Phyton-
International Journal of
Experimental Botany

Tech Science Press

http://doi.org/10.32604/phyton.2026.083016

ARTICLE

Co-Analysis of Metabolomics and Transcriptomics Reveals Key Pathways and Genes Associated with Pigment and Flavonoid Regulation in Mint

Xiangdong Wang[1], Hailong An[2], Yanzhi Ma[3,*], Hong Chen[3,4], Qi Sun[3] and Shaoying Ke[5,*]

[1]School of Physical Science and Technology, Tangshan Normal University, Tangshan, China
[2]Department of Resource Management, Tangshan Normal University, Tangshan, China
[3]Department of Life Sciences, Tangshan Normal University, Tangshan, China
[4]College of Horticulture, Nanjing Agricultural University, Nanjing, China
[5]College of Life Sciences, Hebei Agricultural University, Baoding, China
*Corresponding Authors: Yanzhi Ma. Email: 129026@tstc.edu.cn; Shaoying Ke. Email: Kshy@hebau.edu.cn

Received: 27 March 2026; Accepted: 30 July 2026; Published: 28 August 2026

ABSTRACT: Mint is notably rich in phenolic acids, flavonoids, antioxidants and other bioactive components, and is widely used as food, medicine, spices, and flavoring agents. Thus, metabolite composition serves as a critical indicator for assessing mint quality. In this study, two mint genotypes of *Mentha canadensis* L., were sampled, namely purple mint and green mint. The two genotypes are distinguished by stem color: the purple mint exhibits purple stems, whereas the green mint has green stems. The purple mint exhibited significantly higher anthocyanin and total flavone contents than green mint. Integrated transcriptomic and metabolomic analyses were performed to elucidate the regulatory mechanisms underlying pigment and flavonoid accumulation in mint stems. High-throughput RNA-Seq yielded 167,901 unigenes, of which 34,608 genes were differentially expressed. These differentially expressed genes (DEGs) were mainly involved in the lignin metabolic process and flavonoid biosynthetic process. A total of 143 differentially expressed metabolites (DEMs) were enriched in isoflavonoid, flavonoid biosynthesis, flavone and flavonol biosynthesis, and anthocyanin biosynthesis pathways. Co-analysis of DEGs and DEMs revealed that the flavone and flavonol biosynthesis pathway (ko00944) contained the most DEMs, followed by the flavonoid biosynthesis pathway (ko00941) and the anthocyanin biosynthesis pathway (ko00942). Furthermore, nine key genes and metabolites were identified using the O2PLS model. These findings provide a theoretical basis for understanding the key pathways and genes involved in pigment and flavonoid regulation in mint stems.

KEYWORDS: Mint; pigments; flavones; transcriptomics; metabolomics

1 Introduction

Mint (*Mentha* spp.), a taxonomically complex genus of the *Lamiaceae* family with numerous species, natural hybrids, and cultivars [1], is widely distributed worldwide, especially in temperate and tropical/subtropical regions. Corn mint (*Mentha canadensis* L.), spearmint (*Mentha spicata* L.), and peppermint (*Mentha piperita* L.) are the most important and well-known of the mint species. Valued for its dualutility in flavoring and traditional medicine, *Mentha* species represent economically significant crops within the medicinal and aromatic plant sector [2]. *Mentha* species are widely utilized in the food and flavor industries. Their leaves and stems are traditionally consumed as herbal teas and culinary spices, while fresh or dried plant material, crude extracts, and essential oils are incorporated into confectionery, beverages, and baked goods. Additionally, mint-derived compounds serve as flavor-enhancing agents in toothpaste,

chewing gum, cosmetics, and oral-care products [3–5]. In addition to its food use, mint is well known for its traditional medicinal properties, which are attributed to its richness in antioxidants. For millennia, *Mentha* species have been exploited for their medicinal properties. Traditional applications include topical poultices and balms, as well as inhalation of vaporized essential oils rich in menthol. Mint extracts possess high total phenolic and flavonoid content and are therefore associated with various health benefits in humans, including antimicrobial, anticancer, antiallergenic, analgesic, antidyspeptic, antioxidative, antineuralgic, hypoglycemic, and antidiarrheal properties [6,7].

Mint harbors thousands of bioactive constituents that are non-toxic and serve as highly effective alternatives to synthetic drugs, producing virtually no adverse effects [2]. The predominant bioactive constituents identified across *Mentha* species are terpenoids, with menthol occurring in free form or as ester derivatives. Menthol in peppermint oil has its medicinal properties, while esters such as menthyl acetate contribute to the characteristic minty taste and aroma [8]. The bioactive constituents of *Mentha* species hold substantial economic value. Mint oil serves as a flavoring agent across food, pharmaceutical, and fragrance industries. Volatile secondary metabolite biosynthesis occurs in peltate glandular trichomes, specialized epidermal tissues distributed on leaves, stems, petals or seed coat surfaces, depending on the species [9]. Mint oil is a complex blend of organic chemicals that include volatile components such as carvone (1%), pulegone (0.5–1.6%), β-myrcene (0.1–1.7%), β-caryophyllene (2–4%), limonene (1–7%), isomenthone (2–8%), menthofuran (1–10%), menthyl acetate (2–11%), 1,8-cineole (eucalyptol) (5–13%), menthone (15–32%), and menthol (33–60%) [10,11]. Many essential oil chemotypes show distinct aromatic flavor conferred by different proportions. With the development of high-throughput sequencing technology, we can utilize the datasets generated by transcriptomic sequencing and apply co-expression networks analysis to identify specific gene modules exhibiting coordinated expression. This approach enables exploration of associations between the gene networks and the phenotypes of interest, as well as identification of core genes involved in particular biological processes [12–14].

Metabolomics studies the accumulation and changes of metabolites in specific samples, as metabolites are one of the final products of gene expression in cells, which can directly reflect the physiological state of the organism. Furthermore, the draft genome sequence of *Mentha longifolia* has been completed [15]. These advances laid the foundation for in-depth research on mint. Members of the genus *Mentha* show great variability in phenotype and metabolite composition, both intra- and inter-species, reflecting their specific gene expression and metabolic pathways. Understanding the molecular mechanisms of mint is a prerequisite for its genetic improvement. In the present study, samples of two genotypes of *Mentha canadensis* L. were collected. High-throughput transcriptomics and metabolomics techniques were employed to explore the regulatory mechanisms of anthocyanin and flavone biosynthesis. To our knowledge, this study represents the first comparison of two mint genotypes with different stem colors to reveal the important agronomic genes related to mint quality and inform breeding strategies for mint.

2 Materials and Methods

2.1 Plant Materials

The two experimental materials used in this study belonged to the same species, *Mentha canadensis* L., and exhibited two distinct stem-color genotypes: purple-stem mint and green-stem mint. The plant seedlings were provided by AnGuo Jingxu Seed Station (Hebei, China) and were authenticated as *Mentha canadensis* L. by Researcher Meng Sen at the Research Institute of Tropical Forestry, Chinese Academy of Forestry. Both genotypes were cultivated at the field experimental station of Tangshan Normal University, Tangshan, China.

To minimize environmental variation, stems were harvested uniformly 20 days after planting from both genotypes. For each genotype, fresh stems from 6 independent individual plants were pooled to form one biological replicate. Three independent biological replicates were prepared for each genotype. All samples were immediately frozen in liquid nitrogen and stored at −80°C before subsequent transcriptomics and metabolomics analyses.

2.2 Biochemical Measurements of the Two Mint Genotypes

The content of chlorophyll was determined by the acetone extraction-spectrophotometry method. Based on the absorption characteristics of chlorophyll at specific wavelengths, the absorbance of the extracted liquid at the corresponding wavelength was measured by a spectrophotometer, and the chlorophyll content was calculated according to the Lambert-Beer law. The anthocyanin content was determined using a spectrophotometer (U-3000, HITACHI) method. The optical densities at wavelengths of 530 nm and 657 nm were measured, respectively. Anthocyanin content was calculated as follows: W(content of anthocyanin) = $(OD_{530} - 0.25 \times OD_{657})/m$, where OD_{530} represents the optical density at 530 nm, OD_{657} represents the optical density at 657 nm, and m is the mass of the sample (g). Total flavone content was determined by a colorimetric method. After sample preparation, the absorbance was measured at 415 nm and converted using the rutin standard curve. The unit of total flavone content was mg/g. The content of soluble protein was determined by the Coomassie Brilliant Blue G-250 staining method, and the content of soluble sugar was determined by the anthrone-sulfuric acid method. Each physiological index was measured with 3 biological replicates.

