Metabolome and transcriptome analyses reveal the colouring mechanism of red honeysuckle (*Lonicera japonica* Thunb.)

Xiaodong ZHANG¹a*, Caixia LI²b*, Zhanchao HAO³

¹Xuchang University, Food and Pharmacy College, 88 Bayi Road, Xuchang 461000, Henan, China; zxd95@xcu.edu.cn (corresponding author)
²Xuchang University, Key Laboratory of Micro-Nano Materials for Energy Storage and Conversion of Henan Province, Institute of Surface Micro and Nano Materials, College of Chemical and Materials Engineering, 88 Bayi Road, Xuchang 461000, Henan, China; 20201009@xcu.edu.cn (corresponding author)
³Yuzhou Traditional Chinese Medicine Standardization Center, Yuzhou 461670, Henan, China; yzszyb666@126.com

These authors contributed equally to the work

Abstract

Honeysuckle has been widely used as a medicinal herb and food additive in China for a long time. However, little is known about the pigment composition and colouring mechanism of red honeysuckle, which is a rare germplasm resource. This study aims to investigate the anthocyanin components and colouring mechanism of red honeysuckle, and to identify potential regulatory genes in the anthocyanin biosynthesis pathway. ‘Yujin 1’ and ‘Yujin 2’, with yellow-white and red flower buds, respectively, were selected for the study. Using a metabolomics method, we identified the anthocyanin components, while transcriptomics analysis was used to mine the structural and regulatory genes of the anthocyanin biosynthesis pathway. Additionally, protein-protein interaction analysis was employed to predict the regulation mechanism of anthocyanin biosynthesis. The results revealed that cyanidin-3,5-O-diglucoside, peonidin-3,5-O-diglucoside, and cyanidin-3-O-glucoside were the main pigment components of red honeysuckle. We also constructed a possible anthocyanin biosynthetic pathway and identified MYB and bHLH transcription factors that may play regulatory roles in this pathway. Furthermore, our findings suggest that bHLH23 may regulate anthocyanin biosynthesis by binding to the *DFR* gene promoter. These findings have significant implications for breeding new honeysuckle varieties and developing functional foods and medicines.

**Keywords:** anthocyanin; colouring mechanism; *Lonicera japonica*; metabolome; transcriptome

Introduction

*Lonicera japonica* Thunb. is a climbing plant that belongs to the genus *Lonicera* in the family Caprifoliaceae. It is commonly found in mountainous areas, sparse forests, rocky areas, along mountain roads, and village fences, and its original habitat is East Asia, mainly distributed in China, Japan, and Korea (Editorial Committee of Flora of China, 1988). The medicinal parts of honeysuckle are its dried flower buds and dried stems, which are commonly used to treat infections, fever, ulcers, swelling, and influenza (Shang *et al.*, 2011;
National Pharmacopoeia Committee, 2020). Recent research has shown that honeysuckle can inhibit influenza A virus and COVID-19 (Zhang et al., 2012; Zhou et al., 2015). Modern research indicates that the main active ingredients of honeysuckle are flavonoids, phenolic acids, and iridoids (Wang et al., 2020; Wang et al., 2023). In addition, honeysuckle is also widely used in cosmetics, food additives, beverages, and landscaping due to its medicinal properties and beautiful flowers (Cao et al., 2022; Ge et al., 2022).

Currently, research on honeysuckle mainly focuses on metabolite analysis (Wang et al., 2023), flower development (Yang et al., 2019), colour changes (Li et al., 2019; Xia et al., 2021), carotenoid metabolism (Pu et al., 2020; Yu et al., 2022), stress (Cai et al., 2022), genome sequencing (Xiao et al., 2021), and effective component biosynthesis pathway (Wang et al., 2020). However, there are few reports on the identification of pigment components, anthocyanin biosynthesis pathways, and regulatory mechanisms of red honeysuckle flower (Yuan et al., 2012; Li et al., 2019).

‘Yujin 2’ is a superior cultivar of honeysuckle, whose young branches, leaves, flower buds, and corolla are all red. This cultivar has strong resistance to cold and drought, high yield, early flowering, long flowering period, large flower buds, and a strong, non-bitter fragrance (Li et al., 2019; Li et al., 2022). In contrast, another superior honeysuckle cultivar, ‘Yujin 1’, has green young branches and leaves, and green-white flower buds and corolla. It also has characteristics such as cold and drought resistance, disease and pest resistance, long flowering period, large flower buds, and high yield. Therefore, ‘Yujin 2’ and ‘Yujin 1’, the two contrasting cultivars with different flower bud colours, are excellent materials for studying the pigment components, biosynthetic pathway and regulatory mechanism of honeysuckle anthocyanins. Previous studies have shown that the content of chlorophyll and carotenoids in the flower buds of ‘Yujin 2’ is low, and the high content of anthocyanins is the root cause of its red colour (Yuan et al., 2012; Li et al., 2019).

The anthocyanin biosynthesis and its regulatory mechanism in plants has been extensively studied (LaFountain and Yuan, 2021; Araguirang and Richter, 2022). The pathway originates from the cytosolic phenylpropanoid pathway, where phenylalanine is catalysed by phenylalanine ammonialyase (PAL), cinnamic acid 4-hydroxylase (C4H), and 4-coumarate CoA ligase (4CL) to form p-coumaroyl-CoA. Then, one molecule of p-coumaroyl-CoA and three molecules of malonyl-CoA are catalysed by chalcone synthase (CHS), chalcone isomerase (CHI), flavonoid 3',5'-hydroxylase (F3'5'H), and flavanone-3-hydroxylase (F3H), and flavonoid 3'-hydroxylase (F3'H) to form dihydroquercetin, dihydrokaempferol and dihydromyricetin, respectively. These are the early reaction of anthocyanin synthesis. Then, dihydromyricetin, dihydrokaempferol and dihydroquercetin are catalysed by dihydroflavonol 4-reductase (DFR), anthocyanin synthase (ANS) and anthocyanin: flavonoid glucosyltransferase (UFGT) to generate delphinidin 3-O-glucoside, pelargonidin 3-O-glucoside, and cyanidin 3-O-glucoside, respectively. These anthocyanins are further modified by methyltransferase (MT) and acyltransferase (AT) before being transported by glutathione S-transferase (GST) into the vacuole for storage and accumulation. Studies also have shown that almost all flowering plants regulate anthocyanin biosynthesis through a highly conserved MYB-bHLH-WDR (MBW) protein complex (LaFountain and Yuan, 2021; Liu et al., 2021).

