#### Check for updates

#### **OPEN ACCESS**

EDITED BY Qunfeng Zhang, Chinese Academy of Agricultural Sciences (CAAS), China

REVIEWED BY Qiong Wang, Jiangxi Agricultural University, China Jing Zhang, Beijing Forestry University, China

\*CORRESPONDENCE Wenqing Li I-wqfjyc@163.com Chaoyuan Zheng Zhengcy@cau.edu.cn

RECEIVED 10 March 2025 ACCEPTED 02 June 2025 PUBLISHED 23 June 2025

#### CITATION

Jia K, Shi J, Bai L, Wang X, Wang Y, Li X, Li W and Zheng C (2025) Integrated transcriptomic and metabolomic analysis of flavonoid biosynthesis in cigar tobacco leaves under variable nitrogen regimes. *Front. Plant Sci.* 16:1589215. doi: 10.3389/fpls.2025.1589215

#### COPYRIGHT

© 2025 Jia, Shi, Bai, Wang, Wang, Li, Li and Zheng. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# Integrated transcriptomic and metabolomic analysis of flavonoid biosynthesis in cigar tobacco leaves under variable nitrogen regimes

Kai Jia<sup>1</sup>, Juanjuan Shi<sup>1</sup>, Lingli Bai<sup>1</sup>, Xueren Wang<sup>2</sup>, Yuemin Wang<sup>3</sup>, Xiangzhen Li<sup>1</sup>, Wenqing Li<sup>3\*</sup> and Chaoyuan Zheng<sup>1\*</sup>

<sup>1</sup>College of Resources and Environment, Fujian Agriculture and Forestry University, Fuzhou, Fujian, China, <sup>2</sup>Sanming Tobacco Sciences Institute, Sanming Tobacco Company of Fujian Province, Sanming, Fujian, China, <sup>3</sup>Institute of Tobacco Sciences, Fujian Provincial Tobacco Monopoly Bureau, Fuzhou, Fujian, China

**Introduction:** Flavonoids are the most abundant secondary metabolites in plants and play important roles in plant growth, environmental adaptation, and human health. Tobacco serves as a vital cash crop and key source of flavonoids. Given tobacco's high nitrogen demand and the particular sensitivity of flavonoid synthesis to nitrogen supply, understanding the regulatory mechanisms of flavonoid biosynthesis under varying nitrogen conditions is essential.

**Methods and results:** We systematically investigated the flavonoid biosynthesis pathways in the cigar tobacco cultivar Haiyan 204 under different nitrogen application rates using integrated transcriptomic and metabolomic analyses. This study identified 694 differentially expressed metabolites (DEMs) and 6,114 differentially expressed genes (DEGs). Integrated analyses revealed significant enrichment in the flavonoid biosynthesis and phenylpropanoid biosynthesis pathways. A regulatory network of flavonoid biosynthesis in response to nitrogen was constructed. Weighted gene co-expression network analysis (WGCNA) identified CHS2 as a key gene strongly associated with flavonoid biosynthesis, alongside its potential regulatory transcription factors *MYB7*, *MYB9*, *bHLH14*, and *MYB\_related1*. The reliability of transcriptome data was validated via qPCR.

**Discussion:** These results provide a scientific basis for elucidating the biosynthesis mechanisms of tobacco flavonoids and identifying key nitrogen-responsive genes. The constructed regulatory network and identified transcriptional regulators offer critical insights into nitrogen-modulated flavonoid synthesis in cigar tobacco.

KEYWORDS

nitrogen, transcriptomics, metabolomics, flavonoids, cigar tobacco

## **1** Introduction

Flavonoids are among the most prevalent secondary metabolites in plants and are frequently present in fruits, vegetables and other crops, mainly concentrated in flowers, leaves and seeds. Flavonoids regulate plant growth and development and are integral to plant adaptation to biotic and abiotic stresses (Shen et al., 2022; Shi et al., 2021). As antioxidants or detoxifiers in plants, they have the ability to eliminate reactive oxygen species (ROS), which effectively protects plants from biotic and abiotic stresses such as UV radiation, cold stress, pathogen infection and insect nibbling (Cavaiuolo et al., 2013; Iwashina, 2003; Pourcel et al., 2007; Zhang et al., 2020). Meanwhile, flavonoids also have beneficial biochemical effects on human beings, helping to delay aging and prevent various diseases (Fernandes et al., 2017; Imran et al., 2019; Selvakumar et al., 2020). Flavonoid biosynthesis initiates from the phenylpropanoid pathway and branches into diverse secondary metabolites, including flavonoids, flavonols, isoflavones and anthocyanins, through intermediates such as coumaroyl coenzyme A and chalcone (Dong and Lin, 2021; Liu et al., 2021a; Nabavi et al., 2020). The biosynthesis of flavonoids is jointly regulated by internal factors such as structural genes, transcription factors, MiRNAs and environmental factors such as temperature, light, water, nutrition and hormones, among which transcription factors play a key role in the synthesis process by regulating the expression of structural genes (Liu et al., 2021b; Xu et al., 2015; Ye et al., 2019). The biosynthesis of flavonoids is regulated by a variety of enzymes, including but not limited to: PAL (phenylalanine ammonia-lyase), C4H (cinnamate 4hydroxylase), 4CL (4-coumarate-CoA ligase), CHS (chalcone synthase), CHI (chalcone isomerase), F3H (flavanone 3hydroxylase), FLS (flavonol synthase), ANS (anthocyanidin synthase) etc (Liu et al., 2021a). Song's research on the flavonoid biosynthesis mechanism in camellia seeds found that this process is regulated by transcription factors such as MYC2, bHLH3, bHLH18, etc (Song et al., 2022). In Sophora alopecuroides, flavonoid biosynthesis-related genes are primarily regulated by transcription factors such as MYB, bHLH, and NAC (Zhu et al., 2021). Additionally, the flavonoid biosynthesis in tea plants is also regulated by MYB, bHLH, and WD transcription factors (Wang et al., 2018; Zhao et al., 2013). These findings indicate that the genes and regulatory factors related to flavonoid biosynthesis vary among different plants depending on species, gene families, functions, and experimental setups, and other factors.

Nitrogen is an important nutrient element necessary for the biosynthesis of amino acids and secondary metabolites and a key factor limiting crop yield (Li et al., 2017; Maathuis, 2009). Studies have shown that appropriate nitrogen application rates can improve the agronomic traits of tobacco, increase the accumulation of dry matter in cigar tobacco leaves, and enhance its economic value (Chen et al., 2020; Yang et al., 2022). As carbon-based secondary metabolites, the biosynthesis of flavonoids is strongly influenced by plant nitrogen status (Deng et al., 2019; Liu et al., 2010b). Research has shown that while fertilization increases the yield of medicinal plants, vegetables, and crops, it may also reduce the accumulation

and synthesis of active compounds such as flavonoids and terpenes, thereby affecting their quality (Groenbaek et al., 2016; Tshivhandekano et al., 2013). It has also been shown that nitrogen fertilization can reduce the accumulation of flavonoids in the leaves of *Gynostemma pentaphyllum*, and increased nitrogen nutrition also leads to a decrease in flavonoid content in apple leaves (Cavaiuolo et al., 2013; Leser and Treutter, 2005; Rühmann et al., 2002). Adequate nitrogen levels promote flavonol glycoside biosynthesis through gene regulation and the accumulation of substrate carbohydrates, while imbalanced nitrogen, especially excess nitrogen, can suppress this process (Dong et al., 2019).

Tobacco (Nicotiana tabacum L.) belongs to the Solanaceae family, as a distinct tobacco product, the primary difference between cigars and cigarettes is that cigars are wrapped in tobacco leaves, whereas cigarettes are wrapped in paper (Yang et al., 2022). In addition, tobacco has a high medicinal value and is an important source of medicinal alkaloids and flavonoids, especially rutin, the content of which is closely related to the quality, aroma and flavor of the tobacco leaf (Fathi et al., 2006; Gebhardt, 2016; Williamson and Gwynn, 1982). Although flavonoid biosynthetic pathways have been extensively studied, flavonoid synthesis is specific, and the biosynthetic processes and expression levels of related genes vary depending on plant species, developmental stage and tissue type (Wan et al., 2020). No studies have been seen on the flavonoid biosynthetic pathways in cigar tobacco leaves, especially in response to nitrogen application rates. The aim of this study was to elucidate the effect of nitrogen on flavonoid biosynthetic pathways in cigar tobacco leaves through transcriptomics and metabolomics approaches. This information will provide an important reference for developing optimal cultivation strategies to improve the quality of cigar tobacco leaves.