2.3 Transcriptomics Analysis

The total RNA was isolated using the TRIzol reagent (www.tiangen.com) according to the manufacturer's instructions. RNA quantity, quality, and integrity were assessed using an Agilent 2100 bioanalyzer. Samples with RIN > 7 were used to prepare RNA-seq libraries. One microgram of RNA per sample was used as input material for the RNA-seq library preparation. RNA-seq libraries were generated using NEBNext® Ultra™ RNA Library Prep Kit for Illumina® (NEB, USA) following the manufacturer's protocol. Index codes were added to identify sequences for each sample. The clustering of the index-coded samples was performed on a cBot Cluster Generation System using TruSeq PE Cluster Kit v3-cBot-HS (Illumina) according to the manufacturer's instructions. After cluster generation, the libraries were sequenced on an Illumina platform by Metware Biotechnology Co., Ltd. (Wuhan, China), and 150 bp paired-end reads were generated.

The software fastp was used to filter the raw data and obtain the clean reads [16]. Adapter reads, reads with over 10% unknown bases and low-quality reads, which had over 20% reads with a quality score lower than 15, were removed to obtain clean reads. Transcriptome assembly was performed using Trinity [17]. Corset was used to cluster relevant transcripts into gene clusters (https://github.com/Oshlack/Corset) [18]. TransDecoder (https://github.com/TransDecoder/TransDecoder/wiki) was used to identify candidate coding regions within transcript sequences generated by *de novo* RNA-Seq assembly using Trinity. RSEM (version 1.2.8) was used to calculate expression levels of genes and transcripts [19]. Differentially expressed genes (DEGs) were classified using the official annotation and classification system. Gene expression levels from RNA-Seq were estimated by RSEM, and then FPKM (fragments per kilobase of transcript per million fragments mapped) for each gene was calculated based on gene length. DESeq2 (version 1.26) was employed to analyze differential gene expression, and genes with $|\log_2\text{FoldChange}| > 1$ and FDR (false discovery rate) < 0.05 were identified as DEGs [20]. Gene function was annotated based on the following databases using DIAMOND [21] and HMMER [22]: Nr (NCBI non-redundant protein sequences), Swiss-Prot (A manually

annotated and reviewed protein database), Trembl (A variety of new documentation files and the creation of TrEMBL), KEGG (Kyoto Encyclopedia of Genes and Genomes), GO (Gene Ontology), KOG (eukaryotic Ortholog Groups) and Pfam (Protein family). The statistical power of this experimental design, calculated in RNASeqPower [23], was 0.41.

2.4 Metabolomics Analysis

Mint samples were freeze-dried by a vacuum freeze-dryer (Scientz-100F). The freeze-dried samples were ground using a mixer mill (MM 400, Retsch) with a zirconia bead for 1.5 min at 30 Hz. Fifty milligrams of lyophilized powder was dissolved in 1.2 mL 70% methanol solution, vortexed for 30 s every 30 min for a total of 6 cycles. Following centrifugation at 12,000 rpm for 3 min, the extracts were filtrated (SCAA-104, 0.22 μm pore size; ANPEL, Shanghai, China, http://www.anpel.com.cn/) before UPLC-MS/MS analysis.

The sample extracts were analyzed using a UPLC-ESI-MS/MS system (UPLC, SHIMADZU Nexera X2, Shanghai, China, https://www.shimadzu.com.cn/; MS, Sciex 4500 Q TRAP, https://sciex.com/). The analytical conditions follow Peng et al.'s protocol [24]. ESI source conditions included a source temperature of 550°C and spray voltages of 5500 V (positive) and −4500 V (negative). Nebulizer gases were configured as follows: GS1 at 50 psi, GS2 at 60 psi, and CUR at 25 psi. CAD intensity was set to high. Mass calibration for QQQ and LIT modes employed polypropylene glycol at 10 and 100 μmol/L, respectively. Quantification relied on scheduled MRM with nitrogen collision gas at medium level; DP and CE were individually tuned per transition, and elution-time-specific transition sets were applied.

Metabolite characterization was performed using the MWDB database (https://www.metware.cn/) based on secondary spectral information. During data processing, isotopic signals, duplicate signals containing K+, Na+, NH_4+, and duplicate signals of fragment ions of large molecular weight substances were removed. Metabolite quantification was achieved using multiple reaction monitoring (MRM). Analyst 1.6.3 software was employed for mass spectrometry analysis. Differential metabolites were determined by VIP (Variable importance in projection, VIP ≥ 1) and $|\text{Log}_2\text{FC}| \geq 1.0$. VIP values were extracted from the OPLS-DA results, which included score plots and permutation plots generated using the R package MetaboAnalystR (https://github.com/jsychong/MetaboAnalystR). The data were $\log_2$-transformed and mean-centered before OPLS-DA. To avoid overfitting, a permutation test (200 permutations) was performed. Differentially expressed metabolites (DEMs) were identified and enriched, and then the GO and KEGG enrichment analyses were performed as described in the transcriptomics analysis.

2.5 Data Analyses

The principal component analysis (PCA) was performed using the prcomp function in R software. All data were preprocessed with unit variance scaling before PCA. The hierarchical cluster analysis (HCA) of genes and metabolites was presented, and results were presented as heatmaps with dendrograms. For HCA, normalized signal intensities of metabolites (unit variance scaling) were visualized as a color spectrum. Correlation analysis (Pearson correlation coefficient) of genes and metabolites was performed using the cor function in R, and the O2PLS model was applied for integrated analysis [25]. Model reliability was assessed using R^2X and R^2Y (cumulative variation explained in the $\boldsymbol{X}$ and $\boldsymbol{Y}$ matrices, respectively) and Q^2 (predictive ability estimated via 7-fold cross-validation). A permutation test (n = 200) was performed to validate model significance and assess overfitting; the Q^2 intercept was calculated as the intercept of the regression line fitted to Q^2 values versus the correlation between permuted and original $\boldsymbol{Y}$-vectors.

3 Results

3.1 Physiological Measurements

The purple mint (Z) had higher anthocyanin, total flavone and soluble protein content, whereas green mint (L) had higher chlorophyll content (Table 1).

Table 1: Biochemical data of mint samples.

Materials	Chlorophyll (mg/g)	Anthocyanin (mg/g)	Total Flavone (mg/g)	Soluble Sugar (mg/g)	Soluble Protein (mg/g)
Purple mint	0.215 ± 0.008^{a}	8.493 ± 0.481^{a}	8.531 ± 0.758^{a}	60.861 ± 10.027^{a}	4.126 ± 0.306^{a}
Green mint	0.152 ± 0.009^{b}	0.363 ± 0.035^{b}	5.457 ± 0.276^{b}	70.407 ± 0.439^{a}	1.773 ± 0.104^{b}

Note: Different letters represent significant differences at $p < 0.05$.

Multivariate statistical analyses, including PCA and HCA, were performed to evaluate the overall differences in gene expression and metabolite accumulation between the two mint genotypes. PCA score plots showed distinct separation between purple mint and green mint, indicating remarkable transcriptomic and metabolomic differences associated with stem-color variation. HCA further confirmed that samples of the same genotype were clustered together, demonstrating high biological repeatability and reliable genotype-specific divergence. These multivariate results provided robust support for subsequent screening of DEGs and DEMs and highlighted the unique regulatory characteristics of pigment and flavone biosynthesis in the two mint genotypes.