With the rapid development of sequencing technology, metabolome and transcriptome have been widely used in the identification of secondary metabolites such as longan anthocyanins (Yi et al., 2021), honeysuckle pigments and aroma (Xia et al., 2021; Li et al., 2022), soapberry triterpenoid Saponin (Xu et al., 2022), and blueberry flavonoids (Wu et al., 2022), as well as the analysis of their biosynthetic pathways and regulatory mechanisms. In this study, we identified the anthocyanins present in flower buds of two varieties using ultra-performance liquid chromatography (UPLC) and electrospray ionization mass spectrometry/mass spectrometry (ESI-MS/MS). Then, we conducted transcriptome sequencing analysis to reveal the biosynthesis pathway and possible regulatory mechanism of anthocyanin in red honeysuckle. Our findings offer valuable insights into the biosynthesis and regulation mechanism of anthocyanins in red honeysuckle.
Materials and Methods

Plant materials

The fresh flower buds of *Lonicera japonica* Thunb. cultivar 'Yujin 1' (LjW, accession: SAMN33906972, SAMN33906973, and SAMN33906974) and 'Yujin 2' (LjR, accession: SAMN33906969, SAMN33906970, and SAMN33906971) at harvesting stage were collected from the Taihefulou Village, Huolong Town, Yuzhou City, Henan Province, China (Figure 1).

![Plant materials](image1)

**Figure 1.** Plant materials used for metabolome and transcriptome analysis. (A) and (B) were LjW and LjR samples collected on October 1, 2022, respectively.

Sample preparation and extraction

The sample was freeze-dried, ground into powder (30 Hz, 1.5 min) using ball mill (MM400, Retsch, Germany), and stored at -80 °C until needed. 50 mg powder was weighted using electronic balance (AS 60/220.R2, RADWAG, Poland) and extracted with 0.5 mL methanol/water/hydrochloric acid (500:500:1, V/V/V). Then the extract was vortexed for 5 min and ultrasonically using ultrasonic cleaner (KQ5200E, Shumei, China) for 5 min and centrifuged using centrifuge (5424R, Eppendorf, Germany) at 12,000 g under 4 °C for 3 min. The residue was re-extracted by repeating the above steps again under the same conditions. The supernatants were collected, and filtrated through a membrane filter (0.22 μm, Anpel) before LC-MS/MS analysis.

UPLC conditions

The sample extracts were analyzed using an UPLC-ESI-MS/MS system (UPLC, ExionLC™ AD, https://sciex.com.cn/; MS, Applied Biosystems 6500 Triple Quadrupole, https://sciex.com.cn/). The analytical conditions were as follows, UPLC: column, WatersACQUITY BEH C18 (1.7 μm, 2.1 mm × 100 mm); solvent system, water (0.1% formic acid): methanol (0.1% formic acid); gradient program, 95:5 (V/V) at 0 min, 50:50 (V/V) at 6 min, 5:95 (V/V) at 12 min, hold for 2 min, 95:5 (V/V) at 14 min; hold for 2 min; flow rate, 0.35 mL/min; temperature, 40 °C; injection volume, 2 μL.

ESI-MS/MS conditions

Linear ion trap (LIT) and triple quadrupole (QQQ) scans were acquired on a triple quadrupole-linear ion trap mass spectrometer (QTRAP), QTRAP 6500plus LC-MS/MS System, equipped with an ESI Turbo Ion-Spray interface, operating in positive ion mode and controlled by Analyst 1.6.3 software (Sciex). The ESI source operation parameters were as follows: ion source, ESI+; source temperature 550 °C; ion spray voltage (IS) 5500 V; curtain gas (CUR) was set at 35 psi, respectively. Anthocyanins were analyzed using scheduled multiple reaction monitoring (MRM). Data acquisitions were performed using Analyst v1.6.3 software (Sciex). Multiquant v3.0.3 software (Sciex) was used to quantify all metabolites. Mass spectrometer parameters including the declustering potentials (DP) and collision energies (CE) for individual MRM transitions were...
done with further DP and CE optimization. A specific set of MRM transitions were monitored for each period according to the metabolites eluted within this period.

**RNA extraction and qualification**

The total RNA of LjW and LjR was extracted using the RNAprep Pure Plant Kit (Tiangen, Beijing, China) according to the instructions provided by the manufacturer. RNA concentration and purity were measured using NanoDrop 2000 (Thermo Fisher Scientific, Wilmington, DE). RNA integrity was assessed using the RNA Nano 6000 Assay Kit of the Agilent Bioanalyzer 2100 system (Agilent Technologies, CA, USA).

**Library preparation and transcriptome sequencing**

A total amount of 1 μg RNA per sample was used as input material for the RNA sample preparations. Sequencing libraries were generated using Heif NGS Ultima Dual-mode mRNA Library Prep Kit for Illumina (Yeasen Biotechnology (Shanghai) Co., Ltd.) following manufacturer’s recommendations and index codes were added to attribute sequences to each sample. Briefly, mRNA was purified from total RNA using poly-T oligo-attached magnetic beads. First strand cDNA was synthesized and second strand cDNA synthesis was subsequently performed. Remaining overhangs were converted into blunt ends via exonuclease/polymerase activities. After adenylation of 3’ ends of DNA fragments, NEBNext Adaptor with hairpin loop structure were ligated to prepare for hybridization. The library fragments were purified with AMPure XP system (Beckman Coulter, Beverly, USA). Then 3 μl USER Enzyme (NEB, USA) was used with size-selected, adaptor-ligated cDNA at 37 °C for 15 min followed by 5 min at 95 °C before PCR. Then PCR was performed with Phusion High-Fidelity DNA polymerase, Universal PCR primers and Index (X) Primer. At last, PCR products were purified (AMPure XP system) and library quality was assessed on the Agilent Bioanalyzer 2100 system. The libraries were sequenced on an Illumina NovaSeq6000 platform to generate 150 bp paired-end reads, according to the manufacturer’s instructions.

**Quality control**

Raw reads of fastq format were firstly processed through in-house perl scripts. In this step, clean reads were obtained by removing reads containing adapter, reads containing poly-N and low-quality reads from raw data. At the same time, Q20, Q30, GC-content and sequence duplication level of the clean data were calculated. All the downstream analyses were based on clean data with high quality.

**Reads mapping to the reference genome**

The adaptor sequences and low-quality sequence reads were removed from the data sets. Raw sequences were transformed into clean reads after data processing. These clean reads were then mapped to the reference genome sequence. Only reads with a perfect match or one mismatch were further analysed and annotated based on the reference genome. Hisat2 v2.04 software was used to map with reference genome assembly ASM2146441v1.

**Novel transcripts Prediction**

Reference Annotation Based Transcript (RABT) assembly method was used to construct and identify both known and novel transcripts from Hisat2 alignment results using the StringTie v2.2.1 software.

**Gene functional annotation**

Gene function was annotated based on the following databases: Nr (NCBI non-redundant protein sequences, diamond v2.0.15, parameters: -k 100 -e -eval 1e-5 -f 5); eggNOR (A database of orthology relationships, functional annotation, and gene evolutionary histories, diamond v2.0.15); KOG/COG (Clusters of Orthologous Groups of proteins, diamond v2.0.15); Swiss-Prot (A manually annotated and
reviewed protein sequence database, diamond v2.0.15); KO (KEGG Ortholog database, diamond v2.0.15); Pfam (Protein family, hmmscan v3.3.2, parameters: --noali --cut_nc --acc --notextw); GO (Gene Ontology, InterProScan v5.34-73.0, parameters: -appl Pfam -goterms -iprlookup -pa -f xml -dp -t p).

