


# Deciphering aroma formation during flowering in nectar tree (*Tilia amurensis*): insights from integrated metabolome and transcriptome analysis

Kewei Cai<sup>1,2#</sup>, Qiushuang Zhao<sup>2#</sup>, Hanxi Li<sup>2</sup>, Qinhui Zhang<sup>2</sup>, Yan Li<sup>1</sup>, Rui Han<sup>1</sup>, Tingbo Jiang<sup>2</sup>, Xiaona Pei<sup>1</sup>, Lina Zhang<sup>3\*</sup> and Xiyang Zhao<sup>1,2\*</sup> 

<sup>1</sup> Jilin Provincial Key Laboratory of Tree and Grass Genetics and Breeding, College of Forestry and Grassland Science, Jilin Agricultural University, Changchun, China

<sup>2</sup> State Key Laboratory of Tree Genetics and Breeding, Northeast Forestry University, Harbin, China

<sup>3</sup> School of information technology, Jilin Agricultural University, Changchun, China

# These authors contributed equally: Kewei Cai, Qiushuang Zhao

\* Corresponding authors, E-mail: [zhangln@jlau.edu.cn](mailto:zhangln@jlau.edu.cn); [zhaoxyphd@163.com](mailto:zhaoxyphd@163.com)

## Abstract

*Tilia amurensis* is a significant ornamental and economically-important tree species, known for its fragrant flowers, which are a source of high-quality honey production. However, the regulatory mechanisms involved in aroma formation during flower development in *T. amurensis* remains limited. The current study revealed the detection of plant hormones at every assessed stage of flower development. Among them, auxin and brassinosteroid contents significantly increased at stage 3, potentially regulating crucial functions during *T. amurensis* flower development. Moreover, the study examined the levels and change patterns of secondary metabolites and employed a combination of transcriptomics and metabolomics to comprehensively assess essential structural genes implicated in the biosynthesis pathways of terpenoid and phenylpropanoid. A comprehensive set of 89,526 differentially expressed genes (DEGs) was uncovered, including candidate structural genes *ACAT*, *HDS*, *TPS*, *ACL*, *CAD*, and *CCOAMT*, which are specifically involved in the biosynthesis of terpenoids and phenylpropanoids. Maslinic acid, 2 $\alpha$ ,3 $\alpha$ -dihydroxyursolic acid, and betulinic acid were accumulated in the terpenoid biosynthesis pathway. In contrast, metabolites with differential accumulation, such as phenylalanine, coniferyl alcohol, and cinnamic acid, were specifically enriched in the phenylpropanoid biosynthesis pathway. The *C2H2*, *MYB*, and *NAC* transcription factor families are crucially associated with the terpenoid and phenylpropanoid biosynthesis pathways. Two transcription factors, *C2H2-17* and *MYB-24*, exhibited strong co-expression with structural genes in two networks, and were identified as central regulatory factors. These findings establish a solid groundwork for elucidating the generation of floral fragrance and provide comprehensive genetic and metabolic information for further studies on *T. amurensis*.

**Citation:** Cai K, Zhao Q, Li H, Zhang Q, Li Y, et al. 2023. Deciphering aroma formation during flowering in nectar tree (*Tilia amurensis*): insights from integrated metabolome and transcriptome analysis. *Forestry Research* 3:24 <https://doi.org/10.48130/FR-2023-0024>

## Introduction

*Tilia amurensis* is among the most valuable species of the Tiliaceae family, with an expansive range of uses<sup>[1]</sup>. Currently, it has been classified as a second-level protected plant in China. Industrial oils can be extracted from the seeds of *T. amurensis*, and its wood can be used to make building materials. *T. amurensis* possesses an elegant tree structure, exhibits golden autumn foliage, and showcases remarkable ornamental characteristics, making it an ideal choice as a street tree species<sup>[2]</sup>. Additionally, *T. amurensis* flowers can release odorous fragrances and are an excellent nectar source. *Tilia amurensis* can also be used as medicine and is a source of traditional Chinese medicine for clearing heat and relieving skin<sup>[3]</sup>. Noteworthy, *T. amurensis* has been favored since ancient times because of its characteristic aromatic flavors, and its flower aroma quality has a major impact on the flavors of linden honey. In general, immature flower buds of *T. amurensis* are odorless, and only mature flowers give off their characteristic aroma during flowering, which lasts for about a week<sup>[4]</sup>. Moreover, aromatic