3.2 Transcriptomics

3.2.1 Differentially Expressed Genes

The transcriptomic sequencing of 6 samples (Z1, Z2, Z3, L1, L2 and L3) with three replicates was performed in this study. The statistical power of this experimental design, calculated using RNASeqPower [23] for Z_vs_L, was 0.41. While suboptimal for detecting subtle expression changes, stringent filtering criteria (FDR < 0.05, $|\log_2 FC| \geq 1$) were applied to enhance result reliability and minimize false positives. A total of 48.67 Gb of clean data were obtained in our study. Clean data for each sample were greater than 7 Gb, and the Q30 base percentage exceeded 91%. A total of 282,411 transcripts and 167,901 unigenes were obtained after assembly. Unigene annotation was performed by comparing sequences with the Nr, Swiss-Prot, KEGG, GO, KOG and TrEMBL databases using DIAMOND software. The annotation information of each database is shown in Supplementary Table S1. The genes expressed in between-group samples showed high correlation, and significant differences were observed between groups (Supplementary Fig. S1). DESeq2 was employed to analyze DEGs between purple mint and green mint. In total, 34,608 genes were differentially expressed (Supplementary Table S2). Compared to purple mint, 14,226 genes were downregulated, and 20,382 genes were upregulated in green mint (Supplementary Fig. S2). These downregulated genes likely encompass key structural and regulatory elements of anthocyanin and flavonoid biosynthesis, accounting for the reduced pigment accumulation in this genotype (Table 1). These genes may also encompass photosynthesis-related genes whose reduced expression correlates with the lower chlorophyll and soluble protein content observed in green mint. Conversely, the 20,382 upregulated genes suggest substantial transcriptomic reprogramming, potentially reflecting compensatory activation of alternative metabolic pathways, such as enhanced soluble sugar metabolism or putative stress-responsive pathways, to offset the deficiency in photoprotective pigments and antioxidant capacity. These interpretations are speculative and require experimental validation.

This asymmetric DEG distribution suggests that stem color polymorphism in mint is driven by both the suppression of pigment biosynthesis in green mint and compensatory activation of alternative pathways, rather than by simple transcriptional regulation of a single gene.

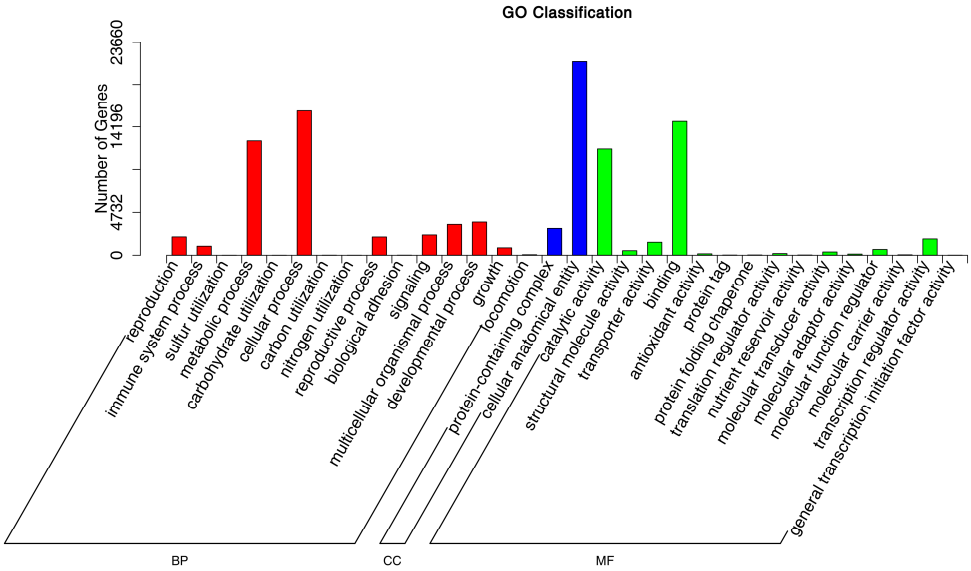
3.2.2 Annotation and Enrichment Analysis of DEGs

After GO annotation of DEGs, the annotated DEGs were classified according to the secondary level of GO classification. The top five classifications were cellular anatomical entity (21,503 DEGs), cellular process (15,966 DEGs), binding (14,791 DEGs), metabolic process (12,629 DEGs) and catalytic activity (11,732 DEGs) (Fig. 1a). The GO enrichment was measured by the rich factor and q-value. The rich factor refers to the ratio of the number of DEGs enriched in the pathway to the number of all genes annotated in the pathway. The 19 most significantly enriched GO terms were displayed in Fig. 1b. The GO enrichment results showed that the top 5 pathways were plant-type secondary cell wall biogenesis (193 DEGs), lignin metabolic process (171 DEGs), flavonoid biosynthetic process (167 DEGs), oxidoreductase activity (120 DEGs) and lignin catabolic process (113 DEGs). The co-enrichment of flavonoid biosynthesis and lignin metabolism pathways is biologically significant, as both pathways share common phenylpropanoid precursors originating from phenylalanine. This metabolic coupling suggests that the purple mint may redirect carbon flux from lignin toward flavonoid and anthocyanin synthesis, contributing to stem pigmentation while potentially altering cell wall properties. The high number of oxidoreductase-related DEGs (120) further supports enhanced redox activity in the phenylpropanoid pathway, which is essential for both pigment production and stress responses. KEGG analysis of DEGs showed that the most enriched genes were characterized by metabolic pathways (4708 DEGs), biosynthesis of secondary metabolites (2625 DEGs) and plant-pathogen interaction (1825 DEGs) (Fig. 1c).

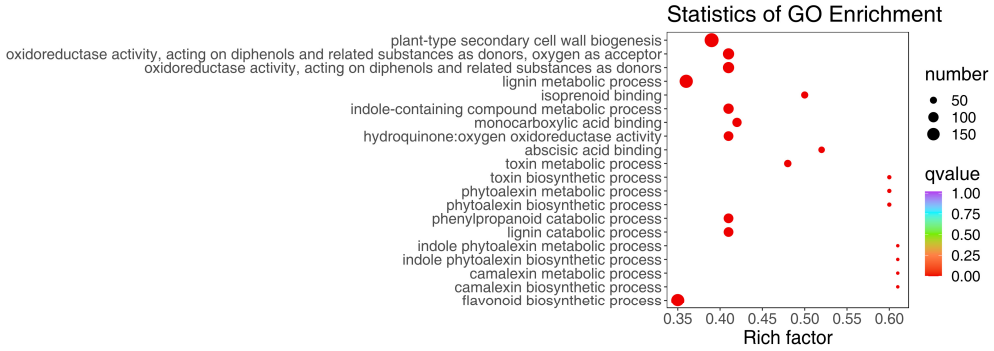
3.3 Metabolomics

3.3.1 Differentially Expressed Metabolites

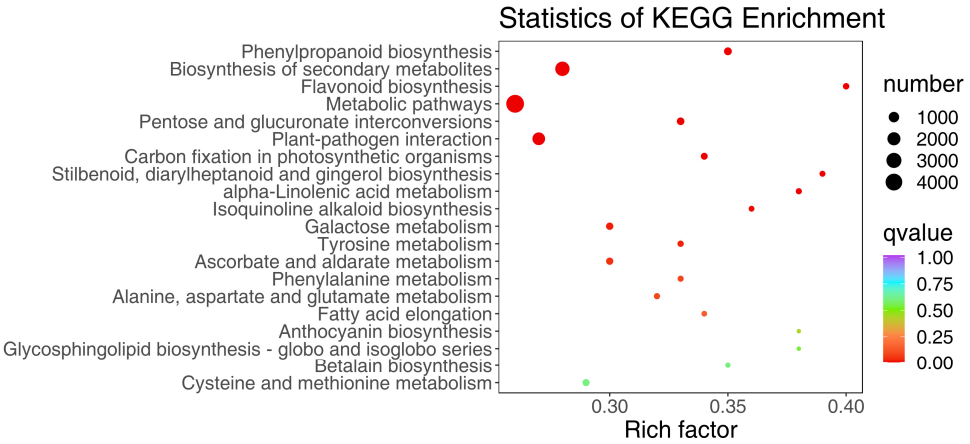
Flavonoids are a class of compounds commonly found in nature, playing an important role in many aspects of plant growth, development, and resistance. Based on the UPLC-MS/MS platform, a total of 229 metabolites were detected in this study. Among them, there were a total of 143 differential metabolites. Compared with purple mint, the content of 38 metabolites in green mint decreased, while the content of 105 metabolites increased (Supplementary Table S3 and Supplementary Fig. S3). The greater number of upregulated metabolites in green mint indicates extensive metabolic reprogramming, likely representing compensatory responses to the deficiency in anthocyanin-derived antioxidant protection. The 38 down-regulated metabolites are enriched in key anthocyanins and flavonoid glycosides that directly contribute to purple pigmentation, consistent with the significantly lower anthocyanin and total flavone contents in green mint (Table 1). Conversely, the 105 upregulated metabolites may include alternative secondary metabolites or stress-responsive compounds that help maintain cellular homeostasis in the absence of robust flavonoid accumulation. In order to display the DEMs clearly, the fold change (FC) values of metabolites between L and Z groups were calculated. The FC values were arranged from large to small, and the top 10 metabolites for upregulation and downregulation were listed in Table 2; a dynamic distribution diagram of metabolite contents was shown in Fig. 2.