## 2 Materials and methods

#### 2.1 Plant materials and treatment

The field trial was conducted in Sanyuan District, Sanming City (E117°16'-117°48', N26°01'-26°25'), a region with a mid-subtropical monsoon climate. The site experiences an annual precipitation of 1,500–1,900 mm, of which over 70% occurs during the cigar tobacco growing season (January to June). The mean annual temperature is 19.4°C, with an average sunshine duration of 1,394.7 hours during the primary growth period and an annual frost-free period ranging from 248 to 352 days (mean 305 days).

The cigar core variety Haiyan 204 was selected for this study. A total of five nitrogen fertilizer gradients (according to the amount of pure nitrogen application rates) were set as follows: N0 (0 kg N/ hm<sup>2</sup>), N1 (120 kg/hm<sup>2</sup>), N2 (180 kg/hm<sup>2</sup>), N3 (210 kg/hm<sup>2</sup>), and N4 (240 kg/hm<sup>2</sup>). The fertilizers for each treatment were prepared in specific proportions using Ca(NO<sub>3</sub>)<sub>2</sub>·4H<sub>2</sub>O, KNO<sub>3</sub>, NH<sub>4</sub>HCO<sub>3</sub>, K<sub>2</sub>SO<sub>4</sub>, KH<sub>2</sub>PO<sub>4</sub>, Mg(OH)<sub>2</sub>, and calcium-magnesium-phosphate. The preparation principle ensured that, except for the differing nitrogen amounts, the phosphorus and potassium application rates remained consistent across treatments. Calcium, magnesium, and

phosphate were mixed with fine soil and applied as hole fertilizers;  $NH_4HCO_3$  along with some  $KNO_3$  and  $K_2SO_4$  were used as topdressing; the remaining fertilizers were applied as base fertilizers in furrows,  $N:P_2O_5:K_2O=1:1:2.64$ . The basic soil properties are as follows: pH 5.75, organic matter 1.23%, total nitrogen 0.072%, available phosphorus 1.44 mg/kg, and available potassium 164.38 mg/kg.

The experiment was conducted in a randomized block design initiated on 21 January 2022. Each treatment comprised 60 plants per plot with three replications, and row spacing was maintained at 120  $\text{cm} \times 35$  cm. Samples were taken on 21 June, when the upper leaves in the field were at maturity, on the 86th day after transplanting. Each treatment consisted of three replicates, so that we obtained samples from fifteen plots (2 to 4 tobacco leaves per plot) based on five N application treatments. The leaf veins were removed and frozen in liquid nitrogen and brought back to the laboratory. Based on this samples were milled in liquid nitrogen for phenotypic data and subsequent transcriptomics and metabolomics assays.

#### 2.2 Chlorophyll, total free amino, superoxide dismutase, total flavonoid, lignin determination

Leaf samples were flash-frozen in liquid nitrogen, lyophilized, and homogenized to a fine powder for subsequent assays. Chlorophyll a and b concentrations were quantified spectrophotometrically at 663 nm and 645 nm, respectively (Porra et al., 1989). Free amino acid extraction followed the ninhydrin reaction protocol (Moore and Stein, 1954), while total flavonoids were assessed via the NaNO<sub>2</sub>–Al(NO<sub>3</sub>)<sub>3</sub>– NaOH colorimetric assay (Lu et al., 2021). Superoxide dismutase (SOD) activity was measured by nitroblue tetrazolium photoreduction (Elavarthi and Martin, 2010), and lignin content was analyzed through the Klason gravimetric procedure (Lupoi et al., 2015).

# 2.3 Determination of agronomic traits figures and phenotypic data analysis

Agronomic traits of tobacco plants were measured 1 week after topping, including plant height, stem circumference, pitch, and number of leaves. Measure with a tape measure. Phenotypic data were analyzed using one-way ANOVA with multiple comparisons conducted using Duncan's method.

#### 2.4 Transcriptome assay

## 2.4.1 RNA extraction, cDNA library construction and sequencing

Using the same sample as 2.1, a total of fifteen (five treatments, each repeated three times), the RNA purification, reverse transcription, library construction, and sequencing were carried out

at Shanghai Majorbio Bio-pharm Biotechnology Co., Ltd. (Shanghai, China). Total RNA was extracted using the MJZol Total RNA Extraction Kit (Shanghai Meiji Junhua Biomedical Technology Co., Ltd., China). The RNA concentration and purity were determined using a NanoDrop 2000 (Thermo Fisher Scientific, USA). RNA integrity was evaluated through agarose gel electrophoresis using Biowest Agarose (Biowest, Spain), and the RIN value was obtained using an Agilent 5300 (Agilent Technologies, USA). For library construction, the RNA requirements included a total amount of 1 µg, a concentration of  $\geq$  30 µg/µL, an RIN > 6.5, and an OD260/280 ratio between 1.8 and 2.2. The library was constructed using the Illumina NovaSeq Reagent Kit method (Illumina, USA)to obtain the final library.

## 2.4.2 Transcriptome data processing and functional annotation of genes

Reference gene source: *Nicotiana\_tabacum*; Reference genome version: GCF\_000715135.1(https://www.ncbi.nlm.nih.gov/genome/ ?term=txid4097[orgn]). In the transcriptome data, raw paired-end reads were trimmed and quality controlled by fastp v0.19.5 (Chen et al., 2018a) with default parameters. Reads containing adapters, those without insert fragments, and those with ambiguous bases (N, more than five) were removed. Low-quality bases at the start and end of the sequences were trimmed, and sequences shorter than 30 bp after adapter removal and quality trimming were discarded. The clean reads were then aligned to the reference genome with orientation mode using HISAT2 v2.1.0 (Kim et al., 2015). The mapped reads of each sample were assembled using StringTie v2.1.2 (Pertea et al., 2015) in a reference-based approach. Gene expression levels in RNA-Seq analysis were determined by quantifying the number of sequences mapped to genomic regions.

#### 2.4.3 Statistical analysis of gene data

The expression level of each transcript was calculated using the transcripts per million reads (TPM) method to identify DEGs between two samples. RSEM software (v1.3.3, University of Wisconsin-Madison, USA) (Li and Dewey, 2011) was used to quantitatively analyze the expression levels of genes and transcripts, enabling the subsequent differential expression analysis between different samples. By combining functional information of the sequences, the regulatory mechanisms of the genes can be revealed. The software used for differential expression analysis is DESeq2 (v1.24.0, Bioconductor, USA) (Love et al., 2014), with the criteria for significant differential expression set at p < 0.05&  $|\log 2FC| \ge 1$ . Statistical significance was set at p < 0.05 and results are expressed as mean ± standard error. Bar graphs were used to show up and down-regulated DEGs between different groups. Comparative analysis of different groups of differential genes using Venn diagrams. Enrichment analyses of major pathways were performed using the KEGG database. Cluster analysis was performed using the R (v4.3.0) package: fastcluster (Müllner, 2013), and gene expression levels (TPM) were clustered after logarithmic transformation (log10), and subclusters were illustrated.

#### 2.5 Metabolite analysis

#### 2.5.1 Metabolite extraction

Using the same sample as 2.1, a total of fifteen (five treatments, each repeated three times) First, accurately weigh  $100 \pm 5$  mg of the sample into a 2 mL centrifuge tube and add a 6 mm diameter grinding bead. Add 400 µL of extraction solvent (Methanol: water=4:1), which includes four internal standards (L-2-chlorophenylalanine at 0.02 mg/mL, e.g.). Use the Wonbio-96C freeze tissue grinder (Shanghai Wanbo Biotechnology Co., Ltd., China) to grind for 6 minutes (-10°C, 50 Hz). Use the SBL-10TD temperature-controlled ultrasonic cleaning machine (Ningbo Xinzhi Biotechnology Co., Ltd., China) for low-temperature ultrasonic extraction for 30 minutes (5°C, 40 KHz), allow the samples to sit at -20°C for 30 minutes. Then, centrifuge using the Centrifuge 5430R (Eppendorf, Germany) at 13,000 g for 15 minutes (13000 g, 4°C) and transfer the supernatant to sampling vials with inserts for analysis.

