 MYB genes were chosen for verification by qRT-PCR. The primers for PCR are listed in Additional file 5: Table S5. Values are presented as means of three biological replicates and three technical replicates. The relative gene expression levels were calculated by the $2^{-\Delta\Delta t}$ method.

Supplementary Information

Additional file 1: Table S1. Primer sequences for qRT-PCR.

Additional file 2: Table S2. Differentially accumulated anthocyanins during the ripening process in *C. humilis* fruit peel.

Additional file 3: Table S3. Summary of transcriptome sequencing data transcriptome assembly.

Additional file 4: Table S4. Anthocyanin biosynthesis related genes.

Additional file 5: Table S5. Correlation analysis of anthocyanin synthesis genes and metabolites in *Cerasus humilis* fruit peel.

Additional file 6: Figure S1. GO classification of unigenes from *Cerasus humilis* fruit peel.

Additional file 7: Figure S2. Classification of unigenes from *Cerasus humilis* fruit peel.

Additional file 8: Figure S3. Functional categorization of genes with significant changes in different stages of *Cerasus humilis* fruit peel development. S2_S1(a), S3_S2 (b) and S4_S3 (c). red indicates up-regulated and gray indicated down-regulated genes.

Additional file 9: Figure S4. Scatterplot of the top 20 enriched KEGG pathway statistics in DEGs. S2_S1(a), S3_S2 (b) and S4_S3 (c). The q value is the multiple hypothesis test-corrected P value. The q value ranges from [0–1]. The closer that number is to 0, the more significant the enrichment is. The greater the rich factor, the greater the degree of enrichment is.

Abbreviations

ANS: Anthocyanidin synthase; CHI: Chalcone isomerase; CHS: Chalcone synthase; C4H: Cinnamic acid 4-hydroxylase; COG: Cluster of Orthologous Group; DEGs: Differentially expressed genes; DFR: Dihydroflavonol 4-reductase; DPA: Days post-anthesis; F3H: Flavanone 3-hydroxylase; GO: Gene ontology; HCA: Hierarchical cluster analysis; HPLC-MS/MS: High-performance liquid chromatography-tandem mass spectrometry; KEGG: Kyoto Encyclopedia of Genes and Genomes; MRM: multiple reaction monitoring; PAL: Phenylalanine ammonia-lyase; PCA: Principal component analysis; qRT-PCR: Real-time quantitative

polymerase chain reaction; RPKM: Reads Per Kilobase Million Mapped Reads; TF: Transcription factor; UFGT: UDP-glucose: flavonoid-3-O-glycosyltranferase; 4CL: 4-coumarate-CoA ligase

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

All data generated or analysed during this study are included in this published article and its Supplementary information files. The raw RNA-seq data are freely available at: www.ncbi.nlm.nih.gov/bioproject/PRJNA636897.

Competing interests

The authors declare that they have no competing interests.

Funding

The analysis of the study was funded by Fundamental Research Funds for the Central Universities (Grant No. 2572018CG02) and the National Natural Science Foundation of China (Grant No. 31901660). The funding bodies had no role but provided the financial support to this research.

Authors' contributions

XJ and XS conceived and designed the experiments. XJ and LZ performed the experiments. XJ, JR, LS, SL and XS analyzed the data. XJ, JR and XS wrote the manuscript. All authors read and approved the manuscript.

Acknowledgments

We thank the Department of Genetics, College of Life Science, Northeast Forestry University in Heilongjiang Province.