(**a**)



(**b**)



(**c**)

Figure 1: GO and KEGG analysis results. (**a**) GO classification histogram. BP, biological process; CC, cellular component; MF, molecular function. (**b**) GO enrichment plot diagram. (**c**) KEGG enrichment plot diagram.

Table 2: List of top 10 metabolites upregulated or downregulated between groups.

Index	Compounds	Class	Log_2FC	Type
Lmzn001572	Scutellarein-7-O-glucuronosyl-(1->2)-glucuronide	flavones	1.97E+01	up
Hmpp003270	Luteolin-4′-O-glucoside	flavones	1.91E+01	up
MWSHY0031	Apigenin-7-O-glucuronide	flavones	1.70E+01	up
MWSHY0050	Kaempferol-3-O-rutinoside	flavonols	1.62E+01	up
pmb0606	Chrysoeriol-7-O-glucuronide	flavones	1.50E+01	up
Lmfn003760	Quercetin-4′-O-glucuronide	flavonols	1.38E+01	up
Lmmp003783	Quercetin-3-O-glucuronide	flavonols	1.35E+01	up
pma0791	Naringenin-7-O-(6″-malonyl) glucoside	flavanones	1.31E+01	up
MWSHY0186	Isosakuranetin (5,7-Dihydroxy-4′-methoxy flavanone)	flavanones	1.19E+01	up
MWSHY0145	Eriodictyol (5,7,3′,4′-Tetrahydroxy flavanone)	flavanones	1.18E+01	up
Lmpn006208	8-Methoxykaempferol-7-O-rhamnoside	flavonols	−2.39E+00	down
pmb3042	Tricin-5-O-glucoside	flavones	−2.42E+00	down
mws2118	Phloretin-2′-O-glucoside	chalcones	−2.88E+00	down
HJAP120	Rhamnetin-3-O-rutinoside	flavonols	−2.98E+00	down
pmb3041	Tricin-7-O-saccharic acid	flavones	−3.00E+00	down
HJAP011	Chrysoeriol-8-C-glucoside	flavones	−3.11E+00	down
HJAP154	Galloylisorhamnetin	flavonols	−3.51E+00	down
Hmmp002121	Isorhamnetin-3-O-gallate	flavonols	−3.68E+00	down
Lmqp002170	Kaempferol-3-O-sophorotrioside	flavonols	−9.91E+00	down
Lmjp003655	6-C-Methyl Kaempferol-3-glucoside	Flavones	−1.38E+01	down

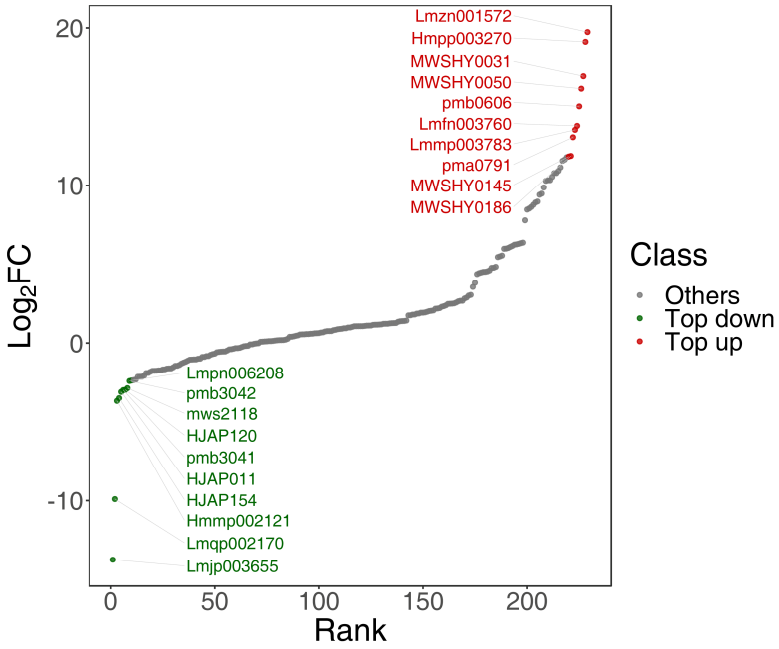


Figure 2: The dynamic distribution diagram of metabolite contents.

3.3.2 Enrichment Analysis of DEMs

A total of 6 significantly correlated pathways were obtained by using the KEGG database annotation. They were metabolic pathways, isoflavonoid biosynthesis, flavonoid biosynthesis, flavone and flavonol biosynthesis, biosynthesis of secondary metabolites and anthocyanin biosynthesis (Fig. 3). The number of metabolites enriched in each pathway is shown in Table 3, which showed that the flavone and flavonol biosynthesis pathway contained the highest number of metabolites, followed by the flavonoid biosynthesis and biosynthesis of secondary metabolites pathways.

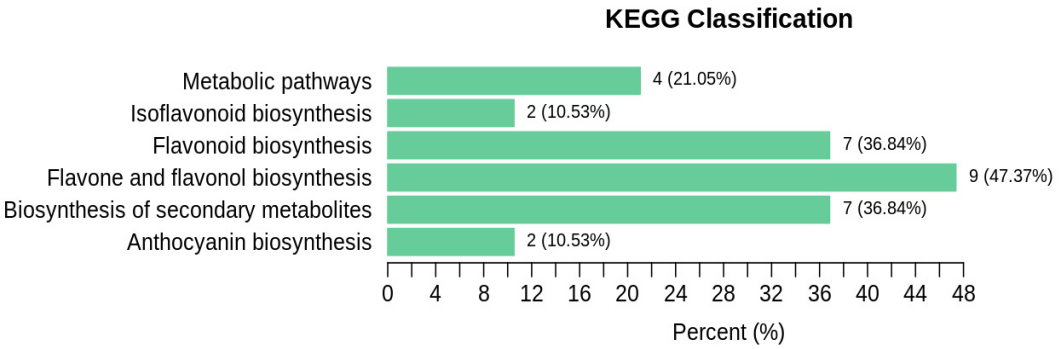


Figure 3: KEGG classification diagram of DEMs.

Table 3: Statistics of annotated metabolites in KEGG pathways.

Ko_ID	Sig_Compound	All_Compound
ko00941	7	12
ko01100	7	9
ko01100	4	5
ko00944	9	16
ko00942	2	2
ko00943	2	2

Note: Sig_compound refers to the number of significant compounds in the pathway; All_compound refers to the number of all compounds in the pathway.

3.4 Joint Analysis of Transcriptomics and Metabolomics

Based on the results of transcriptomics and metabolomics in this study, the DEGs and DEMs were simultaneously mapped onto the KEGG pathway diagram to better understand the relationship between genes and metabolites. By jointly analyzing the KEGG enrichment results of DEGs and DEMs, a total of 6 KEGG pathways were identified (Supplementary Table S4). The results of KEGG enrichment analysis (Table 4) showed that the metabolic pathways (ko01100) and biosynthesis of secondary metabolites (ko01110) pathways contained a large number of DEGs. Four flavonoid-related pathways were annotated. The flavone and flavonol biosynthesis pathway (ko00944) contained the most DEMs, followed by the flavonoid biosynthesis pathway (ko00941) (Fig. 4). The anthocyanin biosynthesis pathway (ko00942), which is related to mint stem color, included 2 DEMs and 36 DEGs. The dominance of flavone and flavonol pathways in the joint analysis, rather than the anthocyanin pathway, indicates that the metabolic divergence between the two genotypes extends beyond pigmentation to encompass broader flavonoid-mediated physiological

functions, including UV protection, antioxidant defense, and signaling. The relatively fewer DEMs but substantial DEGs in the anthocyanin pathway (ko00942) suggest that differential gene expression may contribute to color polymorphism, although whether transcriptional regulation is the primary driver remains to be experimentally verified. While the downstream metabolic flux toward anthocyanin end products appears tightly controlled, this observation is correlational and does not establish causation; stem color variation may involve additional regulatory mechanisms beyond gene expression differences.