## 2.5.2 Metabolite detection and analysis of chromatographic conditions

Use the Thermo Fisher Scientific Ultra High Performance Liquid Chromatography Tandem Fourier Transform Mass Spectrometry (UHPLC-Q Exactive) system (Thermo Fisher Scientific, USA) for LC-MS analysis (Shanghai Majorbio Biopharm Biotechnology Co., Ltd.). The chromatography conditions involve using an ACQUITY UPLC BEH C18 (100 mm\*2.1 mm i.d, 1.7 µm; Waters, Milford, USA); The mobile phase A consists of 2% acetonitrile in water (containing 0.1% formic acid), while mobile phase B is acetonitrile (containing 0.1% formic acid). The injection volume is 3 µL, and the column temperature is maintained at 40 °C. Mass spectrometry conditions: Samples are ionized using electrospray ionization (ESI) and mass spectral signals are collected in both positive and negative ion scanning modes. The scanning mode is set to full scan (SCAN), with a mass scanning range of m/z 50-500. The ion spray voltage is set to 3500 V for positive ion mode and -3000 V for negative ion mode. Normalized collision energy is optimized for the detection of metabolites.

#### 2.5.3 Metabolome data preprocessing and metabolite identification

Using Progenesis QI (Waters Corporation, Milford, USA) for peak extraction, alignment, identification, etc., a data matrix is generated containing retention time, peak area, mass-to-charge ratio, and identification information for further processing and bioinformatics analysis. Afterwards, the software was used for feature peak search library identification, matching MS and MS/ MS mass spectrometry information with the self-built plant specific metabolite database (MJDBPM). The MS mass error is set to less than 10 ppm, and metabolites are identified based on the scores from the MS/MS matching.

#### 2.5.4 Statistical analysis of metabolome data

Significant differential metabolites are selected based on the variable importance in projection (VIP) values obtained from the

OPLS-DA model and the results of Student's t-test, with criteria set at p < 0.05 &  $|\log 2FC| \ge 1$  & VIP > 1.0. The identified DEMs are annotated for metabolic pathways using the KEGG database (https://www.kegg.jp/ke-gg/pathway.html) to determine the pathways in which these metabolites are involved. Metabolomic data analysis involved sample correlation analysis and principal component analysis (PCA) to validate biological replicates and assess similarities among different treatment samples. Bar plots were utilized to display up-regulated and down-regulated differential metabolites across different groups, complemented by Venn diagrams for comparative analysis. The KEGG database was referenced to illustrate secondary metabolic pathways associated with the differential metabolites. Clustering analysis was conducted using the R (v4.3.0) package: fastcluster (Müllner, 2013), with relative expression levels of metabolites log-transformed (log10) for clustering purposes.

# 2.6 Combined transcriptome and metabolome analysis

A joint analysis was conducted on pathways commonly annotated by DEGs and DEMs to display the enriched metabolic pathways across different groups. Each group was selected to show ten enriched pathways, based on the *p* value of DEGs and DEMs as well as the number of associated metabolites. Notably, both DEGs and DEMs were significantly enriched in the plant Flavonoid biosynthesis pathway (pathway KO00941), according to the Kyoto Encyclopedia of Genes and Genomes (KEGG). Weighted gene co-expression network analysis was performed using the R package(v4.3.0): WGCNA (Langfelder and Horvath, 2008) to construct gene co-expression networks. All genes were grouped into different modules based on their expression patterns, with genes within the same module exhibiting strong correlations and similar expression profiles. The correlation and significance between the gene matrix values and core modules of differentially accumulated metabolites were calculated using significance function in the R WGCNA package Relationships between transcription factors, core genes and metabolites were examined using Pearson correlation analysis and visualized using Cytoscape v3.3.0 (Cytoscape Consortium, San Diego, CA, USA).

# 2.7 qPCR verification of flavonoid biosynthesis-related key genes

Total RNA was extracted using the MJZol Total RNA Extraction Kit. The RNA purification, reverse transcription, library construction, and sequencing were carried out at Shanghai Majorbio Bio-pharm Biotechnology Co., Ltd. (Shanghai, China). qPCR was performed using the CFX Connect<sup>TM</sup> real-time PCR system (Bio-Rad, USA) with 2 × ChamQ Universal SYBR qPCR Master Mix (Vazyme, China) and MagicSYBR Mixture (CW3008, CWBIO, China). Actin was used as the internal reference gene. Each reaction included three biological and technical replicates to

ensure accuracy. The amplification conditions were as follows: 95°C for 10 min (pre-denaturation); 40 cycles of 95°C for 10 s (denaturation), 60°C for 30 s (annealing), and 72°C for 30 s (extension). The melt curve analysis was conducted using the instrument's default settings. Gene expression levels were calculated using the  $2^{(-\Delta\Delta Ct)}$  method (Livak and Schmittgen, 2001). The primers employed in this study are listed in Supplementary Table S1.

#### **3** Results

#### 3.1 Changes of tobacco leaves after nitrogen application rates

Generally, nitrogen application significantly increased the biomass of tobacco plants, therefore, preliminary agronomic traits were determined in this study (Supplementary Table S2). Nitrogen fertilizer treatments significantly increased the height, stem circumference and spacing, and number of tobacco leaves.

Phenotypic data and plant condition at maturity in the field (Supplementary Figure S1). Significant changes were also observed in the color of the tobacco leaves, which were lighter in the N0 treatment as compared to the nitrogen application treatment. Preliminary analysis of the tobacco leaves showed a significant increase in both total free amino acid and chlorophyll content with higher N application. Amino acid synthesis and photosynthetic capacity in tobacco leaves were significantly increased by nitrogen application. The activity of SOD decreased significantly with increasing nitrogen application, while nitrogen also significantly reduced the total flavonoid content of the tobacco leaves.

#### 3.2 Differential gene transcriptome analysis

Through high-throughput sequencing, each sample obtained between 68,678,358 ~ 80,461,874 raw reads. After trimming adapters and filtering out low-quality sequences, each sample produced between 68,021,938 ~ 79,599,370 clean reads. The distribution of Q30 bases in the clean reads exceeded 94.15%, with an average GC content of 42.97% (Supplementary Table S3). The alignment rate of the transcriptome data to the reference genome ranged from 93.04% to 95.47% (Supplementary Table S4). A total of 69,782 genes were obtained.

Principal component analysis of the gene expression levels of the 15 samples (Figure 1A) revealed that the differences between the two principal components were significant, accounting for 44.3% and 12.88% of the total variance, respectively. The biological replicates were closely clustered, while samples from different treatments were distinctly separated, indicating significant differences in the transcriptome results of tobacco leaves after nitrogen treatment. However, the N0 and N1 treatments were different compared to the other treatments. There were 1,210, 2,856, 3,736 and 4,107 differentially expressed genes in the comparisons of N0 vs N1, N0 vs N2, N0 vs N3 and N0 vs N4, respectively, of which 421, 636, 783 and 842 were up-regulated and 789, 2,220, 2,953 and 3,265 were down-regulated. The number of DEGs exhibited a dose-dependent increase in response to elevated N supply, and the comparison between N0 *vs* N4 showed the highest number of up-regulated and down-regulated genes (Figure 1B). Venn analysis identified 533 genes co-expressed across all N treatment comparisons, highlighting conserved transcriptional responses (Figure 1C) and the exclusive differential genes increased with increasing nitrogen application, these results indicated that nitrogen application had a significant effect on these differential genes.

A total of 6,114 DEGs were identified by screening for significant with p < 0.05 and  $|\log 2FC| \ge 1$ . Hierarchical cluster analysis (Figure 1E) was performed on this gene set. The results showed significant differences between samples from different treatments, which was consistent with the results of PCA analysis. DEGs from the N0 and N1 treatments showed high similarity, which distinguished them from the other treatments, and most of the DEGs showed an increasing trend with increasing N application. Hierarchical clustering results classified all DEGs into five groups, and all samples showed good reproducibility (Figure 1D<sub>1-5</sub>). Subcluster 1 contained 1,499 DEGs, which showed a decreasing trend with increasing nitrogen content. Subcluster 2 contained 1,725 DEGs, and these DEGs tended to increase with the increase of nitrogen fertilizer application. Subcluster 3 had 297 DEGs, subcluster 5 had 342 DEGs, and subcluster 4 contained 2,251 DEGs, which generally increased with increasing nitrogen application, with the highest expression levels observed in nitrogen treatment 3. Overall, subclusters 1, 2 and 4 showed significant changes in relation to the level of nitrogen application.

#### 3.3 Differential metabolite analysis

Through metabolomic analysis, a total of 694 DEMs were obtained. The screening criteria for differential metabolites were set at  $p < 0.05 \& | \log_2FC| \ge 1 \& VIP > 1.0$ . Among the identified metabolites, 40 were flavonoid compounds (Supplementary Table S4). These includes 2 types of anthocyanins, 1 type of flavanol, 2 types of flavanones, 5 types of flavones, 20 types of flavonoid glycosides, 3 types of flavonols, and 7 types of isoflavones.