flavors play crucial roles in maintaining the ecological associations between flowers and pollinators<sup>[5,6]</sup>. The synthesis of aromatic compounds in flowers is a complex process that encompasses secondary metabolism. The synthesis and degradation of compounds give rise to intricate networks of metabolites and regulatory mechanisms, involving diverse pathways<sup>[7]</sup>. Presently, the application of various metabolite detection platforms has greatly accelerated the exploration of plant secondary metabolite biosynthesis pathways<sup>[8]</sup>. Despite previous investigations into the chemical composition of *T. amurensis* flowers<sup>[9]</sup>, it is imperative to assess the secondary metabolite profile throughout different stages of flower development. This is particularly significant given the rising importance of *T. amurensis* in the food industry.

In recent years, the exploration of plant development has revealed the multifaceted roles of plant hormones in orchestrating the intricate process of flowering. Among the array of hormonal cues, gibberellins, brassinosteroids, auxins, cytokinins, and ethylene have emerged as central players in regulating the initiation and progression of floral development<sup>[10]</sup>.

Gibberellins, known for their influence on stem elongation, have also been implicated in promoting floral meristem development, thus influencing the timing and dimensions of blooming<sup>[11]</sup>. Auxins, in their diverse functions, contribute to the positioning and differentiation of floral organs, sculpting the delicate architecture of blossoms<sup>[12]</sup>. Cytokinins, integral to cell division processes, intricately shape the growth of floral structures, further underscoring their significance in flowering patterns<sup>[13]</sup>. Ethylene, a gaseous hormone, governs not only the acceleration of flower senescence but also influences the emission of alluring floral fragrances<sup>[14]</sup>. Furthermore, by influencing biochemical processes such as cell elongation, differentiation, and division, brassinolide also participates in regulating the formation and development of floral organs<sup>[15]</sup>. Existing studies have indicated that gibberellins, brassinolide, and ethylene signaling were involved in flower differentiation and development in *Nelumbo nucifera*<sup>[16]</sup>. In the ornamental plant *Hedychium coronarium*, research has revealed that under the influence of auxin induction, transcription factors *HcMYB1* and *HcMYB2* are specifically expressed in flowers, regulating the biosynthesis of floral scent compounds<sup>[17]</sup>. Additionally, Janowska et al.<sup>[11]</sup> discovered that cytokinins and gibberellins can promote the flowering and post-harvest longevity of the colorful inflorescence spathes of Calla Lilie. In conclusion, plant hormones play a significant role in flower development, influencing not only the formation and growth of floral organs but also potentially affecting the generation of floral fragrances.

The intricate biochemical pathways associated with the synthesis of floral aroma compounds mainly include pathways of terpenoids, phenylpropanoids, and fatty acid-derived volatile biosynthesis<sup>[18]</sup>. So far, terpenoids, a diverse family of secondary metabolites, have been considered the main aroma components of flowers and have been well-studied in plants. The possible terpenoid biosynthesis pathways have been identified in plants. Monoterpenes (C<sub>10</sub>) and sesquiterpenes (C<sub>15</sub>) constitute the major components of terpenoids. Isopentenyl pyrophosphate (IPP) and dimethylallyl diphosphate (DMAPP) serve as the primary C<sub>5</sub> carbon precursors for the synthesis of these compounds. The mevalonic acid (MVA) pathway and the plastidial 2-C-methylerythritol 4-phosphate (MEP) pathway serve as distinct biosynthetic routes for the synthesis of these compounds, providing the precursors required for their formation<sup>[19]</sup>. IPP is derived from acetyl-CoA, whereas the MVA pathway is responsible for producing precursors for sterols, and sesquiterpenes. The MEP pathway utilizes pyruvate and glyceraldehyde-3-phosphate (G3P) to synthesize IPP and DMAPP, which subsequently contribute to the production of monoterpenes, and diterpenes<sup>[20,21]</sup>. The relationship between terpenoid biosynthesis regulatory genes and aroma production has been established in different plants, such as cucumber<sup>[22]</sup>, jasmine<sup>[23]</sup>, and celery<sup>[24]</sup>. Despite terpenoids being the primary constituents of volatile compounds, the genes responsible for terpenoid biosynthesis in *T. amurensis* flowers remain unidentified.