**Quantification of gene expression levels**
Gene expression levels were estimated by fragments per kilobase of transcript per million fragments mapped (FPKM). The formula is shown as follow:

\[
\text{FPKM} = \frac{\text{cDNA Fragments}}{\text{Mapped Fragments (Millions)} \times \text{Transcript Length (kb)}}
\]

Pearson’s correlation coefficient \(r\) was used to assess the biological replicates’ correlation (Liu et al., 2018).

**Differential expression analysis**
Differential expression analysis of two groups was performed using the DESeq2 v1.30.1 software. DESeq2 provide statistical routines for determining differential expression in digital gene expression data using a model based on the negative binomial distribution. The resulting P values were adjusted using the Benjamini and Hochberg’s approach for controlling the false discovery rate. Genes with an adjusted P-value < 0.01 & Fold Change≥2 found by DESeq2 were assigned as differentially expressed.

**GO enrichment analysis**
Gene Ontology (GO) enrichment analysis of the differentially expressed genes (DEGs) was implemented by the clusterProfiler (Wu et al., 2021) packages based Wallenius non-central hyper-geometric distribution (Young et al., 2010), which can adjust for gene length bias in DEGs.

**KEGG pathway enrichment analysis**
We used KOBAS (Mao et al., 2005) database and clusterProfiler (Wu et al., 2021) software to test the statistical enrichment of DEGs in KEGG pathways.

**Construction of protein protein interaction (PPI)**
The sequences of the DEGs was blast (blastx) to the genome of a related species (the protein-protein interaction of which exists in the STRING database: http://string-db.org/) to get the predicted PPI of these DEGs. Then the PPI of these DEGs were visualized in Cytoscape (Shannon et al., 2003).

**Splicing event quantification**
Different alternative splicing (AS) events were detected and quantified using the rMATS v4.02 software. The number of AS events in each sample was estimated, separately.

**Transcription factors prediction**
The transcription factors (TFs) were predicted with iTAK v1.6 (Zheng et al., 2016).

**Results**

**Appearance characteristics of honeysuckle flower buds**
In general, most varieties of honeysuckle produce yellow-white or green-white flower buds, with reports of red honeysuckle is extremely rare. LjW, one of the more common varieties, produces yellowish-white or
greenish-white, densely villous flower buds (Figure 1A). In contrast, the flower buds of LjR are red and densely pubescent (Figure 1B).

**Determination of the main contents of anthocyanins and other flavonoids**

To determine the types and components of anthocyanins and other flavonoids in the honeysuckle flower buds of LjW and LjR, we used LC-MS/MS technology for quantitative analysis. Results showed that a total of 59 kinds of anthocyanins (including 10 kinds of cyanidin, 12 kinds of delphinidin, 10 kinds of malvidin, 8 kinds of pelargonidin, 11 kinds of peonidin, and 8 kinds of petunidin), and 5 kinds of flavonoids were identified (Table 1). In the flower bud of LjW, the main anthocyanin components were peonidin-3-O-(6-O-p-coumaroyl)-glucoside (9.21 μg/g) and pelargonidin-3-O-(6-O-p-coumaroyl)-glucoside (7.1193 μg/g), and the main flavonoid components were rutin (1131.05 μg/g), quercetin-3-O-glucoside (294.49 μg/g), and kaempferol-3-O-rutinoside (102.11 μg/g) (Table 1). In the flower bud of LjR, the main anthocyanin components were cyanidin-3,5-O-diglucoside (633.50 μg/g), peonidin-3,5-O-diglucoside (73.94 μg/g) and cyanidin-3-O-glucoside (3.75 μg/g), and the main flavonoid components were rutin (971.78 μg/g), quercetin-3-O-glucoside (344.48 μg/g), and kaempferol-3-O-rutinoside (77.29 μg/g) (Table 1). Therefore, the biosynthesis pathway of cyanidin will be studied in detail.