Hierarchical cluster analysis of the metabolite sets showed significant differences among the samples from different treatments. the DEMs between N0 and N1 showed high similarity, which distinguished them from the other treatments, and most of the DEMs showed an increasing trend with increasing N application (Figure 2A). Principal component analysis of the metabolites of the 15 samples (Figure 2B) showed that the two principal components accounted for 52.5% and 12.1% of the total variance, respectively, which were significantly different. The samples showed proximity within replicates, whereas significant separation was observed between treatments. the N0 and N1 treatments were significantly different from the other treatments, suggesting significant changes in tobacco leaf metabolites after N application.



Transcriptome data analysis of tobacco leaves after different nitrogen treatments. (A) PCA analysis of gene expression levels. (B) Venn diagram of common or unique expression of DEGs between the groups. (C) DEGs quantity statistics. ( $D_{1-5}$ ) Hierarchical clustering of five sub clusters. (E) Cluster analysis of DEGs.

Venn analysis showed (Figure 2C) that there are 298 common differential metabolites shared among the four comparison groups. In contrast to genes, the unique metabolites are most abundant in the N0 *vs* N1 group and least abundant in the N0 *vs* N3 group.

Based on KEGG functional annotation, the DEMs were grouped into three categories, with the majority mapped to metabolic pathways: 94 metabolites were annotated to "Biosynthesis of other secondary metabolites" 93 metabolites were annotated to



"Lipid metabolism" and 83 metabolites were annotated to "Amino acid metabolism" (Figure 2D). Most differential metabolites were up-regulated in N0 and N1 treatments, while downregulation was more prevalent in N3 and N4 treatments (Figure 2E). Specifically, the number of genes in the comparisons of N0 *vs* N1, N0 *vs* N2, N0 *vs* N3, and N0 *vs* N4 were 476, 2856, 3736, and 4107, respectively. Overall, nitrogen application significantly influenced the synthesis of metabolic products.

#### 3.4 Flavonoid biosynthesis pathway map

To further determine the metabolic pathways co-enriched with DEGs and DEMs, KEGG pathways co-enriched with DEGs and DEMs in each comparison group were used. By combining the p value of DEGs and DEMs along with the number of metabolite and gene, 10 enriched pathways were selected for each group (Supplementary Table S6). The metabolic pathways co-annotated

across different comparison groups showed a high degree of overlap, with "Arginine and proline metabolism" appearing four times, and pathways such as "Flavonoid biosynthesis", "Diterpenoid biosynthesis", "Phenylpropanoid biosynthesis", "Linoleic acid metabolism", "Arachidonic acid metabolism" appearing three times. Nitrogen treatment significantly influenced the synthesis of amino acids, flavonoids, terpenes, and lipids in tobacco leaves (Figure 3). Flavonoids are produced from phenylalanine through the phenylpropanoid pathway, and our results indicate that nitrogen treatment greatly affects pathways involved in flavonoid biosynthesis (Phenylpropanoid biosynthesis and Flavonoid biosynthesis). To clearly and intuitively display the variations in genes and metabolites in the flavonoid biosynthesis pathway of cigar tobacco leaves after nitrogen treatment, a pathway map was created (Figure 4). This pathway primarily consists of three routes: Phenylpropanoid biosynthesis, Flavonoid biosynthesis, and Flavone and flavonol biosynthesis.

Through KEGG enrichment analysis and gene functional annotation, a total of 39 genes and 10 corresponding enzymes were identified (Supplementary Table S7). The identified enzymes include: PAL (phenylalanine ammonia-lyase [EC:4.3.1.24]), 4CL (4-coumarate-CoA ligase [EC:6.2.1.12]), CYP73A (trans-cinnamate 4-monooxygenase [EC:1.14.14.91]), HCT (shikimate O-hydroxycinnamoyltransferase [EC:2.3.1.133]), CCoAOMT (caffeoyl-

CoA O-methyltransferase [EC:2.1.1.104]), CYP98A (5-O-(4coumaroyl)-D-quin-ate 3'-monooxygenase [EC:1.14.14.96]), CHI (chalcone isomerase [EC:5.5.1.6]), CHS (chalcone synthase [EC:2.3.1.74]), F3H (naringenin 3-dioxygenase [EC:1.14.11.9]), CYP75B1(flavonoid 3'-monooxygenase [EC:1.14.1-4.82]]), FLS (flavonol synthase [EC:1.14.20.6]). Additionally, there are 9 flavonoid metabolites (Supplementary Table S8), including: Rutin, Caffeoyl shikimic acid, Leucocyanidin, Sophoraflavonoloside, Troxerutin, 3,7-Di-O-methylquercetin, 3-O-Methylquercetin, Tricetin, Isoquercitrin, Quercetin. It was observed that, except for gene LOC107831218, the expression levels of the other five genes regulating PAL and 4CL increase with nitrogen application. In the Flavonoid biosynthesis pathway, the expression of the genes HCT, CYP98A and CCoAOMT, except for gene LOC107826706, showed an increase with nitrogen. The metabolite Caffeoyl shikimic acid, on the other hand, showed a decrease with increasing nitrogen. Conversely, the gene expression levels of CHS, CHI, F3H, FLS, and CYP78B1 show a decreasing trend with increasing nitrogen levels. The metabolites Troxerutin and Leucocyanidin display a similar trend. In contrast, the expression of Quercetin expression increases, while the gene CYP78B1, which regulates the conversion of Quercetin, serves as a key gene connecting the Flavonoid biosynthesis and Flavone and flavonol biosynthesis pathways. In the Flavone and flavonol biosynthesis pathway, influenced by the up-stream genes and





CYP78B1, the expression levels of *Quercetin, Isoquercitrin, Rutin, Sophoraflavonoloside, 3,7-Di-O-methylquercetin,* and *3-O-Methylquercetin* exhibit a regular pattern of increase and decrease with increasing nitrogen.

#### 3.5 WGCNA analysis

To further investigate and validate the regulatory process of genes involved in the flavonoid biosynthesis pathway, WGCNA was performed. A total of 69,782 genes detected in the samples were analyzed, filtering for genes with TPM greater than 1, resulting in 5,022 genes. The analysis used the nine DEMs annotated in the flavonoid biosynthesis pathway, nitrogen fertilizer rate (N.rate), and total flavonoid content as phenotypic data. A soft threshold of  $\beta =$  14 was used, at which point the scale-free topological fit index  $R^2$  was greater than 0.8 and the average connectivity converged to zero (Supplementary Figure S2A). The results suggested a significant

similarity between the no nitrogen treatment (N0) and the low nitrogen treatment between the no nitrogen treatment (N0) and the low nitrogen treatment (N1), while the medium to high nitrogen treatments (N2, N3, N4) showed a higher degree of similarity (Supplementary Figure S2B). The results indicated the identification of five modules in the clustering dendrogram (Figure 5A), comprising 1,443 genes in the blue module, 798 genes in the brown module, 568 genes in the grey module, 1,470 genes in the turquoise module, and 743 genes in the yellow module (Supplementary Table S9). Notably, all modules, except for the grey module, showed significant correlation with nitrogen application levels. Additionally, four modules exhibited significant correlations with all metabolites total flavonoid content except for *3,7-Di-O-methylquercetin* (Figure 5B).