Benzenoids, especially phenylpropanoids, are metabolites with key functions in plants and are the main components of floral aroma, contributing to the attraction of pollinators by plants<sup>[25]</sup>. Among the various secondary metabolite pathways in plants, the benzenoid pathway has garnered significant attention and has been extensively studied. Benzenoids and phenylpropanoids (BPs) are synthesized from phenylalanine

(Phe) and tyrosine, two aromatic amino acids in the shikimic acid pathway<sup>[26]</sup>. Phenylpropanoid metabolites contain an aromatic ring that is linked to a 3-C propane side chain<sup>[27]</sup>. The incorporation of diverse aromatic ring substituents and the positioning of the propylene double bond give rise to an extensive array of bioactive compounds with remarkable variations<sup>[28]</sup>. The biosynthesis of phenylpropanoid metabolites, including cinnamic acid, involves a coordinated cascade of enzymatic reactions. The enzymatic modification of the core structure of these metabolites relies on key enzymes such as cinnamate-4 hydroxylase (C4H), cinnamoyl-CoA reductase (CCR), and 4-coumarate-CoA ligase (4CL), which play crucial roles in this process<sup>[27,29]</sup>. Many studies have revealed that aroma production is related to sophisticated mechanisms that regulate the biosynthesis of phenylpropanoid metabolites, such as those reported in tea<sup>[26]</sup> and tomatoes<sup>[30]</sup>.

The regulation of gene expression in terpenoid and phenylpropanoid biosynthesis is significantly influenced by the involvement of transcription factors (TFs), which play crucial roles in these processes<sup>[5,21]</sup>. Numerous TFs examined thus far have been implicated in the biosynthesis of terpenoids. Within *Catharanthus roseus*, the *AP2/ERF* transcription factor known as *ORCA3* governs the biosynthesis of terpenoid indole alkaloids<sup>[31]</sup>. In *Cymbidium faberi*, *MYC* TFs were identified as candidates participating in the biosynthesis of flower scents<sup>[32]</sup>. Furthermore, *R2R3-MYB* genes has been discovered to exert a significant regulatory function in terpenoid biosynthesis, particularly within subgroups 4, 5, and 15<sup>[33,34]</sup>. Several transcriptional regulators have been identified that modulate phenylpropanoid, flavonoid, and benzenoid accumulation. In petunia, *PhMYB4* regulates floral aroma by repressing the transcription of the cinnamate-4-hydroxylase gene, shifting the metabolic flux toward the synthesis of volatile phenylpropanoids<sup>[35]</sup>. Moreover, *ODO1*, an *R2R3-MYB* TF, has emerged as the pioneering TFs identified to regulate floral aroma production. It exerts its regulatory influence by modulating the expression pattern of various genes related to the phenylpropanoid pathway<sup>[36]</sup>. Despite numerous studies have investigated the regulatory role of TFs in the biosynthesis pathways of terpenoids and phenylpropanoids, the specific gene regulatory mechanisms in *T. amurensis* are still not fully understood.

In the current study, fresh flowers at different growing stages of *T. amurensis* were collected. We aimed to investigate the changes in key secondary metabolites of *T. amurensis* during flowering and elucidate the roles and involvement of key genes in the synthesis of key metabolites. We profiled the flower metabolites during different flowering stages by metabolomics to achieve these aims. Subsequently, transcriptome sequencing was conducted to discover crucial structural genes and TFs responsible for regulating the biosynthesis of terpenoids and phenylpropanoids. The phytohormone content was determined during flower development, and the expression profiles of DEGs related to phytohormone biosynthesis and signal transduction were examined. Additionally, the reliability of the transcriptome data was substantiated through RT-qPCR (real time quantitative polymerase chain reaction) analysis. This investigation will provide valuable perspectives on the molecular mechanisms governing terpenoid and phenylpropanoid biosynthesis, thus establishing a basis for targeted regulation of crucial metabolic elements within flowers of *T. amurensis*.

