



OPEN ACCESS

EDITED BY

Ting Wu,
China Agricultural University, China

REVIEWED BY

Nan Wang,
Shandong Agricultural University, China
Abdul Jalal,
Jiangsu University, China

*CORRESPONDENCE

Guofeng Xu

✉ xuguofeng@caas.cn

Jing Li

✉ ljing01@caas.cn

SPECIALTY SECTION

This article was submitted to
Functional and Applied Plant Genomics,
a section of the journal
Frontiers in Plant Science

RECEIVED 28 October 2022

ACCEPTED 29 March 2023

PUBLISHED 20 April 2023

CITATION

Li Y, Li H, Wang S, Li J, Bacha SAS, Xu G
and Li J (2023) Metabolomic and
transcriptomic analyses of the flavonoid
biosynthetic pathway in blueberry
(*Vaccinium* spp.).

Front. Plant Sci. 14:1082245.

doi: 10.3389/fpls.2023.1082245

COPYRIGHT

© 2023 Li, Li, Wang, Li, Bacha, Xu and Li.
This is an open-access article distributed
under the terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Metabolomic and transcriptomic analyses of the flavonoid biosynthetic pathway in blueberry (*Vaccinium* spp.)

Yinping Li¹, Haifei Li¹, Shiyao Wang², Jing Li¹,
Syed Asim Shah Bacha¹, Guofeng Xu^{1*} and Jing Li^{1*}

¹Laboratory of Quality and Safety Risk Assessment for Fruit, Research Institute of Pomology, Chinese Academy of Agricultural Sciences, Xingcheng, Liaoning, China, ²Department of Applied Biosciences, Toyo University, Ora-gun, Japan

As a highly economic small fruit crop, blueberry is enjoyed by most people in terms of color, taste, and rich nutrition. To better understand its coloring mechanism on the process of ripening, an integrative analysis of the metabolome and transcriptome profiles was performed in three blueberry varieties at three developmental stages. In this study, 41 flavonoid metabolites closely related to the coloring in blueberry samples were analyzed. It turned out that the most differential metabolites in the ripening processes were delphinidin-3-O-arabinoside (dpara), peonidin-3-O-glucoside (pnglu), and delphinidin-3-O-galactoside (dpgal), while the most differential metabolites among different varieties were flavonols. Furthermore, to obtain more accurate and comprehensive transcripts of blueberry during the developmental stages, PacBio and Illumina sequencing technology were combined to obtain the transcriptome of the blueberry variety Misty, for the very first time. Finally, by applying the gene coexpression network analysis, the darkviolet and bisque4 modules related to flavonoid synthesis were determined, and the key genes related to two flavonoid 3', 5'-hydroxylase (*F3'5'H*) genes in the darkviolet module and one *bHLH* transcription factor in the bisque4 module were predicted. It is believed that our findings could provide valuable information for the future study on the molecular mechanism of flavonoid metabolites and flavonoid synthesis pathways in blueberries.

KEYWORDS

blueberry, PacBio sequencing, flavonoid, transcriptome, metabolite

Introduction

Blueberry, an economically important small fruit, belongs to the section *Cyanococcus* of the genus *Vaccinium*, family *Ericaceae* (Cho et al., 2017; Hou et al., 2017; Li X et al., 2020; Guo et al., 2021). The year of 1906 saw its beginning of commercial cultivation; since then, blueberry has been widely planted around the world (Brevis et al., 2008; Bian et al., 2014).

It features a unique flavor and rich nutrition (Yuan et al., 2020; Zang et al., 2021). It contains sugar, organic acid, various amino acids, vitamins, unsaturated fatty acids, and minerals. More importantly, it is an excellent source of active substances such as anthocyanin, ascorbic acid, ellagic acid, and pterostilbene, which can be useful when it comes to improving eyesight, softening blood vessels, fighting bacteria and inflammation, inhibiting the growth of cancer cells, and resisting oxidation (Hou et al., 2017; Morita et al., 2017; Munagala et al., 2017; Wang et al., 2017; Lin et al., 2018). The antioxidant capacity of blueberry is mainly attributed to flavonoids, including flavonols, anthocyanins, and proanthocyanidins (Lin et al., 2018). Anthocyanin, one of the key quality factors in blueberries, can increase its commercial value. The common anthocyanins in blueberries are delphinidin, cyanidin, peonidin, petunidin, and malvidin (Chai et al., 2021). Anthocyanin mainly exists in the form of glycoside, and the glycosides include arabinose, glucose, galactose, and xylose (Hosseini and Beta, 2007; Timmers et al., 2017; Zhou et al., 2018).

Flavonoids are synthesized starting from the phenylalanine pathway in the plant cytoplasm (Zifkin et al., 2012; Hassani et al., 2020). The molecular synthesis mechanism of flavonoids has been revealed in depth, including the structural genes directly encoding key enzymes in each step of the flavonoid synthesis pathway (Cho et al., 2016; Zhao et al., 2019; Liu et al., 2020). As a substrate, phenylalanine is catalyzed by phenylalanine ammonia lyase (PAL), cinnamate 4-hydroxylase (C4H), 4-coumarate-CoA ligase (4CL), chalcone synthase (CHS), chalcone isomerase (CHI) enzymes to synthesize flavanones. Under the catalysis of flavanone 3-hydroxylase (F3H) and flavonoid 3'-hydroxylase (F3'H), flavanone forms dihydroflavonol, which is the most important intermediate metabolite in the flavonoid synthesis pathway. Dihydroflavonol 4-reductase (DFR), leucoanthocyanidin dioxygenase (LDOX), anthocyanin synthase (ANS), and UDP-glucose flavonoid 3-O-glucosyltransferase (UFGT) are responsible for dihydroflavonols to anthocyanins. Flavonol synthase (FLS) directs the oxidation of dihydroflavonol to produce flavonol. In addition, leucoanthocyanidins and anthocyanidins produce proanthocyanidins under the catalysis of leucoanthocyanidin reductase (LAR) and anthocyanin reductase (ANR), respectively (Tohge et al., 2017; Karppinen et al., 2021). Flavonoid glycosides are thought to be synthesized in the cytoplasm and transported to the vacuole through transport mechanism (including an ABC transporter) (Liu et al., 2020).

The biosynthesis and inhibition of flavonoids are affected by many factors, such as light, temperature, pH, and phytohormones (Liu et al., 2018; Li G et al., 2019). Flavonoid biosynthesis genes are largely regulated at the transcriptional level. The MYB-bHLH-WD40 (MBW) regulatory complex, composed of the *R2R3-MYB* transcription factor, *bHLH* transcription factor, and WD40 protein, has been shown that it could synergistically activate multiple genes for anthocyanin production in all angiosperms to date (Zhao et al., 2019; Li Y et al., 2020; Sun et al., 2020). In this MBW complex, the *MYB* transcription factor is core to determining the intensity and the pattern of anthocyanin production in plants (Castillejo et al., 2020). Moreover, another type of *R2R3-MYB* transcription factor can act alone with the structural genes of flavonoids (Naik et al.,

2022). The negative regulation of flavonoids involves the participation of a variety of inhibitors, namely, *R2R3-MYB*, *R3-MYB*, *NAC*, *miRNA*, *HD-ZIP*, and so forth. They inhibit the syntheses of flavonoids by inhibiting the expression of the MBW complex or destroying the stability of the MBW complex (LaFountain and Yuan, 2021; Naik et al., 2022).

The genome of blueberry is relatively large. Commercial blueberry varieties cover several species and interspecific hybrids, including tetraploid highbush blueberry, diploid and tetraploid lowbush blueberry (*V. myrtilloides* and *V. angustifolium*), hexaploid rabbiteye blueberry (*V. virgatum*), and hybrids thereof (Rowland et al., 2012; Ge et al., 2019). To further study the molecular synthesis and regulation of flavonoids, transcriptome sequencing is considered to be an effective method. So far, transcriptome sequencing has been used to study the genes of the flavonoid synthesis pathway in potatoes (Cho et al., 2016), grapes (He et al., 2010), mulberry (Zhao et al., 2019), and so forth. PacBio single-molecule long-reads sequencing technology (SMRT) breaks through the reading length restriction of next-generation high-throughput sequencing (NGS) and has been applied to effectively capture full-length sequences of genomes and transcripts (Hoang et al., 2017; Chao et al., 2018). Transcriptome and metabolomics are combined to study the regulatory network between genes and metabolites. In this study, the samples from three developmental stages of the tetraploid southern highbush blueberry Misty were sequenced by SMRT to obtain a more comprehensive and accurate full-length transcriptome sequence of blueberry fruit. Moreover, we used the Illumina sequencing technique to obtain the transcriptomes of Duke and Emerald at three developmental stages and analyzed all the transcriptome sequencing results of the three blueberry varieties to identify the important regulatory genes and possible pathways related to their ripening. Finally, we combined the metabolic data of flavonoids in these three blueberry cultivars to further screen the key genes and verify the potentially related genes, providing a molecular basis for the study of blueberry fruit color.

Materials and methods

Plant samples

Three blueberry varieties Misty, Duke, and Emerald were collected in sheds from Niuyingzi village, Wanghai Township, Xingcheng City, Liaoning Province, China, in May 2018 (Figure 1). Three developmental stages of whole blueberry fruits, namely, green, pink, and blue, were sampled and immediately frozen in liquid N₂ and stored at -80°C for flavonoid analysis and RNA extraction.

The RNAs of nine samples (three biological replicates per developmental stage) of Misty were mixed three replicates of each developmental stage with equal volume and sequenced on the PacBio RS II platform, while the RNAs of 18 samples of Duke and Emerald were subjected to 150 bp paired-end sequencing using the Illumina HiSeq X Ten platform.



FIGURE 1

Green, pink, and blue stages of Misty, Duke, and Emerald. (A) Misty, (B) Duke, and (C) Emerald.