Table 4: The results of KEGG enrichment analysis.

KEGG_Map	Description	*p*-Value_Meta	Count_Meta	*p*-Value_Gene	Count_Gene
ko01100	Metabolic pathways	0.3808	4	0.0000	4708
ko01110	Biosynthesis of secondary metabolites	0.2583	7	0.0000	2625
ko00941	Flavonoid biosynthesis	0.8027	7	0.0000	141
ko00942	Anthocyanin biosynthesis	0.3931	2	0.0031	36
ko00943	Isoflavonoid biosynthesis	0.3931	2	0.0140	64
ko00944	Flavone and flavonol biosynthesis	0.8931	9	0.2014	11

Note: *p*-value_meta: the *p*-value of the metabolomics significance test; Count_meta: the number of DEMs annotated to this pathway; *p*-value_gene: the *p*-value of the transcriptome significance test; Count_gene: the number of DEGs annotated to this pathway.

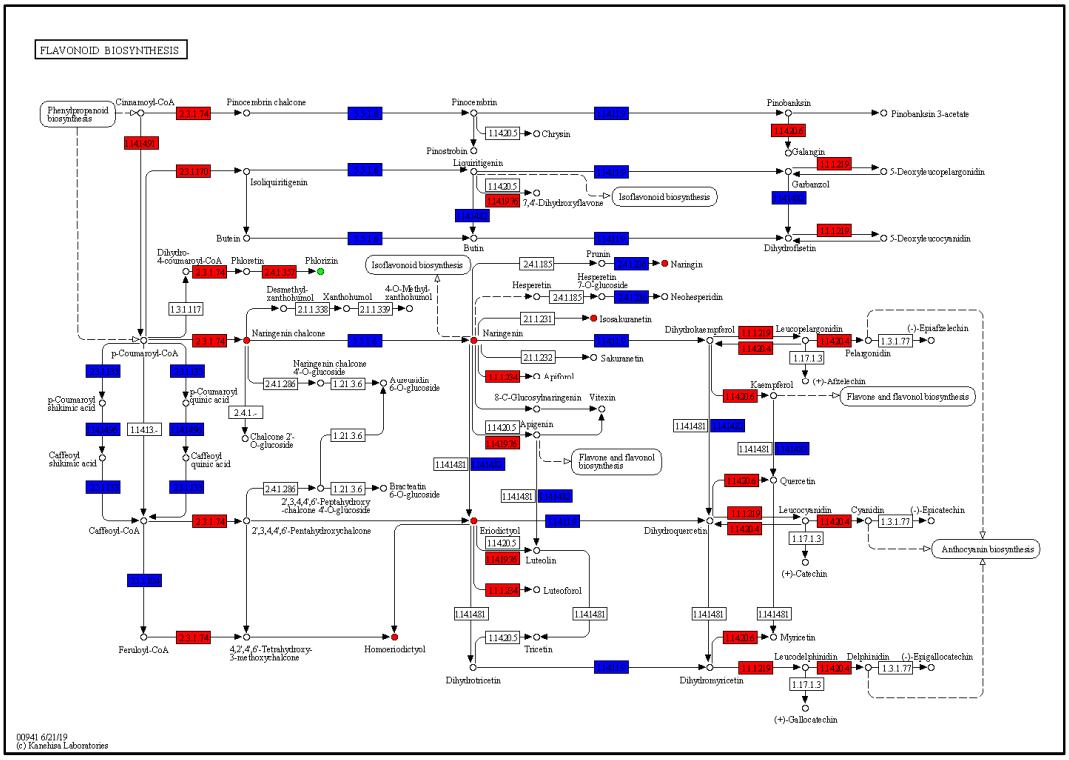


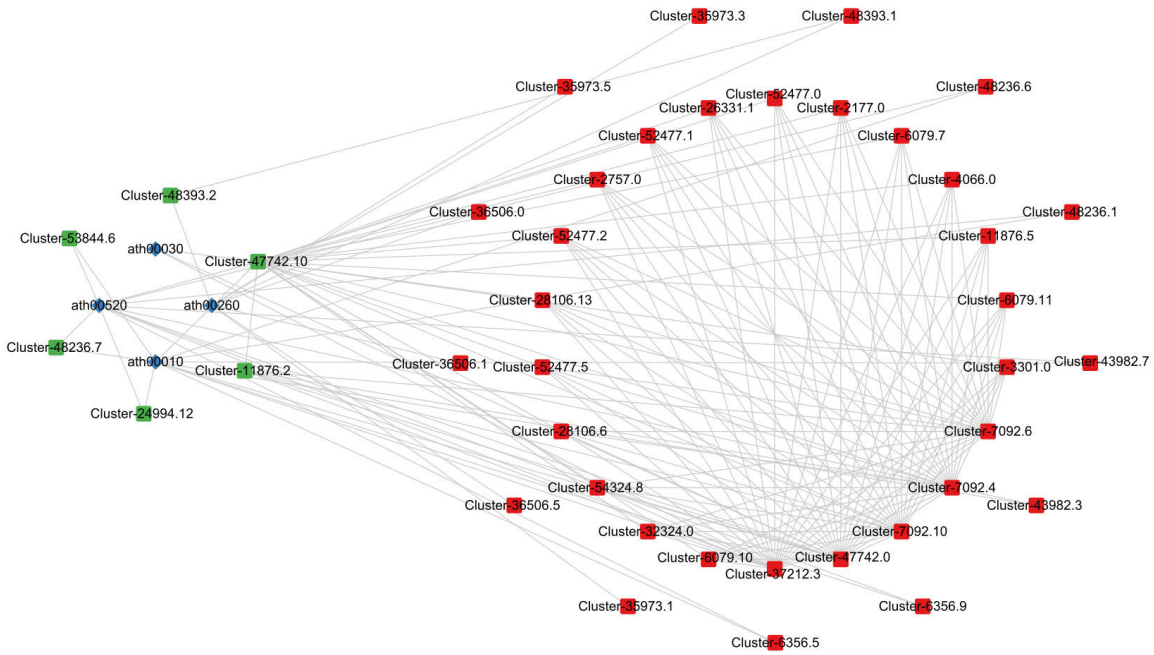
Figure 4: The flavonoid biosynthesis pathway map. The box represents a gene, the dot represents a metabolite, the red color represents upregulation of a gene or metabolite, and the blue color represents both upregulation and downregulation of genes or metabolites.

The O2PLS model was used for the integration analysis between two omics data sets. Model validation parameters were as follows: $R^2X = 0.868$, $R^2Y = 0.998$, $Q^2 > 0.9$, permutation test p-value = 0.0049, and Q^2 intercept = −0.35, indicating excellent predictive performance, robust model stability and negligible overfitting risk. Permutation testing ($n = 200$) confirmed that the observed Q^2 value was significantly higher than those obtained from randomly permuted data, supporting the statistical significance of the integrative model and ruling out overfitting. This model reflects the overall relationships among different data sets and directly indicates the weight of different variables in the model (greater weight indicates greater impact on the other group), thereby enabling accurate identification of key regulatory factors. O2PLS is a bidirectional integrative model that objectively describes correlations between two data sets, avoiding false associations as much as possible, which helps to identify key regulatory factors associated with pigment and flavonoid biosynthesis. In this study, all differentially expressed genes and metabolites were selected to establish an O2PLS model. Loading plots were established to identify the most important genes and metabolites (Table 5).