## Materials and methods

### Plant samples

The adult *T. amurensis* plants were cultivated at Northeast Forestry University in Harbin, China (coordinates: 126°37'57.28" E, 45°43'6.53" N). Flower samples at three distinct stages of flower development (S1, S2, and S3, respectively) were collected. The fresh flower samples from the S1, S2, and S3 stages were collected on June 5, 2022, June 15, 2022, and June 25, 2022, respectively. To minimize the impact of circadian rhythms on the results, each sampling was consistently conducted between 9 am and 10 am. A total of 5 g of each sample was collected for subsequent analysis, with three biological replicates obtained for each sample. Following collection, all samples were promptly placed in liquid nitrogen and preserved at a temperature of  $-80^{\circ}\text{C}$  for subsequent analyses. These samples were used for hormone concentration measurements, transcriptome sequencing (RNA-seq), and metabolite extraction.

### Phytohormone concentration measurements

At various stages of flower development, the concentrations of cytokinin (CTK), auxin (IAA), abscisic acid (ABA), ethylene (ETH), gibberellin (GA), zeatin (ZT), a type of cytokinin, brassinosteroids (BR), and jasmonic acid (JA) were quantified. The phytohormone content in each flowering stage was analyzed in three biological replicates. Each replicate involved the collection of 50 mg of fresh sample, and subsequent analysis of phytohormone concentrations in the extracts was performed using the high-performance liquid chromatography-mass spectrometry (HPLC-MS) and enzyme-linked immunosorbent assay (ELISA) methods<sup>[37–39]</sup>. The statistical analysis of the data was carried out using IBM Crop's SPSS 26.0 software (Armonk, NY, USA) in addition to Microsoft Office Excel software. To evaluate the statistical significance of variations among different samples, the Student–Newman–Keuls (S–N–K) test was employed, with a significance threshold set at  $p < 0.05$ .

### Sampling preparation and metabolite analysis in flowers

Wuhan Metware Biotechnology Co., Ltd. employed ultra-high performance liquid chromatography-mass spectrometry (UPLC-MS) and gas chromatography mass spectrometry (GC/MS) for extraction, and quantification of metabolites in the samples. In a concise manner, each freeze-dried flower sample underwent individual grinding into powder under liquid nitrogen in three biological replicates for every flowering stage. For LC-MS, 3 mL of the sample was used after mixing. It was placed in a 50 mL centrifuge tube, and then the intact sample was immersed in liquid nitrogen. Then the sample was placed into a lyophilizer for freeze-drying. After lyophilization, 100  $\mu\text{L}$  of 70% methanol internal standard extract was added. The samples were centrifuged ( $21,560 \times g$ ,  $4^{\circ}\text{C}$ ) for 3 min. The resulting supernatant was passed through a micropore filter membrane, subsequently stored in a flask. The LC-MS system (UPLC, SHIMADZU Nexera X2) was utilized for the analysis of the sample extracts<sup>[7,40]</sup>. The Metware database (MWDB) was employed for the annotation of all metabolites, while quantification was accomplished using the multiple reaction monitoring (MRM) technique. The obtained metabolite data were analyzed using Analyst software (1.6.3)<sup>[41]</sup>. To conduct GC/MS analysis, 500 mg portion of powder was transferred into a 20

mL headspace vial that contained a NaCl-saturated solution to prevent enzymatic reactions. For the analysis using solid-phase microextraction (SPME), each vial was heated to a temperature of  $60^{\circ}\text{C}$  for a duration of 5 min. Afterward, the sample's headspace was subjected to a  $120\ \mu\text{m}$  DVB/CWR/PDMS fiber at  $100^{\circ}\text{C}$  for a duration of 15 min<sup>[42]</sup>. To conduct the identification and quantification of volatile compounds, an Agilent Model 8890 system coupled with a 7000D mass spectrometer (Agilent Technologies, California, USA) was employed. The MassHunter software was utilized to process the raw data<sup>[43]</sup>.

To explore the relationship between the metabolic profile at different flowering stages, principal component analysis (PCA) was conducted using R software ([www.r-project.org](http://www.r-project.org)). Multiple supervision methods were employed to select an orthogonal partial least-squares discriminant analysis (OPLS-DA) model<sup>[44]</sup>. The OPLS-DA model employed variable importance in projection (VIP) values to evaluate the relative significance of each metabolite across the samples. Metabolites meeting the criteria of  $\text{VIP} \geq 1$  and fold change  $\geq 2$  or  $\leq 0.5$  were classified as differential metabolites. Metabolic pathway assignments were carried out utilizing KOBAS2.0, leveraging the KEGG orthology database accessible at <http://www.genome.ad.jp/kegg/><sup>[45]</sup>. Moreover, the metabolite content data obtained were subjected to unit variance scaling for normalization purposes. Subsequently, a heatmap was generated using the TBtools software<sup>[46]</sup>.