Chemical reagents

HPLC-grade acetonitrile and methanol were purchased from Thermo Fisher Scientific (Fair Lawn, NJ, USA). Formic acid of LC-MS grade was purchased from Anaque Chemicals Supply (Shanghai, China). Flavonoids, including 13 flavonols and 28 anthocyanins, were purchased from Sigma-Aldrich Company Ltd. (Shanghai, China), ChromaDex Corp (Santa Anna, CA, USA), and PhytoLab company (Vestenbergsgreuth, Germany) (Supplementary Tables S1, S2).

Measurement of flavonoid

The extraction of flavonoids was carried out in a way that described in a previous study with minor modifications (Appelhagen et al., 2011; Vrhovsek et al., 2012). The obtained filtrate was dissolved to 5 ml and filtered through a 0.22- μ m organic phase filter membrane before UPLC-MS/MS analysis.

Metabolite profiling was conducted using an Acquity UPLC system. The system was interfaced to a Xevo triple-quadrupole mass spectrometer and was equipped with an electrospray ionization source operating in positive ionization modes (Waters Corp., Milford, MA, USA). Chromatographic separation was performed on a Waters ACQUITY UPLC[®] HSS T3 column (2.1 mm \times 150 mm, 1.8 μ m) with a column temperature of 40°C. The mobile phase was composed of solvent A (methanol: acetonitrile, 7:3) and solvent B (5% formic acid solution) at a flow rate of 0.3 ml/min. The following gradient elution program was run: 0–1 min, 10% A; 1–16 min, 25% A; 16–18 min, 40% A; 18–19 min, 80% A; 19–20 min, 10% A. The injection volume was 2 μ l. The MS/MS parameters were used: ion source temperature of 150°C, desolvation temperature of 500°C, desolvent gas flow of 800 liter/h, cone gas flow of 50 liter/h, collision gas (high-purity argon) flow, and 0.13 ml/min (Supplementary Tables S1, S2).

The acquisition was performed in multiple reaction monitoring (MRM) mode. The metabolite data were analyzed by principal component analysis (PCA) and the orthogonal partial least-

squares-discriminant analysis (OPLS-DA) modules with SIMCA14.1 software. The predictive variable importance for the projection larger than 1 (VIP_{pred} > 1) was considered to be the differential metabolite.

RNA extraction, cDNA library construction, and sequencing

Total RNA was isolated with PureLink[®] Plant RNA reagent (Life Technologies, Carlsbad, CA, USA) following the standard protocol. The extracted RNA was checked on 1% agarose gels for degradation and contamination and then measured using a Qubit[®] RNA Assay Kit in a Qubit[®] 2.0 Fluorometer (Life Technologies, Carlsbad, CA, USA) for its concentration. The RNA was qualified with a NanoPhotometer[®] spectrophotometer (IMPLEN, CA, USA) and an Agilent Bioanalyzer 2100 system (Agilent Technologies, Santa Clara, CA, USA) before library construction.

A total amount of 5 μ g of total RNA was used to synthesize cDNA using the SMARTer[®] PCR cDNA Synthesis Kit (Takara Biotechnology, Dalian, China). Input DNA damage and end repatriation were treated with a SMRTbell Template Prep Kit (Pacific Biosciences, Menlo Park, CA, USA). Amplicons' size selection was performed using the BluePippin[™] Size Selection System (Sage Science, Beverly, MA, USA). SMRTbell template libraries were generated using the Pacific Biosciences DNA Template Prep Kit 2.0 before sequencing. The libraries were generated using the PacBio DNA/Polymerase Kit for sequencing primer annealing and polymerase binding. Subsequently, SMRT sequencing was carried out on a PacBio RSII instrument at the Biomarker Biotechnology Corporation (Beijing, China).

A total amount of 3- μ g RNA per sample was used as an input material, and sequencing libraries were generated using the NEBNext[®] Ultra[™] RNA Library Prep Kit for Illumina[®] (NEB, Ipswich, MA, USA) following the manufacturer's recommendation. Their ligated cDNA library was sequenced using the Illumina HiSeq[™] 2000 system by Biomarker Biotechnology Corporation (Beijing, China).

PacBio data analysis

PacBio raw reads were filtered out of low-quality reads and adaptor sequences using SMRT analysis v2.3.0 software. The reads of insert (ROI) sequences were extracted from original sequences, requiring a minimum read quality of 0.75 and a minimum of one full pass. Full-length (FL) reads and non-full-length reads (NFL reads) were identified using the RS module in SMRT Analysis software by detecting the existence of 3' and 5' primers, as well as the polyA tail in ROLs. FL reads contained all three elements, and full-length non-chimeric (FLNC) reads were regarded as no additional copies of adapter sequence within FL reads. Consensus isoforms were obtained by using the ICE (Iterative Clustering for Error Correction) algorithm to cluster FL reads into a cluster. Then, combined with NFL reads, the consistent sequences in each cluster were polished by the quiver program and, finally, high-quality isoforms with more than 99% postcorrection accuracy were obtained. Additionally, read corrections were carried out by using Illumina RNA-Seq reads of the same experiment with Proovread v2.12 (Rodriguez-Mateos et al., 2012), resulting in the corrected isoforms.

The transcripts were annotated by MISA software for SSR analysis and TransDecoder software for predicting the potential coding sequences and corresponding amino acid sequences of transcripts. The lncRNA identification from isoform sequencing was calculated by the four common used approaches, including CPC analysis (Li et al., 2014; Cho et al., 2017), CNCI analysis (Altschul et al., 1997), Pfam protein domain analysis, and CPAT analysis (Deng et al., 2006). Transcription factors in blueberries were analyzed with iTAK software (Zheng et al., 2016). Using Blast v2.2.26 software (Altschul et al., 1997), the obtained nonredundant transcripts were compared with the NR (Deng et al., 2006), SwissProt (Apweiler, 2004), GO (Ashburner et al., 2000), COG (Tatusov et al., 2000), KOG (Koonin et al., 2004), eggNOG (Huerta-Cepas et al., 2016), and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases (Huerta-Cepas et al., 2016) to obtain annotation information for the transcripts.

Illumina data analysis

RNA-seq and *de novo* transcriptome assembly in blueberries were carried out as described by Huerta-Cepas (Wu et al., 2016). The expression analysis was performed with RESM software based on the FPKM (fragments per kilobases per million fragments) values (Li and Dewey, 2011). The differentially expression analyses between samples from different developmental stages were performed by the DESeq R package (Leng et al., 2013) based on a model of the negative binomial distribution. Genes with FDR < 0.01 and fold change ≥ 2 were identified as differentially expressed genes (DEGs). GO enrichment analysis of DEGs was implemented by the topGO R packages based on the Kolmogorov–Smirnov test. KEGG pathways of DEGs were constructed using KOBAS software (Mao et al., 2005).

WGCNA, KEGG, and candidate hub genes analysis

For coexpression network analysis, we conducted weighted gene coexpression network analysis (WGCNA) with the default settings in BMKCloud (<http://www.biocloud.net>). To obtain the genes related to flavonoid synthesis, all DEGs and metabolomics data detected in the three developmental stages of blueberries were subjected to WGCNA to construct trait-related modules. Additionally, the meaningful modules were annotated with KEGG pathway enrichment analysis by ggplot on the BMKCloud platform. The coexpression network was visualized using Cytoscape software (version 3.7.1). The cytoHubba app was used to find the top 20 hub genes based on the maximal clique centrality (MCC) algorithm.

Results

Metabolic profiling of blueberry

To understand the color formation of blueberry during ripening, UPLC–MS/MS was used to detect 13 flavonols and 28 anthocyanins in the three blueberry cultivars (Figure 1). All metabolites were analyzed by PCA and OPLS-DA modules.

First, the metabolite differences at different developmental stages were studied. The pairwise comparison of the three varieties at different stages showed that the difference of metabolites between the green and blue fruit stages was the most significant (Figures 2A–C; Supplementary Tables S1, S2). In the green and pink fruit stages, 19 differential metabolites (VIPpred > 1) in the three varieties were found (Figure 2A). In the green and blue fruit stages, 18 differential metabolites among the three varieties were discovered (Figure 2B). In the pink and blue fruit stages, there were 19 differential metabolites among the three varieties (Figure 2C). Among the three blueberry varieties, 12 identical metabolites were found in the three developmental stages of fruit ripening, and the highest VIPpred values were delphinidin-3-*O*-arabinoside (dpara), peonidin-3-*O*-glucoside (pnglu), and delphinidin-3-*O*-galactoside (dpgal).

Then, we compared the metabolite differences among blueberry varieties during fruit ripening and found that the main differential metabolite of blueberry fruits was flavonols (Figures 2D–F; Supplementary Tables S1, S2). The results of pairwise comparisons among the three blueberry varieties during the fruit ripening process showed that there were 14 differential metabolites between Misty and Duke, nine differential metabolites between Misty and Emerald, and nine differential metabolites between Duke and Emerald (Figures 2D–F). Additionally, five identical metabolites among the three blueberry cultivars were isorhamnetin-3-*O*-galactoside (isogal), syringetin-3-*O*-pentoside (syphen), quercetin-3-*O*-malonylglycoside (qumal), isorhamnetin-3-*O*-glucoside (isoglu), and quercetin-3-*O*-rhanosyl-galactoside (qurhan).

