



OPEN ACCESS

EDITED BY

Shifeng Cao,
Zhejiang Wanli University, China

REVIEWED BY

Pankaj Kumar Verma,
Ben-Gurion University of the Negev, Israel
Long-Hai Zou,
Zhejiang Agriculture and Forestry University,
China

*CORRESPONDENCE

Changhua Liu
✉ liuchanghua@hlju.edu.cn
Yan Sun
✉ suyan@hlju.edu.cn

RECEIVED 13 December 2024

ACCEPTED 21 February 2025

PUBLISHED 14 March 2025

CITATION

Liu G, Gao H, Song Y, Wang H, Zhang D,
Wang Y, Liu S, Li Z, Liu C and Sun Y (2025)
Multiomic analysis reveals that the flavonoid
biosynthesis pathway is associated with cold
tolerance in *Heracleum moellendorffii* Hance.
Front. Plant Sci. 16:1544898.
doi: 10.3389/fpls.2025.1544898

COPYRIGHT

© 2025 Liu, Gao, Song, Wang, Zhang, Wang,
Liu, Li, Liu and Sun. This is an open-access
article distributed under the terms of the
[Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/).
The use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Multiomic analysis reveals that the flavonoid biosynthesis pathway is associated with cold tolerance in *Heracleum moellendorffii* Hance

Guan Liu^{1,2}, Huan Gao¹, Yu Song², Hanhui Wang²,
Dongye Zhang¹, Yang Wang¹, Shuo Liu¹, Zhonghua Li³,
Changhua Liu^{1*} and Yan Sun^{1*}

¹College of Advanced Agriculture and Ecological Environment, Heilongjiang University, Harbin, China,

²State Key Laboratory of Tree Genetics and Breeding, College of Forestry, Northeast Forestry University, Harbin, China, ³Heilongjiang Greater Hinggan Mountains Region Agriculture Forestry Research Institute, Da Hinggan Ling, China

Heracleum moellendorffii Hance is a perennial herbaceous plant that is adaptable to cold environments and has both edible and medicinal value. Given that no reference genome for this species is available, we constructed a high-quality transcript isoform library using full-length transcriptome sequencing and conducted a comparative genomic analysis. Samples were obtained from plants that had been subjected to cold stress for 12, 24 and 36 hours (Cold_12, Cold_24, and Cold_36, respectively) and from control plants (Cold_0) that were not subjected to cold stress and used in transcriptome and nontargeted metabolome analyses. Compared with the genes expressed in CK (Cold_0), the number of differentially expressed genes (DEGs) in Cold_12, Cold_24, and Cold_36 increased gradually over time; plants subjected to 12, 24 and 36 hours of cold stress displayed 669, 6084, and 24,129 DEGs, respectively. The DEGs were clustered into 8 subclasses by k-means clustering; subclasses 2, 3, 4, and 7 were enriched in pathways related to “flavonoid biosynthesis”. Nontargeted metabolome analysis revealed that 3719 annotated metabolites were shared by all four groups of samples. We identified 1186, 1087, and 1097 differentially accumulated metabolites (DAMs) in three comparisons: Cold_12 vs. CK, Cold_24 vs. CK, and Cold_36 vs. CK, respectively. The DAMs were predominantly enriched in the “flavonoid biosynthesis pathway”. Through WGCNA, we obtained five modules and 29 flavonoid-related metabolites with extremely significant module–metabolite paired relationships (|correlation coefficient| > 0.9, $P < 0.01$). We analysed the DEGs and DAMs of the flavonoid biosynthetic pathway in *H. moellendorffii* Hance under cold stress and constructed a correlation network between transcription factors (TFs) and structural genes in the pathway. RT–qPCR was used to confirm the expression

of four hub genes from the WGCNA, six TFs, and 15 structural genes of the flavonoid biosynthetic pathway. These data provide a foundation for functional genomics studies of *H. moellendorffii* Hance and contribute to the study of the molecular mechanisms and transcriptional regulation of flavonoid accumulation by TFs under cold stress conditions in plants.

KEYWORDS

Heracleum moellendorffii Hance, cold stress, multi-omics, flavonoid biosynthesis, transcription factor

1 Introduction

Heracleum moellendorffii Hance, a perennial herbaceous geophyte belonging to the family Apiaceae and the genus *Heracleum*, is representative of Northeast China's mountain wild vegetables that possess both nutritional and economic value (Li et al., 2019). Although it belongs to the same family as the common culinary vegetable *Apium graveolens* L., *H. moellendorffii* presents a rich nutritional profile in multiple vitamin metabolites, amino acids, mineral substance, as well as flavonoids, coumarins, and saponins (Li et al., 1997; Wang et al., 2017). Thus, *H. moellendorffii* has medicinal and health benefits. Its aerial parts are often used as a source of medications for diabetes, and its roots, which have effects similar to those of *Angelica dahurica*, are commonly used in traditional Chinese medicine. It can be used to treat rheumatism and hypertension and has anti-fatigue, analgesic and anti-inflammatory effects (Liu et al., 2019a; Kim H, et al., 2019). *H. moellendorffii* thrives in environments with high organic matter content, deep soil, abundant groundwater, and dry, loose surfaces and is widely adaptable and susceptible to few pests and diseases. Its strong cold tolerance allows it to survive several hours at -4°C without damage (Li et al., 2019).

Cold stress is one of the most common abiotic stresses. When plants are subjected to cold stress, changes in cell morphology, subcellular structure, and protoplasm occur and are accompanied by disordered respiration, altered enzyme activities, and metabolic imbalances (Chinnusamy et al., 2007; Mahajan and Tuteja, 2015). Earlier studies of the cold stress response focused on physiological aspects such as morphological changes and damage, changes in the concentrations of metabolic products, and the activity of protective enzymes. In recent years, with the commercialization of various omics platforms such as high-throughput sequencing, gas chromatography–mass spectrometry (GC–MS), and liquid chromatography–mass spectrometry (LC–MS), multiomics approaches for dissecting the molecular mechanisms of plant cold stress have become increasingly prevalent (Cheong et al., 2020; Dong et al., 2020; Lee et al., 2020). These studies typically involve cold stress treatment of biological materials, sampling at various time points, and identification of differentially expressed genes (DEGs) through transcriptome sequencing, followed by clustering

of transcription factors (TFs) within the DEGs and GO and KEGG enrichment analysis. These studies provide a landscape of the entire transcriptome and DEGs, and the mechanisms through which organisms respond to cold stress are also described from the perspective of the gene families that are potentially involved in stress resistance (such as CBF, WRKY, and MYB) and the GO and KEGG categories related to stress resistance. Related studies have been conducted on crops such as rice (Pan et al., 2020), corn (Frey et al., 2020), wheat (Kruse et al., 2020), and peanuts (Bonthala et al., 2016). There are also reports on the molecular functions of particular DEGs, such as the ability of ShNAC1 overexpression in tomato plants to reduce cold tolerance; further research suggests that this transcription factor may be involved in ethylene biosynthesis and signal transduction pathways (Wang et al., 2024). Similarly, the cold-induced TF SIGRAS4 in tomato enhances cold stress tolerance by binding to the promoter region of SICBF and activating its expression without inhibiting normal plant growth (Liu et al., 2020). Studies of this type often focus on specific TFs involved in cold stress, such as MYB (An et al., 2020), DREB (Zhang et al., 2020), ICE1 (Tang et al., 2020), WRKY (Zhang et al., 2016), NAC (An et al., 2018), and bZIP (Liu et al., 2019b).

Metabolites are the ultimate manifestation of the transcription of genes into mRNAs, which are further translated into proteins, and they also extensively participate in the response of plants exposed to cold stress. In plants, cold stress may lead to the accumulation of toxic products due to metabolic imbalances. Soluble sugars act as important signalling molecules during cold stress (Rolland et al., 2006). Lipid content changes significantly when plants are exposed to cold stress (Cheong et al., 2020). Cold stress also induces the accumulation of amino acids and carbohydrates (Kou et al., 2018). Other metabolites, such as flavonoids (Song et al., 2024) and polyamines (Kovács et al., 2010), are also involved in the adaptation of plants to cold stress.

In this work, we focused on the involvement of flavonoids in the regulation of cold stress. Flavonoid biosynthesis is an important branch of the phenylpropanoid metabolic pathway and plays a crucial role in plant growth, development, and stress resistance (Daryanavard et al., 2023). Cold temperatures have been reported to activate the phenylpropanoid pathway, leading to the accumulation of flavonoid compounds and thereby reducing oxidative damage in

Poa crymophila Keng (Wang et al., 2021) and bell pepper (Xu et al., 2023). Sun et al. (2021) reported that the chalcone isomerase gene (CHI) is specifically upregulated in cold-tolerant kiwifruit varieties exposed to cold stress, and the flavonoid metabolic pathway is also specifically upregulated, increasing the ability of these plants to scavenge reactive oxygen species (ROS). Similarly, chalcone and stilbene synthases, which participate in the flavonoid biosynthetic pathway, are involved in the cold stress response in tobacco (Hu et al., 2022). Studies of this type often focus on differences in the expression of structural genes and flavonoid-related metabolites in the flavonoid metabolic pathway. In addition, it has been reported that TFs play a role in regulating flavonoid biosynthesis. For example, MYB transcription factors constitute an important class of regulators of the production of flavonoid metabolites. R2R3-MYB transcription factors often act as transcriptional activators or repressors to directly or indirectly regulate the expression of structural genes related to flavonoid biosynthesis (Zhao et al., 2013). The MBW (MYB-bHLH-WD40) complex is also involved in regulation (Xu et al., 2015). Other transcription factors, such as NAC, ZIP, and WRKY, have also been reported to participate in the regulation of the flavonoid biosynthesis pathway (Morishita et al., 2009; Malacarne et al., 2016; Zhang et al., 2023).

With the development of sequencing technology, genome assembly, an important foundation for functional genomics in the postgenomic era, has become increasingly common. The reference genome of celery, which also belongs to the family Apiaceae, has been released (Li et al., 2020; Song et al., 2021). Fundamental research on *H. moellendorffii* is relatively rare; there are only a few reports involving the use of omics methods to study seed dormancy characteristics and disease resistance (Li et al., 2019; Liu et al., 2019a, 2024b). The absence of a high-quality reference genome severely hinders the development of a functional genome that can be used in molecular breeding of *H. moellendorffii* and the in-depth development and utilization of its resources.

H. moellendorffii grows normally in low-temperature environments and adapts well to cold. This study is the first to construct a full-length transcriptome library of *H. moellendorffii* and thereby provide a reference for transcriptomics and other studies. Through time-course transcriptome and metabolome analyses, structural genes and regulatory factors of the flavonoid pathway that respond to exposure of the plant to low temperatures were subsequently identified. This research may lay a molecular foundation for subsequent genetic improvement and benefit the protection of germplasm resources of characteristic vegetables in the Apiaceae family, and the development and utilization of specialty crops.

2 Materials and methods

2.1 Plant material source and sampling

Seeds of *H. moellendorffii* were collected from the wild, cleaned, and air dried for later use. Subsequently, seeds with plumped grains were washed with clean water, disinfected with 75% ethanol for 30

seconds, and soaked in distilled water for 24 hours. After incubation at low temperatures to break dormancy, the seeds of *H. moellendorffii* were sown in pots within the greenhouse at the Horticultural Engineering Center of Northeast China Agricultural University (126°67' latitude, 45°73' longitude) in November 2021. For PacBio full-length sequencing, plant material was collected in 2024 from experimental fields in a greenhouse. We selected the roots, stems, leaves, flowers, and 7-day postpollination fruits of two-year-old *H. moellendorffii*. RNA was extracted from approximately 0.3–0.5 g of each tissue sample and used to construct a full-length transcriptome reference library. When the seedlings had developed two fully expanded true leaves, they were subjected to low-temperature stress at 4°C, which was set empirically (Guy et al., 2008; Jan et al., 2024). Leaf samples were collected after 0 h, 12 h, 24 h, and 36 h of low-temperature stress; these samples were designated CK (0 h), Cold_12 (12 h), Cold_24 (24 h) and Cold_36 (36 h), respectively. The samples were rapidly frozen in liquid nitrogen and stored at -80°C for use in future experiments. For transcriptome sampling, three biological replicates (approximately 0.5 g each) were obtained. For nontargeted metabolomics, six biological replicates were sampled; each sample weighed 0.3–0.5 g. Notably, three of the six metabolomic biological replicates were derived the same source of the RNA-seq samples, and the rest three were obtained from additional separate samples.