### RNA sequencing and sequence assembly

The total RNA was extracted following the manufacturer's instructions, utilizing the plant RNA isolation kit (Tiangen, Beijing, China). Quantification of total RNA was conducted, and its quality was evaluated by the Agilent Bioanalyzer 2100. Nine libraries were prepared by utilizing 1.5  $\mu\text{g}$  of total RNA of each sample as the material for subsequent steps. Sequencing of all libraries was conducted on the Illumina HiSeq™ 2,500 platform, resulting in the generation of 150 bp paired-end reads as raw data<sup>[39]</sup>.

In order to acquire processed reads of high quality, the Fastp software (version 0.12) was employed to filter the raw data from all samples<sup>[47]</sup>. The clean reads were subjected to assembly, where they were clustered into expressed sequence tags and de novo assembled into transcripts. The assembly process was conducted by employing Trinity software (version 2.6.6)<sup>[48]</sup>. The reference sequences for subsequent analysis were selected based on the longest sequence within each cluster. Subsequently, the reference sequences were aligned with sequences from publicly available databases, such as Gene Ontology (GO), eukaryotic Orthologous Groups (KOG), Protein family (Pfam), Kyoto Encyclopedia of Genes and Genomes (KEGG), Swiss-Prot protein database (Swiss-Prot), Non-redundant protein sequence database (Nr), and Translated EMBL Nucleotide Sequence database (TrEMBL) using BLAST<sup>[49]</sup>. This allowed us to acquire the relevant functional annotation information.

### Identification of genes exhibiting differential expression (DEGs)

The alignment of clean reads to the reference sequences was performed by RSEM software (version 1.2.26)<sup>[50]</sup>. The expression level of each gene was represented by the FPKM (fragments per kilobase of transcript per million mapped reads) value<sup>[51]</sup>. The DESeq2 R package (version 1.38.2) was employed to detect DEGs across various comparison groups<sup>[52]</sup>. To select

significantly DEGs, a threshold of adjusted  $P$ -value  $< 0.05$  and  $|\log_2\text{Fold Change (FC)}| \geq 1$  was applied. Functional annotations of the DEGs were performed using GO and KEGG databases. The WEGO 2.0 tool (<https://wego.genomics.cn/>) and ArigGO were utilized for GO classification, and enrichment analysis<sup>[53]</sup>. To perform KEGG enrichment analyses of the DEGs, the KOBAS 2.0 software was utilized. Moreover, DEGs encoding for transcription factors (TFs) were predicted by hmmscan alignment using the iTAK software<sup>[54]</sup>. The Cytoscape software was utilized to visualize the network of connections between TFs and DEGs<sup>[55]</sup>.

### RT-qPCR analysis

To ensure the reliability of the RNA-seq data, the expression profiles of DEGs were validated using real time quantitative polymerase chain reaction (RT-qPCR) analysis. RNA extraction from various samples was carried out using Tiangen Plant RNA Extraction Kit (Tiangen, Beijing, China). The initial cDNA synthesis was conducted utilizing the PrimeScript RT reagent Kit with gDNA Eraser (Takara, Dalian, China). For RT-qPCR analysis, the ABI 7500 RT PCR system was utilized. The specific primers for the genes were designed using TBtools software<sup>[46]</sup> (Supplemental Table S1). The following conditions were used for RT-qPCR analysis: an initial denaturation at 94 °C for 30 s, followed by 45 cycles of denaturation at 94 °C for 5 s, annealing at 60 °C for 35 s, and extension at 95 °C for 15 s. A final extension step was performed at 60 °C for 1 min, followed by a melt curve analysis at 95 °C for 15 s. The  $2^{-\Delta\Delta\text{CT}}$  method was employed to calculate the relative expression levels of the selected genes<sup>[56]</sup>.