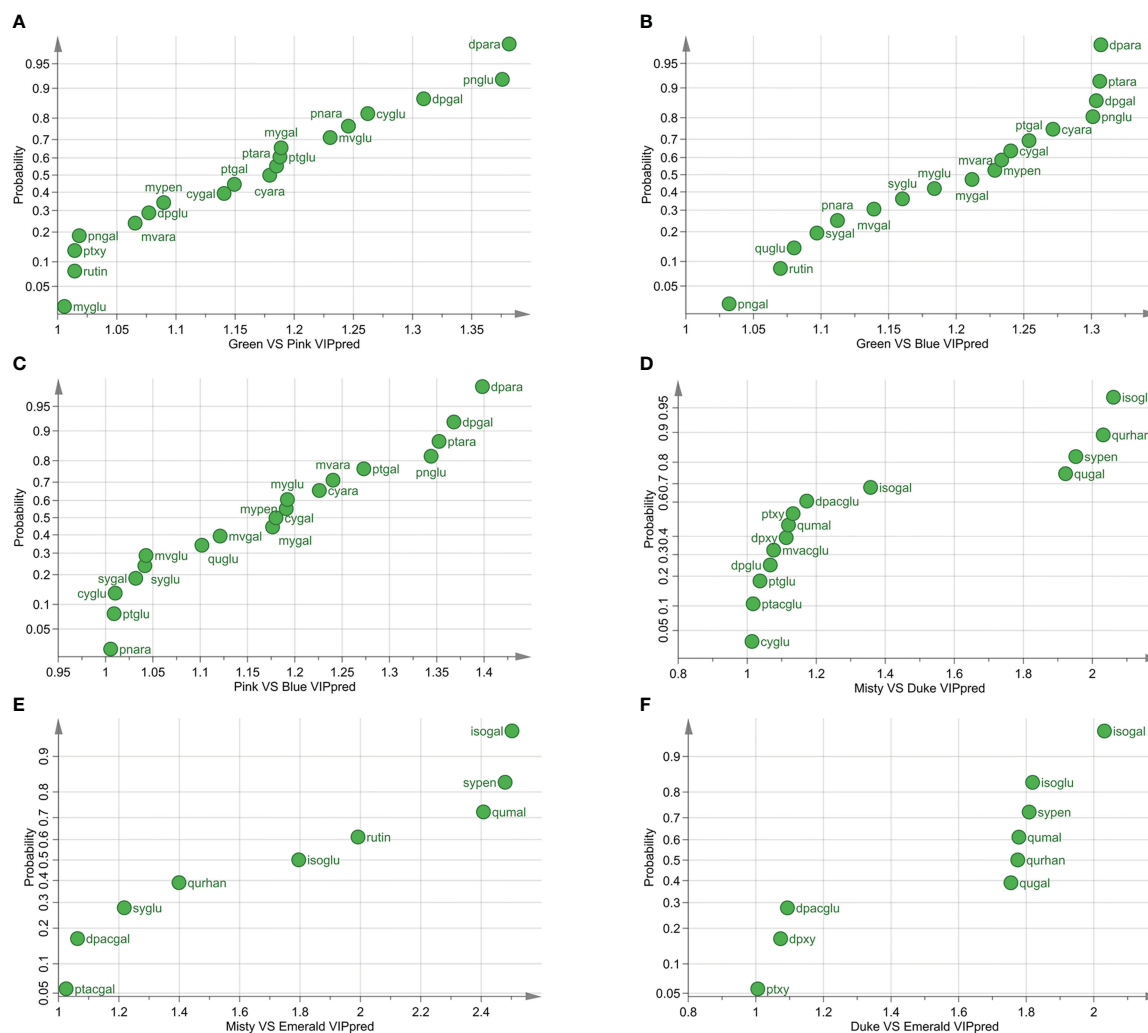


FIGURE 2

Different metabolites of three varieties at different developmental stages. (A) Differential metabolites of three blueberry varieties at the green and pink stages (VIP > 1). (B) Differential metabolites of three blueberry varieties at the green and blue stages (VIP > 1). (C) Differential metabolites of three blueberry varieties at the pink and blue stages (VIP > 1). (D) Differential metabolites of Misty and Duke in three developmental stages (VIP > 1). (E) Differential metabolites of Misty and Emerald in three developmental stages (VIP > 1). (F) Differential metabolites of Duke and Emerald in three developmental stages (VIP > 1). VIPpred > 1, predictive variable importance for the projection larger than 1.

PacBio sequencing and functional annotation

To obtain a comprehensive catalog of the southern highbush blueberry transcriptome, PacBio technology was employed to sequence the transcriptomes of the three developmental stages of Misty fruits. A pooled sample representing high-quality RNAs from different developmental stages of Misty fruits was sequenced, and three size-fractionated libraries (1–2, 2–3, and 3–6 kb) were constructed to avoid loading bias. Size-fractionated libraries (1–2, 2–3, and 3–6 kb) were constructed and sequenced using the PacBio RS II platform with nine cells for each mixed sample (Supplementary Figure S1A), yielding 15.24 Gb data. A total of 675,483 ROIs were generated, including 232,368 in 1–2 kb libraries, 264,200 in 2–3 kb libraries, and 178,915 in 3–6 kb libraries. As expected, the mean length of the ROIs was consistent with each size-selected library, with an average length of 3435.33 bp. Of these

ROIs, 285,909 FL reads were identified based on the presence of 5' primers, 3' primers, and poly (A) tails. Additionally, 284,329 FLNC reads were identified with 0.55% artificial connectors (Supplementary Figure S1B). In total, 158,832 consensus isoforms (27,102 polished high-quality isoforms and 131,718 polished low-quality isoforms) were obtained from isoform-level clustering and quiver polishing, which were further corrected with Illumina RNA-seq data (Supplementary Figure S1C).

After removing the redundant sequences for all SMRT subreads using Cogent software, 57,220 nonredundant reads were produced. Using transcoder software to predict the coding region, the results showed that a total of 55,143 ORFs were obtained, of which 34,313 were complete ORFs. After the sequence-structure analysis of nonredundant reads, 43,868 SSRs and 34,313 complete CDS regions were predicted (Supplementary Figures S2A, S2B). After filtering through the four approaches (CPC, CNCI, PLEK, and Pfam), we finally obtained 1,295 long noncoding RNAs (lncRNAs)

(Supplementary Figure S2C). Then, the LncTar target gene prediction tool was used to predict the target genes of lncRNAs, and a total of 1,293 novel lncRNAs were targeted. Moreover, a total of 5,492 transcription factors were identified from the Iso-Seq reads, including 2,223 transcription factors, 883 transcript regulators, and 2,454 protein kinases, which belong to 20 families of transcription factors (Supplementary Figure S2D). To analyze the functions of the transcripts, we annotated 54,585 nonredundant transcripts using the data from seven nucleotide and protein databases (Supplementary Figures S3).

Illumina sequencing and identification of the differentially expressed genes

In this study, to understand the molecular basis of the metabolic differences detected in the three different developmental stages, transcriptome sequencing of 27 samples was carried out using Illumina sequencing technique. After removing the adaptor and low-quality sequence, a total of 178.03 G of clean data was generated, and the Q30 values of each sample were not less than 92.53%. In this project, nonredundant transcripts obtained by SMRT sequencing were used as references for sequence alignment and subsequent analysis. Clean reads were sequence-aligned with transcripts by using STAR to obtain location information on the transcripts. Using RSEM software (Li and Dewey, 2011), FPKM was used as the index to measure the expression level of transcripts or genes through the location information of mapped reads on the third-generation transcripts. Taking Pearson correlation coefficient (R) as the evaluation index of biological repeat correlation, the correlation coefficients of the biological replicates at the same developmental stage of the same blueberry cultivar were higher than 0.95, indicating that the reproducibility of the biological replicates was satisfied. These results showed that the quantity and quality of the sequencing data were high enough to guarantee further analysis.

Pairwise comparisons of each sample detected 205,930 DEGs by DESeq software (Supplementary Figure S4). First, the DEGs of the same blueberry variety at different developmental stages were analyzed. In the comparison of MGvsMP, MGvsML, and MPvsML, 2,539, 4,125, and 633 significant DEGs were identified. In the comparison of DGvsDP, DGvsDL, and DPvsDL, 5,546, 7,619, and 677 DEGs were found out. In the comparison of EGvsEP, EGvsEL, EPvsEL, 4,389, 6,342, and 887 DEGs were discovered. We also identified DEGs among different varieties at the same developmental stage. In the comparison of DGvsEG, DGvsMG, and EGvsMG, 20,011, 18,605, and 18,774 DEGs were obtained. In the comparison of DPvsEP, DPvsMP, and EPvsMP, 20,450, 19,086, and 18,146 DEGs were obtained. In the comparison of DBvsEB, DBvsMB, and EBvsMB, 20,795, 19,205, and 18,101 DEGs were obtained.

Analyses of coexpression networks revealed flavonoid-related DEGs

To identify flavonoid biosynthesis-related transcripts, we performed WGCNA on all DEGs and identified 12 modules

(Figure 3; Supplementary Figure S5). In the module-trait relationships, the modules showing the highest correlations with the synthesis of most flavonoid metabolites were selected, namely, darkviolet and bisque4 modules. Darkviolet was positively correlated with 36 metabolites, and bisque4 was negatively correlated with 34 metabolites. To identify gene expression patterns in the two modules, their expression patterns were shown in the heatmap (Figure 4). The heatmap and expression of key genes indicated that the two modules were divided into two groups. The expression patterns of Misty and Emerald's three developmental stages were similar in these two modules, and different from the expression patterns of Duke's three periods in these two modules, which implies a difference in the bioactive product contents between the southern and northern highbush blueberry varieties.

KEGG analysis was carried out to further determine the gene function of these two modules (Figure 5). The results showed that genes in the darkviolet and bisque4 modules were mainly related to the synthesis and metabolism of primary and secondary metabolites. From Figure 5A, genes in the darkviolet module were mainly enriched in phenylpropanoid, flavone, flavonol, and anthocyanin biosynthesis pathways. The abundant flavonoid synthesis structural genes in the darkviolet module further proved that the coexpression genes in this module were involved in the accumulation of flavonoids in blueberry fruits. From Figure 5B, genes in the bisque4 module were significantly enriched for photosynthesis-antenna proteins and carbon fixation in photosynthetic organization pathways.