2.2 PacBio-based full-length transcriptome sequencing and data analysis

Total RNA was extracted from the root, stem, leaf, flower, and fruit of *H. moellendorffii*. The concentration and purity of the RNA were measured using a Nanodrop 2000 spectrophotometer. The integrity of the RNA was confirmed by agarose gel electrophoresis, and the RNA integrity number (RIN) was determined on an Agilent 2100 bioanalyzer, ensuring that all the samples had RIN values ≥ 8 . RNA samples that met the quality criteria were pooled in equal amounts and used to construct full-length transcriptome sequencing libraries. The final libraries each contained $\geq 5 \mu\text{g}$ total RNA at $\geq 300 \text{ ng}/\mu\text{L}$ and had OD_{260/280} ratios ranging from 1.8 to 2.2. Library construction was carried out using the SMARTerTM PCR cDNA Synthesis Kit (Pacific Biosciences, USA), with the following steps: 1) synthesis of full-length cDNA from mRNA via the SMARTerTM PCR cDNA Synthesis Kit (Pacific Biosciences, USA); 2) PCR amplification of the synthesized cDNA; 3) purification of the amplified full-length cDNA using PB magnetic beads; 4) end-repair of the purified full-length cDNA; 5) ligation of SMRT bell-shaped adapters to the repaired cDNA; 6) nuclease exonuclease digestion to prepare the cDNA for sequencing; and 7) further purification on PB magnetic beads to obtain the final sequencing library. Sequencing of the libraries to generate long-read sequences for transcriptome analysis was performed on the PacBio Sequel II platform.

The raw sequencing data were subjected to quality control. Sequences obtained from polymerase read fragments shorter than 50 base pairs (bp) or with accuracies less than 0.90 were discarded.

The adapter sequences were cleaved and filtered out to obtain subreads. Subreads shorter than 50 bp were filtered out, resulting in clean data. Circular consensus sequences (CCS) were extracted from the original sequences on the basis of the criteria of ≥ 3 full passes and sequence accuracy greater than 0.90, followed by statistical analysis and assessment of the data. Full-length consensus sequences were obtained via Iso-Seq3 in PacBio SMRT Analysis software. The analysis process consisted of three main stages: full-length sequence identification, isoform-level clustering to obtain consensus sequences, and polishing of the consensus sequences. The detailed steps were as follows: 1) the reads of insert (ROI) sequences were extracted from the original raw sequences, cDNA primers and polyA were filtered out, and the sequences were categorized into full-length and non-full-length sequences and chimeric and nonchimeric sequences on the basis of the presence of 3' primers, 5' primers, and poly A; 2) clustering of full-length sequences from the same isoforms was performed using the iterative isoform-clustering (ICE) algorithm, which forms clusters of similar sequences, with each cluster yielding a single consensus sequence; 3) clustering of non-full-length sequences was performed using the Quiver algorithm, followed by polishing of the consensus sequences to further select high-quality sequences. The Cogent software package was used to reconstruct “fake contigs”, and the consensus sequences obtained were aligned with the reconstructed “fake contigs” using pbmm2 v1.2.1 (<https://github.com/PacificBiosciences/pbmm2>) with default parameter settings. The collapse tool from the Iso-Seq3 pipeline was used to remove redundancy from the alignment results, ultimately yielding a reference transcriptome library for *H. moellendorffii*. The completeness of the final transcripts was assessed via BUSCO (eukaryota_odb10) (Simão et al., 2015). The full-length transcripts were functionally annotated using five major databases (NR, UniProt, GO, KEGG, and Pfam), and the transcription factors were annotated using the Plant Transcription Factor Database (PlantTFDB) (Jin et al., 2017). MISA was used to call simple sequence repeats (SSRs) (<https://webblast.ipk-gatersleben.de/misa/>). LncRNAs were predicted using the CPC, CNCI, PLEK and Pfam databases (Chang et al., 2021).

2.3 Comparative genome analysis

Protein coding prediction to identify protein isoforms was performed using TransDecoder v.5.7.1 (<https://github.com/TransDecoder/TransDecoder>). The longest protein sequence for each gene was selected for further analysis. We selected eight species from the same family (umbelliferae) or genus to *H. moellendorffii* for comparative genomic analysis, and *Solanum lycopersicum* was set as an outgroup. Orthofinder v2.5.5 software (Emms and Steven, 2019) was employed to classify protein sequences from the ten species into families, and the PANTHER-18.0 database (<https://pantherdb.org>) was utilized to annotate the identified gene families. Gene Ontology (GO) and KEGG enrichment analyses were subsequently conducted for gene families specific to *H. moellendorffii*. All single-copy orthologous genes were extracted and used to construct a phylogenetic tree.

Specifically, MAFFT v7.520 (Katoh et al., 2009) was used to align single-copy orthologue gene sequences with parameters set to `-localpair -maxiterate 1000`. Gblocks v0.91b (Talavera and Jose, 2007) was then used to trim the sequences (parameters: `-b5=h`). The concatenated sequences of all the aligned and trimmed sequences were assembled into a supergene. The best-fit model (JTT+F+R4) was determined using the ModelFinder tool within IQ-TREE (Nguyen et al., 2015). Finally, the maximum likelihood (ML) method was employed to construct the evolutionary tree using the optimal model, with bootstrap=1000 and *S. lycopersicum* designated as the outgroup. Divergence time estimation was performed using the MCMCTREE module in PAML v4.10 (Yang, 2007), with fossil calibration times obtained from the TimeTree website (<http://www.timetree.org/>). The contraction and expansion of gene families relative to the ancestor in *H. moellendorffii* were predicted using CAFE v5.1 (Han et al., 2013).

2.4 Illumina-based transcriptome sequencing and gene expression analysis

Total RNA was extracted from leaf tissue samples collected at 0 h, 12 h, 24 h, and 36 h following the beginning of exposure to cold stress. The RNA quality control methods used were the same as those used for the full-length transcriptome. High-quality RNA samples were treated with fragmentation buffer, and the mRNAs were randomly fragmented into fragments of approximately 300 base pairs (bp) using a Covaris Sonicator (Covaris, USA). Under the action of reverse transcriptase and with the aid of random primers, first-strand cDNA was synthesized using mRNA as a template, followed by synthesis of the complementary strand. The next step involved the addition of End Repair Mix (Illumina, San Diego, CA) to create blunt ends from the sticky ends, followed by the addition of an A base at the 3' end to facilitate the ligation of adapter sequences. The products obtained after adapter ligation were purified and size selected. The selected products were subjected to PCR amplification and further purified to obtain the final sequencing library. After quantification using Qubit 4.0, the library was sequenced on the NovaSeq X Plus platform.

The raw sequencing data were processed using fastp to obtain clean data (Chen et al., 2018), with data filtering criteria consistent with those of Liu et al. (2024a). The clean data were then aligned to the full-length transcriptome reference using HISAT2 (Kim D, et al., 2019) to generate mapped data (reads) for subsequent transcript assembly and quantification of gene expression. RSEM (Li and Dewey, 2011) was used to perform TPM (transcripts per million) quantification analysis for both genes and transcripts. Genes whose TPM expression was less than 1 were filtered out before PCA (Marini and Binder, 2019) and differential gene expression analysis were conducted via DESeq2 (Love et al., 2014). The criteria for selecting differentially expressed genes (DEGs) were $FDR < 0.05$ and $|\log_2FC| \geq 1$. GO and KEGG enrichment analyses were subsequently performed on the DEGs. The *k*-means clustering algorithm was employed to classify the aforementioned DEGs (Yang et al., 2022). For WGCNA, all expressed genes were used

as a base gene set. WGCNA was performed via TBtools II (power = 12, minModuleSize = 30, and MEDissThres = 0.25) (Chen et al., 2023).

2.5 Identification of CBF-dependent and flavonoid biosynthesis-associated pathway genes in *H. moellendorffii*

The methods used to identify genes in CBF-dependent pathway and associated with flavonoid biosynthesis were similar to published methods (Liu et al., 2024a). Briefly, the protein sequences of reported flavonoid biosynthesis-associated genes and CBF-dependent pathway genes were downloaded from NCBI databases. These BLASTP searches were used to call putative protein sequences using the full-length protein sequences of *H. moellendorffii* as queries. Finally, putative proteins lacking the same Pfam domains (Mistry et al., 2021) as the query genes were excluded from the analysis.

2.6 Metabolomics assay and data analysis

After the leaf samples had been dried at -50°C , 50 mg of each powdered sample was extracted with 400 μL of extraction solvent (methanol:water = 4:1, v/v) containing an internal standard (L-2-chlorophenylalanine at 0.02 mg/mL). The sample mixture was ground for 6 minutes at -10°C and 50 Hz in a cryogenic tissue grinder, followed by ultrasonic extraction at 5°C and 40 kHz for 30 minutes. The samples were then allowed to stand at -20°C for 30 minutes and then centrifuged at 4°C and $13000 \times g$ for 15 minutes; the resulting supernatant was transferred to a vial with an insert for instrumental analysis. The samples were analysed on a Thermo Fisher Scientific UHPLC-Q Exactive HF-X system (Shanghai Meiji Biopharma Technology Co., Ltd.). A 3- μL sample was separated on an HSS T3 column (100 mm \times 2.1 mm i.d., 1.8 μm) and introduced into the mass spectrometer. Mobile phase A consisted of 95% water and 5% acetonitrile (with 0.1% formic acid), and mobile phase B consisted of 47.5% acetonitrile, 47.5% isopropanol, and 5% water (with 0.1% formic acid). The flow rate was set to 0.40 mL/min, and the column temperature was maintained at 40°C . Mass spectrometry signals were acquired in both positive and negative ionization modes, with a mass scan range of 70–1050 m/z. The sheath gas flow rate was 50 psi, the auxiliary gas flow rate was 13 psi, the auxiliary gas heater temperature was 425°C , the capillary temperature was 325°C , the positive mode ion spray voltage was set at 3500 V, and the negative mode ion spray voltage was set at -3500 V, with a normalized collision energy of 20–40–60 eV in a cyclic manner. The resolution for the first mass spectrometer was 60000, and that for the second mass spectrometer was 7500; data-dependent acquisition (DDA) mode was used.

The raw LC–MS data obtained in the analysis were imported into the metabolomics processing software Progenesis QI (Waters Corporation, Milford, MA, USA) for baseline filtering, peak identification, integration, retention time correction, and peak

alignment, resulting in a data matrix of retention time, mass-to-charge ratio, and peak intensity. To obtain information on metabolites, the MS and MS/MS spectral information was matched against the public metabolite databases HMDB (<http://www.hmdb.ca/>) and Metlin (<https://metlin.scripps.edu/>) as well as against an in-house database built by Meiji. The data matrix was then preprocessed as follows: missing values were removed using the 80% rule, which retains variables with nonzero values in at least 80% of the samples, followed by imputation of missing values with the minimum value in the original matrix. To reduce errors caused by sample preparation and instrument instability, the response intensities of the MS peaks were normalized using the total sum normalization method, resulting in a normalized data matrix. Variables with relative standard deviation (RSD) $>30\%$ in the QC samples were removed, and the data were log₁₀ transformed to obtain the final data matrix for subsequent analysis. The preprocessed data matrix was subjected to principal component analysis (PCA) and orthogonal partial least squares-discriminant analysis (OPLS-DA) via the R package ropls (version 1.6.2). The stability of the model was evaluated using cross-validation. Significantly different metabolites were selected on the basis of the variable importance in the projection (VIP) values obtained from the OPLS-DA model and the *P* values from Student's *t* test, with $\text{VIP} > 1$ and $P < 0.05$ indicating significant differences. The differentially abundant metabolites were annotated using the KEGG database to identify the pathways in which they are involved. Pathway enrichment analysis was performed via the Python package scipy.stats, and significant biological pathways were identified through the use of Fisher's exact test.