**Table 1.** Anthocyanins and other flavonoids detected in LjW and LjR flower buds

<table>
<thead>
<tr>
<th>Class</th>
<th>Compounds</th>
<th>Q1 (Da)</th>
<th>Q3 (Da)</th>
<th>Molecular weight</th>
<th>Ion mode</th>
<th>LjR (μg/g)</th>
<th>LjW (μg/g)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cyanidin</td>
<td>Cyanidin-3,5-O-diglucoside</td>
<td>611.2</td>
<td>287.1</td>
<td>611.1612</td>
<td>Positive</td>
<td>633.5035</td>
<td>0.6263</td>
</tr>
<tr>
<td></td>
<td>Cyanidin-3-O-glucoside</td>
<td>449.1</td>
<td>287.1</td>
<td>449.1084</td>
<td>Positive</td>
<td>3.7547</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>Cyanidin-3-O-rutinoside</td>
<td>595.37</td>
<td>287.1</td>
<td>595.1663</td>
<td>Positive</td>
<td>0.2263</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>Cyanidin-3-(6-O-p-cafeoyl)-glucoside</td>
<td>611.14</td>
<td>287.1</td>
<td>611.1401</td>
<td>Positive</td>
<td>0.0793</td>
<td>0.0201</td>
</tr>
<tr>
<td></td>
<td>Cyanidin-3-O-arabinoside</td>
<td>419.1</td>
<td>287.1</td>
<td>419.0978</td>
<td>Positive</td>
<td>0.0174</td>
<td>0.0129</td>
</tr>
<tr>
<td></td>
<td>Cyanidin-3-O-(6-O-malonyl-beta-D-glucoside)</td>
<td>535.1</td>
<td>287.1</td>
<td>535.1088</td>
<td>Positive</td>
<td>0.0079</td>
<td>0.0072</td>
</tr>
<tr>
<td></td>
<td>Cyanidin-3-O-sambubioside</td>
<td>581.1</td>
<td>287.1</td>
<td>581.1506</td>
<td>Positive</td>
<td>0.0037</td>
<td>0.0042</td>
</tr>
<tr>
<td></td>
<td>Cyanidin-3,5-O-(6-O-p-coumaroyl)-diglucoside</td>
<td>757.2</td>
<td>287.1</td>
<td>757.198</td>
<td>Positive</td>
<td>0.0024</td>
<td>0.0012</td>
</tr>
<tr>
<td></td>
<td>Cyanidin-3,5,3’-O-triglucoside</td>
<td>773.21</td>
<td>287.1</td>
<td>773.214</td>
<td>Positive</td>
<td>NA</td>
<td>0.0007</td>
</tr>
<tr>
<td></td>
<td>Cyanidin-3-O-rutinoside-5-O-glucoside</td>
<td>757.22</td>
<td>287.1</td>
<td>757.2191</td>
<td>Positive</td>
<td>NA</td>
<td>0.0014</td>
</tr>
<tr>
<td>Delphinidin</td>
<td>Delphinidin-3-O-glucoside</td>
<td>461.5</td>
<td>303.1</td>
<td>465.1033</td>
<td>Positive</td>
<td>0.5803</td>
<td>0.1966</td>
</tr>
<tr>
<td></td>
<td>Delphinidin-3-O-galactoside</td>
<td>461.5</td>
<td>303.1</td>
<td>465.1033</td>
<td>Positive</td>
<td>0.2212</td>
<td>0.0715</td>
</tr>
<tr>
<td></td>
<td>Delphinidin-3-O-rutinoside-5-O-glucoside</td>
<td>773.21</td>
<td>303.1</td>
<td>773.214</td>
<td>Positive</td>
<td>0.1329</td>
<td>0.0862</td>
</tr>
<tr>
<td></td>
<td>Delphinidin-3-O-sambubioside</td>
<td>627.15</td>
<td>303.1</td>
<td>627.1561</td>
<td>Positive</td>
<td>0.0889</td>
<td>0.0429</td>
</tr>
<tr>
<td></td>
<td>Delphinidin-3-O-sambubioside</td>
<td>611.1</td>
<td>303.1</td>
<td>611.1612</td>
<td>Positive</td>
<td>0.0426</td>
<td>0.0252</td>
</tr>
<tr>
<td>Malvidin</td>
<td>Malvidin-3-O-galactoside</td>
<td>461.5</td>
<td>303.1</td>
<td>465.1033</td>
<td>Positive</td>
<td>0.2212</td>
<td>0.0715</td>
</tr>
<tr>
<td></td>
<td>Malvidin-3-O-arabinoside</td>
<td>461.5</td>
<td>303.1</td>
<td>465.1033</td>
<td>Positive</td>
<td>0.2212</td>
<td>0.0715</td>
</tr>
<tr>
<td></td>
<td>Malvidin-3-O-rutinoside</td>
<td>639.06</td>
<td>331.1</td>
<td>639.1925</td>
<td>Positive</td>
<td>0.0124</td>
<td>0.0386</td>
</tr>
<tr>
<td></td>
<td>Malvidin-3-O-(6-O-p-cafeoyl)-glucoside</td>
<td>639.17</td>
<td>331.1</td>
<td>639.1714</td>
<td>Positive</td>
<td>NA</td>
<td>0.0343</td>
</tr>
<tr>
<td></td>
<td>Malvidin-3-O-(6-O-acetylglucoside)-5-glucoside</td>
<td>697.2</td>
<td>331.1</td>
<td>697.198</td>
<td>Positive</td>
<td>NA</td>
<td>0.0263</td>
</tr>
<tr>
<td></td>
<td>Malvidin-3-O-sambubioside</td>
<td>625.19</td>
<td>331.1</td>
<td>625.1769</td>
<td>Positive</td>
<td>0.0003</td>
<td>0.0113</td>
</tr>
<tr>
<td></td>
<td>Malvidin-3-O-glucoside</td>
<td>655.4</td>
<td>331.1</td>
<td>655.1874</td>
<td>Positive</td>
<td>0.0077</td>
<td>0.0592</td>
</tr>
<tr>
<td></td>
<td>Malvidin-3-O-glucoside</td>
<td>463.3</td>
<td>331.06</td>
<td>463.124</td>
<td>Positive</td>
<td>NA</td>
<td>0.0437</td>
</tr>
<tr>
<td>Pelargonidin</td>
<td>Pelargonidin-3-O-(6-O-p-cafeoyl)-glucoside</td>
<td>639.17</td>
<td>331.1</td>
<td>639.1714</td>
<td>Positive</td>
<td>NA</td>
<td>0.0343</td>
</tr>
<tr>
<td></td>
<td>Pelargonidin-3-O-(6-O-acetylglucoside)-5-glucoside</td>
<td>697.2</td>
<td>331.1</td>
<td>697.198</td>
<td>Positive</td>
<td>NA</td>
<td>0.0263</td>
</tr>
<tr>
<td></td>
<td>Pelargonidin-3-O-sambubioside</td>
<td>787.4</td>
<td>331.1</td>
<td>787.2297</td>
<td>Positive</td>
<td>0.0000</td>
<td>0.0113</td>
</tr>
<tr>
<td></td>
<td>Pelargonidin-3-O-galactoside</td>
<td>787.4</td>
<td>331.1</td>
<td>787.2297</td>
<td>Positive</td>
<td>0.0000</td>
<td>0.0113</td>
</tr>
<tr>
<td></td>
<td>Pelargonidin-3-O-sambubioside</td>
<td>801.22</td>
<td>331.1</td>
<td>801.2242</td>
<td>Positive</td>
<td>0.0014</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>Pelargonidin-3-O-galactoside</td>
<td>787.4</td>
<td>331.1</td>
<td>787.2297</td>
<td>Positive</td>
<td>0.0000</td>
<td>0.0113</td>
</tr>
<tr>
<td></td>
<td>Pelargonidin-3-O-sambubioside</td>
<td>787.4</td>
<td>331.1</td>
<td>787.2297</td>
<td>Positive</td>
<td>0.0000</td>
<td>0.0113</td>
</tr>
<tr>
<td></td>
<td>Pelargonidin-3-O-glucoside</td>
<td>493.5</td>
<td>331.1</td>
<td>493.1346</td>
<td>Positive</td>
<td>NA</td>
<td>0.0060</td>
</tr>
<tr>
<td></td>
<td>Pelargonidin-3-O-(6-O-p-cafeoyl)-glucoside</td>
<td>579.15</td>
<td>271.1</td>
<td>579.1503</td>
<td>Positive</td>
<td>NA</td>
<td>7.1193</td>
</tr>
<tr>
<td></td>
<td>Pelargonidin-3-O-(6-O-acetylglucoside)-5-glucoside</td>
<td>565.2</td>
<td>271.1</td>
<td>565.1557</td>
<td>Positive</td>
<td>0.0030</td>
<td>0.0223</td>
</tr>
<tr>
<td></td>
<td>Pelargonidin-3-O-sambubioside</td>
<td>559.14</td>
<td>271.1</td>
<td>559.1663</td>
<td>Positive</td>
<td>0.0008</td>
<td>0.0021</td>
</tr>
<tr>
<td></td>
<td>Pelargonidin-3-O-sambubioside</td>
<td>757.22</td>
<td>271.1</td>
<td>757.2191</td>
<td>Positive</td>
<td>0.0008</td>
<td>0.0008</td>
</tr>
</tbody>
</table>

Zhang X et al. (2023). Not Bot Horti Agrobo 51(3):13150
accession number: PRJNA948551.

5.80 to 6.70 Gb (Table 2), which have been deposited in the Short Reads Archive (SRA) database under public databases including Nr, eggNOR, GO, Pfam, Swiss-Prot, KOG, and KEGG. A total of 35,687 assembled sequences were mapped to the reference genome (Table S1). These results suggested that our data was reliable. We used the Stringtie software (Pertea et al., 2015) to assemble the transcripts, and unigenes that were detected in the sequencing results but were absent in the reference genome were classified as novel genes. Our analysis identified a total of 7,609 novel genes, with 6,313 genes (82.97% of novel genes) detected in LjW, and 7,006 genes (92.08% of novel genes) detected in LjR.