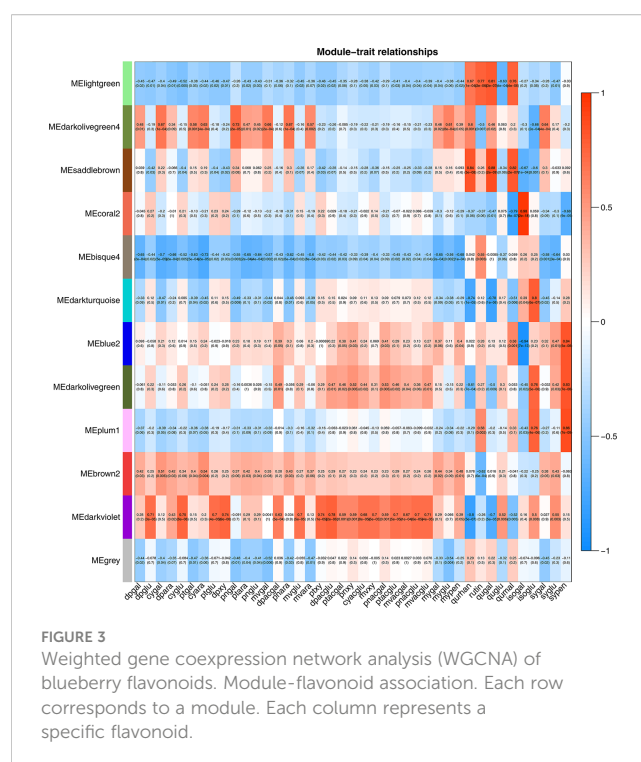


FIGURE 3
Weighted gene coexpression network analysis (WGCNA) of blueberry flavonoids. Module-flavonoid association. Each row corresponds to a module. Each column represents a specific flavonoid.

Structural genes involved in flavonoid synthesis

As shown in Figures 4, 6, the structural and regulatory genes play an important role in flavonoid synthesis process. Structural genes are directly involved in the flavonoid pathway through the synthesis of specific enzymes (Figure 6). The structural genes in the two modules were analyzed respectively (Figure S6; Supplementary Tables S3, S4). In the darkviolet module, eight structural genes in the flavonoid pathway were revealed, namely, *PAL*, *CAH*, *CHS*, *F3H*, *F3'H*, *F3'5'H*, *LDOX*, and *UFGT*, while in the bisque4 module, only one flavonoid synthesis structural gene was disclosed, namely, *LAR*.

Transcription factors involved in the synthesis of flavonoids

Transcription factors are involved in regulating various pathways of plant growth and development, and can also regulate the expression of specific genes in the flavonoid synthesis pathway (Li Y et al., 2020; Karppinen et al., 2021). At present, the reported transcription factors include *MYB*, *bHLH*, *bZIP*, *AP2/ERF*, *NAC*, and zinc finger (LaFountain and Yuan, 2021). In this study, the transcription factors related to flavonoid synthesis in darkviolet module include 1 *MYB*, 2 *MYB*-related, 1 *bZIP*, 4 *AP2/ERF-ERF*, 3 *NAC*, and 3 zinc fingers (Supplementary Figure S6A; Tables S3, S4). Through phylogenetic analysis, it is found that the *MYB* transcription factor named *VcMYBM1* belongs to the *R2R3-MYB* SG6 subfamily, which mainly regulates the synthesis of anthocyanins (Supplementary Figure S7; Table S5). More than 10 repressors to flavonoid synthesis were found, mainly the inhibitors of corresponding environmental signals or hormones, which directly or indirectly regulate MBW complex and structural genes in the flavonoid synthesis pathway through protein-protein interaction or post-transcriptional gene silencing (Xiao et al., 2022). In bisque4 module, there are 4 *bHLH*, 1 *bZIP*, 2 *DBB*, 1 *B3*, 1 *AUX/IAA*, 7 *HB-HD-ZIP*, 1 *AP2/ERF-ERF*, 7 zinc fingers, 2 *GARP-ARR-B*, and 1 *GARS* (Supplementary Figure S6B). Blasted against the TAIR database, *PB.15903.1*, *PB.22598.1*, and *PB.30746.1*, belonging to *bHLH* transcription factor, may involve in light regulated flavonoid biosynthesis. *PB.4531.3* belonging to the *DBB* transcription factor, is related to the negative regulation of light signal transduction. *PB.24525.1*, *PB.4604.1* and *PB.13554.1* related to the hormone regulation of blueberry, belong to *AUX/IAA*, zinc fingers and GRAS transcription factor, respectively. *PB.8456.1* belonging to *GARP-ARR-B* translation factor is presumably related to nitrogen synthesis of blueberry. These transcription factors may inhibit the expression of structural genes involved in flavonoid synthesis by participating in blueberry response to light, hormones, nitrogen, and other signals.

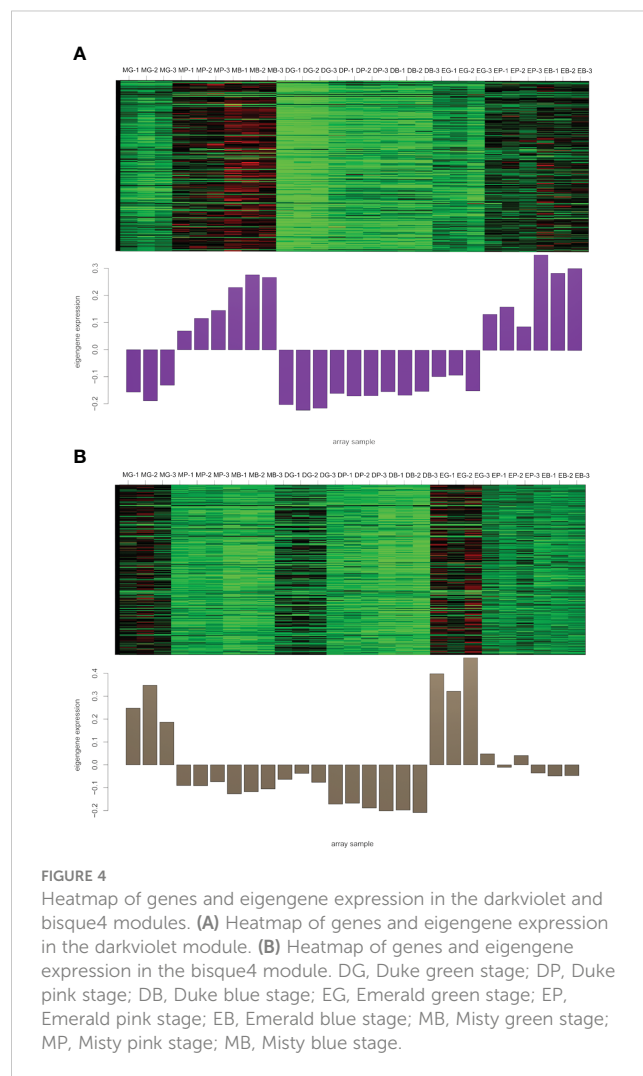
Candidate hub genes related to flavonoid synthesis

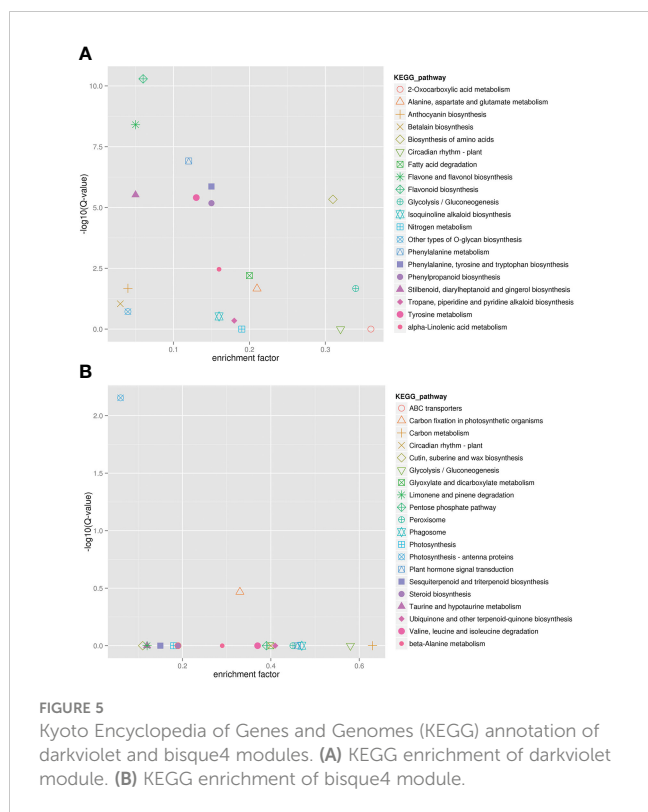
To study the hub genes in darkviolet and bisque4 modules, we obtained the first 20 genes of the two modules using MCC method

in the cytoHubba plugin and annotated them (Figure 7; Supplementary Table S6). Two *F3'5'H* genes (*PB.12741.1* and *PB.12745.3*) in the darkviolet module and one *bHLH* transcription factor (*PB.15903.1*) in the bisque4 module were key synthetic and regulated genes in the flavonoid synthesis pathway (Table S7). Blasting against TAIR database, *PB.15903.1* is the homologous sequence of *CIB2* interacting with *CRY2*, which may be related to the light regulation of flavonoid synthesis. Other genes may also be related to the synthesis and regulation of flavonoids, which needs further verification.