2.7 Cojoint analysis of the transcriptome and metabolome

The transcriptome and metabolome data obtained from the same three biological replicate samples were used for conjoint analysis. Correlation analysis of the WGCNA modules with flavonoid metabolites was performed on the MetWare cloud platform (<https://cloud.metware.cn>) (parameters: powerEstimate=18, mergeCutHeight=0.25, minModuleSize=50). The correlation between transcription factors and flavonoid biosynthesis structural genes was calculated via the rcorr function of the R language Hmisc software package (Harrell and Harrell, 2019), and a correlation clustering heatmap was plotted via the heatmap plugin in TBtools II (Chen et al., 2023).

2.8 Reverse-transcription quantitative PCR assay

To confirm the expression of flavonoid biosynthesis-related genes, a total of 19 genes were selected for RT–qPCR. These included two hub genes from WGCNA, four potential interacting TFs and 13 structural genes associated with the flavonoid biosynthesis pathway. The primers used were designed using

Primer3 (<https://bioinfo.ut.ee/primer3-0.4.0/>) (Supplementary Table S5). The samples were subjected to RNA-seq. RT-qPCR was performed on a Roche Light Cycler 96 (Roche, Switzerland) using actin as an internal reference. The PCR system consisted of 10 μ L of Power SYBR[®] Green PCR Master Mix (Applied Biosystems, Foster City, CA, USA), 1 μ L of forward primer (10 μ M), 1 μ L of reverse primer (10 μ M), 1 μ L of cDNA template, and 2 μ L of nuclease-free H₂O. The reaction conditions are described in Liu et al. (Liu et al., 2018). The relative expression levels of the target genes were determined via the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001).

3 Results

3.1 Full-length sequencing and annotation

Pooled full-length transcriptome sequencing yielded approximately 29 Gb of clean subreads with an average read length of 1755 bp. After data filtering, a total of 340,273 full-length nonchimeric (FLNC) CCS reads were obtained, with an N50 value of 2032 bp. These reads were primarily distributed in the 500 bp–4000 bp range, with the highest frequency in the 1501 bp–2000 bp interval (Figure 1A). After redundant reads were removed, a total of 82,673 high-quality isoforms were obtained. The BUSCO evaluation value was 83.9%, suggesting that the transcript reference library has potential value. The full-length transcript sequences were aligned with five databases (NR, UniProt, GO, KEGG, and Pfam), and 77,075 transcripts were found to be annotated in at least one of these databases. The majority of the isoforms (~93%) were annotated via either UniProt or NR (Figure 1B). Subsequently, 29,187 SSR loci, primarily of types p1 and p2, were identified (Figure 1C). We continued to predict potential long noncoding RNAs (lncRNAs) and ultimately identified 1041 potential lncRNAs that were supported by all four prediction results (Figure 1D). Given the crucial role of transcription factors (TFs) in stress responses, we also predicted TFs for this transcriptome library. A total of 2,020 TFs belonging to 49 classes were identified; those whose abundance represented more than 5% of the total were MYB (9.80%), bHLH (7.82%), NAC (6.14%), bZIP (5.59%), GRAS (5.59%), C2H2 (5.64%), and C3H (5.30%) (Figure 1E).

3.2 Gene family and phylogenetic evolution analysis

Among the 82,673 isoforms obtained for *H. moellendorffii*, a total of 58,675 were predicted to encode proteins. Further gene family clustering analysis revealed 35,992 families; *H. moellendorffii* contained 15,238 of these families, and 1,210 of those were species-specific, encompassing 3,518 genes (Figure 2A; Supplementary Figure S1C). The species-specific gene families enriched in GO or KEGG terms were related to “fatty acid metabolic process”, “starch and sucrose metabolism”, and “biosynthesis of amino acids” (Supplementary Figures S1A, B). Among all the gene families, 581

single-copy orthologous genes were identified. These were used to construct a phylogenetic tree for nine species within the family Apiaceae and the outgroup *S. lycopersicum* (Figure 2B). The results indicated that *H. moellendorffii* is most closely related to *H. sosnowskyi*, with a divergence time of approximately 6 million years ago (Mya); its divergence occurred after the divergence of the common ancestor of *A. graveolens* and *H. moellendorffii* (~14 Mya). In the genome of *H. moellendorffii*, 547 genes underwent expansion, and 633 genes experienced contraction. The contracted genes were enriched in the biosynthesis of secondary metabolites, plant hormone signal transduction, and starch and sucrose metabolism (Figure 2C), whereas the expanded genes were enriched in metabolic pathways, the biosynthesis of secondary metabolites and carbon metabolism (Figure 2D).

3.3 Differentially expressed genes identified through RNA-seq

To further assess the performance of the full-length transcriptome library and investigate the mechanisms of the cold stress response in *H. moellendorffii*, we conducted transcriptome sequencing under cold stress conditions. As shown in Figure 3A, compared with those of the control at 0 h, the leaves of plants subjected to cold stress presented evident wilting, with increased severity as the duration of the treatment increased. Transcriptome sequencing generated a total of 78.37 Gb of clean data, with each sample exceeding 6.15 Gb and with a Q30 value greater than 96% (Supplementary Table S1). PCA revealed a high degree of similarity among the biological replicates (Figure 3B). Violin plots indicated that samples from the plants subjected to cold treatment for three different amounts of time (Cold_12, Cold_24, and Cold_36) presented higher gene expression levels than did the control CK, particularly Cold_24 and Cold_36 (Figure 3C). We performed DEG analysis using six different pairwise comparisons. In the comparisons between Cold_12 vs. CK, Cold_24 vs. CK, Cold_36 vs. CK, Cold_24 vs. Cold_12, Cold_36 vs. Cold_12 and Cold_36 vs. Cold_24, we identified 669 (Supplementary Figure S2A), 6084 (Supplementary Figure S2B), 24,129 (Supplementary Figure S2C), 1928 (Supplementary Figure S2D), 18,927 (Supplementary Figure S2E), and 18,177 (Supplementary Figure S2F) DEGs, respectively. The shared and unique DEGs identified in the six comparisons are shown in Figure 3D, resulting only five commonly shared DEGs and 54, 1042, 4358, 180, 1619, 2138 unique DEGs in Cold_12 vs. CK, Cold_24 vs. CK, Cold_36 vs. CK, Cold_24 vs. Cold_12, Cold_36 vs. Cold_12 and Cold_36 vs. Cold_24, respectively. The number of DEGs increased progressively with time in the plants that were subjected to cold stress, indicating a progressively increased response to cold stress in *H. moellendorffii*. In all the pairwise comparisons, the number of upregulated genes exceeded the number of downregulated genes (Figure 3E). The numbers of upregulated genes in Cold_12 vs. CK and in Cold_24 vs. Cold_12 were 328 and 984 respectively, whereas the number of upregulated genes observed in the other pairwise comparisons was 3400 minimally in Cold_12 vs. CK, reaching a maximum of 13,250 in Cold_36 vs. CK. This suggests