References

1. Mu XP, Wang PF, Du JJ, Gao YG, Zhang JC. Comparison of fruit organic acids and metabolism-related gene expression between *Cerasus humilis* (Bge.) Sok and Cerasus glandulosa (Thunb.) Lois. Plos One. 2018; 13(4).
2. Wang F, Wang JW, Sun LJ, Song XS. The molecular cloning and functional characterization of ChNAC1, a NAC transcription factor in *Cerasus humilis*. Plant Growth Regul. 2019;89(3):331–43.
3. Ye L, Yang C, Li W, Hao J, Sun M, Zhang J, Zhang Z. Evaluation of volatile compounds from Chinese dwarf cherry (*Cerasus humilis* (Bge.) Sok.) germplasms by headspace solid-phase microextraction and gas chromatography-mass spectrometry. Food Chem. 2017;217:389–97.
4. Wang PF, Mu XP, Du JJ, Gao YG, Bai DH, Jia LT, Zhang JC, Ren HY, Xue XF. Flavonoid content and radical scavenging activity in fruits of Chinese dwarf cherry (*Cerasus humilis*) genotypes. J Forestry Res. 2018;29(1):55–63.
5. Wang PF, Mu XP, Gao YG, Zhang JC, Du JJ. Successful induction and the systematic characterization of tetraploids in *cerasus humilis* for subsequent breeding. Sci Hortic-Amsterdam. 2020; 265.
6. Chau CF, Wu SH. The development of regulations of Chinese herbal medicines for both medicinal and food uses. Trends Food Sci Tech. 2006;17(6):313–23.
7. Xu W, Dubos C, Lepiniec L. Transcriptional control of flavonoid biosynthesis by MYB-bHLH-WDR complexes. Trends Plant Sci. 2015;20(3):176–85.
8. Jia L, Wu Q, Ye N, Liu R, Shi L, Xu W, Zhi H, Rahman AN, Xia Y, Zhang J. Proanthocyanidins inhibit seed germination by maintaining a high level of abscisic acid in *Arabidopsis thaliana*. J Integr Plant Biol. 2012;54(9):663–73.
9. Dong T, Han R, Yu J, Zhu M, Zhang Y, Gong Y, Li Z. Anthocyanins accumulation and molecular analysis of correlated genes by metabolome and transcriptome in green and purple asparaguses (*Asparagus officinalis*, L.). Food Chem. 2019;271:18–28.
10. Solfanelli C, Poggi A, Loreti E, Alpi A, Perata P. Sucrose-specific induction of the anthocyanin biosynthetic pathway in *Arabidopsis*. Plant Physiol. 2006;140(2):637–46.
11. Takos AM, Jaffe FW, Jacob SR, Bogs J, Robinson SP, Walker AR. Light-induced expression of a MYB gene regulates anthocyanin biosynthesis in red apples. Plant Physiol. 2006;142(3):1216–32.
12. 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 Communications. 2020;1(1):100006.
13. Bai S, Tao R, Tang Y, Yin L, Ma Y, Ni J, Yan X, Yang Q, Wu Z, Zeng Y, et al. BBX16, a B-box protein, positively regulates light-induced anthocyanin accumulation by activating MYB10 in red pear. Plant

Biotechnol J. 2019;17(10):1985–97.

14. Zhang XB, Liu CJ. Multifaceted Regulations of Gateway Enzyme Phenylalanine Ammonia-Lyase in the Biosynthesis of Phenylpropanoids. Mol Plant. 2015;8(1):17–27.
15. Liu XY, Yu HN, Gao S, Wu YF, Cheng AX, Lou HX. The isolation and functional characterization of three liverwort genes encoding cinnamate 4-hydroxylase. Plant Physiol Biochem. 2017;117:42–50.
16. Li Y, Kim JI, Pysh L, Chapple C. Four Isoforms of Arabidopsis 4-Coumarate: CoA Ligase Have Overlapping yet Distinct Roles in Phenylpropanoid Metabolism. Plant Physiol. 2015;169(4):2409–21.
17. Sun W, Meng X, Liang L, Jiang W, Huang Y, He J, Hu H, Almqvist J, Gao X, Wang L. Molecular and Biochemical Analysis of Chalcone Synthase from Freesia hybrid in flavonoid biosynthetic pathway. PLoS One. 2015;10(3):e0119054.
18. Kang JH, McRoberts J, Shi F, Moreno JE, Jones AD, Howe GA. The flavonoid biosynthetic enzyme chalcone isomerase modulates terpenoid production in glandular trichomes of tomato. Plant Physiol. 2014;164(3):1161–74.
19. Flachowsky H, Halbwirth H, Treutter D, Richter K, Hanke MV, Szankowski I, Gosch C, Stich K, Fischer TC. Silencing of flavanone-3-hydroxylase in apple (*Malus* x *domestica* Borkh.) leads to accumulation of flavanones, but not to reduced fire blight susceptibility. Plant Physiol Biochem. 2012;51:18–25.
20. Luo P, Ning G, Wang Z, Shen Y, Jin H, Li P, Huang S, Zhao J, Bao M. Disequilibrium of Flavonol Synthase and Dihydroflavonol-4-Reductase Expression Associated Tightly to White vs. Red Color Flower Formation in Plants. Front Plant Sci. 2015;6:1257.
21. Liu Y, Shi Z, Maximova S, Payne MJ, Guiltinan MJ. Proanthocyanidin synthesis in *Theobroma cacao*: genes encoding anthocyanidin synthase, anthocyanidin reductase, and leucoanthocyanidin reductase. BMC Plant Biol. 2013;13:202.
22. Lin-Wang K, Micheletti D, Palmer J, Volz R, Lozano L, Espley R, Hellens RP, Chagne D, Rowan DD, Troggio M, et al. High temperature reduces apple fruit colour via modulation of the anthocyanin regulatory complex. Plant Cell Environ. 2011;34(7):1176–90.
23. Dubos C, Stracke R, Grotewold E, Weisshaar B, Martin C, Lepiniec L. MYB transcription factors in *Arabidopsis*. Trends Plant Sci. 2010;15(10):573–81.
24. Zhou H, Lin-Wang K, Wang H, Gu C, Dare AP, Espley RV, He H, Allan AC, Han Y. Molecular genetics of blood-fleshed peach reveals activation of anthocyanin biosynthesis by NAC transcription factors. Plant J. 2015;82(1):105–21.
25. Xie Y, Tan H, Ma Z, Huang J. DELLA Proteins Promote Anthocyanin Biosynthesis via Sequestering MYBL2 and JAZ Suppressors of the MYB/bHLH/WD40 Complex in Arabidopsis thaliana. Mol Plant. 2016;9(5):711–21.
26. An JP, Wang XF, Zhang XW, Xu HF, Bi SQ, You CX, Hao YJ. An apple MYB transcription factor regulates cold tolerance and anthocyanin accumulation and undergoes MIEL1-mediated degradation. Plant Biotechnol J. 2020;18(2):337–53.
27. Illa E, Eduardo I, Audergon JM, Barale F, Dirlewanger E, Li XW, Moing A, Lambert P, Le Dantec L, Gao ZS, et al. Saturating the *Prunus* (stone fruits) genome with candidate genes for fruit quality. Mol