KEGG enrichment analysis of these four modules (Supplementary Figure S3, S4) revealed distinct biological functions. The blue module was significantly enriched in fundamental plant metabolic pathways, including Plant-pathogen interaction, Plant hormone signal



(A) Dendrogram from hierarchical clustering based on topological overlap. (B) Heatmap of correlations between module and sample traits. (C) protein interaction analysis of candidate genes.

transduction, and Sesquiterpenoid and triterpenoid biosynthesis. The brown module was enriched in pathways such as Flavonoid biosynthesis, Isoquinoline alkaloid biosynthesis, and Pentose and glucuronate interconversions. The turquoise module showed significant enrichment in Arginine and proline metabolism, Photosynthesis antenna proteins, Phenylpropanoid biosynthesis, and Flavonoid biosynthesis. The yellow module was significantly enriched in Glutathione metabolism, Cyanoamino acid metabolism, and Alanine, aspartate, and Glutamate metabolism. Notably, both the brown and turquoise modules were significantly enriched in Phenylpropanoid and Flavonoid biosynthesis pathways. Therefore, we consider the brown and turquoise modules to be the core modules in this study. The genes in these modules were not only significantly enriched in the flavonoid biosynthesis pathways but also exhibited strong correlations with phenotypic data in the WGCNA analysis. To further screen for key genes in the flavonoid biosynthesis pathway, we selected a total of 39 genes, including genes from the brown and turquoise modules of the WGCNA results, as well as DEGs in the flavonoid biosynthesis pathway. Using the STRING database (https:// cn.string-db.org/), with *Nicotiana tabacum* as the species, it was found that two of the inputted 39 genes had no proteins associated with that name in *Nicotiana tabacum*. Additionally, there are two outlier genes and two genes matching to corresponding proteins: *NtF3H-LOC107770893(F3H1)* (Zhang et al., 2021c) and *NtFLS*-

LOC107814657(FLS2) (Wang et al., 2019) (Supplementary Table S7). The exported data was visualized using Cytoscape, resulting in a network with 37 nodes and 348 edges, demonstrating a dense protein-protein interaction network. To further investigate the critical genes associated with nitrogen response within flavonoid biosynthesis, we use the cytoNCA plugin to calculate degree centrality (DC) values and create a visual representation (Figure 5C). It was found that four CHS genes had the highest DC value of 32, followed by 4CL3, 4CL1 and HCT1, CCoAOMT1 and CCoAOMT2, which had a DC of 29, 27, 23, 21, 21. Thus, these genes are preliminarily identified as key regulators in the nitrogen response of flavonoid biosynthesis in cigar tobacco leaves. Flavonoid biosynthesis is primarily regulated by structural genes, while transcription factors mainly influence the synthesis of flavonoids by affecting these structural genes. Therefore, the role of transcription factors in flavonoid biosynthesis should not be underestimated. To clearly and intuitively display the relationships between potential transcription factors (genes), candidate genes, and flavonoid components outside the flavonoid biosynthesis pathway. Using the PlantTFDB database and focusing on the species Nicotiana tabacum, we identified a total of 104 genes containing 212 transcription factors, including ARF, B3, bHLH, bZIP, WRKY, MYB, and MIKC (Supplementary Table S10). We selected 20 bHLH and 20 MYB transcription factors (including 7 MYB-related factors) related to flavonoid biosynthesis (Supplementary Table S11). We analyzed these 40 transcription factors along with 39 core genes and their relationship with 9 metabolites. The analysis method employed was Pearson correlation analysis, using expression data and filtering for p value < 0.01 (Supplementary Table S12). Additionally, we utilized the cytoNCA plugin to calculate DC values (Supplementary Table S13) and arranged the genes according to their DC values (Figure 6).

The results indicate a highly significant relationship among transcription factors, regulated genes, and metabolite expression. To further identify core transcription factors and genes within this network, we filtered the data for p < 0.01 and correlation coefficients > 0.95. In the relationship between genes and metabolites, there is a significant correlation between PAL1, PAL2, CHS2, FLS2, CCoAOMT1, CCoAOMT5, Caffeoyl shikimic acid, Tricetin, Quercetin, Isoquercitrin, Rutin. Although no correlation greater than 0.95 was found between CCoAOMT2 and the metabolites, we discovered that it has significant correlations with nine metabolites in flavonoid biosynthesis. At the same time, CCoAOMT1, CCoAOMT2, FLS2, and CHS2 also have multiple correlations with transcription factors. Additionally, in correlation network analysis of genes, transcription factors, metabolites, CCoAOMT1, CCoAOMT2, and FLS2 have a DC value of 38, while CHS2 has a DC value of 40, indicating that they are core genes linking transcription factors and metabolites. CCoAOMT1, CCoAOMT2, FLS2, and CHS2 are also genes with high DC value in the protein interaction network results. In the relationship between transcription factors and structural genes, MYB7, MYB8, MYB9, MYB10, bHLH14, MYB\_related1, and MYB\_related7 show significant correlations with CCoAOMT1, CCoAOMT2, FLS2, and *CHS2* (correlation > 0.77, p < 0.01). Therefore, we believe that *CCoAOMT1*, *CCoAOMT2*, *FLS2*, and *CHS2* are core genes in the nitrogen-regulated flavonoid biosynthesis pathway of cigar tobacco leaves, while *MYB7*, *MYB8*, *MYB9*, *MYB10*, *bHLH14*, *MYB\_related1*, and *MYB\_related7* are core transcription factors.

### 4 Discussion

Previous studies demonstrate that nitrogen supplementation enhances photosynthetic efficiency, stimulates amino acid biosynthesis, accelerates biomass accumulation, and optimizes agronomic performance in tobacco (Ripullone et al., 2003; Trápani et al.,1999). In this study, it is possible that the rapid plant growth with increased nitrogen fertilizer inputs led to leaf senescence, allowing for a decrease in SOD activity. Previous studies have also shown that leaf senescence leads to reduced SOD activity (Dhindsa et al., 1981). In the current study, the total flavonoid content decreased gradually with the increase in nitrogen fertilizer application. This observation is consistent with the general understanding that nitrogen reduces the accumulation of flavonoids in plants (Liu et al., 2010a; Nguyen and Niemeyer, 2008; Strissel et al., 2005). Flavonoids have antioxidant properties (D'Amelia et al., 2018), and SOD activity decreased with increasing nitrogen application, a trend consistent with changes in total flavonoid content. Previous research have also indicate that the total flavonoid content in tea was higher under no and low nitrogen treatments and lower under other nitrogen treatments (Liu et al., 2010a). This observation aligns with our findings, wherein total flavonoid content inversely correlated with N availability.

#### 4.1 Structural genes and metabolites of flavonoid biosynthesis

Flavonoids are regulated by various structural genes involved in pathways such as phenylpropanoid, flavonoid, isoflavonoid, and flavonol biosynthesis (Borreda et al., 2022; Yi et al., 2022). This study identified two core modules related to the flavonoid biosynthesis pathway through WGCNA analysis and further screened the core genes CHS2 and FLS2 through protein-protein interaction network analysis. Both genes showed significant correlations with eight annotated metabolites and had high DC values in the protein interaction network. CHS is a key and ratelimiting enzyme in the flavonoid biosynthetic pathway (Deng et al., 2014; Zhang et al., 2017). In tomato (Solanum lycopersicum), the suppression of CHS mediated by RNA interference leads to a reduction in total flavonoid levels (Schijlen et al., 2007). Upregulation of FLS can boost the metabolic flux towards flavonoid synthesis (Zhang et al., 2021b). In the current study, the expression of CHS2 and FLS2 was down-regulated under medium and high nitrogen treatments (N2, N3, and N4), and CHI and F3H showed a consistent expression trend between these two genes, which may be one of the reasons for the decrease in total flavonoid content. High nitrogen treatment reduces the activity of CHS and FLS enzymes,



which in turn leads to a decrease in flavonoid content in apple leaves (Strissel et al., 2005), which aligns with the findings of this study.

In the analysis of metabolites, the levels of most metabolites showed a decreasing trend with increasing nitrogen application, consistent with the trend in total flavonoid content. The contents of *rutin, quercetin,* and *isoquercitrin* increased significantly with increasing nitrogen application, reaching the highest levels under high nitrogen treatment. Previous studies (Li et al., 2016; Yang et al., 2008) have shown that these compounds have strong antioxidant effects, and their accumulation may be a plant response to alleviate the oxidative stress induced under high nitrogen conditions. This was also corroborated with the increase in SOD activity with increasing nitrogen application. Lignin biosynthesis begins with the general phenylpropanoid pathway, which can also produce precursors for flavonoids. And the same substrates may be competed for by flavonoid synthesis and lignin biosynthesis (Zhang et al., 2021a; Zhu et al., 2019). A critical role of CCoAOMT in lignin deposition was evidenced by its enzymatic inhibition resulting in drastically decreased lignin content (Zhong et al., 2000). Meanwhile, PAL is also a key gene in lignin biosynthesis (Lu et al., 2019), and its expression trend is consistent with CCoAOMT. In this study, we discovered that nitrogen application promotes the expression of CCoAOMT. Therefore, we believe that the increase in CCoAOMT expression may enhance lignin synthesis, competing for precursor substances with flavonoids, which in turn leads to a decrease in total flavonoid content. The results of lignin content determination also support this viewpoint (Supplementary Figure S5). Plants may increase carbon flow to lignin synthesis by regulating carbon and nitrogen metabolism, thereby enhancing the mechanical strength of cell walls and improving stress resistance. This may be the reason why N0 treatment also has a higher lignin (Chen et al., 2018b).