Discussion

Blueberry is the most popular small blue fruit and its unique color is closely related to the type and content of flavonoids (Dong et al., 2019). Flavonoids in blueberries show excellent physiological activities, including antioxidant, anticancer, and antidiabetic activities, and reduce cardiovascular disease (Cappai et al., 2018). The high content of flavonoids, especially anthocyanins, is believed to be responsible for the health benefits of these blueberries





(Rodriguez-Mateos et al., 2012). Therefore, the determination of flavonoids at different developmental stages of blueberry is not only helpful to analyze the coloring mechanism of blueberry but is also of great significance for the development of health products containing blueberry flavonoids. In this study, the UPLC-MS/MS method was used to assess the flavonoid profiles of three blueberry varieties at three typical developmental stages. The results showed that there were significant differences in the types and contents of flavonoids in different varieties and developmental stages. Among them, the highest metabolite content and the largest differential metabolites of blueberry varieties at different developmental stages were delphinidin glycosides, which is consistent with previous reports (Timmers et al., 2017; Chai et al., 2021). It is noteworthy that myricetin-3-*O*-galactoside (mvgal) has the highest flavonoid content in Duke, while delphinidin-3-*O*-galactoside has the highest flavonoid content in Misty and Emerald. According to Günther's study, malvidin-3-*O*-glycosides were predominant in northern highbush blueberry Nui and rabbiteye Velluto Blue (Günther et al., 2020). Meanwhile, the least differential metabolites were found between Misty and Emerald among the three blueberry varieties. The flavonoid content of blueberries is affected by variety, light, temperature, environment, and other factors (Li X et al., 2019). It owes to the fact that Misty and Emerald both belong to southern highbush blueberry, and the number of differential metabolites between them was lowest compared with the remaining pairwise comparisons.

Among the new blueberry varieties announced in the past decade, the southern highbush blueberry has developed the largest number of new varieties compared with other species, which leads the main breeding direction of blueberry breeding at

present. Although, the tetraploid northern highbush blueberry *V. corymbosum* has been published (Colle et al., 2019), and the transcriptome sequencing results of some blueberry varieties have been completed successively (Zifkin et al., 2012; Tang et al., 2021), which has greatly promoted the identification of candidate genes and pathways for excellent fruit quality and accelerated the molecular breeding of northern *Vaccinium* species. However, compared with the northern highbush blueberry, the southern highbush blueberry has a complex genetic background, its tetraploid genome and transcriptome resources are insufficient, and there are few researches on key quality factor genes such as anthocyanin and their regulatory factors, which seriously restricts the cultivation of the southern highbush blueberry in China (Rowland et al., 2012; Ge et al., 2019). To better understand the molecular mechanism of flavonoid synthesis, PacBio sequencing was used for the very first time in tetraploid southern highbush blueberry Misty at three fruit development stages, which provides a new interpretation of FL transcripts, variable shear, lncRNAs, and transcription factors of tetraploid southern highbush blueberry. We combined the Illumina data of the two other blueberry varieties from three different periods to analyze the DEGs of the three different varieties and three different periods. The results showed that the maximum number of DEGs was identified when compared with the green fruit stage in different varieties or different developmental stages, which is consistent with previous results (Lin et al., 2018). During the development of blueberry fruit from the green to blue fruit stage, significant changes have taken place in sugar and organic acids (Li X et al., 2020), fruit firmness (Cappai et al., 2018), fruit size, single fruit weight (Yang et al., 2021), and flavonoids (Spinardi et al., 2019). The early development of blueberry is not only involved in the development process of blueberry but also involved in the quality change of blueberry. Therefore, the differentially expressed genes are the most abundant when compared with the green fruit stage in different varieties or different developmental stages.

The synthesis of flavonoids includes phenylalanine, flavonoid, anthocyanin, flavone, and flavonol biosynthesis pathways, involving a variety of structural and regulated genes (Yi et al., 2021). Through WGCNA of differential gene and flavonoid data, two modules darkviolet and bisque4 were obtained. Nine structural genes for flavonoid synthesis were disclosed in the two modules, of which the *F3'5'H* gene in darkviolet module was identified as the key gene for flavonoid synthesis. In the present study, the expression levels of nine structural genes in Misty were significantly higher than those in Duke. Dihydroquercetin is an important precursor and node in different types of anthocyanin biosynthesis. The *F3'5'H* gene hydroxylates the B-rings of anthocyanidins (Zhang et al., 2014; Yi et al., 2021). Therefore, the high expression of *F3'5'H* genes promotes the accumulation of delphinidin and its derivatives, which are the precursors of blue and reddish-purple pigments (Zhang et al., 2014; Yi et al., 2021). Additionally, the differential gene expression of *LDOX* and *UFGT* in southern and northern highbush blueberry varieties contributed to the enrichment of delphinidin-3-*O*-galactoside in southern highbush varieties and malvidin-3-*O*-galactoside in northern highbush blueberry varieties. The glycosylation and side-chain modification of

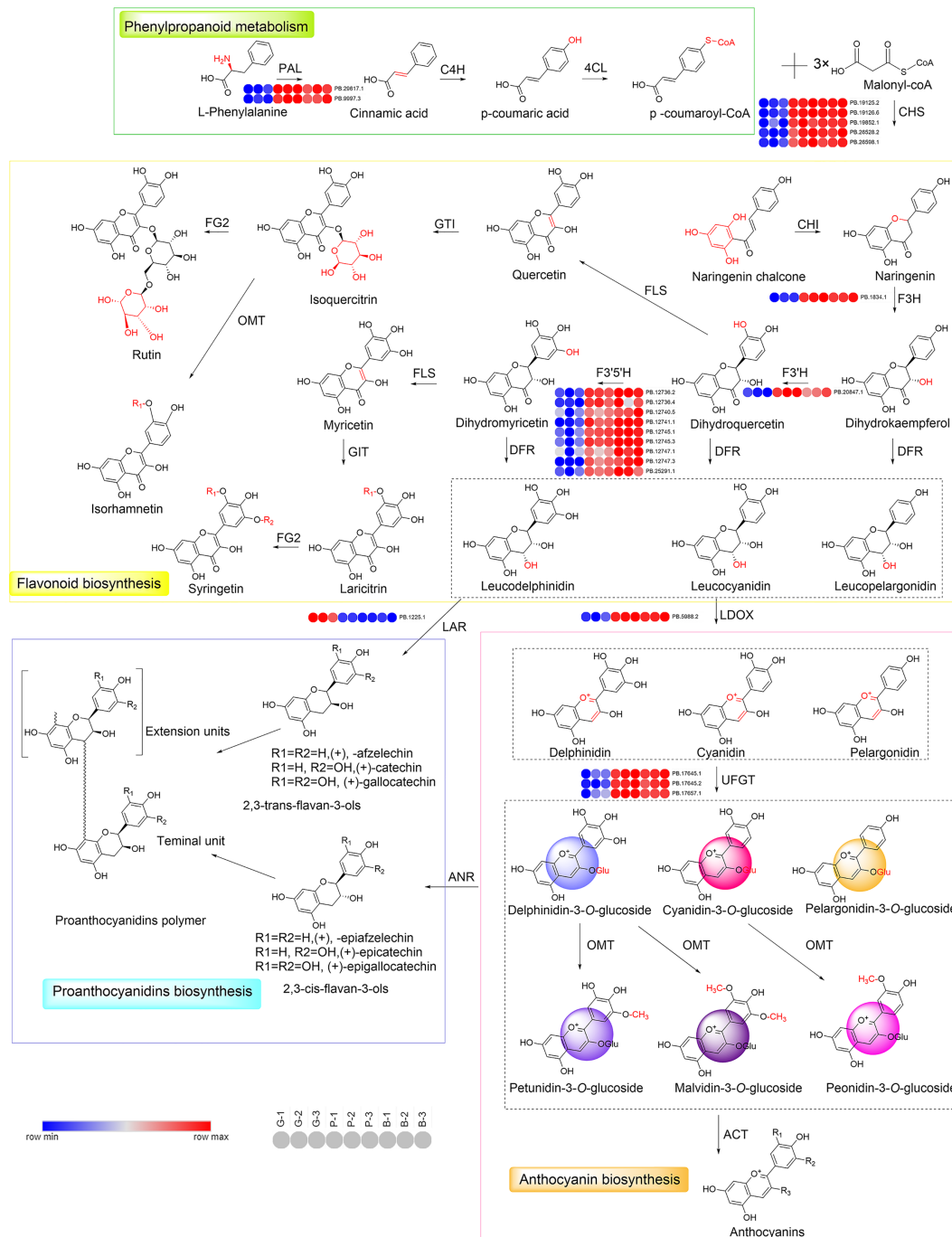
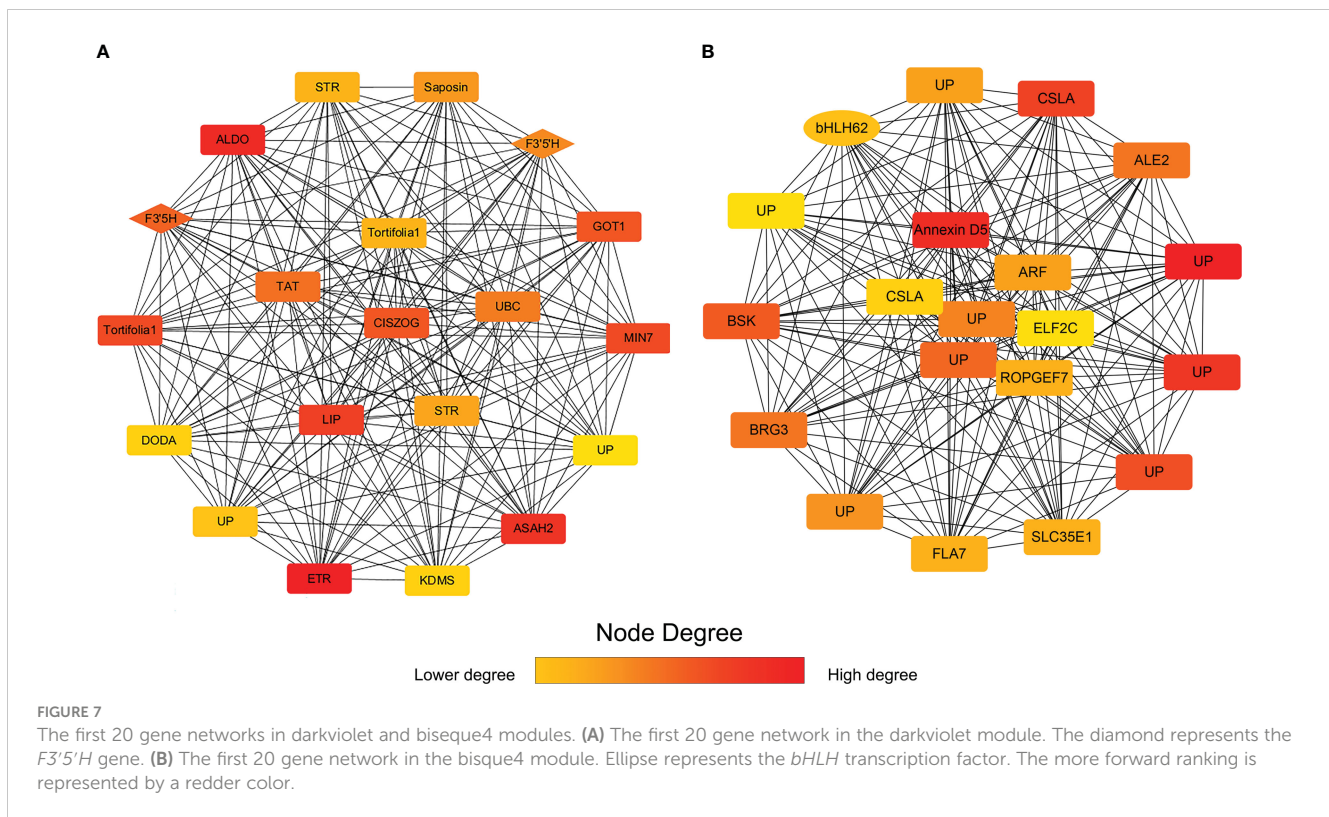