Breeding. 2011;28(4):667–82.

28. Bai S, Sun Y, Qian M, Yang F, Ni J, Tao R, Li L, Shu Q, Zhang D, Teng Y. Transcriptome analysis of bagging-treated red Chinese sand pear peels reveals light-responsive pathway functions in anthocyanin accumulation. Sci Rep. 2017;7(1):63.
29. Zhuang H, Lou Q, Liu H, Han H, Wang Q, Tang Z, Ma Y, Wang H. 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).
30. Meng Y, Wang Z, Wang Y, Wang C, Zhu B, Liu H, Ji W, Wen J, Chu C, Tadege M, et al. The MYB Activator WHITE PETAL1 Associates with MtTT8 and MtWD40-1 to Regulate Carotenoid-Derived Flower Pigmentation in *Medicago truncatula*. Plant Cell. 2019;31(11):2751–67.
31. Zhang Q, Wang L, Liu Z, Zhao Z, Zhao J, Wang Z, Zhou G, Liu P, Liu M. Transcriptome and metabolome profiling unveil the mechanisms of *Ziziphus jujuba* Mill. peel coloration. Food Chem. 2020;312:125903.
32. Wang S, Tu H, Wan J, Chen W, Liu X, Luo J, Xu J, Zhang H. Spatio-temporal distribution and natural variation of metabolites in citrus fruits. Food Chem. 2016;199:8–17.
33. Zhou C, Mei X, Rothenberg DO, Yang Z, Zhang W, Wan S, Yang H, Zhang L. Metabolome and Transcriptome Analysis Reveals Putative Genes Involved in Anthocyanin Accumulation and Coloration in White and Pink Tea (*Camellia sinensis*) Flower. Molecules. 2020; 25(1).
34. Steyn WJ, Holcroft DM, Wand SJE, Jacobs G. Regulation of pear color development in relation to activity of flavonoid enzymes. J Am Soc Hortic Sci. 2004;129(1):6–12.
35. Luo F, Cai JH, Kong XM, Zhou Q, Zhou X, Zhao YB, Ji SJ. Transcriptome profiling reveals the roles of pigment mechanisms in postharvest broccoli yellowing. Hortic Res. 2019;6:74.
36. Liu JY, Osbourn A, Ma PD. MYB Transcription Factors as Regulators of Phenylpropanoid Metabolism in Plants. Mol Plant. 2015;8(5):689–708.
37. Fraser LG, Seal AG, Montefiori M, McGhie TK, Tsang GK, Datson PM, Hilario E, Marsh HE, Dunn JK, Hellens RP, et al. An R2R3 MYB transcription factor determines red petal colour in an *Actinidia* (kiwifruit) hybrid population. BMC Genom. 2013;14:28.
38. Tian J, Peng Z, Zhang J, Song T, Wan H, Zhang M, Yao Y. McMYB10 regulates coloration via activating *McF3'H* and later structural genes in ever-red leaf crabapple. Plant Biotechnol J. 2015;13(7):948–61.
39. Fraga CG, Clowers BH, Moore RJ, Zink EM. Signature-Discovery Approach for Sample Matching of a Nerve-Agent Precursor Using Liquid Chromatography-Mass Spectrometry, XCMS, and Chemometrics. Anal Chem. 2010;82(10):4165–73.
40. Chen W, Gong L, Guo Z, Wang W, Zhang H, Liu X, Yu S, Xiong L, Luo J. A novel integrated method for large-scale detection, identification, and quantification of widely targeted metabolites: application in the study of rice metabolomics. Mol Plant. 2013;6(6):1769–80.
41. Wishart DS, Feunang YD, Marcu A, Guo AC, Liang K, Vazquez-Fresno R, Sajed T, Johnson D, Li C, Karu N, et al. HMDB 4.0: the human metabolome database for 2018. Nucleic Acids Res.