# 4.2 Transcription factors in flavonoid biosynthesis

After constructing the flavonoid biosynthesis pathway map for cigar tobacco leaves, WGCNA analysis identified that MYB and bHLH are the most prevalent transcription factor families under different nitrogen treatments. Specifically, 20 MYB (including 7 MYB-related) transcription factors and 20 bHLH transcription factors were found to be associated with 37 structural genes involved in flavonoid biosynthesis. In addition, other transcription factor families such as ARF and B3 were also identified. When filtering for correlation > 0.95, MYB7, MYB8, MYB9, MYB10, bHLH14, and MYB\_related1, MYB\_related7 were all significantly correlated with CCoAOMT1, CCoAOMT2, FLS2, and CHS2 (correlations > 0.77). MYB and bHLH are two crucial transcription factor families in plants, and their interactions govern various enzymatic stages in flavonoid biosynthesis (Liu et al., 2021a; Totsuka et al., 2018; Xu et al., 2018). Previous studies have shown that the Arabidopsis transcription factor AtMYB11 increases the content of beneficial flavonols in tobacco (Pandey et al., 2015). Additionally, the buckwheat gene FtMYB31, which encodes an R2R3-MYB transcription factor, promotes flavonoid accumulation in tobacco (Sun et al., 2020). We suggest that these transcription factors interact with each other to up-regulate the expression of CCoAOMT1 and CCoAOMT2 and repress the expression of FLS2 and CHS2 with increasing nitrogen application, which ultimately leads to a decrease in flavonoid content of tobacco leaves with increasing nitrogen application. Li's study (Li et al., 2022) identified MYB as a prominent transcription factor family among those associated with nitrogen utilization. Furthermore, bHLH transcription factors have also been shown to enhance wheat's tolerance to phosphorus and nitrogen deprivation (Yang et al., 2016). Therefore, we conclude that MYB7, MYB8, MYB9, MYB10, bHLH14, and MYB\_related1, MYB\_related7 are key transcription factors responding to different nitrogen treatments in the flavonoid biosynthesis of tobacco.

In summary, the selected key genes and transcription factors are significantly correlated with the biosynthesis of flavonoids under nitrogen-responsive conditions. Nitrogen application may reduce the expression of *CHS2* and *FLS2* by affecting transcription factors, while upregulating the expression of *CCoAOMT1* and *CCoAOMT2*, thereby promoting lignin synthesis, which competes with flavonoid biosynthesis for common precursors, resulting in a decrease in total flavonoid content, which is consistent with our hypothesis. However, there were some differences between the qPCR results and transcriptome data for certain genes. Except for *CHS2*, the expression changes of *MYB7*, *MYB9*, *bHLH14*, and *MYB-related 1* 

observed in the transcriptome data were consistent with those in the qPCR results. Possible reasons may include sequencing bias, transcript isoform effects, or experimental technique errors. Future research should further explore how nitrogen application regulates transcription factors and associated metabolic pathways, further clarify the competitive relationship between lignin and flavonoid metabolism, and thus improve our understanding of the physiological mechanisms underlying nitrogen-regulated secondary metabolism.

## 5 Conclusions

By integrating transcriptomics and metabolomics, we constructed a regulatory network for flavonoid biosynthesis in cigar tobacco leaves and identified *CHS2* as a key gene, revealing the potential regulatory mechanism of nitrogen responsive flavonoid biosynthesis in cigar tobacco. Research has found that transcription factors *MYB7*, *MYB9*, *bHLH14*, and *MYB-related 1* play important roles in flavonoid biosynthesis. The nitrogen application treatment mainly reduced the expression of *CHS2* by regulating the expression of these transcription factors, affecting the accumulation of flavonoids and thus reducing the total flavonoid content. Our research provides a scientific basis for a deeper understanding of nitrogen-responsive flavonoid biosynthesis in cigar tobacco leaves, as well as important references for screening key genes related to nitrogen regulation and exploring plant growth regulatory networks.

## Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

## Author contributions

KJ: Conceptualization, Data curation, Visualization, Writing – original draft, Writing – review & editing. JS: Data curation, Writing – review & editing. LB: Data curation, Writing – review & editing. XW: Data curation, Writing – review & editing. YW: Data curation, Writing – review & editing. XL: Conceptualization, Methodology, Writing – review & editing. WL: Conceptualization, Methodology, Writing – review & editing. CZ: Conceptualization, Methodology, Supervision, Writing – original draft, Writing – review & editing.

## Funding

The author(s) declare that financial support was received for the research and/or publication of this article. The Research Foundation of Fujian Provincial Tobacco Monopoly Bureau (2021350000240082,

2024350000240042) and the Major Program Foundation of National Tobacco Monopoly Bureau (110202201039 XJ-10).

### Conflict of interest

Author XW was employed by company Sanming Tobacco Company of Fujian Province.

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

#### Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

#### Publisher's note

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

### Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2025.1589215/ full#supplementary-material

#### References

Borreda, C., Perez-Roman, E., Talon, M., and Terol, J. (2022). Comparative transcriptomics of wild and commercial Citrus during early ripening reveals how domestication shaped fruit gene expression. *BMC Plant Biol.* 22, 123. doi: 10.1186/s12870-022-03509-9

Cavaiuolo, M., Cocetta, G., and Ferrante, A. (2013). The antioxidants changes in ornamental flowers during development and senescence. *Antioxidants* 2, 132–155. doi: 10.3390/antiox2030132

Chen, Y., Ren, K., He, X., Chen, Y., Hu, B., Hu, X., et al. (2020). The response of fluecured tobacco cultivar K326 to nitrogen fertilizer rate in China. *J. Agric. Sci.* 158, 371– 382. doi: 10.1016/j.fcr.2018.08.019

Chen, X., Wang, J., Wang, Z., Li, W., Wang, C., Yan, S., et al. (2018b). Optimized nitrogen fertilizer application mode increased culms lignin accumulation and lodging resistance in culms of winter wheat. *Field Crops Res.* 228, 31–38. doi: 10.1017/s0021859620000738

Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018a). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34, i884–i890. doi: 10.1093/bioinformatics/bty560

D'Amelia, V., Aversano, R., Chiaiese, P., and Carputo, D. (2018). The antioxidant properties of plant flavonoids: their exploitation by molecular plant breeding. *Phytochem. Rev.* 17, 611–625. doi: 10.1007/s11101-018-9568-y

Deng, X., Bashandy, H., Ainasoja, M., Kontturi, J., Pietiäinen, M., Laitinen, R. A., et al. (2014). Functional diversification of duplicated chalcone synthase genes in anthocyanin biosynthesis of Gerbera hybrida. *New Phytol.* 201, 1469–1483. doi: 10.1111/nph.12610

Deng, B., Li, Y., Xu, D., Ye, Q., and Liu, G. (2019). Nitrogen availability alters flavonoid accumulation in Cyclocarya paliurus via the effects on the internal carbon/ nitrogen balance. *Sci. Rep.* 9, 2370. doi: 10.1038/s41598-019-38837-8

Dhindsa, R. S., Plumb-Dhindsa, P., and Thorpe, T. A. (1981). Leaf senescence: correlated with increased levels of membrane permeability and lipid peroxidation, and decreased levels of superoxide dismutase and catalase. *J. Exp. Bot.* 32, 93–101. doi: 10.1093/jxb/32.1.93

Dong, F., Hu, J., Shi, Y., Liu, M., Zhang, Q., and Ruan, J. (2019). Effects of nitrogen supply on flavonol glycoside biosynthesis and accumulation in tea leaves (*Camellia sinensis*). *Plant Physiol. Biochem.* 138, 48–57. doi: 10.1016/j.plaphy.2019.02.017

Dong, N. Q., and Lin, H. X. (2021). Contribution of phenylpropanoid metabolism to plant development and plant–environment interactions. *J. Integr. Plant Biol.* 63, 180–209. doi: 10.1111/jipb.13054

Elavarthi, S., and Martin, B. (2010). Spectrophotometric assays for antioxidant enzymes in plants. *Methods Mol. Biol.* 639, 273–280. doi: 10.1007/978-1-60761-702-0\_16

Fathi, A. F., Delazar, A., and Amiri, R. (2006). Extraction of flavonoids and quantification of rutin from waste tobacco leaves. *Iran. J. Pharm. Res.* 3, 222–227. doi: 10.22037/ijpr.2010.680

Fernandes, I., Perez-Gregorio, R., Soares, S., Mateus, N., and de Freitas, V. (2017). Wine flavonoids in health and disease prevention. *Molecules* 22, 292. doi: 10.3390/molecules22020292

Gebhardt, C. (2016). The historical role of species from the Solanaceae plant family in genetic research. *Theor. Appl. Genet.* 129, 2281–2294. doi: 10.1007/s00122-016-2804-1