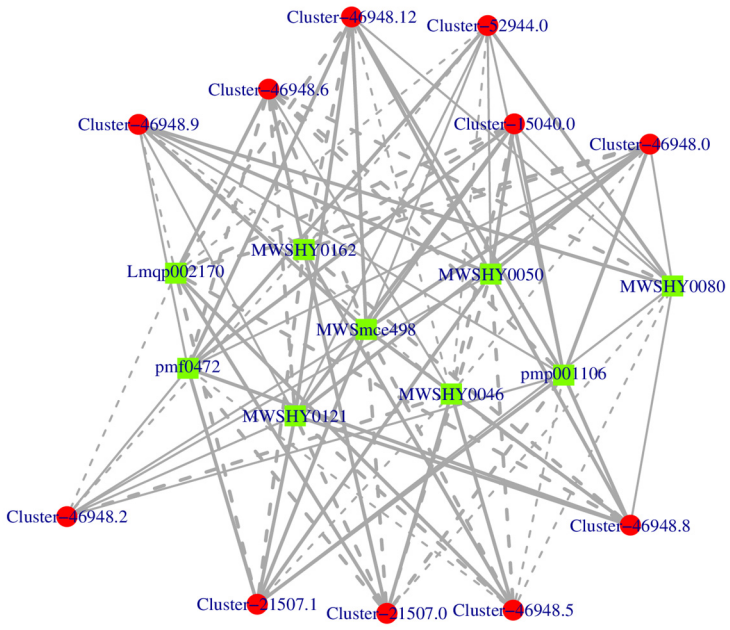
Table 5: The most important genes and metabolites screened using the O2PLS model.

Genes	Metabolites	Description_Meta
Cluster-51396.2	Lmmp003783	Quercetin-3-O-glucuronide
Cluster-7658.8	Lmmn004625	Dihydrokaempferol-7-O-glucoside
Cluster-16235.0	HJN086	Eriodictyol-3′-O-glucoside
Cluster-52100.1	pmb2976	Chrysoeriol-8-C-arabinoside-7-O-sophoroside
Cluster-25073.0	Hmpp003270	Luteolin-4′-O-glucoside
Cluster-42022.2	MWSHY0031	Apigenin-7-O-glucuronide
Cluster-53874.0	Lmzn001572	Scutellarein-7-O-glucuronosyl-(1->2)-glucuronide
Cluster-1116.5	pma0791	Naringenin-7-O-(6″-malonyl) glucoside
Cluster-1676.2	MWSHY0050	Kaempferol-3-O-rutinoside (Nicotiflorin)

The KGML (KEGG markup language) results contain not only the relationship among graphic objects in the KEGG pathway, but also the information on orthologous genes in the KEGG database. This information reveals the network relationship among genes and metabolites, which is convenient for studying the interaction between transcriptome and metabolome. The KGML result (Fig. 5a) showed that 41 genes were closely related to 4 pathways (ath00520, ath00030, ath00260 and ath00010). Among the 41 genes, 6 genes were downregulated in expression, while the other 35 genes were upregulated in expression. Cluster-47742.10, Cluster-7092.10, Cluster-7092.4, Cluster-7092.6, Cluster-37212.3 and Cluster-47742.0 were more closely related to other genes. We also analyzed the correlation of genes and metabolites through network diagrams by selecting differentially expressed genes and metabolites in the pathway with a Pearson correlation coefficient greater than 0.80 and a p-value less than 0.05. For example, the result (Fig. 5b) visually displays the correlation between genes and metabolites in the ko00944 pathway.



(**a**)



(**b**)

Figure 5: Interaction networks between genes and metabolites. (**a**) KGML results. The square represents the gene, and the diamond represents the pathway. Red indicates gene upregulation, and green indicates gene downregulation. (**b**) Correlation network diagram. The red circle represents genes, and the green square represents metabolites. A solid line indicates positive correlation, while a dashed line indicates negative correlation.

4 Discussion

Mentha species are of great economic value and widely used in food and pharmaceutical industries because they are rich in phenolic compounds, especially phenolic acids and flavonoids. In this study,

the purple mint had much higher anthocyanin (8.493 ± 0.481) and total flavone (8.531 ± 0.758) contents compared with green mint (0.363 ± 0.035 and 5.457 ± 0.276, respectively), suggesting that purple mint may have greater commercial potential. However, the molecular mechanisms of anthocyanin and flavone biosynthesis remain unclear in purple and green mint. Transcriptomic sequencing is a powerful tool for screening gene expression patterns and identifying candidate genes, and it has also been successfully used in gene mining in the genus *Mentha* [26]. For example, integrated transcriptome and metabolome analyses have revealed the mechanisms underlying chemical diversity in *Lamiaceae* [27]. Transcriptome profiling also identified numerous DEGs related to light signaling and monoterpene biosynthesis during mint oil accumulation [28]. Other studies have demonstrated widespread asymmetric gene duplication and ancient polyploidy events in *Lamiaceae*, which help distinguish gene duplicates from other genetic variations [29]. Furthermore, high-throughput sequencing has been used to analyze gene expression responses to *V. dahliae* inoculation in mint, and the most abundant DEGs were detected in roots of resistant accessions at early stages [30]. Six important genes involved in the biosynthesis of triterpenic acids in the mint family were identified using transcriptomics analysis. Based on the results of in-depth data analysis, genes encoding squalene epoxidase and oxidosqualene cyclases were proposed as targets for boosting triterpene production [31]. Therefore, transcriptomics analysis can bridge the knowledge gap and facilitate the identification of key genes in mint.

In the current study, high-throughput RNA-seq data were generated to analyze the transcriptomic differences between purple and green mint. A total of 167,901 unigenes were obtained, and 34,608 genes were differentially expressed. The annotated DEGs were mainly classified as cellular anatomical entity (21,503 DEGs), cellular process (15,966 DEGs), binding (14,791 DEGs), metabolic process (12,629 DEGs) and catalytic activity (11,732 DEGs). The top 5 pathways involved by DEGs were plant-type secondary cell wall biogenesis (193 DEGs), lignin metabolic process (171 DEGs), flavonoid biosynthetic process (167 DEGs), oxidoreductase activity (120 DEGs) and lignin catabolic process (113 DEGs). These results indicate that the four pathways (plant-type secondary cell wall biogenesis, lignin metabolic process, flavonoid biosynthetic process, and lignin catabolic process) were closely related. Lignin is a complex phenylpropanoid polymer deposited in the secondary cell walls of plants, and its biosynthesis is tightly coordinated with flavonoid metabolism through shared phenylpropanoid pathway intermediates [32]. The biosynthesis of both flavonoids and lignin originates from phenylalanine, and they are tightly connected in a common biosynthetic network. Flavonoids such as naringenin chalcone, naringenin, dihydrotricin, and tricin are lignin monomers. These compounds are incorporated into the lignin polymer. They were released from the lignin by Derivatization Followed by Reductive Cleavage (DFRC), indicating that at least a fraction of each was integrated into the lignin as ether-linked structures [33]. Defects in the early flavonoid biosynthetic genes encoding chalcone synthase (CHS), chalcone isomerase (CHI), and CHI-like (CHIL) can disrupt tricin-lignin formation, accompanied by marked changes in lignin composition and cell wall digestibility [34]. Therefore, in this study, the purple mint with higher total flavone content also showed significantly different gene expression patterns of secondary cell wall biogenesis and lignin metabolic processes. The results are useful for identifying the key genes in the flavone biosynthetic pathway.

Metabolomics is the systematic identification and quantification of all metabolites in an organism or biological sample [35]. The combination of chromatography and mass spectrometry enables the entire process from substance separation using chromatography to substance identification using mass spectrometry. The UPLC-MS/MS platform can perform accurate qualitative and quantitative analysis of plant metabolites. Plants, as a direct or indirect source of nutrition, energy, and medicine for humans, can synthesize a large number of metabolic substances with diverse biological functions. Therefore, metabolomics

plays an important role in plant research. Given that mint contains numerous secondary metabolites, the metabolomic analysis provides an effective approach to detect and screen metabolites with significant biological significance, and to elucidate the metabolic processes and mechanisms of mint. In the present study, a total of 143 differentially expressed metabolites were detected. Compared with purple mint, the content of 38 metabolites in green mint decreased, while 105 metabolites increased (Supplementary Table S3 and Supplementary Fig. S3). These DEMs were enriched in isoflavonoid biosynthesis, flavonoid biosynthesis, flavone and flavonol biosynthesis, and anthocyanin biosynthesis. The results indicated that flavone biosynthesis in purple mint differed significantly from that in green mint.