2018;46(D1):D608–17.

42. Zhu ZJ, Schultz AW, Wang J, Johnson CH, Yannone SM, Patti GJ, Siuzdak G. Liquid chromatography quadrupole time-of-flight mass spectrometry characterization of metabolites guided by the METLIN database. Nat Protoc. 2013;8(3):451–60.

Figures

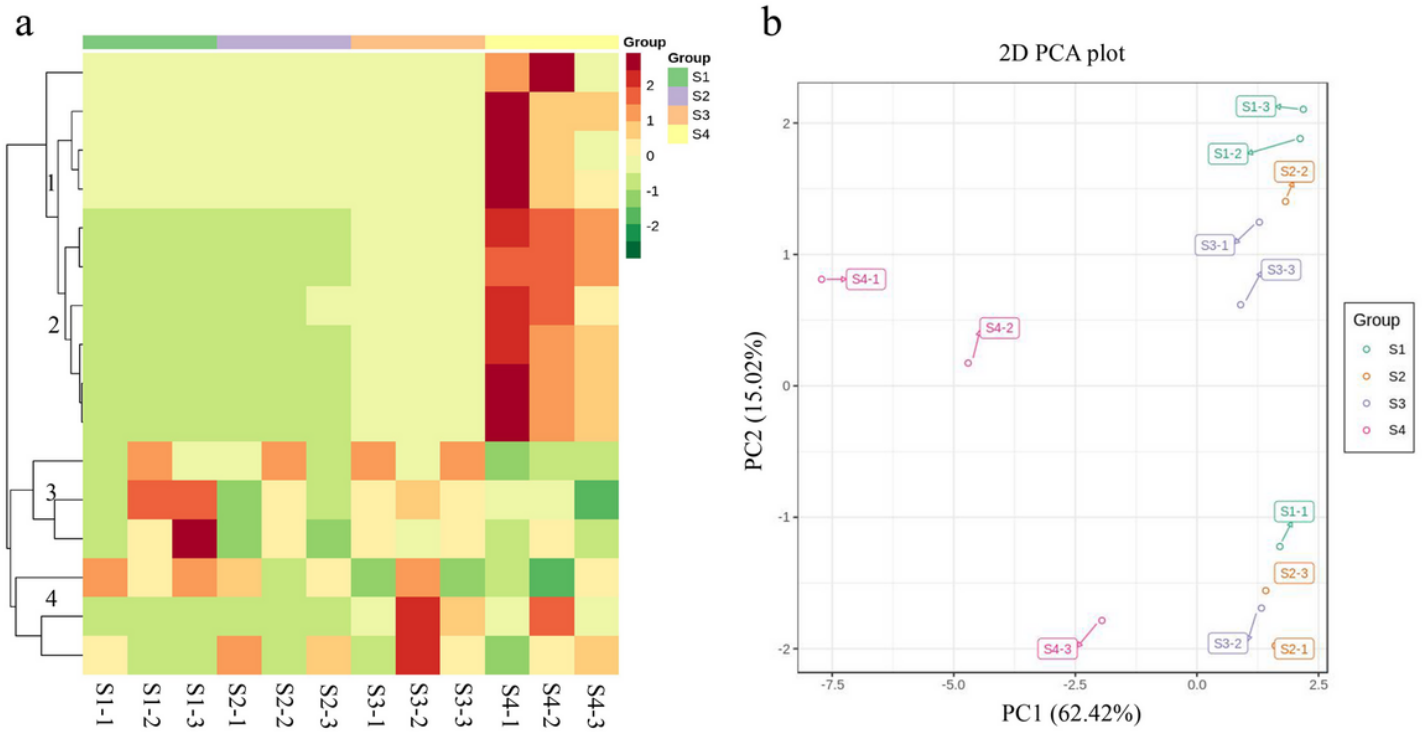

Figure 1

Heat map and PCA of the relative differences in anthocyanins in different fruit development stages in Cerasus humilis. a: Heat map visualization. Each sample is visualized in a single column and each metabolite is represented by a single row. Red indicates high abundance and green indicates low abundance. b: Score plots for PCA. young fruit (S1), green fruit (S2), slightly red fruit (S3), and red fruit (S4)