## Results

### Morphological characteristics and hormone concentration changes during flowers development

In this study, three developmental stages of the flowers (S1, S2, and S3) were assessed. The phenotypes of the different flower development stages are shown in Fig. 1. To observe the microstructure of the flowers, a representative sample was selected from each stage for analysis. The observation results indicated that the S1 stage corresponded to the early bud stage, with small and completely closed petals which were still green. In the S2 flower development stage, the color and size of

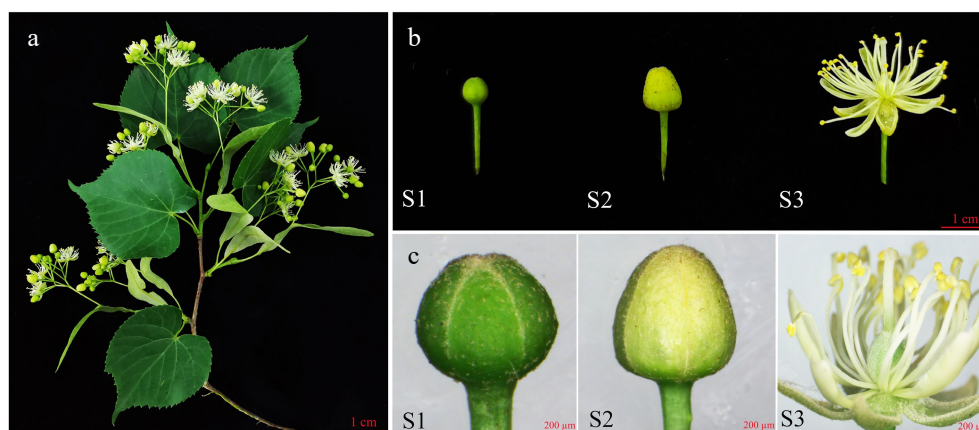
petals exhibited obvious changes compared to S1. S2 corresponded to the middle bud stage, with large and slightly closed petals that were light yellow. Subsequently, for S3, it corresponded to the full bloom stage, the petals were completely open, and the anthers and stigmata were formed. The findings indicated significant alterations in the flower morphological characteristics during the transition to flowering.

The concentrations of eight endogenous hormones in the flowers were further determined to investigate their roles across the flowering stages. As shown in Fig. 2, the CTK concentration displayed significant differences ( $p < 0.05$ ) between S1 and S2, decreasing progressively from S1 to S3. The GA and JA concentration decreased significantly from S1 to S2 but increased rapidly from S2 to S3 and reached their highest levels in S3. The ABA, ETH, and ZT concentration increased gradually from S1 to S2, showed relatively high levels in S2, and decreased sharply from S2 to S3. Finally, the concentration of BR and IAA increased gradually from S1 to S3, with significant differences observed between S1, S2, and S3 (Fig. 2).