To identify key candidate genes of anthocyanin biosynthesis pathway, RNA sequencing was performed using the flower buds LjW (LjW1, LjW2, and LjW3) and LjR (LjR1, LjR2, and LjR3). Six cDNA libraries were sequenced, and after data filtering, each sample generated over 19,668,207 clean reads, with Q20 and Q30 values greater than 97.45% and 92.90%, separately. The clean reads obtained from each sample ranged from 5.80 to 6.70 Gb (Table 2), which have been deposited in the Short Reads Archive (SRA) database under accession number: PRJNA948551.

Table 2. Quality assessment of the sequencing data

<table>
<thead>
<tr>
<th>Sample</th>
<th>Clean reads</th>
<th>Clean bases</th>
<th>GC (%)</th>
<th>Q20 (%)</th>
<th>Q30 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LjR1</td>
<td>22,430,399</td>
<td>6,715,212,210</td>
<td>44.36%</td>
<td>97.68</td>
<td>93.42%</td>
</tr>
<tr>
<td>LjR2</td>
<td>21,382,613</td>
<td>6,401,061,708</td>
<td>44.57%</td>
<td>97.70</td>
<td>93.47%</td>
</tr>
<tr>
<td>LjR3</td>
<td>21,026,606</td>
<td>6,293,408,696</td>
<td>44.38%</td>
<td>97.72</td>
<td>93.49%</td>
</tr>
<tr>
<td>LjW1</td>
<td>21,879,624</td>
<td>6,548,751,990</td>
<td>44.51%</td>
<td>97.59</td>
<td>93.23%</td>
</tr>
<tr>
<td>LjW2</td>
<td>20,531,504</td>
<td>6,146,687,976</td>
<td>44.33%</td>
<td>97.59</td>
<td>93.15%</td>
</tr>
<tr>
<td>LjW3</td>
<td>19,668,207</td>
<td>5,888,194,678</td>
<td>44.52%</td>
<td>97.45</td>
<td>92.90%</td>
</tr>
</tbody>
</table>

All clean reads were aligned to the reference genome of *L. japonica*, and a majority of the LjR (92.67-94.13%) and LjW (95.90-96.32%) sequences were mapped to the reference genome (Table S1). These results suggested that our data was reliable. We used the Stringtie software (Pertea et al., 2015) to assemble the transcripts, and unigenes that were detected in the sequencing results but were absent in the reference genome were classified as novel genes. Our analysis identified a total of 7,609 novel genes, with 6,313 genes (82.97% of novel genes) detected in LjW, and 7,006 genes (92.08% of novel genes) detected in LjR.

To annotate the assembled genes, we utilized BLASTX with an E-value threshold of 10^-5; against several public databases including Nr, eggNOR, GO, Pfam, Swiss-Prot, KOG, and KEGG. A total of 35,687 assembled unigenes were successfully annotated (Table S2). Based on sequence similarity, 29,296 (32.40%) unigenes were
classified into three groups (biological process, cellular component and molecular function) using GO term analysis (Figure S1). The biological processes were predominantly associated with "cellular process (12835, 43.81%)" and "metabolic process (12069, 41.20%)". The cellular components were primarily related to "cellular anatomical entity (15195, 51.87%)". The molecular functions were mainly classified into "binding (15320, 52.29%)" and "catalytic activity (14108, 48.16%)". KEGG enrichment analysis was utilized to identify the functional pathways of the unigenes. Of the total unigenes, 24,501 (59.00%) unigenes were significantly enriched in 137 KEGG pathways, with "plant-pathogen interaction (1613, 6.58%)" being the largest group, followed by "plant hormone signal transduction (878, 3.58%)", "protein processing in endoplasmic reticulum (609, 2.49%)", "MAPK signaling pathway-plant (594, 2.42%)", "starch and sucrose metabolism (551, 2.25%)" and "phenylpropanoid biosynthesis (534, 2.18%)" (Figure S2).

**Analysis of gene expression**

The correlation of gene expression levels among samples is a crucial indicator for testing the repeatability of biological experimental operations, evaluating the reliability of DEGs, and assisting in the screening of abnormal samples. To assess the biological replicates' correlation, we employed Pearson's correlation coefficient r as the evaluation index (Liu et al., 2018). In this study, we observed that the correlation coefficient between the LjR samples was greater than 0.93, while it was greater than 0.82 for the LjW samples, indicating stronger correlation among duplicate samples (Figure 2).

![Figure 2](image)

**Figure 2.** Correlation heat map of expression quantity of two samples. The values on each colour block on the heatmap represent the correlation between the two samples on the horizontal and vertical axis corresponding to the colour block.

The higher the value, the stronger the correlation.

Principal component analysis (PCA) can reduce multiple variables into a few independent variables (i.e. principal components). In transcriptome analysis, the gene expressions are reduced to find the rule of sample distribution from complex data, allowing for the evaluation of sample dispersion. The closer the points in the PCA diagram, the more similar the composition. As shown in Figure 3, PCA results showed that the three samples of LjW had similar compositions, so did the samples of LjR.
Figure 3. The results of PCA analysis. The coordinates indicate different principal components, while the corresponding percentage values represent the contribution of these principal components to the differences observed within the samples. Each point on the graph represents a sample, with those belonging to the LjR and LjW groups denoted by red squares and blue dots, respectively.

Identification and analysis of DEGs

Transcriptomic analyses were performed on honeysuckles to identify the key DEGs between the W (LjW) and R (LjR) groups. A total of 4,436 DEGs were obtained, consisting of 1,481 upregulated genes and 2,955 downregulated genes in W vs. R. The DEGs were identified using a threshold of |Fold Change| ≥2 and a false discovery rate (FDR) of less than 0.01 (Figure 4).

Figure 4. Histogram of DEGs statistics. The x-axis indicates the various sets of DEGs, with blue representing all DEGs, orange representing upregulated genes, and green representing downregulated genes. The y-axis indicates the number of DEGs in each set.

To investigate potential functional relationships between DEGs, hierarchical clustering analysis was performed on all screened DEGs. Genes with similar expression patterns are often associated with similar functions. The analysis aimed to display expression differences among the genes in different groups and identify novel functional genes. As shown in Figure 5, the upregulated genes in LjR were clustered in the upper right region of the heatmap, indicating a close association with the red phenotype.
Enrichment analysis of DEGs

To gain insights into the potential functions of the DEGs, we performed differential GO clustering analysis focusing on three major categories: cellular component, biological process, and molecular function. Our analysis revealed that 1329 unigenes were associated with the metabolic process in the biological process category, while 1445 unigenes showed catalytic activity in the molecular function category (Figure 6).