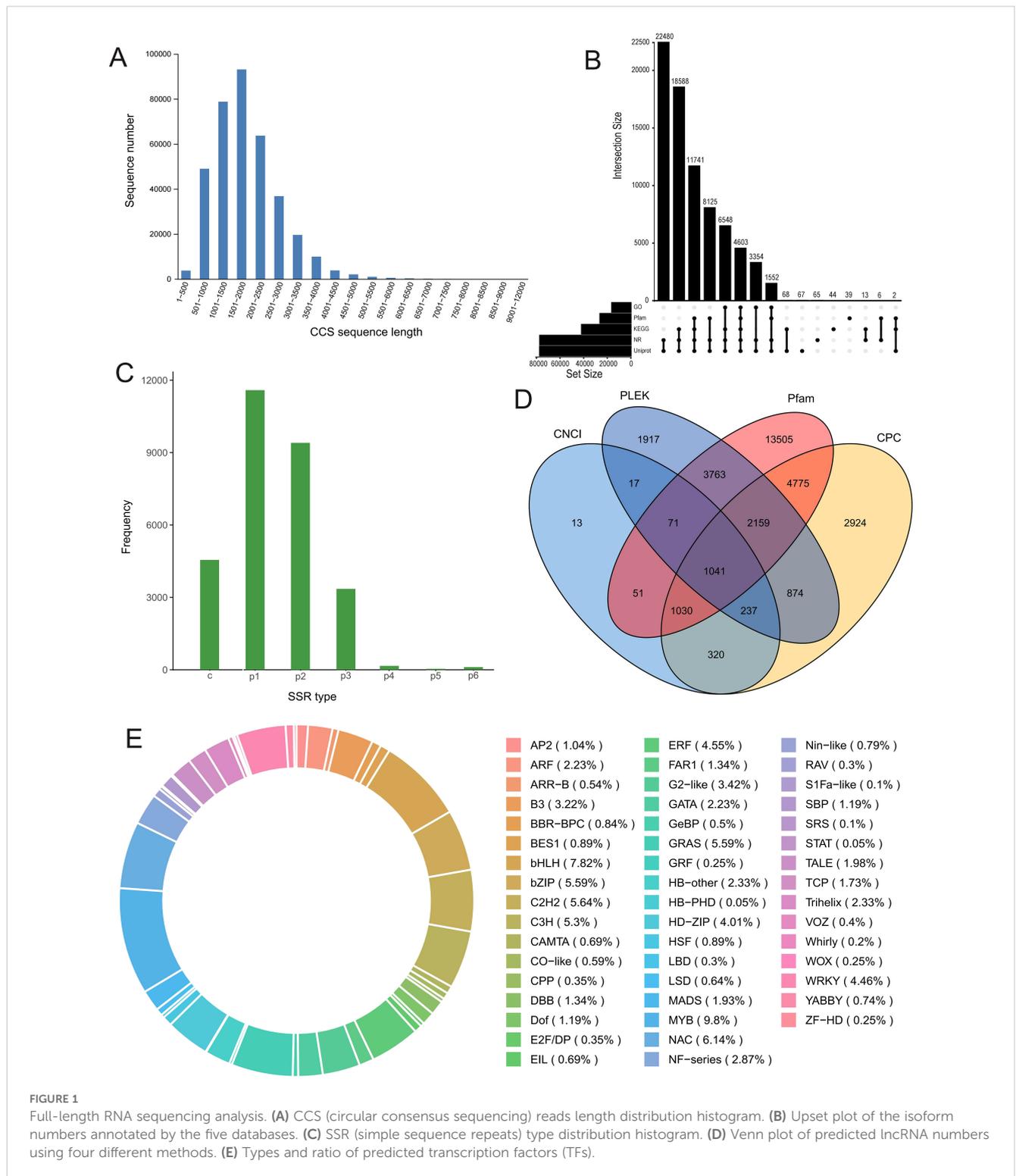


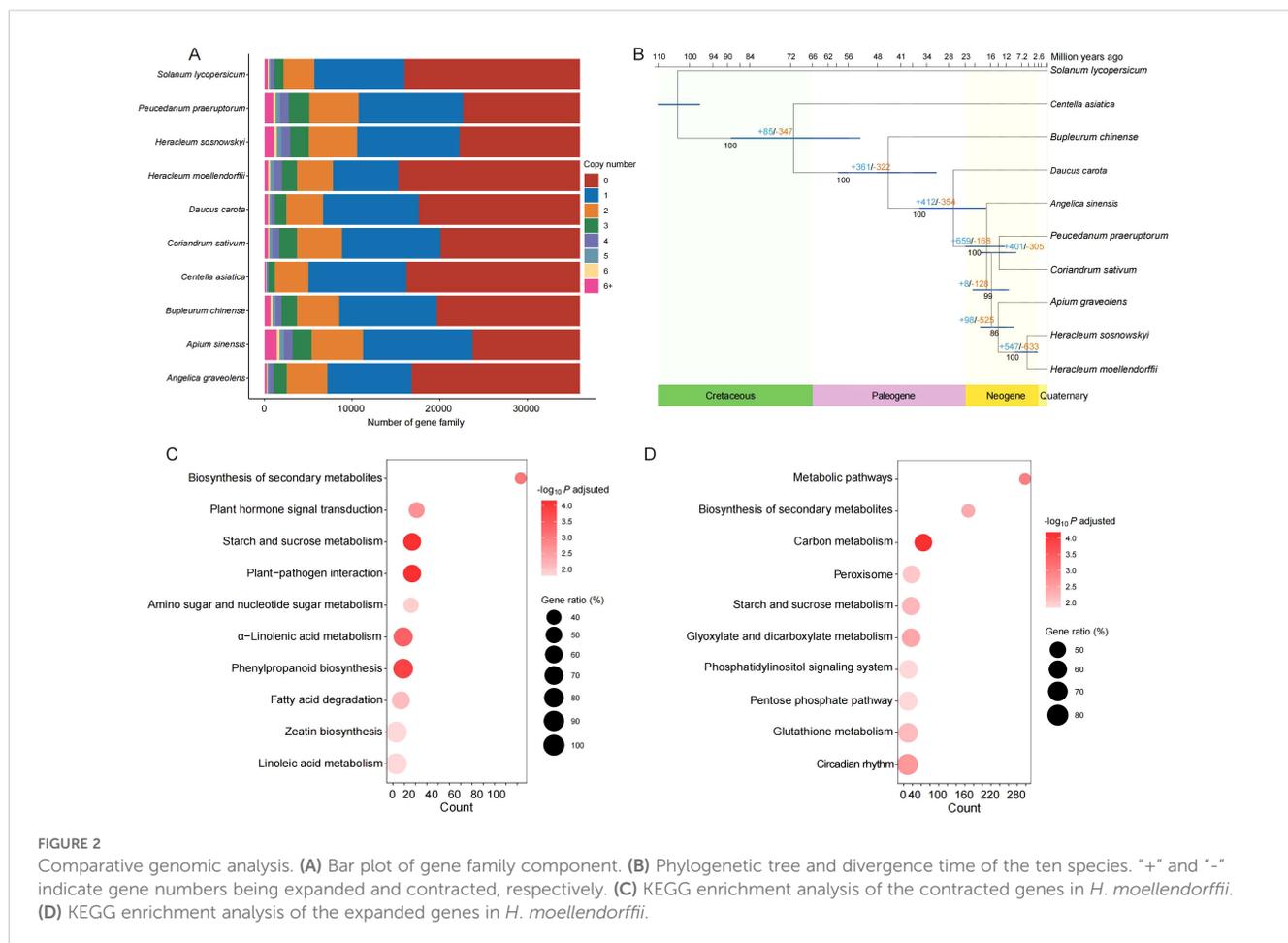
FIGURE 1

Full-length RNA sequencing analysis. (A) CCS (circular consensus sequencing) reads length distribution histogram. (B) Upset plot of the isoform numbers annotated by the five databases. (C) SSR (simple sequence repeats) type distribution histogram. (D) Venn plot of predicted lncRNA numbers using four different methods. (E) Types and ratio of predicted transcription factors (TFs).

that a milder transcriptional response occurred at the first two time points (12 h and 24 h) and that a more intense response occurred at 36 h.

To further analyse the expression trends of different DEGs, we performed K-means clustering analysis on all DEGs across the comparison groups. This resulted in 8 subclasses, representing 8 distinct expression patterns (Figure 3F); KEGG enrichment analysis

was conducted for each subclass gene set (Supplementary Figure S3). The top enriched pathways included “oxidative phosphorylation” (subclass 1), “photosynthesis” (subclass 2), “ribosome biogenesis in eukaryotes” (subclasses 4 and 6), “autophagy” (subclass 5), “protein processing in the endoplasmic reticulum” (subclasses 7 and 8), “flavonoid biosynthesis” (subclass 3), and the typical “plant hormone signal transduction” pathway (subclasses 6, 7, and 8).



Notably, subclass 2, subclass 3, subclass 4, and subclass 6 enriched pathways are related to flavonoid biosynthesis, implying that flavonoid biosynthesis plays a significant role in the cold stress response of *H. moellendorffii*. Additionally, genes involved in the CBF-dependent pathway, including the CBF-type genes Hmo.29573, Hmo.121080, and Hmo.34904 were rapidly upregulated in response to cold stress, whereas BTF (Hmo.21414), BYBC1 (Hmo.114249), and PIF (Hmo.112347) were downregulated (Supplementary Figure S4).

3.4 Metabolomic analysis and identification of differentially accumulated metabolites

Using samples from the same time points as those used for RNA-seq, we also conducted metabolomic experiments. PCA revealed minor differences among the six biological replicates, whereas obvious differences were observed among the samples obtained at different time points (Figure 4A). The nontargeted metabolomic assay identified 3775 annotated metabolites, 3719 of which were detectable in samples from all four time points (Figure 4B). As in the transcriptome analysis, DAMs were identified for six pairwise comparisons: Cold₁₂ vs. CK, Cold₂₄ vs. Cold₁₂, Cold₂₄ vs. CK, Cold₃₆ vs. Cold₂₄, Cold₃₆ vs. Cold₁₂, and Cold₃₆ vs. CK; 1186, 1062, 1087, 1174, 1234, and

1097 DAMs, respectively, were identified, in these comparisons (Figure 4C). Excluding the category classified as "others", the top 5 categories of DAMs in all the comparison groups were carboxylic acids and derivatives, organooxygen compounds, fatty acids, prenol lipids, and flavonoids (Figures 4D-I). Compared with those in CK, the DAMs in Cold₁₂, Cold₂₄, and Cold₃₆ were enriched predominantly in the following pathways: "nucleotide metabolism", "flavonoid biosynthesis", "pyrimidine metabolism", and "purine metabolism". Additionally, DAMs that differed among Cold₁₂, Cold₂₄, and Cold₃₆ were enriched in "flavonoid biosynthesis" (Supplementary Figure S5), indicating that flavonoids may be important metabolites in *H. moellendorffii*'s response to cold stress.

3.5 Correlations between modules identified by WGCNA and flavonoid metabolites

To further explore the associations between gene expression and flavonoid metabolites, we conducted weighted gene coexpression network analysis (WGCNA) and calculated the correlations between expression modules and flavonoid metabolites. Co-expression module was detected by hierarchical cluster tree (Supplementary Figure S6A), and a total of 22 modules

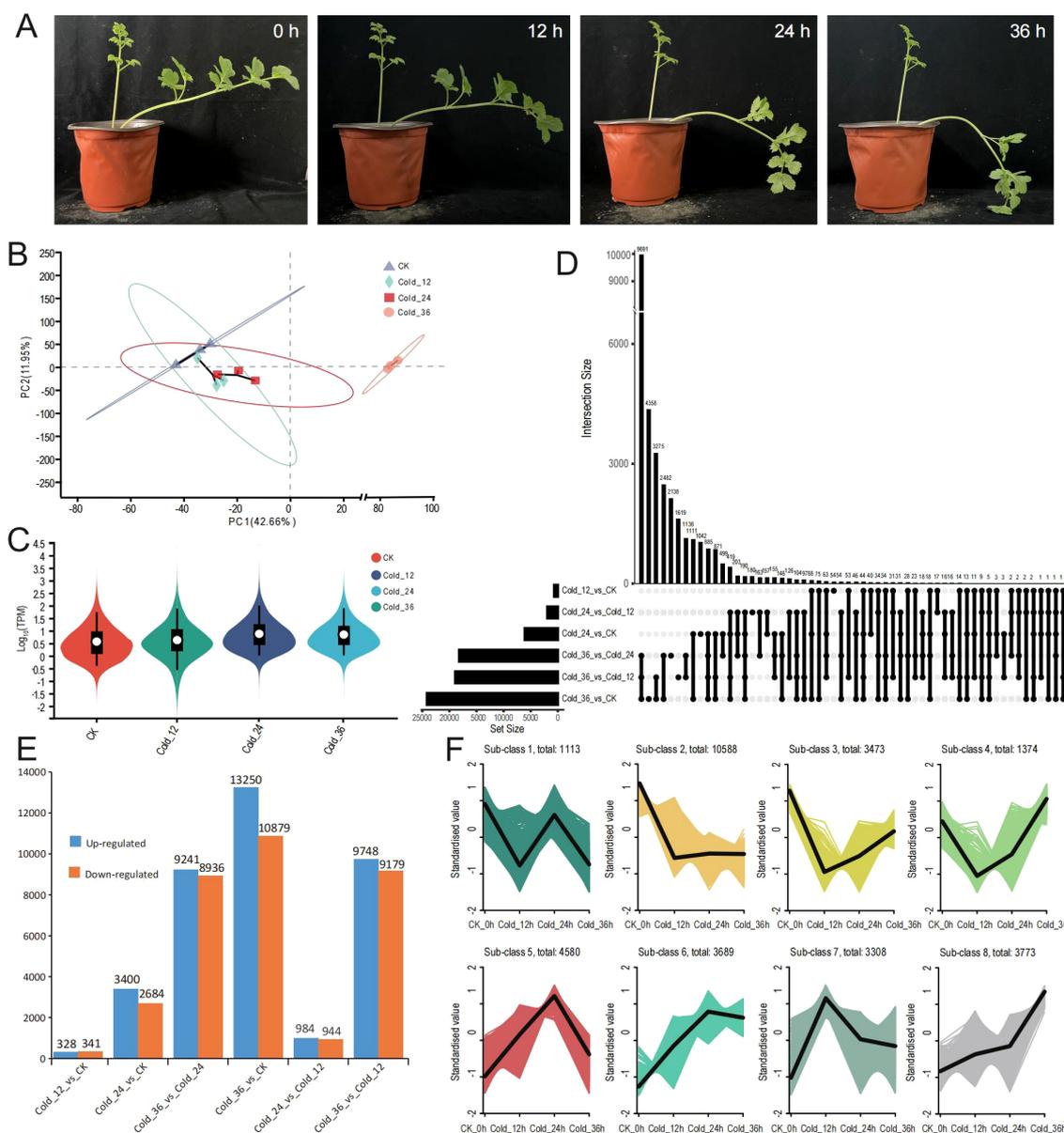
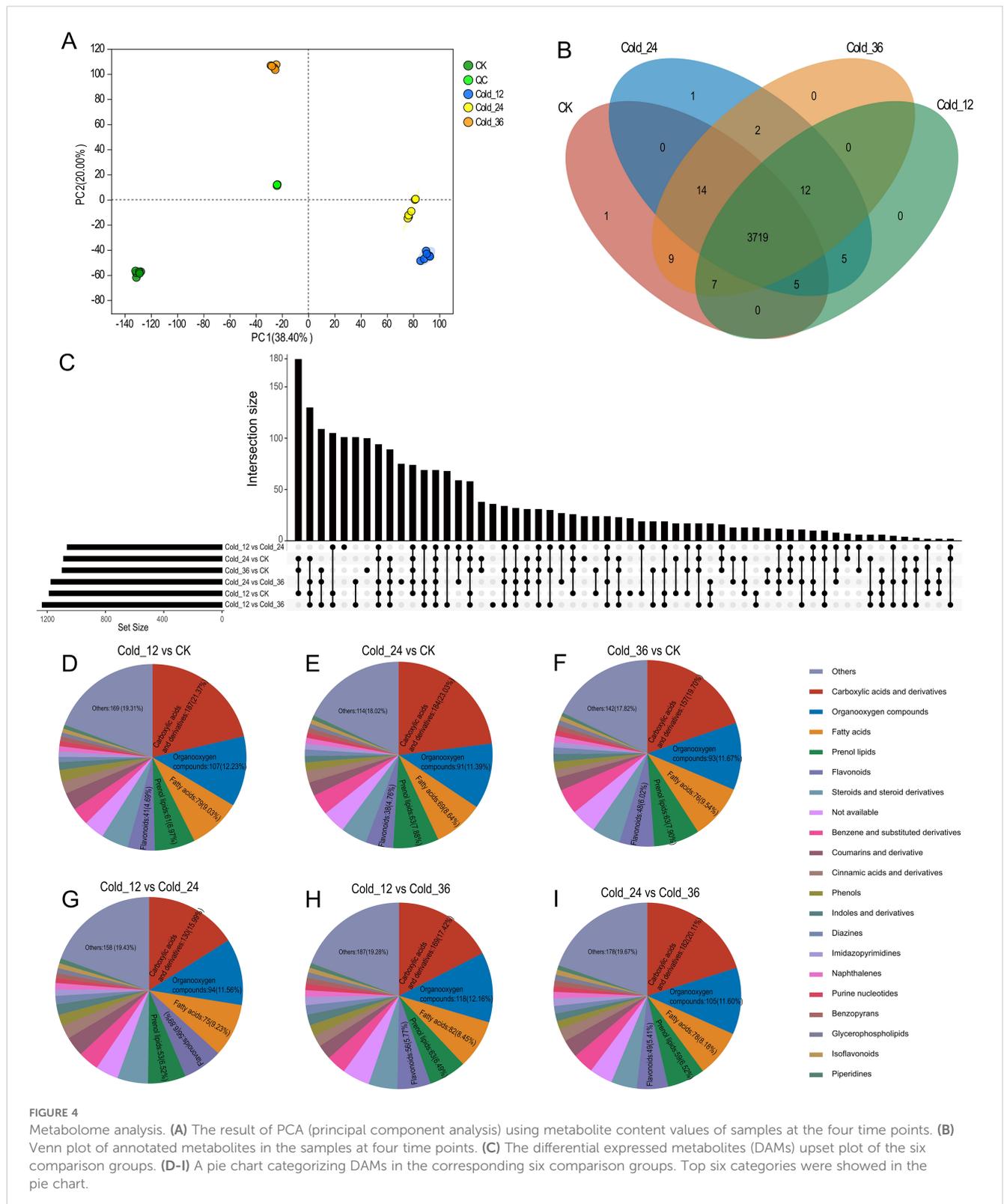