### Sequence assembly and functional annotation

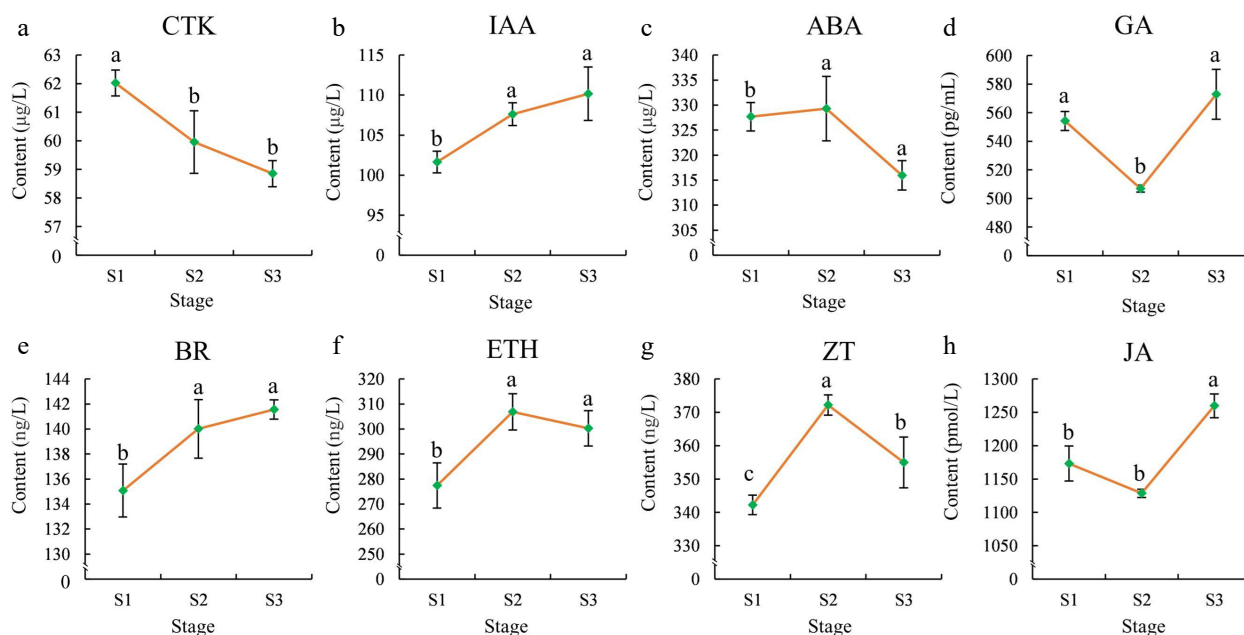
To investigate the alterations in gene expression patterns throughout the three stages of flower development, RNA-seq analysis was implemented. From the RNA-seq analysis, an initial set of 508,883,428 raw reads was obtained. After filtering, 490,842,144 clean reads were retained for further analysis. The clean reads had a cumulative length ranging from 6.59 Gb to 9.24 Gb, averaging at 8.18 Gb in total. The average GC content was determined to be 44.09%, with average percentages of Q20 and Q30 at 96.31% and 91.06% (Supplemental Table S2). After assembly, a total of 407,456 unigenes were obtained, comprising 298,619,887 bp. The N50 value was 904 bp, with an average unigene length of 733 bp (Supplemental Table S3). Furthermore, 229,545 (56.3%) unigenes had a length exceeding 500 bp, and a total of 15,032 (20.5%) unigenes surpassed 1,000 bp in length (Supplemental Table S4). On average, the alignment rate of clean reads to the reference sequence was approximately 82.34% (Supplemental Table S5).

To further investigate the functional annotations of the unigenes after assembly, alignments were performed with sequences from seven public functional databases. Out of the total 307,094 unigenes, at least one function was annotated in



**Fig. 1** Photographs of *T. amurensis* flowers at different developmental stages. (a) Image of a *T. amurensis* branch with flowers. (b) S1: early bud stage. S2: middle bud stage. S3: fully bloomed stage. (c) S1: microstructure at the early bud stage. S2: microstructure at the middle bud stage. S3: microstructure at the fully bloomed stage. The scale is 200  $\mu\text{m}$ .

## Aroma formation of flowers in nectar tree



**Fig. 2** Phytohormone contents in flowers at different stages. (a) CTK, (b) IAA, (c) ABA, (d) GA, (e) BR, (f) ETH, (g) ZT, (h) JA. The error bars in the plot represent the standard error, and the presence of distinct letters on the error bars indicates statistically significant differences at a significance level of  $p < 0.05$ .

the public databases. Among the databases used for annotation, the Nr database contained the largest count of annotated unigenes (305,313), followed by TrEMBL (301,034). In contrast, the KOG database had the lowest number of annotated unigenes (165,110) (Fig. 3a and Supplemental Table S6). The unigenes were strongly matched to gene sequences from *Durio zibethinus* (121,768), *Theobroma cacao* (59,489), and *Herrania umbratica* (28,259) (Fig. 3b). Moreover, 165,110 unigenes were annotated to 25 categories of the KOG protein database, mainly to the (R) "General function prediction only" (34,109), (T) "Signal transduction mechanisms" (19,910), and (O) "Posttranslational modification, protein turnover, chaperones" (18,479) categories (Fig. 3c). The unigenes were annotated to three categories, namely biological process (BP), molecular function (MF), and cellular component (CC), for the purpose of conducting GO analysis. Regarding the BP categories, unigenes were mainly associated with the cellular process (148,751), metabolic process (126,677), and biological regulation (71,579) categories. Regarding the MF categories, binding (148,687), catalytic activity (124,492), and transporter activity (16,531) were the most represented functions. Finally, in the CC categories, unigenes were mainly associated with the cell (178,918), cell part (178,486), and organelle (137,042) ontologies (Fig. 3d).