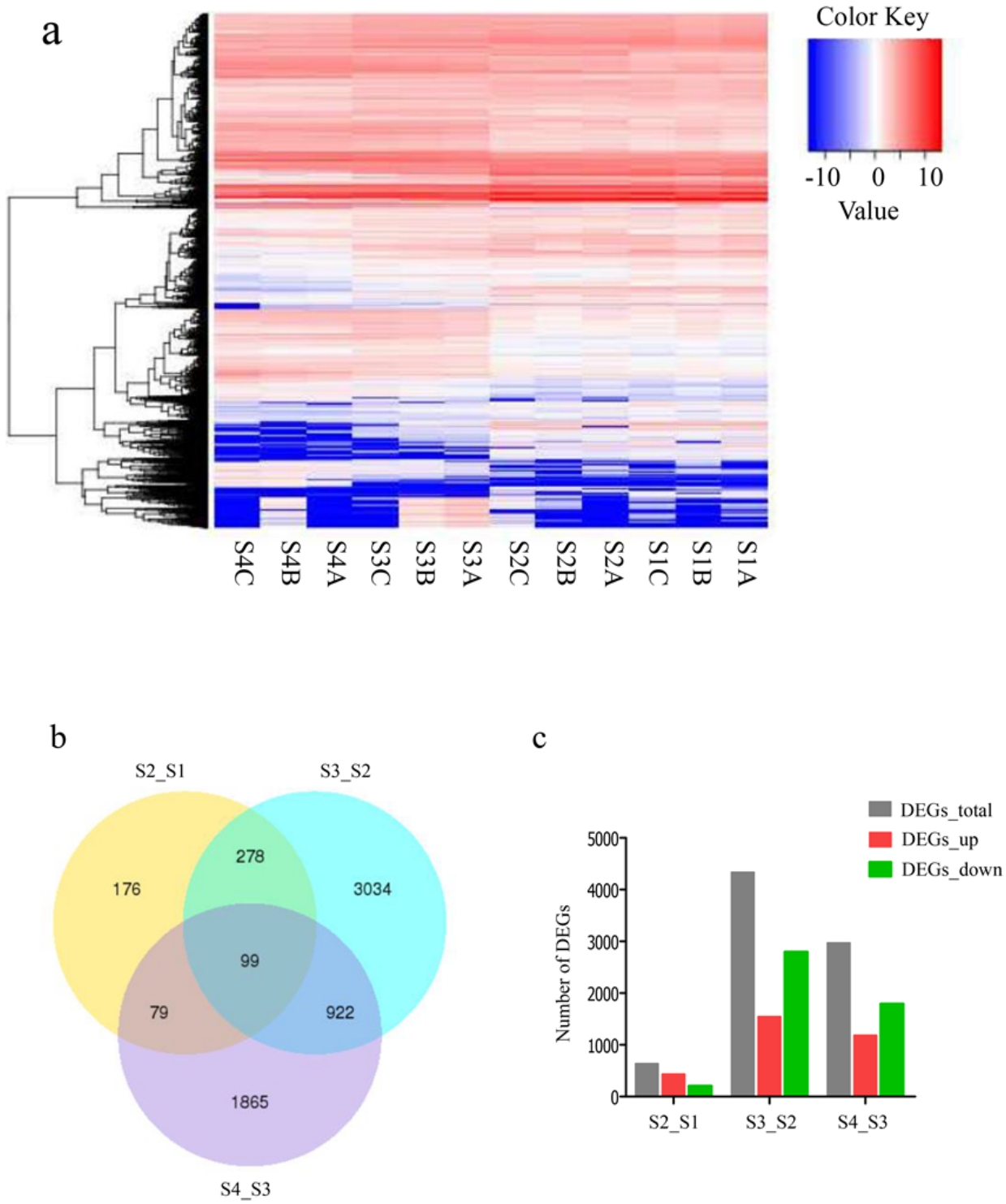

Figure 2

Analysis of transcriptomic data. (a) Heat map of DEGs in different developmental stages of Cerasus humilis peel. (b) Venn diagram of DEGs revealed in pair-wise comparisons. (c) The number of DEGs in pair-wise comparisons. S1–S4, young, green, slightly red, and red, respectively. A, B, and C are three biological replicates