In total, we annotated 2612 genes for 125 KEGG metabolic pathways in the two cultivars. Among them, 810 unigenes (31.01%) were found to be involved in the metabolic pathway. We identified 8 significantly enriched metabolic pathways with a q value < 0.05, including “photosynthesis-antenna proteins”, “starch and sucrose metabolism”, “ascorbate and aldarate metabolism”, “cyanobacteria acid metabolism”, “photosynthesis”, “flavonoid biosynthesis”, “phenylpropanoid biosynthesis”, and “flavone and flavonol biosynthesis”. Notably, “flavonoid biosynthesis (ko00941, 36 DEGs)” and “flavone and flavonol biosynthesis (ko00944, 9 DEGs)” showed significant enhancement (Figure 7).

Figure 5. Cluster diagram of DEGs. The x-axis indicates the sample name, while the y-axis indicates the clustering result of the DEGs. Each column in the figure represents a different sample, and each row represents a different gene. The colour scheme represents the level of gene expression in log10(FPKM+1) for each sample.
**Figure 6.** Histogram of GO enrichment of DEGs. The y-axis represents the number of genes annotated in each GO entry, while the x-axis indicates the different GO entries. The colour of each column indicates the q value of the hypergeometric test for GO enrichment.

**Figure 7.** Bubble diagram of KEGG enrichment of DEGs. Each circle represents a KEGG pathway. The pathway name is on the ordinate, and the abscissa indicates the rich factor, representing the ratio of the proportion of genes annotated to a pathway in the differential gene to the proportion of genes annotated to the pathway in all genes. The larger the enrichment factor, the more significant the enrichment level of DEGs in this pathway. The colour of the circle represents the q-value, which is the p-value corrected by multiple hypothesis tests. The smaller the q-value, the more reliable the enrichment significance of DEGs in the pathway. The size of the circle indicates the number of genes enriched in the pathway, with a larger circle indicating more genes in the pathway. The circle’s position in the upper right corner indicates its higher reference value.
Analysis of structural genes involved in anthocyanin biosynthesis

We constructed an anthocyanin biosynthesis pathway in honeysuckle based on the pathway reported in plants (Khan et al., 2022) (Figure 8). The pathway comprises of structural genes, such as PAL, C4H, and 4CL, which are involved in phenylpropanoid biosynthesis, and CHS, CHI, F3H, F3’H, F3’5’H, DFR, ANS, UFGT, MT, and AT, which are involved in anthocyanin biosynthesis. Multiple copies of these genes were found, and some exhibited significantly higher expression levels in LjR than in LjW. Specifically, the expression levels of 4CL1 and 4CL4, CHS2, CHS3, CHS4, F3H1, F3’H, DFR1, ANS2, ANS3, UFGT4 and UFGT8 genes were much higher in LjR flower buds than in LjW (Figure 8). Enzyme genes at the branching point, FLS4 and ANR4, also showed significantly higher expression levels in LjR than in LjW (Figure 8).

Figure 8. Proposed pathway of anthocyanin biosynthesis and structural gene expression in honeysuckle
(A) The putative biosynthetic pathway of anthocyanin. The arrows denote the deduced reactions of anthocyanin. (B) Heatmap based on expression level of genes involved in anthocyanin biosynthesis between LjR and LjW. The expression level was calculated as FPKM value for each gene, and log10 (FPKM + 1) was used to plot the heatmap.

The conversion of naringenin to dihydrokaempferol, a necessary step in the production of various anthocyanins, is facilitated by the F3H gene. We identified two F3H genes using the transcriptome data, with F3H1 showing strong expression in LjR while F3H2 was expressed at low levels in both LjR and LjW (Figure 8). The F3’5’H and F3’H genes determine the type of anthocyanin that will be formed (Yi et al., 2021), and we found two F3’5’H genes expressed similarly in LjR and LjW, while one F3’H gene was upregulated in LjR.
Therefore, the differential expression of F3'H may be a crucial factor in determining the type of anthocyanins formed.

DFR is a critical enzyme in anthocyanin production with varying catalytic properties for diverse substrates (Wang et al., 2022). We discovered one DFR gene (DFR1) and two ANS genes (ANS2 and ANS3) that were significantly expressed in LjR (Figure 8). The final essential enzymes in anthocyanin biosynthesis are UFGTs, which catalyze the conversion of unstable anthocyanins into stable ones. Two UFGT genes (UFGT4 and UFGT8) were upregulated in LjR, but the expression of the other five UFGT genes remained unchanged in both LjR and LjW (Figure 8). GSTs prevent flavonoid oxidation or direct them into vacuoles (Lu et al., 2021). We observed that six GST genes were downregulated in LjR, and only one GST gene (GST2) was upregulated (Figure 8).

Identification of TFs related to anthocyanin biosynthesis

Anthocyanin biosynthesis in plants is primarily regulated by the MBW protein complex and other TFs (Cui et al., 2021; LaFountain and Yuan, 2021). In this study, we identified several TFs in honeysuckle, with the top four in quantity being ERF (156), MYB (137), NAC (125), and bHLH (111), as shown in Figure S3. We constructed phylogenetic trees of the MYB and bHLH TFs to compare upregulated and downregulated candidate genes with known genes related to anthocyanin biosynthesis in other plants (Figure 9A and 9B). Based on our analysis, we retrieved 12 MYB, 5 MYB-related, and 6 bHLH TFs (Figure 9C). We proposed that LjMYB112, MYB-related66, LjbHLH20, and LjbHLH23 were potential positive regulators of anthocyanin biosynthesis, while LjMYB49, LjMYB70, LjMYB107, LjMYB75, LjMYB111, LjMYB13, LjMYB54, MYB-related34, MYB-related36, and MYB-related42 were negative regulators (Figure 9). These findings suggest that a complex regulatory network controls anthocyanin biosynthesis in honeysuckle flower buds.

In addition to MYB and bHLH TFs, other TFs such as NAC (Zhou et al., 2015; Sun et al., 2019; Zhang et al., 2020), MADS (Li et al., 2021; Qi et al., 2022), ERF (Ni et al., 2019; An et al., 2020), WRKY (Zhou et al., 2015; An et al., 2019), and bZIP (Wang et al., 2022; Liu et al., 2023) were also identified from the DEGs (Figure 9D), consistent with previously reports.19−22 These TFs may also play a role in regulating the structural and regulatory genes involved in anthocyanin biosynthesis. Our identification of these TFs provides valuable information for future studies on the regulatory mechanisms underlying anthocyanin biosynthesis in honeysuckle.
Figure 9. Phylogenetic analysis and expression level of TFs in honeysuckle flower buds
(A) and (B) The phylogenetic analysis of MYBs and bHLHs, respectively. (C) Heatmap based on expression levels of MYB, bHLH, and WD40 in flower buds of LjR and LjW. (D) Heatmap based on expression levels of bZIP, ERF, MADS, NAC, and WRKY in flower buds of LjR and LjW. The expression level for each gene was represented by the FPKM value, and \( \log_{10}(\text{FPKM} + 1) \) was used to plot the heatmap. Candidate genes were selected based on the annotations. The sequences used for analysis are listed in Supplementary file 1 and Supplementary file 2.
Analysis of PPI network

The PPI network analysis revealed that LjDFR1 is the central hub protein in anthocyanin biosynthesis (Figure 10). LjbHLH23, in conjunction with the bHLH binding protein, was found to act as a positive regulator for the expression of DFR1 (Figure 10). Moreover, LjCHS4 and LjF3H2 were identified as the substrates suppliers for LjDFR1, while LjDFR1 could provide substrates for LjANS2 and LjANS3.