FIGURE 6 Reconstruction of the phenylpropanoid-flavonoid-anthocyanin biosynthesis pathway in Misty blueberry. The typical colors of anthocyanidins are also shown. *PAL*, phenylalanine ammonia lyase; *C4H*, cinnamate 4-hydroxylase; *4CL*, 4-coumarate-CoA ligase; *CHS*, chalcone synthase; *CHI*, chalcone isomerase; *F3H*, flavanone 3-hydroxylase; *F3'H*, flavonoid 3'-hydroxylase; *F3'5'H*, flavonoid 3', 5'-hydroxylase; *FLS*, flavonol synthase; *GTI*, Anthocyanidin 5, 3-O-glucosyltransferase; *FG2*, flavonol-3-O-glucoside L-rhamnosyltransferase; *DFR*, dihydroflavonol 4-reductase; *OMT*, o-methyltransferase; *LDOX*, Leucoanthocyanidin Dioxygenase; *UFGT*, UDP-glucose flavonoid 3-O-glucosyltransferase; *LAR*, leucoanthocyanidin reductase; *ANS*, Anthocyanidin synthase; *ACT*, anthocyanin acyltransferase. G, green stage; P, pink stage; B, blue stage.

anthocyanins not only affect the stability of anthocyanins in a natural state but are also related to the color presentation of anthocyanins (Zhang et al., 2014). In this experiment, the structural gene in the negative regulation module is *LAR*, which promotes the synthesis of proanthocyanidins. It should be noted that there is a competitive relationship between proanthocyanidin

and anthocyanin synthesis since they share common precursors (Karppinen et al., 2021).

The MBW complex directly regulates the transcription of flavonoid structural genes. The *MYB* transcription factor is the core transcription factor in MBW complex. The *MYB* transcription factor can regulate specific structural genes, such as the *R2R3-MYB* transcription factor,



which is responsible for these specificities through conserved N-terminal R2 and R3 DNA-binding domains. In our study, *PB.9559.1* was a homologous gene of *MYB1*. In litchi and apple, the targeted regulatory gene of *MYB1* is *UFGT* (Lai et al., 2014). In garlic shoot tissue, the co-bombardment with *MYB1* and *bHLH* (*Zm-Lc*) constructs allowed earlier development of anthocyanin pigmentation (Schwinn et al., 2016). Moreover, some transcription factors are participated in signaling pathways from the environment and hormone perception to flavonoid activation or inhibition. As the most important environmental signal, light plays an important role in the synthesis of flavonoids. In this study, we found several transcription factors might involve in this pathway. The hub gene *PB.15903.1* was found to involve in the photoreceptor *CRY2* signaling pathway. *CRYs* detecting blue light are related to the expression levels of *F3H*, *F3'5'H*, *CHS*, and *CHI* (Zheng et al., 2019). Blasted against the TAIR database, *PB.22598.1* and *PB.30746.1* act as negative regulators interacting with photoreceptors *phyA* and *phyB*. It is reported that *PHY* absorbs red light and far-red light, and *CHS* is not expressed in the *PHY* mutant of *Arabidopsis* (Li et al., 2018). *PB.5700.2* encodings a homolog of *HY5* (*HYH*) may participate in *phyB* signaling pathway. *HY5* has been reported to antagonize the light-regulated central inhibitor *COPI* or directly regulate the expression of *CHS*, *CHI*, *F3H*, *F3'H*, *DFR*, and *LDOX* genes (Jaegle et al., 2016). *PB.4531*, belonging to the DBB family,

has directly or indirectly undergone *COPI*-mediated ubiquitination and degradation (Xiong et al., 2019). The regulation of jasmonate, auxin, trigonolactone, and gibberellin in the synthesis of flavonoids is relatively clear (LaFountain and Yuan, 2021). These plant hormones all involve a Skp1/Cullin/F-box (SCF) E3 ubiquitin ligase complex, and trigger the degradation of specific hormone receptors, negatively regulating anthocyanin accumulation by interacting with the MBW complex. In this study, *PB.4604.1*, *PB.24525.1*, and *PB.19524.2* were related to jasmonate, auxin, and brassinosteroid signaling pathways, respectively. It is reported that gibberellin enhances and inhibits the biosynthesis of anthocyanins in different species and tissues through different types of DELLA protein (LaFountain and Yuan, 2021). As a gibberellin signaling repressor in *Arabidopsis*, DELLA proteins bind to MYBL2 and JAZ proteins, resulting in higher MBW complex activities (Xie et al., 2016). *PB.13554.1* and *PB.3980.1* might participate in the gibberellin signaling pathway. In addition, some transcription factors, such as *AP2/ERF* (Zhao et al., 2021), *HD-ZIP* (Wang et al., 2020), *MYB*-related (Nguyen and Lee, 2016), and *NAC* (LaFountain and Yuan, 2021), were also found to play a role in transcriptional activation or inhibition in flavonoid synthesis. These transcription factors were also identified in our study. In this sense, this study will help to elaborate the molecular mechanism of color transformation over the course of blueberry development.

Data availability statement

The datasets presented in this study can be found the NCBI Sequence Read Archive (SRA) platform (<http://www.ncbi.nlm.nih.gov/sra/>) under accession numbers from SRX14097933 to SRX14097959 of Bioproject PRJNA804018.

Author contributions

YL, GX, and JL (7th author) conceived and designed the experiment. JL (7th author) detected the content of flavonoids. YL, HL, JL (4th author), SW, and SB analyzed the data. YL wrote the manuscript and all authors reviewed the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the Study of Quality Improvement and Advantage Brand Creation of Yunnan Characteristic Fruits (2019ZG002-1-02), Agricultural Science and Technology Innovation Program of Chinese Academy of Agricultural Sciences (CAAS-ZDRW202011), and the scholarship from China Scholarships Council (NO. 202103250038).

References

- Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402. doi: 10.1093/nar/25.17.3389
- Appelhagen, I., Lu, G., Huop, G., Schmelzer, E., Weisshaar, B., and Sagasser, M. (2011). TRANSPARENT TESTA1 interacts with R2R3-MYB factors and affects early and late steps of flavonoid biosynthesis in the endothelium of *Arabidopsis thaliana* seeds. *Plant J.* 67, 406–419. doi: 10.1111/j.1365-313X.2011.04603.x
- Apweiler, R. (2004). UniProt: the universal protein knowledgebase. *Nucleic Acids Res.* 32, D115–D119. doi: 10.1093/nar/gkh131
- Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., et al. (2000). Gene ontology: tool for the unification of biology. *Nat. Genet.* 25, 25–29. doi: 10.1038/75556
- Bian, Y., Ballington, J., Raja, A., Brouwer, C., Reid, R., Burke, M., et al. (2014). Patterns of simple sequence repeats in cultivated blueberries (*Vaccinium* section cyanococcus spp.) and their use in revealing genetic diversity and population structure. *Mol. Breed.* 34 (2), 675–689. doi: 10.1007/s11032-014-0066-7
- Brevis, P. A., Bassil, N. V., Ballington, J. R., and Hancock, J. F. (2008). Impact of wide hybridization on highbush blueberry breeding. *J. Am. Soc. Hortic. Sci.* 133, 427–437. doi: 10.21273/JASHS.133.3.427
- Cappai, F., Benevenuto, J., Ferrão, L., and Munoz, P. (2018). Molecular and genetic bases of fruit firmness variation in blueberry—a review. *Agronomy* 8, 174. doi: 10.3390/agronomy8090174
- Castillejo, C., Waurich, V., Wagner, H., Ramos, R., Oiza, N., Muñoz, P., et al. (2020). Allelic variation of *MYB10* is the major force controlling natural variation in skin and flesh color in strawberry (*Fragaria* spp.) fruit. *Plant Cell* 32, 3723–3749. doi: 10.1105/tpc.20.00474
- Chai, Z., Herrera-Balandrano, D. D., Yu, H., Beta, T., Zeng, Q., Zhang, X., et al. (2021). A comparative analysis on the anthocyanin composition of 74 blueberry cultivars from China. *J. Food Compos. Anal.* 102, 104051. doi: 10.1016/j.jfca.2021.104051
- Chao, Y., Yuan, J., Li, S., Jia, S., Han, L., and Xu, L. (2018). Analysis of transcripts and splice isoforms in red clover (*Trifolium pratense* L.) by single-molecule long-read sequencing. *BMC Plant Biol.* 18, 300. doi: 10.1186/s12870-018-1534-8
- Cho, K., Cho, K., Sohn, H., Ha, I. J., Hong, S. Y., Lee, H., et al. (2016). Network analysis of the metabolome and transcriptome reveals novel regulation of potato pigmentation. *J. Exp. Bot.* 67, 1519–1533. doi: 10.1093/jxb/erv549
- Cho, K. H., Park, S. J., Kim, S. J., Kim, S. H., Lee, H. C., Kim, M. Y., et al. (2017). Blueberry cultivar identification using random amplified polymorphic DNA and