Biological processes are highly complex. Integrating multi-omics data can reduce false positives caused by single-omics analysis. The joint analysis of multiple omics data facilitates the study of phenotype and biological process regulation mechanisms in biological models [36]. Such integration can not only cross-validate findings but also provide insights into biological processes [37]. In this study, by establishing data relationships between mint transcriptomics and metabolic pathway enrichment, we systematically and comprehensively analyzed the regulatory mechanisms of anthocyanin and flavone, ultimately achieving a comprehensive understanding of the biological changes and identifying key genes and metabolic pathways for subsequent in-depth analysis. A total of 6 KEGG pathways were identified by jointly analyzing the KEGG enrichment results of DEGs and DEMs. The flavone and flavonol biosynthesis pathway (ko00944) contained the most DEMs, followed by the flavonoid biosynthesis pathway (ko00941) (Fig. 4). The anthocyanin biosynthesis pathway (ko00942), which is related to mint stem color, included 2 DEMs and 36 DEGs. Furthermore, the most important genes and metabolites were screened using the O2PLS model (Table 5). The identification of glucuronides as key compounds may have functional significance, as flavonoid glucuronidation is known to enhance water solubility and stability, which could facilitate vacuolar sequestration and potentially contribute to visible pigmentation in purple stems. Similarly, sophorosides and nicotiflorin may contribute to co-pigmentation effects, which could stabilize anthocyanin chromophores and modulate color intensity. These specific metabolite-gene associations suggest a possible scenario in which the purple phenotype results not merely from anthocyanin accumulation but from potential coordinated remodeling of the entire flavonoid glycosylation network to achieve both pigmentation and metabolic homeostasis. These results suggest a potential explanation for why purple mint accumulates higher levels of flavonoids and anthocyanins. The higher contents were correlated with the upregulation of key structural genes involved in flavonoid and anthocyanin biosynthesis pathways. This transcriptional activation is hypothesized to reflect increased carbon allocation from the phenylpropanoid pool toward flavonoid branches, potentially driven by the differential expression of pathway-specific enzymes such as chalcone synthase (CHS), flavanone 3-hydroxylase (F3H), and dihydroflavonol 4-reductase (DFR). The enhanced metabolic flux in these pathways may have promoted the accumulation of corresponding metabolites, which could contribute to stronger flavonoid production and purple pigmentation in mint stems. The coordinated upregulation of both biosynthetic genes and glycosyltransferases generates a testable hypothesis that the purple mint may possess an integrated regulatory system coupling pigment synthesis with vacuolar storage and stabilization. However, these interpretations are based on correlative evidence and require functional validation through gene knockout, overexpression, or enzymatic assays.

Environmental factors, including light, temperature, and abiotic stresses, are known to influence anthocyanin and flavonoid accumulation in plants [38–41]. However, these environmental effects were not investigated in the present study. Both purple and green mint genotypes were cultivated simultaneously under identical field conditions at the same location and developmental stage, thereby minimizing environ-

mental variability. Consequently, the observed differences in anthocyanin and flavonoid contents in this study are attributed to genotypic variation.

This study specifically focused on the molecular mechanisms underlying stem color variation between two mint genotypes. Since the significant phenotypic differences were only observable in stem tissues, stem samples were selected for omics analysis to investigate the regulatory mechanisms of flavonoid and anthocyanin accumulation related to stem pigmentation. However, it is well known that flavonoids and pigments are mainly synthesized and accumulated in leaf tissues in most plant species, which may reduce the biological relevance and tissue representativeness of our data to a certain extent. Therefore, in future studies, both leaf and stem tissues should be included to more comprehensively reveal the tissue-specific regulatory mechanisms of flavonoid and anthocyanin biosynthesis in mint. We acknowledge that a statistical power of 0.41 is below the ideal threshold, which may potentially affect the reliability of transcriptomic results. To mitigate this issue, we adopted stringent DEG screening criteria: we used FDR < 0.05 to control the false positive rate, together with a biological effect threshold of $|\log_2\text{FoldChange}| > 1$ for dual filtering. Despite the limited statistical power and sample size, these strict cutoff settings have minimized the risk of false-positive DEGs as much as possible. We also acknowledge that this study lacks further experimental validation, including qRT-PCR verification of key candidate genes as an essential next step, and absolute quantification of major flavonoid and anthocyanin metabolites using authentic standards. The DEGs reported herein should be considered exploratory findings derived from correlative omics analysis; their differential expression patterns require independent validation before functional inference can be drawn. Further experimental verification will be conducted in our subsequent work to validate the regulatory roles of these candidate genes and metabolites.

5 Conclusions

In this study, two genotypes of *Mentha canadensis* L. with purple and green stems were used to investigate pigment and flavone biosynthesis in stem tissues using integrated transcriptomic and metabolomic analyses. The purple mint exhibited significantly higher contents of anthocyanin and total flavone than the green mint. A total of 167,901 unigenes and 34,608 DEGs were obtained, which were mainly enriched in lignin metabolism and flavonoid biosynthesis. A total of 143 DEMs were distributed in flavonoid and anthocyanin biosynthesis pathways. The joint analysis confirmed that flavonoid and anthocyanin biosynthesis pathways accounted for the main compositional differences between the two genotypes.

However, this study has several limitations. Only two genotypes and a single sampling stage were included, and the functions of key candidate genes remain to be verified. The DEGs identified in this study are exploratory results derived from correlative omics analysis; their differential expression patterns require confirmation by qRT-PCR, and the regulatory roles of key candidate genes need to be functionally validated through gene knockout, overexpression, or enzymatic assays. Metabolomics data provide complementary correlative evidence but cannot independently validate transcriptomic results or substitute for statistical validation of gene expression. Follow-up experiments will include: (i) qRT-PCR validation of key candidate genes identified in the transcriptomic analysis; (ii) absolute quantification of major flavonoid and anthocyanin metabolites using authentic standards; and (iii) functional characterization of core genes related to flavonoid and anthocyanin biosynthesis through transgenic or gene-editing technologies. It should be emphasized that these findings are specific to stem tissues and should not be extrapolated to leaves or other organs without further investigation. Future perspectives include using the identified key genes and metabolites as important genetic resources. These further studies will contribute to molecular breeding for quality improvement in mint.

Acknowledgement: Not applicable.

Funding Statement: This work was partly supported by the Hebei Province Science and Technology Support Plan (21327509D); and the Tangshan Talent Support Project (A202110006, A202202019).

Author Contributions: The authors confirm contribution to the paper as follows: Conceptualization, Yanzhi Ma and Xiangdong Wang; methodology, Xiangdong Wang and Hailong An; software, Xiangdong Wang; validation, Xiangdong Wang, Hailong An and Qi Sun; formal analysis, Xiangdong Wang and Hailong An; investigation, Xiangdong Wang; resources, Yanzhi Ma; data curation, Xiangdong Wang; writing—original draft preparation, Xiangdong Wang and Hailong An; writing—review and editing, Qi Sun, Hong Chen and Shaoying Ke; visualization, Xiangdong Wang; supervision, Yanzhi Ma; project administration, Yanzhi Ma; funding acquisition, Yanzhi Ma. All authors reviewed and approved the final version of the manuscript.

Availability of Data and Materials: The data presented in this study are available in the article and Supplementary Material. The raw data from the RNA sequencing were deposited in the Sequence Read Archive (SRA) database under accession No. PRJNA1262660. All metabolomic data used in this publication have been deposited in the MetaboLights database with the identifier MTBLS13345. The complete dataset can be accessed at https://www.ebi.ac.uk/metabolights/MTBLS13345.

Ethics Approval: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

Supplementary Materials: The supplementary material is available online at https://www.techscience.com/doi/10.32604/phyton.2026.083016/s1.