FIGURE 3
 The sample appearance and RNAseq analysis. **(A)** The performance of *H. moellendorffii* at 0 h, 12 h, 24 h and 36 h after 4°C treatment. **(B)** The result of PCA (principal component analysis) using gene expression TPM (transcripts per million) values of samples at the four time points. **(C)** Violin plot of gene expression TPM (transcripts per million) values. Different colors represent different samples. **(D)** Six comparison groups' upset plot of differentially expressed genes (DEGs). **(E)** Six comparison groups' column diagram of up-regulated and down-regulated gene numbers. **(F)** Eight sub-class DEGs using *k*-means cluster method.

were identified through WGCNA (Supplementary Figure S6B); each module identified five potential hub genes (Supplementary Table S2). A total of 42 flavonoid metabolites were involved in the correlation analysis with the modules. There were 36 module–metabolite paired relationships with correlation coefficients greater than 0.8 or less than -0.8 with $P < 0.01$ (14 negative correlations and 22 positive correlations), covering five modules (pink, blue, light green, gray60, midnight blue, and red) and 29 metabolites. Notably, the pink module was highly significantly positively or negatively correlated with 13 flavonoid-

related metabolites, and the blue module was highly significantly positively or negatively correlated with 11 flavonoid-related metabolites, suggesting that genes in these two modules play important roles in flavonoid metabolism in *H. moellendorffii* under cold stress conditions. The highest positive correlation coefficient, 0.93, was found for the relation pairs pink & cyanidin ($P = 1.2e-05$) and blue & quercetin ($P = 1.2e-05$). The highest negative correlation coefficient, -0.93, was found for the relationship between blue and lucidenic acid M ($P = 1.2e-05$) (Figure 5).



3.6 Dissection of the flavonoid metabolic pathway involved in cold stress

The flavonoid biosynthetic pathway has been demonstrated to participate in the cold stress response of plants. Therefore, we

utilized the isoforms from the full-length transcriptome to identify genes involved in the flavonoid biosynthesis pathway in *H. moellendorffii*. A total of 108 genes were identified; 90 of these encode constitutive enzymes, 5 encode transcription factors (TFs), and 13 encode transporters (Supplementary Table S3). We

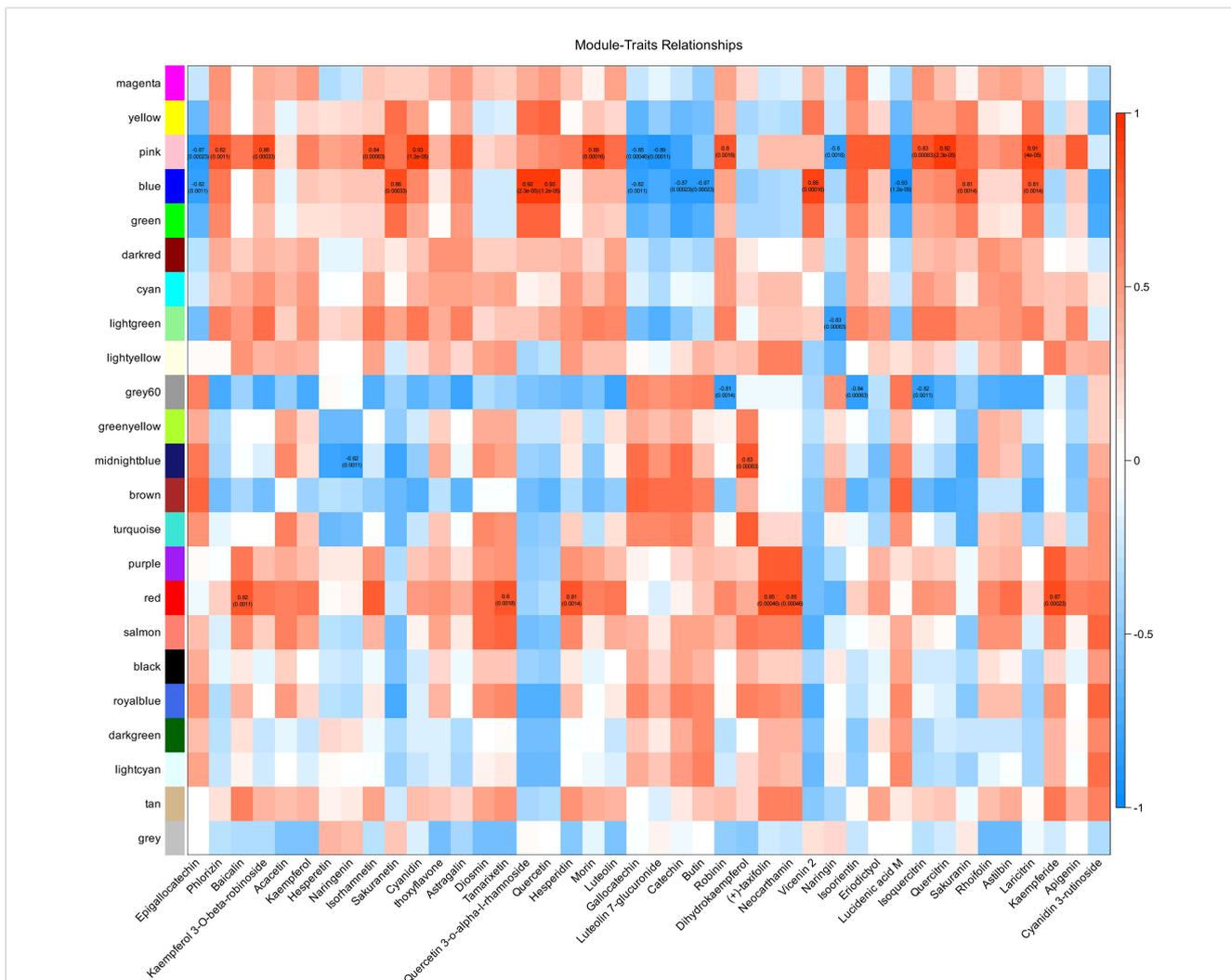


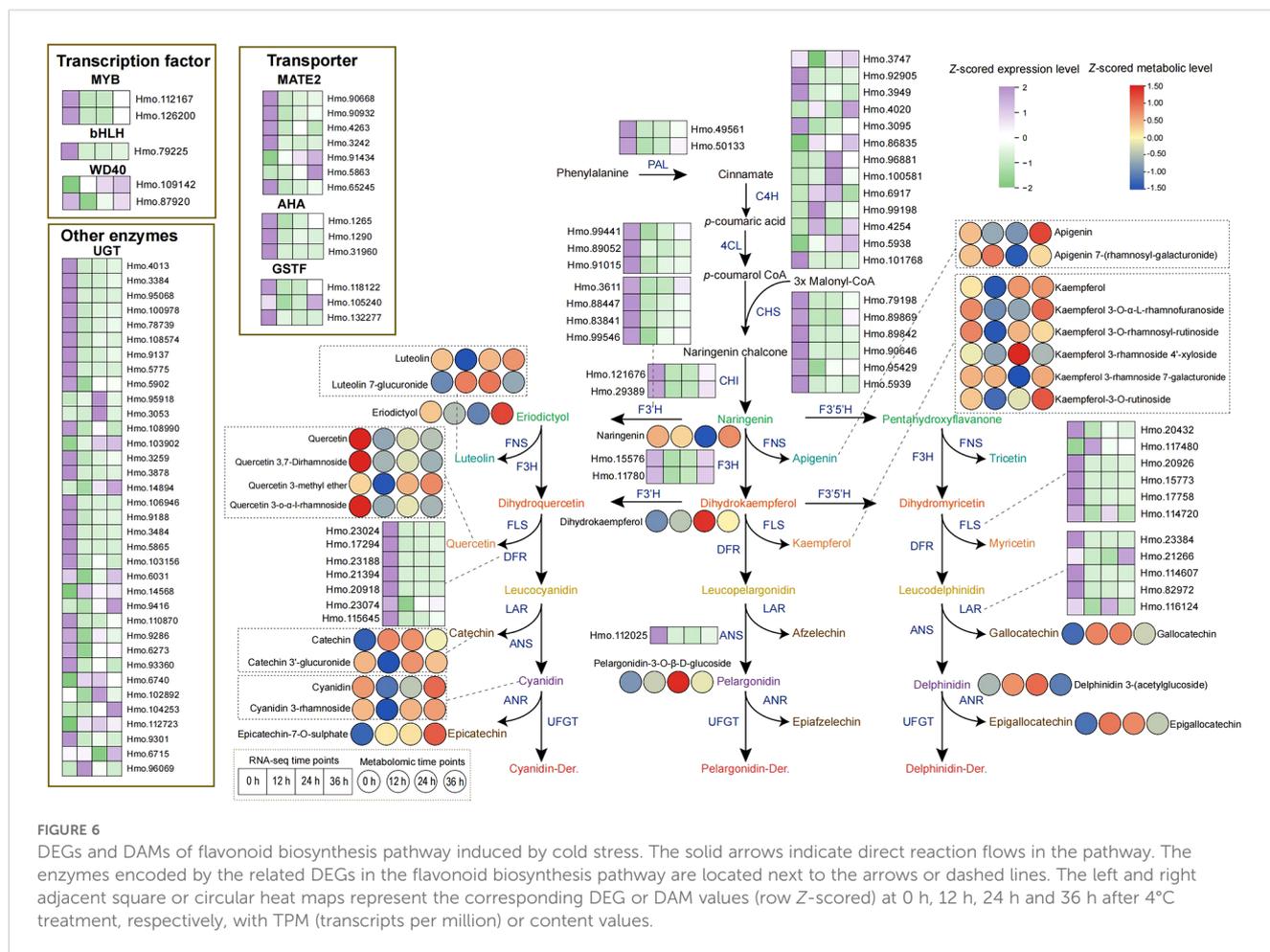
FIGURE 5 Correlation analysis between gene modules by weighted correlation network analysis (WGCNA) and flavonoid metabolites. The left different colors represent different modules. In each correlation block, the number at the top indicates the correlation values, and the corresponding *P* values are indicated in parentheses below.

subsequently integrated the differentially expressed flavonoid biosynthesis genes and the differentially expressed flavonoid-related metabolites into a pathway diagram (Figure 6). The DEGs were found to include the structural genes PAL, C4H, 4CL, CHS, CHI, F3H, DFR, ANS, FLS, and LAR. Genes such as Hmo.49561 and Hmo.50133 (belonging to PAL), Hmo.121676 and Hmo.29389 (belonging to CHI), and Hmo.15576 and Hmo.11780 (belonging to F3H) showed decreased expression at 12 h, followed by sharply increased expression at 24 h and 36 h. The C4H genes Hmo.3747, Hmo.4020, and Hmo.96881, as well as the UGT-type genes Hmo.95918 and Hmo.3053, presented relatively high expression levels after 24 h and 36 h of treatment. Five transcription factors (including MYB, bHLH, and WD40) and 13 transporters were also found to be differentially expressed. Like the expression patterns of structural genes of the PAL, CHI, and F3H types, expression of the AHA-type transporter Hmo.1265 also tended to decrease under cold stress conditions, followed by a gradual increase. Metabolite accumulation also exhibited a variety of patterns. Galliccatechin,

luteolin 7-glucuronide, and pelargonidin content rapidly increased after cold stress began and then decreased at 36 h but remained higher than in the control. Epicatechin-7-O-sulfate content increased after cold stress, and this compound accumulated gradually until reached a peak level at 36 h. The content of epigallocatechin, however, initially increased sharply under cold stress conditions and then gradually decreased. Overall, the accumulation of most flavonoid metabolites at 36 h of cold stress was greater than that in the control.