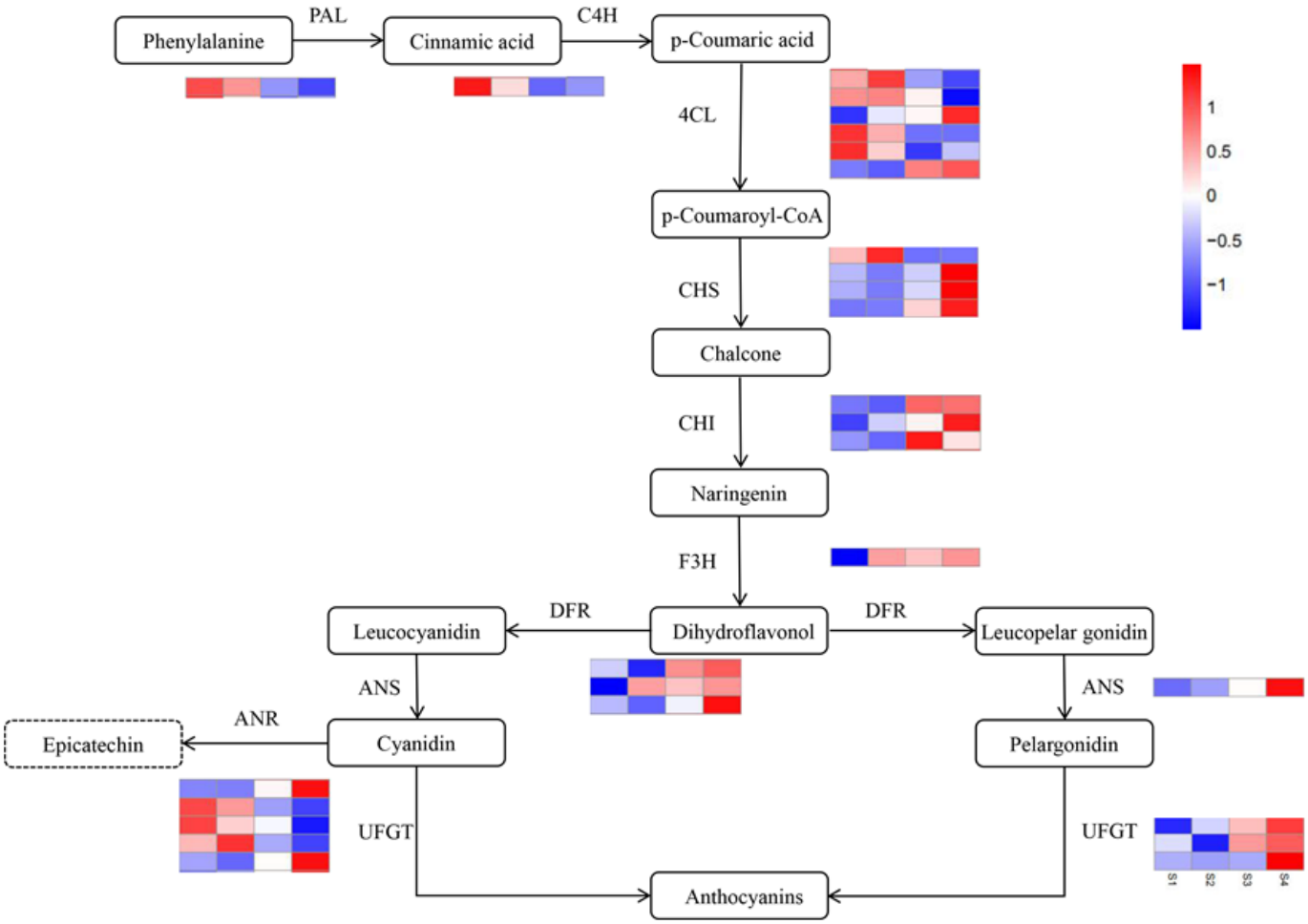

Figure 3

Anthocyanin biosynthetic pathway and their genes expression levels during C. humilis ripening. Blue and red in the heatmap represent the down- and up-regulated structural genes, respectively. Gene names are listed in Table S4