References

1. Yousefian S, Esmaeili F, Lohrasebi T. A comprehensive review of the key characteristics of the genus *Mentha*, natural compounds and biotechnological approaches for the production of secondary metabolites. Iran J Biotechnol. 2023;21(4):e3605. [CrossRef].
2. Ahmad Wani S, Naik H, Wagay J, Ahmad Ganie N, Mulla MZ, Dar BN. *Mentha*: a review on its bioactive compounds and potential health benefits. Qual Assur Saf Crops Foods. 2022;14(4):154–68. [CrossRef].
3. Yilmaztekin M, Lević S, Kalušević A, Cam M, Bugarski B, Rakić V, et al. Characterisation of peppermint (*Mentha piperita* L.) essential oil encapsulates. J Microencapsul. 2019;36(2):109–19. [CrossRef].
4. Tang HP, Zhu EL, Bai QX, Wang S, Wang ZB, Wang M, et al. *Mentha* haplocalyx Briq. (Mint): a comprehensive review on the botany, traditional uses, nutritional value, phytochemistry, health benefits, and applications. Chin Med. 2024;19(1):168. [CrossRef].
5. Salehi B, Stojanović-Radić Z, Matejić J, Sharopov F, Antolak H, Kręgiel D, et al. Plants of genus *Mentha*: from farm to food factory. Plants. 2018;7(3):70. [CrossRef].
6. Biltekin SN, Karadağ AE, Demirci F, Demirci B. *In vitro* anti-inflammatory and anticancer evaluation of *Mentha spicata* L. and *Matricaria chamomilla* L. essential oils. ACS Omega. 2023;8(19):17143–50. [CrossRef].
7. Elansary HO, Szopa A, Kubica P, Ekiert H, Klimek-Szczykutowicz M, El-Ansary DO, et al. Polyphenol profile and antimicrobial and cytotoxic activities of natural *Mentha* × piperita and *Mentha longifolia* populations in northern Saudi Arabia. Processes. 2020;8(4):479. [CrossRef].
8. Hudz N, Kobylinska L, Pokajewicz K, Horčinová Sedláčková V, Fedin R, Voloshyn M, et al. *Mentha piperita*: essential oil and extracts, their biological activities, and perspectives on the development of new medicinal and cosmetic products. Molecules. 2023;28(21):7444. [CrossRef].
9. Tiwari P. Recent advances and challenges in trichome research and essential oil biosynthesis in *Mentha arvensis* L. Ind Crops Prod. 2016;82:141–8. [CrossRef].
10. Dimandja JD, Stanfill SB, Grainger J, Patterson DG Jr. Application of comprehensive two-dimensional gas chromatography (GC × GC) to the qualitative analysis of essential oils. J High Resol Chromatogr. 2000;23(3):208–14. [CrossRef].

11. Gherman C, Culea M, Cozar O. Comparative analysis of some active principles of herb plants by GC/MS. Talanta. 2000;53(1):253–62. [CrossRef].
12. Wang Y, Cui T, Niu K, Ma H. Co-expression analyses reveal key Cd stress response-related metabolites and transcriptional regulators in Kentucky bluegrass. Chemosphere. 2024;363:142937. [CrossRef].
13. Sun Y, Ruan X, Wang Q, Zhou Y, Wang F, Ma L, et al. Integrated gene co-expression analysis and metabolites profiling highlight the important role of ZmHIR3 in maize resistance to *Gibberella* stalk rot. Front Plant Sci. 2021;12:664733. [CrossRef].
14. Kulkarni SR, Vandepoele K. Inference of plant gene regulatory networks using data-driven methods: a practical overview. Biochim Biophys Acta Gene Regul Mech. 2020;1863(6):194447. [CrossRef].
15. Vining KJ, Johnson SR, Ahkami A, Lange I, Parrish AN, Trapp SC, et al. Draft genome sequence of *Mentha longifolia* and development of resources for mint cultivar improvement. Mol Plant. 2017;10(2):323–39. [CrossRef].
16. Chen S, Zhou Y, Chen Y, Gu J. Fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018;34(17):i884–90. [CrossRef].
17. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol. 2011;29(7):644–52. [CrossRef].
18. Davidson NM, Oshlack A. Corset: enabling differential gene expression analysis for de novoassembled transcriptomes. Genome Biol. 2014;15(7):410. [CrossRef].
19. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinform. 2011;12:323. [CrossRef].
20. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550. [CrossRef].
21. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. Nat Methods. 2015;12(1):59–60. [CrossRef].
22. Eddy SR. Accelerated Profile HMM Searches. PLoS Comput Biol. 2011;7(10):e1002195. [CrossRef].
23. Hart SN, Therneau TM, Zhang Y, Poland GA, Kocher J-P. Calculating sample size estimates for RNA sequencing data. J Comput Biol. 2013;20(12):970–8. [CrossRef].
24. Peng L, Gao W, Song M, Li M, He D, Wang Z. Integrated metabolome and transcriptome analysis of fruit flavor and carotenoids biosynthesis differences between mature-green and tree-ripe of cv. "golden *Phoenix*" mangoes (*Mangifera indica* L.). Front Plant Sci. 2022;13:816492. [CrossRef].
25. Bouhaddani SE, Houwing-Duistermaat J, Salo P, Perola M, Jongbloed G, Uh HW. Evaluation of O2PLS in omics data integration. BMC Bioinform. 2016;17(Suppl 2):11. [CrossRef].
26. López-Hernández F, Cortés AJ. Whole transcriptome sequencing unveils the genomic determinants of putative somaclonal variation in mint (*Mentha* L.). Int J Mol Sci. 2022;23(10):5291. [CrossRef].
27. Boachon B, Buell CR, Crisovan E, Dudareva N, Garcia N, Godden G, et al. Phylogenomic mining of the mints reveals multiple mechanisms contributing to the evolution of chemical diversity in Lamiaceae. Mol Plant. 2018;11(8):1084–96. [CrossRef].
28. Yu X, Qi X, Li S, Fang H, Bai Y, Li L, et al. Transcriptome analysis of light-regulated monoterpenes biosynthesis in leaves of *Mentha canadensis* L. Plants. 2021;10(5):930. [CrossRef].
29. Godden GT, Kinser TJ, Soltis PS, Soltis DE. Phylotranscriptomic analyses reveal asymmetrical gene duplication dynamics and signatures of ancient polyploidy in mints. Genome Biol Evol. 2019;11(12):3393–408. [CrossRef].
30. Vining KJ, Pandelova I. Dynamic tissue-specific transcriptome changes in response to *Verticillium dahliae* in wild mint species *Mentha longifolia*. Plants. 2022;11(5):674. [CrossRef].
31. Wang P, Wei G, Feng L, Al E. Research advances in oxidosqualene cyclase in plants. Forests. 2022;13(9):1382.
32. Ninkuu V, Aluko OO, Yan J, Zeng H, Liu G, Zhao J, et al. Phenylpropanoids metabolism: recent insight into stress tolerance and plant development cues. Front Plant Sci. 2025;16:1571825. [CrossRef].
33. Rencoret J, Rosado MJ, Kim H, Timokhin VI, Gutiérrez A, Bausch F, et al. Flavonoids naringenin *Chalcone*, naringenin, dihydrotricin, and tricin are lignin monomers in *Papyrus*. Plant Physiol. 2022;188(1):208–19. [CrossRef].
34. Lam PY, Wang L, Lui ACW, Liu H, Takeda-Kimura Y, Chen MX, et al. Deficiency in flavonoid biosynthesis genes *CHS*, *CHI*, and *CHIL* alters rice flavonoid and lignin profiles. Plant Physiol. 2022;188(4):1993–2011. [CrossRef].

35. Shen S, Zhan C, Yang C, Fernie AR, Luo J. Metabolomics-centered mining of plant metabolic diversity and function: past decade and future perspectives. Mol Plant. 2023;16(1):43–63. [CrossRef].
36. Georgii E, Jin M, Zhao J, Kanawati B, Schmitt-Kopplin P, Albert A, et al. Relationships between drought, heat and air humidity responses revealed by transcriptome-metabolome co-analysis. BMC Plant Biol. 2017;17(1):120. [CrossRef].
37. Cho K, Cho KS, Sohn HB, Ha IJ, Hong SY, Lee H, et al. Network analysis of the metabolome and transcriptome reveals novel regulation of potato pigmentation. J Exp Bot. 2016;67(5):1519–33. [CrossRef].
38. Ma Y, Ma X, Gao X, Wu W, Zhou B. Light induced regulation pathway of anthocyanin biosynthesis in plants. Int J Mol Sci. 2021;22(20):11116. [CrossRef].
39. Kuang L, Chen J, Bao X, Zhang D, Liu J, Wang W, et al. Environmental and phytohormonal factors regulating anthocyanin biosynthesis in fruits. Horticulturae. 2025;11(6):681. [CrossRef].
40. Wang Y, Chen J, He G, Yin L, Liao Y. Unlocking the potential of flavonoid biosynthesis through integrated metabolic engineering. Front Plant Sci. 2025;16:1597007. [CrossRef].
41. Zhou B, Zheng B, Wu W. The ncRNAs involved in the regulation of abiotic stress-induced anthocyanin biosynthesis in plants. Antioxidants. 2023;13(1):55. [CrossRef].