3.7 Interaction network between structural genes associated with flavonoid biosynthesis and TFs

TFs that interact with structural genes are widely involved in the regulation of metabolite synthesis. Therefore, we conducted a correlation analysis between the differentially expressed TFs and

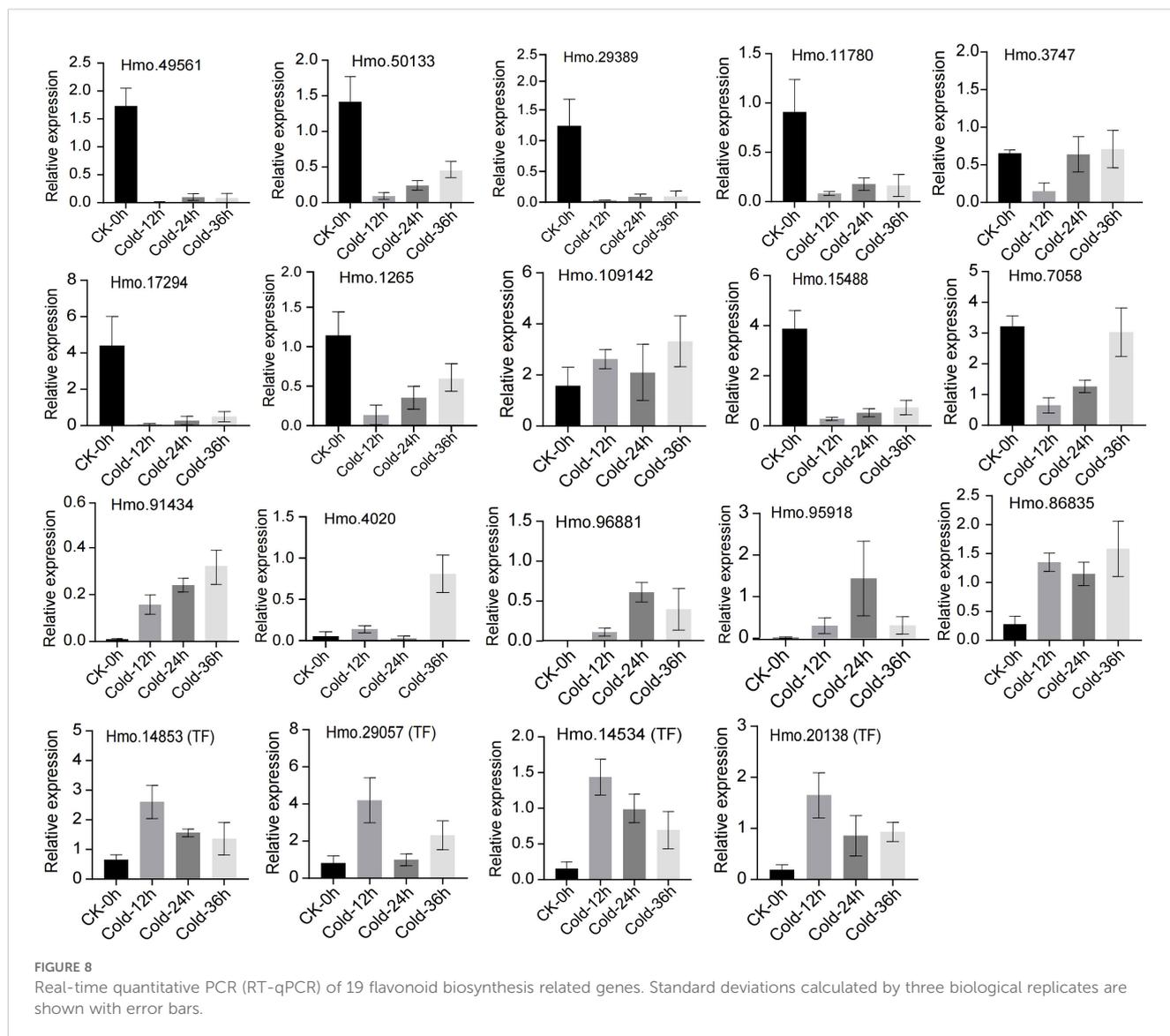


DEGs in the flavonoid biosynthesis pathway and used the results to construct an interaction network (Figure 7; Supplementary Table S4). TFs that potentially interact with C4H (Hmo.86835) and TT12 MATE2 (Hmo.91434) were the most numerous (33 and 22 TFs, respectively), and most of these interactions were positive, suggesting that these interactions may serve as regulatory hubs in response to cold stress. Additionally, the structural genes Hmo.92226 (F3'H), Hmo.103902 (UGT), and Hmo.99198 (C4H) were significantly positively correlated with multiple TFs, including TFs belonging to the NAC, WRKY, and MYB types. In contrast, Hmo.17294 (DFR) was significantly negatively correlated with bHLH (Hmo.23102), GRAS (Hmo.132591, Hmo.131475), ERF (Hmo.29707), and Dof (Hmo.108251, Hmo.92477) expression. Importantly, a single TF can interact with multiple structural genes. For example, expression of the C2H2 TF Hmo.109860 was significantly correlated with expression of Hmo.112167 (MYB), Hmo.23188 (DFR), Hmo.5865 and Hmo.8859 (UGT), and Hmo.99546 (F3'5'H). The expression levels of two MYB-type TFs, Hmo.112167 and Hmo.29251, correlated significantly with those of four flavonoid-related genes (Hmo.23188, Hmo.5865, Hmo.8859, and Hmo.99546) and six flavonoid-related genes (Hmo.3259, Hmo.49561, Hmo.50133, Hmo.9188, Hmo.92226, and Hmo.99441), respectively. Hmo.29759 expression was correlated with that of three UGTs (Hmo.3259, Hmo.5775 and Hmo.9188),

one PAL gene (Hm0.49561), and one F3'H gene (Hm0.92226). Hmo.49415 (TCP-type TF) expression was correlated with the expression of six different types of structural genes (Hmo.115645, Hmo.1290, Hmo.49561, Hmo.5939, Hmo.9188, and Hmo.92226).

3.8 Real-time quantitative PCR confirmation of the expression of flavonoid biosynthesis pathway genes

To confirm the expression of flavonoid biosynthesis pathway genes and their potential regulators, we selected 19 genes for RT-qPCR (Figure 8). The expression correlation (R^2) between RNA-seq and RT-qPCR was 0.8954 (Supplementary Figure S7). The expression of most structural genes, including Hmo.49561 (PAL), Hmo.50133 (PAL), Hmo.29389 (CHI), Hmo.11780 (F3'H), Hmo.3747 (C4H), Hmo.17294 (DFR) and Hmo.1265 (AHA), decreased sharply after 12 h of treatment and then increased with increasing exposure time until 36 h. Hmo.15488 and Hmo.7058 presented similar expression patterns. In contrast, the structural genes Hmo.4020 (C4H), Hmo.96881 (C4H), Hmo.86835 (C4H) and Hmo.95918 (UGT) were strongly upregulated after cold treatment, especially at 24 h and 36 h. Notably, Hmo.91434, which encodes a TT12 MATE2 transporter, showed consistently



exposed to cold stress, we characterized the DEGs (Figure 3) and DAMs (Figure 4) of *H. moellendorffii*. The number of DEGs increased gradually over time, with 19 transcription factors differentially expressed at Cold_12h compared with CK, 189 at Cold_24h, and 784 at Cold_32h. The results suggest that TFs sequentially initiated the response of *H. moellendorffii* to cold stress. The clustered DEGs were enriched in classic stress response pathways, such as “plant hormone signal transduction” (Supplementary Figure S3) (Tian et al., 2022). Among these DEGs, we also observed upregulation and downregulation of genes encoding proteins that participate in the CBF-dependent pathway (Supplementary Figure S4) (Hwarari et al., 2022). However, the hub genes identified through WGCNA did not include genes related to the CBF pathway, implying that although the CBF pathway in *H. moellendorffii* responds to cold stress, genes that are not dependent on the CBF pathway may play a critical role. Similarly, the DAMs that are found in plants exposed to

cold stress can be divided into several different categories, mainly carboxylic acids and derivatives, organooxygen compounds, and fatty acids (Figures 4D–I) (Xu et al., 2022). Cold stress induced DAMs (Cold_12 vs. CK, Cold_24 vs. CK, Cold_36 vs. CK) were also enriched in KEGG pathways such as “nucleotide metabolism” and “flavonoid biosynthesis”, consistent with previous reports (Sun et al., 2021; Song et al., 2024) (Supplementary Figure S5). Strong correlations were detected between different expression modules and flavonoid metabolites via WGCNA. In addition, the *H. moellendorffii* specific gene families enriched in “fatty acid metabolic process”, “starch and sucrose metabolism”, and “biosynthesis of amino acids” (Supplementary Figure S1A, B), which may play roles in cold stress adaptation (Yue et al., 2015; Golizadeh and Kumleh, 2019). These results highlight the complexity of the transcriptional and metabolic responses of *H. moellendorffii* to cold stress. The transcriptome sequencing and metabolome data reported in this study provide an

opportunity to elucidate the regulatory mechanisms involved in the cold response of *H. moellendorffii*.

4.3 Flavonoid biosynthesis and its regulation by interactions with TFs

Structural genes can interact with TFs to regulate metabolic pathways, thereby increasing plant resistance to abiotic stress. The flavonoid pathway has been shown to be involved in various stress responses, including the response to cold stress (Song et al., 2024). However, TFs related to flavonoid biosynthesis in *H. moellendorffii* have not been identified. Moreover, the relationships among the TFs that regulate flavonoid accumulation have not been elucidated. The transcriptome and metabolome results obtained in this study both indicate that the flavonoid pathway plays a significant role in the response of *H. moellendorffii* to cold stress. We identified 108 DEGs (including MYB, bHLH, and WD40) and 26 DAMs in the flavonoid biosynthesis pathway (Figure 6). Flavonoid metabolism is regulated by TFs such as MYBs and WRKYs (Naik et al., 2022; Wang et al., 2022b; Zhao et al., 2024). Interestingly, we observed these TFs among the DEGs (Figure 3), and in the correlation network analysis, we identified 196 TF–structural gene interaction pairs (correlation coefficient > 0.9, $P < 0.01$); these included MYB (Hm0.112167, Hm0.29251), bHLH (Hm0.29759), and WRKY (Hm0.88273, Hm0.38135, and others), and all of them correlated significantly with structural genes of the flavonoid pathway (Figure 7). These TFs could be potential targets for genetic modification; their identification provides a foundation for future research on the transcriptional regulation of flavonoid metabolism in *H. moellendorffii*.