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2023.1082245/full#supplementary-material>

sequence-characterized amplified region markers. *HortScience* 52, 1483–1489. doi: 10.21273/HORTSCI12105-17

Colle, M., Leisner, C. P., Wai, C. M., Ou, S., Bird, K. A., Wang, J., et al. (2019). Haplotype-phased genome and evolution of phytonutrient pathways of tetraploid blueberry. *GigaScience* 8, 1–15. doi: 10.1093/gigascience/giz012

Deng, Y. Y., Li, J. Q., Wu, S. F., Zhu, Y. P., Chen, Y. W., and He, F. C. (2006). Integrated NR database in protein annotation system and its localization. *Comput. Eng.* 32, 71–74. doi: 10.1109/infocom.2006.241

Dong, T., Han, R., Yu, J., Zhu, M., Zhang, Y., Gong, Y., et al. (2019). Anthocyanins accumulation and molecular analysis of correlated genes by metabolome and transcriptome in green and purple asparagus (*Asparagus officinalis*, L.). *Food Chem.* 271, 18–28. doi: 10.1016/j.foodchem.2018.07.120

Ge, Y., Li, X., Li, C., Tang, Q., Duan, B., Cheng, Y., et al. (2019). Effect of sodium nitroprusside on antioxidative enzymes and the phenylpropanoid pathway in blueberry fruit. *Food Chem.* 295, 607–612. doi: 10.1016/j.foodchem.2019.05.160

Günther, C. S., Dare, A. P., McGhie, T. K., Deng, C., Lafferty, D. J., Plunkett, B. J., et al. (2020). Spatiotemporal modulation of flavonoid metabolism in blueberries. *Front. Plant Sci.* 11. doi: 10.3389/fpls.2020.00545

Guo, X., Li, S., Wang, D., Huang, Z., Sarwar, N., Mubeen, K., et al. (2021). Effects of water and fertilizer coupling on the physiological characteristics and growth of rabbiteye blueberry. *PLoS One* 16, e254013. doi: 10.1371/journal.pone.0254013

Hassani, D., Fu, X., Shen, Q., Khalid, M., Rose, J. K., Tang, K., et al. (2020). Parallel transcriptional regulation of artemisinin and flavonoid biosynthesis. *Trends Plant Sci.* 25, 466–476. doi: 10.1016/j.tplants.2020.01.001

He, F., Mu, L., Yan, G. L., Liang, N. N., Pan, Q. H., Wang, J., et al. (2010). Biosynthesis of anthocyanins and their regulation in colored grapes. *Molecules* 15, 9057–9091. doi: 10.3390/molecules15129057

Hoang, N. V., Furtado, A., Mason, P. J., Marquardt, A., Kasirajan, L., Thirugnanasambandam, P. P., et al. (2017). A survey of the complex transcriptome from the highly polyploid sugarcane genome using full-length isoform sequencing and *de novo* assembly from short read sequencing. *BMC Genomics* 18, 395. doi: 10.1186/s12864-017-3757-8

Hosseini, F. S., and Beta, T. (2007). Saskatoon And wild blueberries have higher anthocyanin contents than other manitoba berries. *J. Agric. Food Chem.* 55, 10832–10838. doi: 10.1021/jf072529m