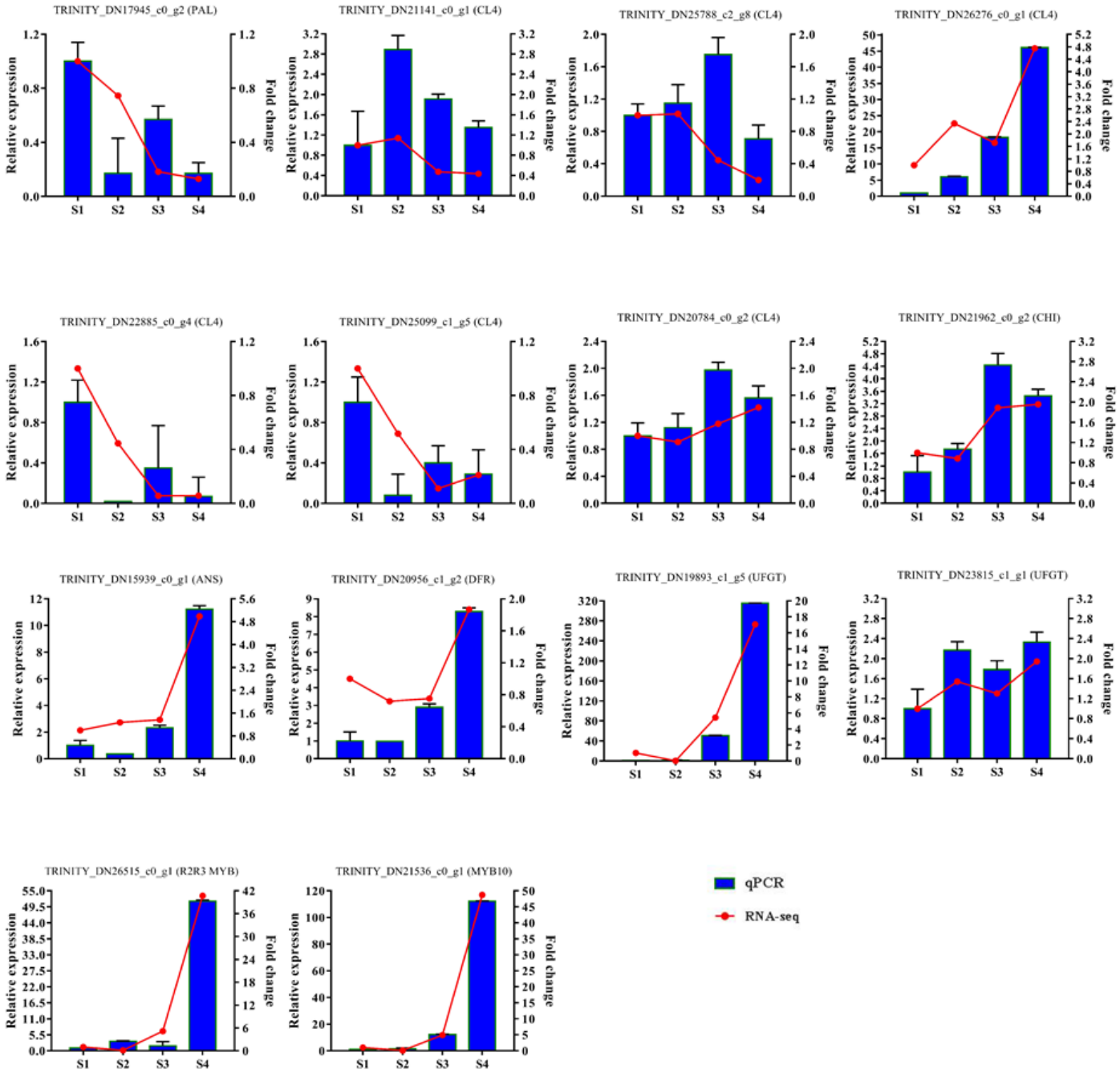

Figure 4

qRT-PCR validation of representative color-related genes in fruit peel development. Values for mean expression and standard errors indicated by vertical bars