Groenbaek, M., Jensen, S., Neugart, S., Schreiner, M., Kidmose, U., and Kristensen, H. L. (2016). Nitrogen split dose fertilization, plant age and frost effects on phytochemical content and sensory properties of curly kale (*Brassica oleracea L.* var. *sabellica*). *Food Chem.* 197, 530–538. doi: 10.1016/j.foodchem.2015.10.108

Imran, M., Rauf, A., Abu-Izneid, T., Nadeem, M., Shariati, M. A., Khan, I. A., et al. (2019). Luteolin, a flavonoid, as an anticancer agent: A review. *Biomed. Pharmacother.* 112, 108612. doi: 10.1016/j.biopha.2019.108612

Iwashina, T. (2003). Flavonoid function and activity to plants and other organisms. *Biol. Sci. Space.* 17, 24–44. doi: 10.2187/bss.17.24

Kim, D., Langmead, B., and Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* 12, 357–360. doi: 10.1038/nmeth.3317

Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinf.* 9, 1–13. doi: 10.1186/1471-2105-9-559

Leser, C., and Treutter, D. (2005). Effects of nitrogen supply on growth, contents of phenolic compounds and pathogen (scab) resistance of apple trees. *Physiol. Plant* 123, 49–56. doi: 10.1111/j.1399-3054.2004.00427.x

Li, B., and Dewey, C. N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinf*. 12, 1–16. doi: 10.1186/1471-2105-12-323

Li, H., Hu, B., and Chu, C. (2017). Nitrogen use efficiency in crops: lessons from *Arabidopsis* and rice. J. Exp. Bot. 68, 2477–2488. doi: 10.1093/jxb/erx101

Li, X., Jiang, Q., Wang, T., Liu, J., and Chen, D. (2016). Comparison of the antioxidant effects of quercitrin and isoquercitrin: Understanding the role of the 6 "-OH group. *Molecules* 21, 1246. doi: 10.3390/molecules21091246

Li, Y., Wang, M., Teng, K., Dong, D., Liu, Z., Zhang, T., et al. (2022). Transcriptome profiling revealed candidate genes, pathways and transcription factors related to nitrogen utilization and excessive nitrogen stress in perennial ryegrass. *Sci. Rep.* 12, 3353. doi: 10.1038/s41598-022-07329-7

Liu, W., Feng, Y., Yu, S., Fan, Z., Li, X., Li, J., et al. (2021a). The flavonoid biosynthesis network in plants. *Int. J. Mol. Sci.* 22, 12824. doi: 10.3390/ijms222312824

Liu, D., Liu, W., Zhu, D., Geng, M., Zhou, W., and Yang, T. (2010a). Nitrogen effects on total flavonoids, chlorogenic acid, and antioxidant activity of the medicinal plant *Chrysanthemum morifolium. J. Plant Nutr. Soil Sci.* 173, 268–274. doi: 10.1002/jpln.200900229

Liu, W., Zheng, T., Yang, Y., Li, P., Qiu, L., Li, L., et al. (2021b). Meta-analysis of the effect of overexpression of MYB transcription factors on the regulatory mechanisms of anthocyanin biosynthesis. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.781343

Liu, W., Zhu, D. W., Liu, D. H., Geng, M. J., Zhou, W. B., Mi, W. J., et al. (2010b). Influence of nitrogen on the primary and secondary metabolism and synthesis of flavonoids in *Chrysanthemum morifolium Ramat. J. Plant Nutr.* 33, 240–254. doi: 10.1080/01904160903434287

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the  $2^{-\Delta\Delta CT}$  Method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 1–21. doi: 10.1186/ s13059-014-0550-8 Lu, J., Shi, Y., Li, W., Chen, S., Wang, Y., He, X., et al. (2019). *RcPAL*, a key gene in lignin biosynthesis in *Ricinus communis* L. *BMC Plant Biol.* 19, 1–11. doi: 10.1186/s12870-019-1777-z

Lu, S., Wang, J., Zhuge, Y., Zhang, M., Liu, C., Jia, H., et al. (2021). Integrative analyses of metabolomes and transcriptomes provide insights into flavonoid variation in grape berries. *J. Agric. Food Chem.* 69, 12354–12367. doi: 10.1021/acs.jafc.1c02703

Lupoi, J. S., Singh, S., Parthasarathi, R., Simmons, B. A., Henry, R. J. J. R., and Reviews, S. E. (2015). Recent innovations in analytical methods for the qualitative and quantitative assessment of lignin. *Renewable Sustain. Energy Rev.* 49, 871–906. doi: 10.1016/j.rser.2015.04.091

Maathuis, F. J. M. (2009). Physiological functions of mineral macronutrients. Curr. Opin. Plant Biol. 12, 250–258. doi: 10.1016/j.pbi.2009.04.003

Moore, S., and Stein, W. H. (1954). A modified ninhydrin reagent for the photometric determination of amino acids and related compounds. J. Biol. Chem. 211, 907–913. doi: 10.1016/S0021-9258(18)71178-2

Müllner, D. (2013). fastcluster: Fast hierarchical, agglomerative clustering routines for R and Python. J. Stat. Software 53, 1–18. doi: 10.18637/jss.v053.i09

Nabavi, S. M., Šamec, D., Tomczyk, M., Milella, L., Russo, D., Habtemariam, S., et al. (2020). Flavonoid biosynthetic pathways in plants: Versatile targets for metabolic engineering. *Biotechnol. Adv.* 38, 107316. doi: 10.1016/j.bioteChadv.2018.11.005

Nguyen, P. M., and Niemeyer, E. D. (2008). Effects of nitrogen fertilization on the phenolic composition and antioxidant properties of basil (*Ocimum basilicum L.*). J. Agric. Food Chem. 56, 8685–8691. doi: 10.1021/jf801485u

Pandey, A., Misra, P., and Trivedi, P. K. (2015). Constitutive expression of *Arabidopsis* MYB transcription factor, *AtMYB11*, in tobacco modulates flavonoid biosynthesis in favor of flavonol accumulation. *Plant Cell Rep.* 34, 1515–1528. doi: 10.1007/s00299-015-1803-z

Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T. C., Mendell, J. T., and Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* 33, 290–295. doi: 10.1038/nbt.3122

Porra, R. J., Thompson, W. A., and Kriedemann, P. E. (1989). Determination of accurate extinction coefficients and simultaneous equations for assaying chlorophylls a and b extracted with four different solvents: verification of the concentration of chlorophyll standards by atomic absorption spectroscopy. *Biochim. Biophys. Acta. Bioenerg*, 975, 384–394. doi: 10.1016/S0005-2728(89)80347-0

Pourcel, L., Routaboul, J. M., Cheynier, V., Lepiniec, L., and Debeaujon, I. (2007). Flavonoid oxidation in plants: from biochemical properties to physiological functions. *Trends Plant Sci.* 12, 29–36. doi: 10.1016/j.tplants.2006.11.006

Ripullone, F., Grassi, G., Lauteri, M., and Borghetti, M. (2003). Photosynthesisnitrogen relationships: interpretation of different patterns between *Pseudotsuga menziesii* and *Populus x euroamericana* in a mini-stand experiment. *Tree Physiol.* 23, 137–144. doi: 10.1093/treephys/23.2.137

Rühmann, S., Leser, C., Bannert, M., and Treutter, D. (2002). Relationship between growth, secondary metabolism, and resistance of apple. *Plant Biol.* 4, 137–143. doi: 10.1055/s-2002-25727

Schijlen, E. G. W. M., de Vos, C. H. R., Martens, S., Jonker, H. H., Rosin, F. M., Molthoff, J. W., et al. (2007). RNA interference silencing of Chalcone synthase, the first step in the flavonoid biosynthesis pathway, leads to parthenocarpic tomato fruits. *Plant Physiol*. 144, 1520–1530. doi: 10.1104/pp.107.100305

Selvakumar, P., Badgeley, A., Murphy, P., Anwar, H., Sharma, U., Lawrence, K., et al. (2020). Flavonoids and other polyphenols act as epigenetic modifiers in breast cancer. *Nutrients* 12, 761. doi: 10.3390/nu12030761

Shen, N., Wang, T., Gan, Q., Liu, S., Wang, L., and Jin, B. (2022). Plant flavonoids: Classification, distribution, biosynthesis, and antioxidant activity. *Food Chem.* 383, 132531. doi: 10.1016/j.foodchem.2022.132531

Shi, L., Chen, X., Wang, K., Yang, M., Chen, W., Yang, et al. (2021). MrMYB6 from chinese bayberry (*Myrica rubra*) negatively regulates anthocyanin and proanthocyanidin accumulation. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.685654