Figure 10. PPI network diagram of DEGs

The nodes in the graph represent proteins, and the edges represent their interactions. The diagram includes several proteins, such as SPL9: Squamosa promoter-binding-like protein 9; SQD1: UDP-sulfoquinovose synthase; UXS1: UDP-glucuronic acid decarboxylase 1; PAR: Phenylacetaldehyde reductase; TKPR1: Tetraketide alpha-pyronate reductase 1; CCR2: Cinnamoyl-CoA reductase 2.

Discussion

The flower buds of honeysuckle contain high levels of flavonoids, phenolic acids, and iridoid compounds, which exhibit anti-inflammatory, antiviral, antioxidant, anticancer, and analgesic effects (Oboh et al., 2013; Han et al., 2016; Kou et al., 2022). These compounds make honeysuckle a valuable resource for various applications, including medicines, food additives, beverages, cosmetics, and landscaping (Xiao et al., 2021). Among the different subspecies of honeysuckle, red honeysuckle has been found to contain higher levels of active substances such as chlorogenic acid, luteoloside, caffeic acid, quercetin, and octylacetate when compared to other varieties (Yuan et al., 2012). Hence, red honeysuckle has greater potential for development and utilization in the food, medicine, cosmetics, and health products industries. In this study, anthocyanin metabolomics analysis identified a total of 59 anthocyanins from two species of honeysuckle, including cyanidin, delphinidin, malvidin, pelargonidin, peonidin, and petunidin. The primary components found in red honeysuckle were cyanidin-3,5-O-diglucoside (633.50 μg/g), peonidin-3,5-O-diglucoside (73.94 μg/g), and cyanidin-3-O-glucoside (3.75 μg/g) (Table 1). These findings provide a scientific basis for further research and development of functional products derived from red honeysuckle.

There are three major types of plant pigments, namely anthocyanins, carotenoids, and chlorophylls. Among these, anthocyanins are the primary pigments responsible for the colour of flowering plants (Khan et al., 2022). Research has shown that anthocyanins possess antioxidant, anti-inflammatory, and other health-promoting properties (Reis et al., 2016). Despite of their potential benefits, there have been limited studies on the red honeysuckle flower (Yuan et al., 2012; Li et al., 2019). To address this gap in knowledge, we utilized metabolomic techniques to obtain detailed information on the six anthocyanin components and their concentrations in the flower buds of two honeysuckle species. Our findings revealed that the cyanidin derivative cyanidin-3,5-O-diglucoside was a common pigment found in both honeysuckle flower buds, with a
content that was approximately 1000 times higher in the LjR flower bud than in the LjW flower bud. Furthermore, cyanidin-3-O-glucoside (3.75 μg/g), cyanidin-3-O-rutinoside (0.22 μg/g), and peonidin-3,5-O-diglucoside (73.93 μg/g) were present only in LjR. Similarly, pelargonidin-3,5-O-diglucoside (0.90 μg/g) was present only in LjR, while pelargonidin-3-O-(6-O-p-coumaroyl)-glucoside (7.12 μg/g) was found only in LjW. Delphinidin-3-O-glucoside was a common pigment in both LjR (0.58 μg/g) and LjW (0.20 μg/g), but the content of the former was about three times higher than that of the latter. In contrast, malvidin derivatives could only be detected in trace amounts in both LjR and LjW. Anthocyanins, the pigments responsible for the red colour in plants, consist of various components that differ from plant to plant. For example, in red apples and red pears, the main pigment component is cyanidin-3-O-galactoside, accounting for more than 80% (Zhang et al., 2012; Chen et al., 2021). Meanwhile, pelargonidin-3-O-glucoside and pelargonidin-3-O-rutinoside are the primary pigment components found in strawberries (da Silva et al., 2007). In red peel longan, cyanidin 3-O-glucoside, cyanidin 3-O-6”-malonyl-glucoside, and cyanidin O-symmetric acid are the primary pigment components, as reported by another study (Yi et al., 2021). Thus, the vibrant red colour of LjR honeysuckle is due to the high concentration of cyanidin-3,5-O-diglucoside, which is its primary pigment component.

The anthocyanin biosynthetic pathway is one of the most clearly studied plant secondary metabolic pathways so far. This pathway includes three enzymes in the phenylpropanoid pathway (PAL, C4H and 4CL), early biosynthetic enzymes of anthocyanins (CHS, CHI, F3H, F3′H, F3′5′H) and late biosynthetic enzymes (DFR, ANS, UFTG, MT, AT). In this study, we found that the expression levels of F3′H, DFR, ANS2, ANS3, UFGT4, and UFGT8 genes in LjR were significantly higher than those in LjW. The F3′H gene, which encodes flavonoid 3′-hydroxylase, is a key gene in the anthocyanin biosynthesis pathway. It catalyses the conversion of dihydrokaempferol to dihydroquercetin. Studies on Arabidopsis mutants have shown that a decrease in F3′H gene expression leads to a decrease in anthocyanin content in leaves, stems and other tissues. Overexpression of apple MdF3′HI and MdF3′HIIb in Arabidopsis tt7-1 mutant restore its wild type and in wild-type tobacco leads to enhanced red flower colour intensity, respectively (Han et al., 2010). Furthermore, a 6 bp insertion in the F3′H gene disrupted its enzymatic activity and hindered cyanidin-/peonidin-type anthocyanin biosynthesis, resulting in a severe reduction of pigment in tree peony (Zhang et al., 2020). Therefore, the high expression of F3′H gene in the LjR observed in our study can provide more raw material dihydroquercetin for the biosynthesis of cyanidin. Dihydrokaempferol is an important precursor and a key branch point of different types of anthocyanin biosynthesis. Thus, our findings suggest that the high expression of the F3′H gene in LjR plays a crucial role in determining the higher concentration of anthocyanins in LjR compared to LjW.