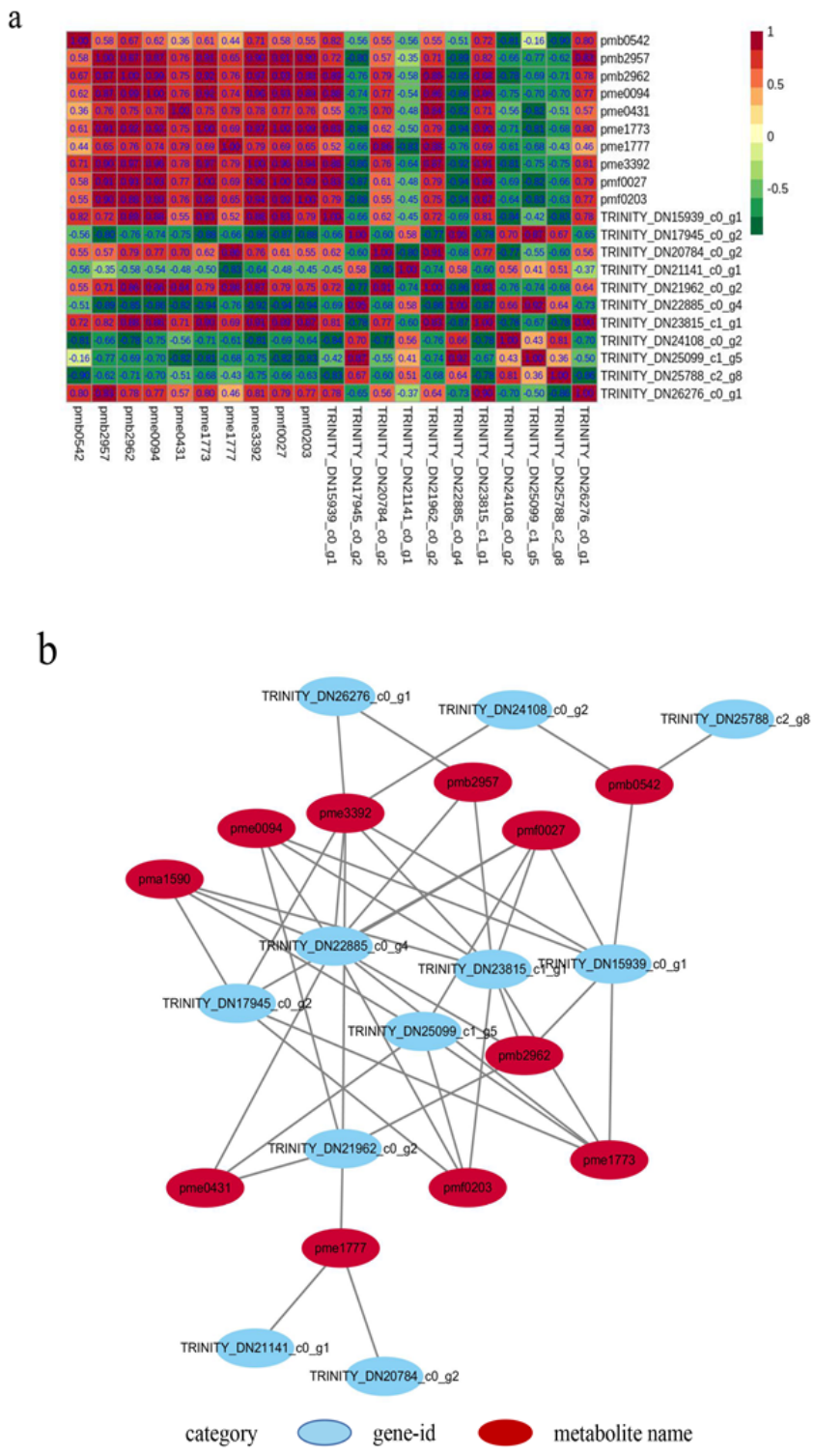

Figure 5

qRT-PCR validation of representative color-related genes in fruit peel development. Values for mean expression and standard errors indicated by vertical bars

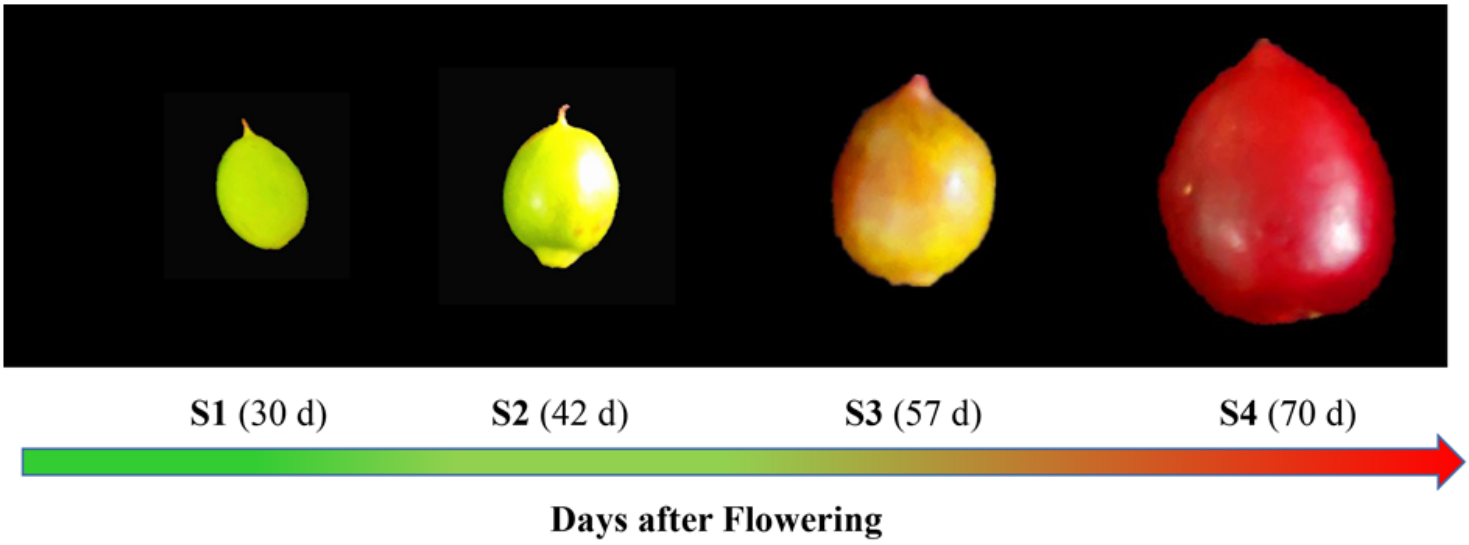

Figure 6

Pigmentation and phenotypes of Cerasus humilis fruit peel during different stages. fruit phenotypes: young fruit (S1), green fruit (S2), slightly red fruit (S3), and red fruit (S4)

Supplementary Files

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

- Supplementaryinformation.doc