5 Conclusions

To investigate the cold response of *H. moellendorffii*, a species for which no reference genome is currently available, PacBio-based full-length transcriptome sequencing was performed, a transcript isoform library was established, the evolutionary status of *H. moellendorffii* was clarified, and multiomics analysis was conducted via RNA-seq and metabolome assays during four key stages of cold stress treatment. A high-quality transcript isoform library was successfully constructed. Multiomic analyses revealed that different transcriptomes and metabolite expression patterns are found in the leaves of *H. moellendorffii* after cold treatment. Subclasses of both DEGs and DAMs were enriched in the flavonoid biosynthetic pathway. WGCNA also revealed expression modules harbouring hub genes that are highly correlated with cold stress, and some of these modules are strongly correlated with flavonoid metabolites. The deciphering of the flavonoid biosynthetic pathway in *H. moellendorffii* identified differentially expressed structural genes and metabolites. The roles of TFs in the flavonoid biosynthesis process were identified, and several potential TFs associated with flavonoid-related genes were

screened. These results provide a foundation for elucidation of the functional genome of *H. moellendorffii* and the regulatory mechanisms involved in flavonoid biosynthesis in that species.

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

GL: Writing – original draft, Writing – review & editing. HG: Formal Analysis, Writing – review & editing. YSo: Writing – review & editing. HW: Methodology, Software, Writing – original draft. DZ: Formal Analysis, Resources, Writing – review & editing. YW: Project administration, Validation, Writing – review & editing. SL: Methodology, Resources, Writing – review & editing. ZL: Resources, Writing – review & editing. CL: Supervision, Validation, Writing – review & editing. YSu: Supervision, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This research was sponsored by the Basic Scientific Research Foundation of Heilongjiang Provincial Universities (2022-KYYWF-1072), 2024 Heilongjiang Province Higher Education Teaching Reform Project (Title: Reform and Practice of Talent Cultivation in Seed Science and Engineering with Deep Integration of Industry and Education, Project Leader: GL).

Acknowledgments

Special thanks to Liu Chengxue from the Heilongjiang Greater Hinggan Mountains Region Agriculture Forestry Research Institute for providing the materials.

Conflict of interest

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

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.2025.1544898/full#supplementary-material>

SUPPLEMENTARY FIGURE 1

Gene family analysis. (A) KEGG enrichment analysis of *H. moellendorffii* specific family. (B) GO enrichment analysis of *H. moellendorffii* specific family. (C) Upset plot of the gene clusters using the ten species.

SUPPLEMENTARY FIGURE 2

Volcano plot of differentially expressed genes (DEGs) in the six comparison groups. (A) Cold₁₂ vs. CK; (B) Cold₂₄ vs. CK; (C) Cold₃₆ vs. CK; (D) Cold₂₄ vs. Cold₁₂; (E) Cold₃₆ vs. Cold₁₂; (F) Cold₃₆ vs. Cold₂₄.

SUPPLEMENTARY FIGURE 3

KEGG enrichment analysis of 8 sub-class differentially expressed genes (DEGs) by *k*-means analysis.

SUPPLEMENTARY FIGURE 4

Clustering heat map of CBF-dependent differentially expressed genes (DEGs).

SUPPLEMENTARY FIGURE 5

KEGG enrichment analysis of differentially accumulated metabolites (DAMs).

SUPPLEMENTARY FIGURE 6

Weighted correlation network analysis (WGCNA) analysis. (A) Hierarchical clustering tree (dendrogram) of all genes with TPM>1. (B) Sample dendrogram and module heatmap. Each row color corresponds to the modules. The right panel represents the minimum (blue color) and maximum (red color) correlation coefficient.

SUPPLEMENTARY FIGURE 7

Pearson's correlation between RT-qPCR and RNA-seq expression quantification.

References

- An, J. P., Li, R., Qu, F. J., You, C. X., Wang, X. F., and Hao, Y. J. (2018). An apple NAC transcription factor negatively regulates cold tolerance via CBF-dependent pathway. *J. Plant Physiol.* 221, 74–80. doi: 10.1016/j.jplph.2017.12.009
- An, J. P., Wang, X. F., Zhang, X. W., Xu, H. F., Bi, S. Q., You, C. X., et al. (2020). An apple MYB transcription factor regulates cold tolerance and anthocyanin accumulation and undergoes MIEL1-mediated degradation. *Plant Biotechnol. J.* 18, 337–353. doi: 10.1111/pbi.13201
- Bonthala, V. S., Mayes, K., Moreton, J., Blythe, M., Wright, V., May, S. T., et al. (2016). Identification of gene modules associated with low temperatures response in baramba groundnut by network-based analysis. *PLoS One* 11, e0148771. doi: 10.1371/journal.pone.0148771
- Chang, T., An, B., Liang, M., Duan, X., Du, L., Cai, W., et al. (2021). PacBio single-molecule long-read sequencing provides new light on the complexity of full-length transcripts in cattle. *Front. Genet.* 12. doi: 10.3389/fgene.2021.664974
- Chen, C., Wu, W., Li, J., Wang, X., Zeng, Z., Xu, J., et al. (2023). TBtools-II: A “one for all, all for one” bioinformatics platform for biological big-data mining. *Mol. Plant* 16, 1733–1742. doi: 10.1016/j.molp.2023.09.010
- Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34, i884–i890. doi: 10.1093/bioinformatics/bty560
- Cheong, B. E., Onyemaobi, O., Wing Ho Ho, W., Biddulph, T. B., Rupasinghe, T. W. T., Roessner, U., et al. (2020). Phenotyping the chilling and freezing responses of young microspore stage wheat spikes using targeted metabolome and lipidome profiling. *Cells* 9, 1309. doi: 10.3390/cells9051309
- Chinnusamy, V., Zhu, J., and Zhu, J. K. (2007). Cold stress regulation of gene expression in plants. *Trends Plant Sci.* 12, 444–451. doi: 10.1016/j.tplants.2007.07.002
- Daryanavard, H., Postiglione, A. E., Mühlemann, J. K., and Muday, G. K. (2023). Flavonols modulate plant development, signaling, and stress responses. *Curr. Opin. Plant Biol.* 72, 102350. doi: 10.1016/j.pbi.2023.102350
- Dong, W., Ma, X., Jiang, H., Zhao, C., and Ma, H. (2020). Physiological and transcriptome analysis of *Poa pratensis* var. *anceps* cv. Qinghai in response to cold stress. *BMC Plant Biol.* 20, 362. doi: 10.1186/s12870-020-02559-1
- Emms, D. M., and Steven, K. (2019). OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol.* 20, 238. doi: 10.1186/s13059-019-1832-y
- Fernandez-Pozo, N., Metz, T., Chandler, J. O., Gramzow, L., Mérai, Z., Maumus, F., et al. (2021). *Aethionema arabicum* genome annotation using PacBio full-length transcripts provides a valuable resource for seed dormancy and Brassicaceae evolution research. *Plant J.* 106, 275–293. doi: 10.1111/tpj.15161
- Frey, F. P., Pitz, M., Schön, C. C., and Hochholdinger, F. (2020). Transcriptomic diversity in seedling roots of European flint maize in response to cold. *BMC Genomics* 21, 300. doi: 10.1186/s12864-020-6682-1
- Gladman, N., Goodwin, S., Chougule, K., McCombie, W. R., and Ware, D. (2023). Era of gapless plant genomes: innovations in sequencing and mapping technologies revolutionize genomics and breeding. *Curr. Opin. Biotech.* 79, 102886. doi: 10.1016/j.copbio.2022.102886
- Golizadeh, F., and Kumleh, H. H. (2019). Physiological responses and expression changes of fatty acid metabolism-related genes in wheat (*Triticum aestivum*) under cold stress. *Plant Mol. Biol. Rep.* 37, 224–236. doi: 10.1007/s11105-019-01150-9
- Guy, C., Kaplan, F., Kopka, J., Selbig, J., and Hinch, D. K. (2008). Metabolomics of temperature stress. *Physiol. Plantarum* 132, 220–235. doi: 10.1111/j.1399-3054.2007.00999.x
- Han, M. V., Thomas, G. W. C., Lugo-Martinez, J., and Hahn, M. W. (2013). Estimating gene gain and loss rates in the presence of error in genome assembly and annotation using CAFE 3. *Mol. Biol. Evol.* 30, 1987–1997. doi: 10.1093/molbev/mst100
- Harrell, F. E. Jr., and Harrell, M. F. E. Jr. (2019). *Package 'Hmisc'; CRAN2018* (Vienna, Australia: The R Foundation), 235–236.
- Hu, Z., Yan, W., Yang, C., Huang, X., Hu, X., Li, Y., et al. (2022). Integrative analysis of transcriptome and metabolome provides insights into the underlying mechanism of cold stress response and recovery in two tobacco cultivars. *Environ. Exp. Bot.* 200, 104920. doi: 10.1016/j.envexpbot.2022.104920
- Hwarari, D., Guan, Y., Ahmad, B., Movahedi, A., Min, T., Hao, Z., et al. (2022). ICE-CBF-COR signaling cascade and its regulation in plants responding to cold stress. *Int. J. Mol. Sci.* 23, 1549. doi: 10.3390/ijms23031549
- Jan, S., Rustgi, S., Barmukh, R., Shikari, A. B., Leske, B., Bekuma, A., et al. (2024). Advances and opportunities in unraveling cold-tolerance mechanisms in the world's primary staple food crops. *Plant Genome* 17, e20402. doi: 10.1002/tpg2.20402
- Jin, J., Tian, F., Yang, D. C., Meng, Y. Q., Kong, L., Luo, J., et al. (2017). PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic Acids Res.* 45, D1040–D1045. doi: 10.1093/nar/gkw982
- Katoh, K., Asimenos, G., and Toh, H. (2009). Multiple alignment of DNA sequences with MAFFT. *Bioinform. DNA Seq. Anal.* 537, 39–64. doi: 10.1007/978-1-59745-251-9_3
- Kim, H. N., Kim, J. D., Yeo, J. H., Son, H. J., Park, S. B., Park, G. H., et al. (2019). *Heracleum moellendorffii* roots inhibit the production of pro-inflammatory mediators through the inhibition of NF- κ B and MAPK signaling, and activation of ROS/Nrf2/HO-1 signaling in LPS-stimulated RAW264.7 cells. *BMC Complem. Altern. M.* 19, 310. doi: 10.1186/s12906-019-2735-x
- Kim, D., Paggi, J. M., Park, C., Bennett, C., and Salzberg, S. L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* 37, 907–915. doi: 10.1038/s41587-019-0201-4
- Kou, S., Chen, L., Tu, W., Scossa, F., Wang, Y., Liu, J., et al. (2018). The arginine decarboxylase gene *ADC1*, associated to the putrescine pathway, plays an important role in potato cold-acclimated freezing tolerance as revealed by transcriptome and metabolome analyses. *Plant J.* 96, 1283–1298. doi: 10.1111/tpj.14126
- Kovács, Z., Simon-Sarkadi, L., Szűcs, A., and Kocsy, G. (2010). Differential effects of cold, osmotic stress and abscisic acid on polyamine accumulation in wheat. *Amino Acids* 38, 623–631. doi: 10.1007/s00726-009-0423-8
- Kruse, E. B., Revolinski, S., Aplin, J., Skinner, D., Murray, T. D., Edwards, C. G., et al. (2020). Carbohydrate accumulation and differential transcript expression in winter