DFR is also a key enzyme gene in the plant anthocyanin biosynthesis pathway, which plays an important role in flower colour development. DFR belongs to the family of reduced coenzyme II (nicotinamide adenine dinucleotide phosphate, NADPH)-dependent short-chain reductases, encoded by single or multiple genes. This enzyme reduces dihydromyricetin flavones, dihydroquercetin, and dihydrokaempferol to their corresponding leucocyanidins by NADPH (Petit et al., 2007). These three substrates of DFR are very similar in structure, differing only in the number of hydroxyl groups on the B benzene ring, which is not the site of enzymatic action. Thus, DFRs from many species utilize all three substrates (Johnson et al., 2001). In Gerbera jamesonii, a single amino acid change in the GjDFR enzyme can alter the substrate specificity of DFR (Johnson et al., 2001). In purple potatoes, whose IbDFR is able to complement the Arabidopsis tt3 (dfr) mutant, downregulation of its expression using RNAi approach diverted the metabolic pathway to the flavonol pathway, resulting in reduced anthocyanin accumulation in young leaves, stems and storage roots (Wang et al., 2013). Therefore, differences in the expression of DFR genes and their substrate specificity will lead to flower colour changes. In this study, two DFR genes were found, the expression level of DFR1 gene in LjR was much higher than that in LjW, and the expression level of DFR2 gene in both was lower. Therefore, we speculate that DFR1 plays an important role in the coloration of LjR.
The ANS gene is essential for the later stages of anthocyanin biosynthesis, as it encodes an enzyme that converts leucocyanidins to coloured anthocyanins using Fe$^{2+}$ and 2-oxoglutarate. Previous studies have demonstrated that ANS is typically encoded by a small gene family in many plants (Zhao and Tao, 2015). For instance, in pomegranate, insertional mutation in the coding region of the ANS gene can result in enzyme inactivation, leading to the emergence of a white pomegranate phenotype without any anthocyanins present in young leaves, flowers, pericarp, and pulp (Ben-Simhon et al., 2015). In this study, three ANS genes were identified, and it was observed that the expression levels of ANS2 and ANS3 genes were significantly higher in LjR than in LjW. Moreover, the expression levels of the ANS1 gene were lower in both. These findings suggest that genetic differences exist in anthocyanin biosynthesis between these plant species and imply that the regulation of ANS gene expression may play a crucial role in determining the colour variation between LjR and LjW.

UFGT is a vital enzyme gene in the late stage of anthocyanin biosynthesis and is responsible for glycosylating anthocyanins to enhance their stability (Muhammad et al., 2022). UFGT has been shown to regulate anthocyanin biosynthesis during litchi fruit coloration (Zhao et al., 2012). Studies on Japanese apricot have revealed that the UFGT activity is higher in safflower than that in white flowers and that it increases as the petals turn red, in parallel with anthocyanins accumulation (Wu et al., 2017). In this study, multiple UFGT genes were identified, and the expression levels of UFGT4 and UFGT8 genes were found to be much higher in LjR than in LjW. Conversely, GST act as an anthocyanin transporter in both fruits and leaves. This study identified seven GST genes, it was observed that the expression levels of GST1 and GST2 genes were much higher in LjR than in LjW. Similarly, the expression levels of GST3, GST5, GST6, and GST7 were higher in LjW, indicating that GST alone cannot account for the observed colour variation between LjW and LjR.

Aside from structural genes in the anthocyanin biosynthetic pathway, TFs can also play a critical role in flower colour development by controlling the spatial and temporal expression of structural genes. The MBW complex is known to regulate plant anthocyanin biosynthesis (Albert et al., 2014; Lloyd et al., 2017). In this study, 12 MYB, 5 MYB-related, and 6 bHLH TFs were identified among the DEGs through expression level and phylogenetic analyses. Positive regulators identified were LjMYB112, MYB-related66, LjbHLH20, and LjbHLH23, while negative regulators included LjMYB49, LjMYB70, LjMYB107, LjMYB75, LjMYB111, LjMYB13, LjMYB54, MYB-related34, MYB-related36, and MYB-related42 (Figure 9). In kiwifruit, the AcMYBF110-AcbHLH1-AcWDR1 complex regulates anthocyanin biosynthesis by directly binding to the promoters of anthocyanin synthesis genes (Liu et al., 2021), whereas AmMYB308 in Antirrhinum majus, AtMYBL2 in Arabidopsis thaliana, GrMYB1R1 and GrMYB1R9 in Gentiana triflora, PhMYB27 and PhMYBx in Petunia hybrida have been reported as negative regulators in anthocyanin biosynthesis (Yan et al., 2021). Similarly, bHLH92 in Leymus chinensis (Zhao et al., 2019), CpbHLH1 in Chimonanthus praecox (Zhao et al., 2020), and SmbHLH1 in Solanum melongena(Duan et al., 2021) has been reported as repressors in the anthocyanin biosynthesis. Additionally, bZIP39, ERF50, MADS-M-type28, MADS-MIKC9, NAC74, NAC81, NAC17, WRKY15, WRKY39, and WRKY49 were differentially expressed between LjR and LjW, suggesting that these TFs may also regulate or influence the structural and regulatory genes involved in anthocyanin production. These findings provide insights into the regulatory networks of anthocyanin biosynthesis in honeysuckle and offer a biological basis for developing new honeysuckle cultivars. The future research will be focused on the identification of candidate transcription factor functions.
Conclusions

Red honeysuckle is a valuable germplasm resource, yet little is known about its red pigment composition, anthocyanin biosynthesis pathway, and regulation mechanism. This study revealed that the major red pigment components in red honeysuckle were cyanidin-3,5-O-diglucoside (633.5 μg/g), peonidin-3,5-O-diglucoside (73.935 μg/g), and cyanidin-3-O-glucoside (3.75 μg/g). The high expression of \( F3H1, F3'H, DFR1, ANS2, ANS3, UFGT4 \) and \( UFGT8 \) genes in red honeysuckle probably accounted for its red coloration. Additionally, bHLH23 may regulate anthocyanins biosynthesis by binding to the \( DFR1 \) gene promoter and promoting its expression. These findings can facilitate the cultivation of new honeysuckle varieties and the development of functional foods and medicines.

Authors’ Contributions

Conceptualization, X.Z. and C.L.; Methodology, X.Z. and C.L.; Formal analysis, X.Z. and C.L.; Investigation, X.Z. and C.L.; Resources, Z.H.; Data curation, Z.X.; Writing–original draft preparation, X.Z. and C.L.; Writing–review and editing, X.Z. and C.L.; Funding acquisition, X.Z. and C.L. All authors read and approved the final manuscript.

Ethical approval (for researches involving animals or humans)

Not applicable.

Acknowledgements

This work was supported by the General Project of Henan Provincial Natural Science Foundation, grant number: 232300420027, the Excellent Research Support and Entrepreneurship Startup Project for Overseas Chinese Scholars at Department of Human Resources and Social Security of Henan Province, grant number: 2023039, and the Start-up Fee for Doctoral Scientific Research at Xuchang University, grant numbers: 20201009 and 20201010.

Conflict of Interests

The authors declare that there are no conflicts of interest related to this article.

References


Zhang X et al. (2023). Not Bot Horti Agrobo 51(3):13150


Maps and affiliations: The publisher stay neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Responsibilities: The editors, editorial board and publisher do not assume any responsibility for the article's contents and for the authors' views expressed in their contributions. The statements and opinions published represent the views of the authors or persons to whom they are credited. Publication of research information does not constitute a recommendation or endorsement of products involved.