- Hou, Y., Zhai, L., Li, X., Xue, Y., Wang, J., Yang, P., et al. (2017). Comparative analysis of fruit ripening-related miRNAs and their targets in blueberry using small RNA and degradome sequencing. *Int. J. Mol. Sci.* 18, 2767. doi: 10.3390/ijms18122767
- Huerta-Cepas, J., Szklarczyk, D., Forslund, K., Cook, H., Heller, D., Walter, M. C., et al. (2016). EggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res.* 44, D286–D293. doi: 10.1093/nar/gkv1248
- Jaegle, B., Uroic, M. K., Holtkotte, X., Lucas, C., Termath, A. O., Schmalz, H. G., et al. (2016). A fast and simple LC-MS-based characterization of the flavonoid biosynthesis pathway for few seed(ling)s. *BMC Plant Biol.* 16, 190. doi: 10.1186/s12870-016-0880-7
- Karppinen, K., Lafferty, D. J., Albert, N. W., Mikkola, N., McGhie, T., Allan, A. C., et al. (2021). MYBA and MYBA transcription factors co-regulate anthocyanin biosynthesis in blue-coloured berries. *New Phytol.* 232, 1350–1367. doi: 10.1111/nph.17669
- Koonin, E. V., Fedorova, N. D., Jackson, J. D., Jacobs, A. R., Krylov, D. M., Makarova, K. S., et al. (2004). A comprehensive evolutionary classification of proteins encoded in complete eukaryotic genomes. *Genome Biol.* 5, R7. doi: 10.1186/gb-2004-5-2-r7
- LaFountain, A. M., and Yuan, Y. W. (2021). Repressors of anthocyanin biosynthesis. *New Phytol.* 231, 933–949. doi: 10.1111/nph.17397
- Lai, B., Li, X. J., Hu, B., Qin, Y. H., Huang, X. M., Wang, H. C., et al. (2014). *LcMYB1* is a key determinant of differential anthocyanin accumulation among genotypes, tissues, developmental phases and ABA and light stimuli in *Litchi chinensis*. *PLoS One* 9, e86293. doi: 10.1371/journal.pone.0086293
- Leng, N., Dawson, J. A., Thomson, J. A., Ruotti, V., Rissman, A. I., Smits, B. M., et al. (2013). EBSeg: an empirical bayes hierarchical model for inference in RNA-seq experiments. *Bioinformatics* 29, 1035–1043. doi: 10.1093/bioinformatics/btt087
- Li, B., and Dewey, C. N. (2011). RSEM: accurate transcript quantification from RNA-seq data with or without a reference genome. *BMC Bioinf.* 12, 323. doi: 10.1186/1471-2105-12-323
- Li, J., He, Y., Zhou, L., Liu, Y., Jiang, M., Ren, L., et al. (2018). Transcriptome profiling of genes related to light-induced anthocyanin biosynthesis in eggplant (*Solanum melongena* L.) before purple color becomes evident. *BMC Genomics* 19, 201. doi: 10.1186/s12864-018-4587-z
- Li, X., Jin, L., Pan, X., Yang, L., and Guo, W. (2019). Proteins expression and metabolite profile insight into phenolic biosynthesis during highbush blueberry fruit maturation. *Food Chem.* 290, 216–228. doi: 10.1016/j.foodchem.2019.03.115
- Li, X., Li, C., Sun, J., and Jackson, A. (2020). Dynamic changes of enzymes involved in sugar and organic acid level modification during blueberry fruit maturation. *Food Chem.* 309, 125617. doi: 10.1016/j.foodchem.2019.125617
- Li, Y., Shan, X., Gao, R., Han, T., Zhang, J., et al. (2020). MYB repressors and MBW activation complex collaborate to fine-tune flower coloration in *Freesia hybrida*. *Commun. Biol.* 3, 396. doi: 10.1038/s42003-020-01134-6
- Li, A., Zhang, J., and Zhou, Z. (2014). PLEK: a tool for predicting long non-coding RNAs and messenger RNAs based on an improved *k*-mer scheme. *BMC Bioinform.* 15, 311. doi: 10.1186/1471-2105-15-311
- Li, G., Zhao, J., Qin, B., Yin, Y., An, W., Mu, Z., et al. (2019). ABA mediates development-dependent anthocyanin biosynthesis and fruit coloration in *Lycium* plants. *BMC Plant Biol.* 19, 317. doi: 10.1186/s12870-019-1931-7
- Lin, Y., Wang, Y., Li, B., Tan, H., Li, D., Li, L., et al. (2018). Comparative transcriptome analysis of genes involved in anthocyanin synthesis in blueberry. *Plant Physiol. Bioch.* 127, 561–572. doi: 10.1016/j.plaphy.2018.04.034
- Liu, Y., Lv, J., Liu, Z., Wang, J., Yang, B., Chen, W., et al. (2020). Integrative analysis of metabolome and transcriptome reveals the mechanism of color formation in pepper fruit (*Capsicum annuum* L.). *Food Chem.* 306, 125629. doi: 10.1016/j.foodchem.2019.125629
- Liu, Y., Tikunov, Y., Schouten, R. E., Marcelis, L. F., Visser, R. G., and Bovy, A. (2018). Anthocyanin biosynthesis and degradation mechanisms in *Solanaceae* vegetables: a review. *Front. Chem.* 6, doi: 10.3389/fchem.2018.00052
- Mao, X., Cai, T., Olyarchuk, J. G., and Wei, L. (2005). Automated genome annotation and pathway identification using the KEGG orthology (KO) as a controlled vocabulary. *Bioinformatics* 21, 3787–3793. doi: 10.1093/bioinformatics/bti430
- Morita, M., Naito, Y., Yoshikawa, T., and Niki, E. (2017). Antioxidant capacity of blueberry extracts: peroxyl radical scavenging and inhibition of plasma lipid oxidation induced by multiple oxidants. *J. Berry Res.* 7, 1–9. doi: 10.3233/JBR-170152
- Munagala, R., Aqil, F., Jeyabalan, J., Agrawal, A. K., Mudd, A. M., Kyakulaga, A. H., et al. (2017). Exosomal formulation of anthocyanidins against multiple cancer types. *Cancer Lett.* 393, 94–102. doi: 10.1016/j.canlet.2017.02.004
- Naik, J., Misra, P., Trivedi, P. K., and Pandey, A. (2022). Molecular components associated with the regulation of flavonoid biosynthesis. *Plant Sci.* 317, 111196. doi: 10.1016/j.plantsci.2022.111196
- Nguyen, N. H., and Lee, H. (2016). MYB-related transcription factors function as regulators of the circadian clock and anthocyanin biosynthesis in *Arabidopsis*. *Plant Signal. Behav.* 11, e1139278. doi: 10.1080/15592324.2016.1139278
- Rodriguez-Mateos, A., Cifuentes-Gomez, T., Tabatabaee, S., Lecras, C., and Spencer, J. P. E. (2012). Procyanidin, anthocyanin, and chlorogenic acid contents of highbush and lowbush blueberries. *J. Agric. Food Chem.* 60, 5772–5778. doi: 10.1021/jf203812w
- Rowland, L. J., Alkharouf, N., Darwish, O., Ogdan, E. L., Polashock, J. J., Bassil, N. V., et al. (2012). Generation and analysis of blueberry transcriptome sequences from leaves, developing fruit, and flower buds from cold acclimation through deacclimation. *BMC Plant Biol.* 12, 46. doi: 10.1186/1471-2229-12-46
- Schwinn, K. E., Ngo, H., Kenel, F., Brummell, D. A., Albert, N. W., McCallum, J. A., et al. (2016). The onion (*Allium cepa* L.) *R2R3-MYB* gene *MYB1* regulates anthocyanin biosynthesis. *Front. Plant Sci.* 7, doi: 10.3389/fpls.2016.01865
- Spinardi, A., Cola, G., Gardana, C. S., and Mignani, I. (2019). Variation of anthocyanin content and profile throughout fruit development and ripening of highbush blueberry cultivars grown at two different altitudes. *Front. Plant Sci.* 10, doi: 10.3389/fpls.2019.01045
- Sun, B., Zhou, X., Chen, C., Chen, C., Chen, K., Chen, M., et al. (2020). Coexpression network analysis reveals an MYB transcriptional activator involved in capsaicinoid biosynthesis in hot peppers. *Hortic. Res.* 7, 162. doi: 10.1038/s41438-020-00381-2
- Tang, Q., Chi, F., Liu, H., Zhang, H., and Song, Y. (2021). Single-molecule real-time and illumina sequencing to analyze transcriptional regulation of flavonoid synthesis in blueberry. *Front. Plant Sci.* 12, doi: 10.3389/fpls.2021.754325
- Tatusov, R. L., Galperin, M. Y., Natale, D. A., and Koonin, E. V. (2000). The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res.* 28, 33–36. doi: 10.1093/nar/28.1.33
- Timmers, M. A., Grace, M. H., Yousef, G. G., and Lila, M. A. (2017). Inter- and intra-seasonal changes in anthocyanin accumulation and global metabolite profiling of six blueberry genotypes. *J. Food Compos. Anal.* 59, 105–110. doi: 10.1016/j.jfca.2017.02.019
- Tohge, T., de Souza, L. P., and Fernie, A. R. (2017). Current understanding of the pathways of flavonoid biosynthesis in model and crop plants. *J. Exp. Bot.* 68, 4013–4028. doi: 10.1093/jxb/erx177
- Vrhovsek, U., Masuero, D., Palmieri, L., and Mattivi, F. (2012). Identification and quantification of flavonol glycosides in cultivated blueberry cultivars. *J. Food Compos. Anal.* 25, 9–16. doi: 10.1016/j.jfca.2011.04.015
- Wang, H., Guo, X., Hu, X., Li, T., Fu, X., and Liu, R. H. (2017). Comparison of phytochemical profiles, antioxidant and cellular antioxidant activities of different varieties of blueberry (*Vaccinium* spp.). *Food Chem.* 217, 773–781. doi: 10.1016/j.foodchem.2016.09.002
- Wang, Z., Wang, S., Xiao, Y., Li, Z., Wu, M., Xie, X., et al. (2020). Functional characterization of a HD-ZIP IV transcription factor *NtHDG2* in regulating flavonols biosynthesis in *Nicotiana tabacum*. *Plant Physiol. Bioch.* 146, 259–268. doi: 10.1016/j.plaphy.2019.11.033
- Wu, Q., Wu, J., Li, S. S., Zhang, H. J., Feng, C. Y., Yin, D. D., et al. (2016). Transcriptome sequencing and metabolite analysis for revealing the blue flower formation in waterlily. *BMC Genomics* 17, 897. doi: 10.1186/s12864-016-3226-9
- Xiao, Y., Chu, L., Zhang, Y., Bian, Y., Xiao, J., and Xu, D. (2022). HY5: a pivotal regulator of light-dependent development in higher plants. *Front. Plant Sci.* 12, doi: 10.3389/fpls.2021.800989
- Xie, Y., Tan, H., and Ma, Z. (2016). DELLA proteins promote anthocyanin biosynthesis via sequestering MYB2L and JAZ suppressors of the MYB/bHLH/WD40 complex in *Arabidopsis thaliana*. *Mol. Plant* 9, 711–721. doi: 10.1016/j.molp.2016.01.014
- Xiong, C., Luo, D., Lin, A., Zhang, C., Shan, L., He, P., et al. (2019). A tomato b-box protein *S/BBX20* modulates carotenoid biosynthesis by directly activating *PHYTOENE SYNTHASE 1*, and is targeted for 26S proteasome-mediated degradation. *New Phytol.* 221, 279–294. doi: 10.1111/nph.15373
- Yang, L., Liu, L., Wang, Z., Zong, Y., Yu, L., Li, Y., et al. (2021). Comparative anatomical and transcriptomic insights into *Vaccinium corymbosum* flower bud and fruit throughout development. *BMC Plant Biol.* 21, 289. doi: 10.1186/s12870-021-03067-6
- Yi, D., Zhang, H., Lai, B., Liu, L., Pan, X., Ma, Z., et al. (2021). Integrative analysis of the coloring mechanism of red longan pericarp through metabolome and transcriptome analyses. *J. Agric. Food Chem.* 69, 1806–1815. doi: 10.1021/acs.jafc.0c05023
- Yuan, J., Li, H., Tao, W., Han, Q., Dong, H., Zhang, J., et al. (2020). An effective method for extracting anthocyanins from blueberry based on freeze-ultrasonic thawing technology. *Ultrason. Sonochem.* 68, 105192. doi: 10.1016/j.ulsonch.2020.105192
- Zang, Z., Chou, S., Tian, J., Lang, Y., Shen, Y., Ran, X., et al. (2021). Effect of whey protein isolate on the stability and antioxidant capacity of blueberry anthocyanins: a mechanistic and *in vitro* simulation study. *Food Chem.* 336, 127700. doi: 10.1016/j.foodchem.2020.127700
- Zhang, Y., Butelli, E., and Martin, C. (2014). Engineering anthocyanin biosynthesis in plants. *Curr. Opin. Plant Biol.* 19, 81–90. doi: 10.1016/j.climinews.2015.01.008
- Zhao, C., Liu, X., Gong, Q., Cao, J., Shen, W., Yin, X., et al. (2021). Three AP2/ERF family members modulate flavonoid synthesis by regulating type IV chalcone isomerase in citrus. *Plant Biotechnol. J.* 19, 671–688. doi: 10.1111/pbi.13494
- Zhao, S., Park, C. H., Yang, J., Yeo, H. J., Kim, T. J., Kim, J. K., et al. (2019). Molecular characterization of anthocyanin and betulinic acid biosynthesis in red and white mulberry fruits using high-throughput sequencing. *Food Chem.* 279, 364–372. doi: 10.1016/j.foodchem.2018.11.101
- Zheng, Y., Jiao, C., Sun, H., Rosli, H. G., Pombo, M. A., Zhang, P., et al. (2016). ITAK: a program for genome-wide prediction and classification of plant transcription factors, transcriptional regulators, and protein kinases. *Mol. Plant* 9, 1667–1670. doi: 10.1016/j.molp.2016.09.014

Zheng, C., Ma, J., Ma, C., Shen, S., Liu, Y., and Chen, L. (2019). Regulation of growth and flavonoid formation of tea plants (*Camellia sinensis*) by blue and green light. *J. Agric. Food Chem.* 67, 2408–2419. doi: 10.1021/acs.jafc.8b07050

Zhou, F., Wang, T., Zhang, B., and Zhao, H. (2018). Addition of sucrose during the blueberry heating process is good or bad? evaluating the changes of anthocyanins/

anthocyanidins and the anticancer ability in HepG-2 cells. *Food Res. Int.* 107, 509–517. doi: 10.1016/j.foodres.2018.02.071

Zifkin, M., Jin, A., Ozga, J. A., Zaharia, L. I., Scherthner, J. P., Gesell, A., et al. (2012). Gene expression and metabolite profiling of developing highbush blueberry fruit indicates transcriptional regulation of flavonoid metabolism and activation of abscisic acid metabolism. *Plant Physiol.* 158, 200–224. doi: 10.1104/pp.111.180950