- wheat lines with different levels of snow mold and freezing tolerance after cold treatment. *Plants* 9, 1416. doi: 10.3390/plants9111416
- Lee, J. G., Yi, G., Choi, J. H., and Lee, E. J. (2020). Analyses of targeted/untargeted metabolites and reactive oxygen species of pepper fruits provide insights into seed browning induced by chilling. *Food Chem.* 332, 127406. doi: 10.1016/j.foodchem.2020.127406
- Li, B., and Dewey, C. N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinf.* 12, 323. doi: 10.1186/1471-2105-12-323
- Li, M. Y., Feng, K., Hou, X. L., Jiang, Q., Xu, Z. S., Wang, G. L., et al. (2020). The genome sequence of celery (*Apium graveolens* L.), an important leaf vegetable crop rich in apigenin in the Apiaceae family. *Hortic. Res.* 7, 9. doi: 10.1038/s41438-019-0235-2
- Li, A., Wang, Y., Zhao, S., Zhao, J., and Wang, Y. (1997). Analysis of nutritional components in several wild celeries. *Spec. Wild Econ. Anim. Plant Res.* 3, 18–19. doi: 10.16720/j.cnki.tcjy.1997.03.007
- Li, F. H., Yu, P., Song, C. H., Wu, J. J., Tian, Y., Wu, X. F., et al. (2019). Differential protein analysis of *Heracleum moellendorffii* Hance seeds during stratification. *Plant Physiol. Biochem.* 145, 10–20. doi: 10.1016/j.plaphy.2019.10.002
- Liu, S., Jiang, X., Liu, Z., Cheng, Y., Sun, T., and Yu, X. (2019a). Mechanism of the breaking of seed dormancy by flower thinning in *Heracleum moellendorffii* Hance. *J. Plant Growth Regul.* 38, 870–882. doi: 10.1007/s00344-018-9898-4
- Liu, G., Liu, F., Pan, L., Wang, H., Lu, Y., Liu, C., et al. (2024a). Agronomic, physiological and transcriptional characteristics provide insights into fatty acid biosynthesis in yellowhorn (*Xanthoceras sorbifolium* Bunge) during fruit ripening. *Front. Genet.* 15. doi: 10.3389/fgene.2024.1325484
- Liu, G., Liu, J., Zhang, C., You, X., Zhao, T., Jiang, J., et al. (2018). Physiological and RNA-seq analyses provide insights into the response mechanism of the Cf-10-mediated resistance to *Cladosporium fulvum* infection in tomato. *Plant Mol. Biol.* 96, 403–416. doi: 10.1007/s11103-018-0706-0
- Liu, C., Schläppi, M. R., Mao, B., Wang, W., Wang, A., and Chu, C. (2019b). The bZIP73 transcription factor controls rice cold tolerance at the reproductive stage. *Plant Biotechnol. J.* 17, 1834–1849. doi: 10.1111/pbi.13104
- Liu, Y., Shi, Y., Zhu, N., Zhong, S., Bouzayen, M., and Li, Z. (2020). SIGRAS4 mediates a novel regulatory pathway promoting chilling tolerance in tomato. *Plant Biotechnol. J.* 18, 1620–1633. doi: 10.1111/pbi.13328
- Liu, H., Wang, Y., Chang, Q. Z., Li, Q., Fang, J., Cao, N., et al. (2024b). Combined metabolome and transcriptome reveal *HmF6/H1* regulating simple coumarin accumulation against powdery mildew infection in *Heracleum moellendorffii* Hance. *BMC Plant Biol.* 24, 507. doi: 10.1186/s12870-024-05185-3
- 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, 550. doi: 10.1186/s13059-014-0550-8
- Mahajan, S., and Tuteja, N. (2015). Cold, salinity and drought stresses: An overview. *Arch. Biochem. Biophys.* 444, 139–158. doi: 10.1016/j.abb.2005.10.018
- Malacarne, G., Coller, E., Czemplak, S., Vrhovsek, U., Engelen, K., Goremykin, V., et al. (2016). The grapevine VvZIP22 transcription factor is involved in the regulation of flavonoid biosynthesis. *J. Exp. Bot.* 67, 3509–3522. doi: 10.1093/jxb/erw181
- Marini, F., and Binder, H. (2019). pcaExplorer: an R/Bioconductor package for interacting with RNA-seq principal components. *BMC Bioinf.* 20, 331. doi: 10.1186/s12859-019-2879-1
- Mehrotra, S., Verma, S., Kumar, S., Kumari, S., and Mishra, B. N. (2020). Transcriptional regulation and signalling of cold stress response in plants: An overview of current understanding. *Environ. Exp. Bot.* 180, 104243. doi: 10.1016/j.envexpbot.2020.104243
- Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G. A., Sonnhammer, E. L. L., et al. (2021). Pfam: The protein families database in 2021. *Nucleic Acids Res.* 49, D412–D419. doi: 10.1093/nar/gkaa913
- Morishita, T., Kojima, Y., Maruta, T., Nishizawa-Yokoi, A., Yabuta, Y., and Shigeoka, S. (2009). Arabidopsis NAC transcription factor, ANAC078, regulates flavonoid biosynthesis under high-light. *Plant Cell Physiol.* 50, 2210–2222. doi: 10.1093/pcp/pcp159
- Naik, J., Misra, P., Trivedi, P. K., and Pandey, A. (2022). Molecular components associated with the regulation of flavonoid biosynthesis. *Plant Sci.* 317, 111196. doi: 10.1016/j.plantsci.2022.111196
- Nguyen, L. T., Schmidt, H. A., von Haeseler, A., and Minh, B. Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* 32, 268–274. doi: 10.1093/molbev/msu300
- Pan, Y., Liang, H., Gao, L., Dai, G., Chen, W., Yang, X., et al. (2020). Transcriptomic profiling of germinating seeds under cold stress and characterization of the cold-tolerant gene LTG5 in rice. *BMC Plant Biol.* 20, 371. doi: 10.1186/s12870-020-02569-z
- Rolland, F., Baena-Gonzalez, E., and Sheen, J. (2006). Sugar sensing and signaling in plants: conserved and novel mechanisms. *Annu. Rev. Plant Biol.* 57, 675–709. doi: 10.1146/annurev.arplant.57.032905.105441
- Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., and Zdobnov, E. M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31, 3210–3212. doi: 10.1093/bioinformatics/btv351
- Song, W., Lin, S. Q., Yin, Q., Liu, T. H., Gan, L. Z., Qi, J. J., et al. (2024). A multi-omics approach reveals low temperature inhibition of flavones and flavonols accumulation in postharvest bananas via downregulation of MabHLH363. *Postharvest Biol. Tec.* 218, 113152. doi: 10.1016/j.postharvbio.2024.113152
- Song, X., Sun, P., Yuan, J., Gong, K., Li, N., Meng, F., et al. (2021). The celery genome sequence reveals sequential paleo-polyploidizations, karyotype evolution and resistance gene reduction in apiales. *Plant Biotechnol. J.* 19, 731–744. doi: 10.1111/pbi.13499
- Sun, S., Fang, J., Lin, M., Hu, C., Qi, X., Chen, J., et al. (2021). Comparative metabolomic and transcriptomic studies reveal key metabolism pathways contributing to freezing tolerance under cold stress in kiwifruit. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.628969
- Talavera, G., and Jose, C. (2007). Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst. Biol.* 56, 564–577. doi: 10.1080/10635150701472164
- Tang, K., Zhao, L., Ren, Y., Yang, S., Zhu, J., and Zhao, C. (2020). The transcription factor ICE1 functions in cold stress response by binding to the promoters of CBF and COR genes. *J. Integr. Plant Biol.* 62, 258–263. doi: 10.1111/jipb.12918
- Tian, J., Ma, Y., Chen, Y., Chen, X., and Wei, A. (2022). Plant hormone response to low-temperature stress in cold-tolerant and cold-sensitive varieties of *Zanthoxylum bungeanum* Maxim. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.847202
- Wang, L., Guo, D., Zhao, G., Wang, J., Zhang, S., Wang, C., et al. (2022b). Group IIc WRKY transcription factors regulate cotton resistance to *Fusarium oxysporum* by promoting GhMCK2-mediated flavonoid biosynthesis. *New Phytol.* 236, 249–265. doi: 10.1111/nph.18329
- Wang, C., and Han, B. (2022c). Twenty years of rice genomics research: from sequencing and functional genomics to quantitative genomics. *Mol. Plant* 4, 593–619. doi: 10.1016/j.molp.2022.03.009
- Wang, M., Huang, J., Liu, S., Liu, X., Li, R., Luo, J., et al. (2022a). Improved assembly and annotation of the sesame genome. *DNA Res.* 29, dsac041. doi: 10.1093/dnares/dsac041
- Wang, Y., Li, X. Y., Li, C. X., He, Y., Hou, X. Y., and Ma, X. R. (2021). The regulation of adaptation to cold and drought stresses in *poa crymophila keng* revealed by integrative transcriptomics and metabolomics analysis. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.631117
- Wang, T., Ma, X., Chen, Y., Wang, C., Xia, Z., Liu, Z., et al. (2024). SINAC3 suppresses cold tolerance in tomatoes by enhancing ethylene biosynthesis. *Plant Cell Environ.* 48, 3132–3146. doi: 10.1111/pce.14933
- Wang, J., Yu, X., Jiang, X., Sun, D., and Liu, S. (2017). Comparative analysis to different ecological conditions on main growth period and nutritional components of *Heracleum dissectum*. *China Cucurbit Veg.* 30, 14–17. doi: 10.3969/j.issn.1673-2871.2017.06.004
- 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., Huang, C., Jiang, X., Zhu, J., Gao, X., and Yu, C. (2022). Impact of cold stress on leaf structure, photosynthesis, and metabolites in *Camellia weiningensis* and *C. oleifera* seedlings. *Horticulturae* 8, 494. doi: 10.3390/horticulturae8060494
- Xu, D., Yuan, S., Chen, B., Shi, J., Sui, Y., Gao, L., et al. (2023). A comparative proteomic and metabolomic analysis of the low-temperature response of a chilling-injury sensitive and a chilling-injury tolerant cultivar of green bell pepper. *Sci. Hortic.* 318, 112092. doi: 10.1016/j.scienta.2023.112092
- Yang, Z. (2007). PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* 24, 1586–1591. doi: 10.1093/molbev/msm088
- Yang, C., Shen, S., Zhou, S., Li, Y., Mao, Y., Zhou, J., et al. (2022). Rice metabolic regulatory network spanning the entire life cycle. *Mol. Plant* 15, 258–275. doi: 10.1016/j.molp.2021.10.005
- Yue, C., Cao, H. L., Wang, L., Zhou, Y. H., Huang, Y. T., Hao, X. Y., et al. (2015). Effects of cold acclimation on sugar metabolism and sugar-related gene expression in tea plant during the winter season. *Plant Mol. Biol.* 88, 591–608. doi: 10.1007/s11103-015-0345-7
- Zhang, Z., Li, W., Gao, X., Xu, M., and Guo, Y. (2020). DEAR4, a member of DREB/CBF family, positively regulates leaf senescence and response to multiple stressors in *Arabidopsis thaliana*. *Front. Plant Sci.* 11. doi: 10.3389/fpls.2020.00367
- Zhang, Y., Yu, H., Yang, X., Li, Q., Ling, J., Wang, H., et al. (2016). CsWRKY46, a WRKY transcription factor from cucumber, confers cold resistance in transgenic-plant by regulating a set of cold-stress responsive genes in an ABA-dependent manner. *Plant Physiol. Biochem.* 108, 478–487. doi: 10.1016/j.plaphy.2016.08.013
- Zhang, J., Zhao, H., Chen, L., Lin, J., Wang, Z., Pan, J., et al. (2023). Multifaceted roles of WRKY transcription factors in abiotic stress and flavonoid biosynthesis. *Front. Plant Sci.* 14. doi: 10.3389/fpls.2023.1303667
- 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. *Func. Integr. Genomic* 13, 75–98. doi: 10.1007/s10142-012-0301-4
- Zhao, J., Xu, Y., Li, H., An, W., Yin, Y., Wang, B., et al. (2024). Metabolite-based genome-wide association studies enable the dissection of the genetic bases of flavonoids, betaine and spermidine in wolfberry (*Lycium*). *Plant Biotechnol. J.* 22, 1435–1452. doi: 10.1111/pbi.14278