Song, Q., Ji, K., Yu, X., Chen, L., Wang, L., Gong, W., et al. (2022). Dynamic metabolic and transcriptomic profiling reveal synthetic characters and regulators of flavonoid biosynthesis in Camellia oleifera seeds. *Ind. Crops Prod.* 186, 115295. doi: 10.1016/j.indcrop.2022.115295

Strissel, T., Halbwirth, H., Hoyer, U., Zistler, C., Stich, K., and Treutter, D. (2005). Growth-promoting nitrogen nutrition affects flavonoid biosynthesis in young apple (*Malus domestica Borkh.*) leaves. *Plant Biol.* 7, 677–685. doi: 10.1055/s-2005-872989

Sun, Z., Linghu, B., Hou, S., Liu, R., Wang, L., Hao, Y., et al. (2020). Tartary buckwheat *FtMYB31* gene encoding an R2R3-MYB transcription factor enhances flavonoid accumulation in Tobacco. *J. Plant Growth Regul.* 39, 564–574. doi: 10.1007/s00344-019-10000-7

Totsuka, A., Okamoto, E., Miyahara, T., Kouno, T., Cano, E. A., Sasaki, N., et al. (2018). Repressed expression of a gene for a basic helix-loop-helix protein causes a white flower phenotype in carnation. *Breed. Sci.* 68, 139–143. doi: 10.1270/jsbbs.17072

Trápani, N., Hall, A. J., and Weber, M. (1999). Effects of constant and variable nitrogen supply on sunflower (*Helianthus annuus* L.) leaf cell number and size. *Ann. Bot.* 84, 599–606. doi: 10.1006/anbo.1999.0954

Tshivhandekano, I., Mudau, F. N., Soundy, P., and Ngezimana, W. (2013). Effect of cultural practices and environmental conditions on yield and quality of herbal plants: Prospects leading to research on influence of nitrogen fertilization, planting density and eco-physiological parameters on yield and quality of field-grown bush tea (*Athrixia phylicoides DC.*). *J. Med. Plants Res.* 7, 2489–2493. doi: 10.5897/JMPR2013.4466

Wan, L., Lei, Y., Yan, L., Liu, Y., Pandey, M. K., Wan, X., et al. (2020). Transcriptome and metabolome reveal redirection of flavonoids in a white testa peanut mutant. *BMC Plant Biol.* 20, 1–19. doi: 10.1186/s12870-020-02383-7

Wang, W. L., Wang, Y. X., Li, H., Liu, Z. W., Cui, X., and Zhuang, J. (2018). Two MYB transcription factors (CsMYB2 and CsMYB26) are involved in flavonoid biosynthesis in tea plant [*Camellia sinensis* (L.) O. Kuntze. *BMC Plant Biol.* 18, 1–15. doi: 10.1186/s12870-018-1502-3

Wang, Z., Wang, S., Wu, M., Li, Z., Liu, P., Li, F., et al. (2019). Evolutionary and functional analyses of the 2-oxoglutarate-dependent dioxygenase genes involved in the flavonoid biosynthesis pathway in tobacco. *Planta* 249, 543–561. doi: 10.1007/s00425-018-3019-2

Williamson, R., and Gwynn, G. (1982). Variation of polyphenols in flue-cured tobacco cultivars attributed to location, stalk position, and year<sup>1</sup>. *Crop Sci.* 22, 144–146. doi: 10.2135/cropsci1982.0011183X002200010030x

Xu, W., Dubos, C., and Lepiniec, L. (2015). Transcriptional control of flavonoid biosynthesis by MYB-bHLH-WDR complexes. *Trends Plant Sci.* 20, 176–185. doi: 10.1016/j.tplants.2014.12.001

Xu, H., Yang, G., Zhang, J., Wang, Y., Zhang, T., Wang, N., et al. (2018). Overexpression of a repressor MdMYB15L negatively regulates anthocyanin and cold tolerance in red-fleshed callus. *Biochem. Biophys. Res. Commun.* 500, 405–410. doi: 10.1016/j.bbrc.2018.04.088

Yang, J., Guo, J., and Yuan, J. F. (2008). *In vitro* antioxidant properties of rutin. *LWT-Food. Sci. Technol.* 41, 1060–1066. doi: 10.1016/j.lwt.2007.06.010

Yang, T., Hao, L., Yao, S., Zhao, Y., Lu, W., and Xiao, K. (2016). *TabHLH1*, a bHLHtype transcription factor gene in wheat, improves plant tolerance to Pi and N deprivation via regulation of nutrient transporter gene transcription and ROS homeostasis. *Plant Physiol. Biochem.* 104, 99–113. doi: 10.1016/ j.plaphy.2016.03.023

Yang, R., Yang, J., Yu, J., Wang, S., Yang, C., and Xu, F. (2022). Effects of different nitrogen application rates on the quality and metabolomics of cigar tobacco. *Agron. J.* 114, 1155–1167. doi: 10.1002/agj2.20983

Ye, J., Cheng, S., Zhou, X., Chen, Z., Kim, S. U., Tan, J., et al. (2019). A global survey of full-length transcriptome of *Ginkgo biloba* reveals transcript variants involved in flavonoid biosynthesis. *Ind. Crops Prod.* 139, 111547. doi: 10.1016/j.indcrop.2019.111547

Yi, S. N., Mao, J. X., Zhang, X. Y., Li, X. M., Zhang, Z. H., and Li, H. (2022). *FveARF2* negatively regulates fruit ripening and quality in strawberry. *Front. Plant Sci.* 13, 1023739. doi: 10.3389/fpls.2022.1023739

Zhang, X., Abrahan, C., Colquhoun, T. A., and Liu, C. J. (2017). A proteolytic regulator controlling chalcone synthase stability and flavonoid biosynthesis in *arabidopsis. Plant Cell.* 29, 1157–1174. doi: 10.1105/tpc.16.00855

Zhang, P., Du, H., Wang, J., Pu, Y., Yang, C., Yan, R., et al. (2020). Multiplex CRISPR/Cas9-mediated metabolic engineering increases soya bean isoflavone content and resistance to soya bean mosaic virus. *Plant Biotechnol. J.* 18, 1384–1395. doi: 10.1111/pbi.13302

Zhang, Y., Liu, J., Yu, J., Zhang, H., and Yang, Z. (2021b). Relationship between the Phenylpropanoid Pathway and Dwarfism of *Paspalum seashore* Based on RNA-Seq and iTRAQ. *Int. J. Mol. Sci.* 22, 9568. doi: 10.3390/ijms22179568

Zhang, Y., Xu, S., Ma, H., Duan, X., Gao, S., Zhou, X., et al. (2021c). The R2R3-MYB gene PsMYB58 positively regulates anthocyanin biosynthesis in tree peony flowers. *Plant Physiol. Biochem.* 164, 279–288. doi: 10.1016/j.plaphy.2021.04.034

Zhang, S., Yang, J., Li, H., Chiang, V. L., and Fu, Y. (2021a). Cooperative regulation of flavonoid and lignin biosynthesis in plants. *Crit. Rev. Plant Sci.* 40, 109–126. doi: 10.1080/07352689.2021.1898083

Zhao, L., Gao, L., Wang, H., Chen, X., Wang, Y., Yang, H., et al. (2013). The R2R3-MYB, bHLH, WD40, and related transcription factors in flavonoid biosynthesis. *Funct. Integr. Genomics* 13, 75–98. doi: 10.1007/s10142-012-0301-4

Zhong, R., Morrison, W. H., Himmelsbach, D. S., Poole, F. L., and Ye, Z. H. (2000). Essential role of caffeoyl coenzyme AO-methyltransferase in lignin biosynthesis in woody poplar plants. *Plant Physiol.* 124, 563–578. doi: 10.1104/ pp.124.2.563

Zhu, Y., Wang, Q., Wang, Y., Xu, Y., Li, J., Zhao, S., et al. (2021). Combined transcriptomic and metabolomic analysis reveals the role of phenylpropanoid biosynthesis pathway in the salt tolerance process of *sophora alopecuroides. Int. J. Mol. Sci.* 22, 2399. doi: 10.3390/jims22052399

Zhu, C., Zhang, S., Fu, H., Zhou, C., Chen, L., Li, X., et al. (2019). Transcriptome and phytochemical analyses provide new insights into long non-coding RNAs modulating characteristic secondary metabolites of oolong tea (*Camellia sinensis*) in solar-withering. *Front. Plant Sci.* 10. doi: 10.3389/fpls.2019.01638