### Identification and enrichment analysis of DEGs

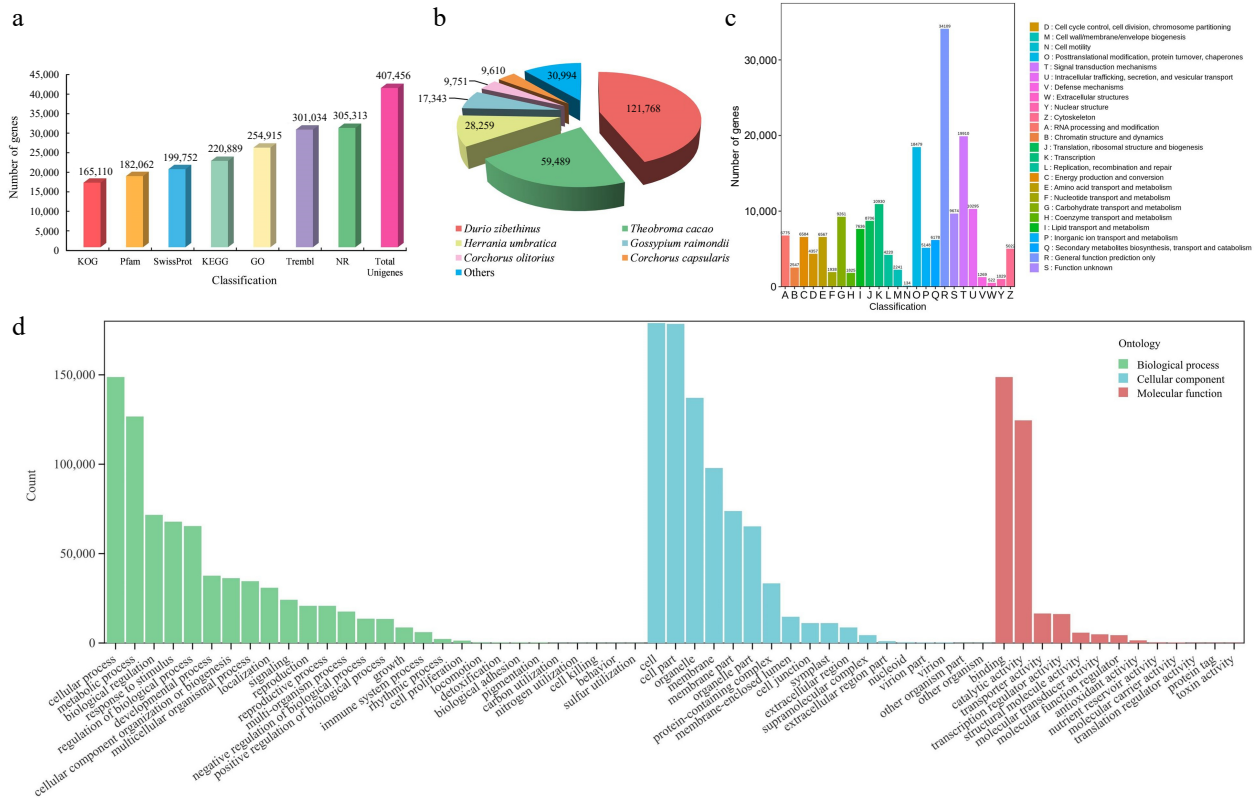
In the pursuit of identifying DEGs during the course of flower development, a comparison was made between the transcript levels of each unigene across various stages of flower development. The comparison between stages S1 and S3 revealed the highest number of DEGs (72,739), including 35,427 that were upregulated and 37,312 that were downregulated. In the S1 vs S2 comparison, 59,806 DEGs (29,228 upregulated and 30,578 downregulated) were identified, and in the S2 vs S3 comparison, 27,717 DEGs were identified (14,577 upregulated and 13,140 downregulated) (Fig. 4a & b). Among the different developmental stages, there were 7,680 DEGs that were

commonly expressed. To illustrate the overlap of these DEGs among the three groups, a Venn diagram was employed for visualization (Fig. 4a). The results revealed the genes that exhibited changes in expression levels during flower development.

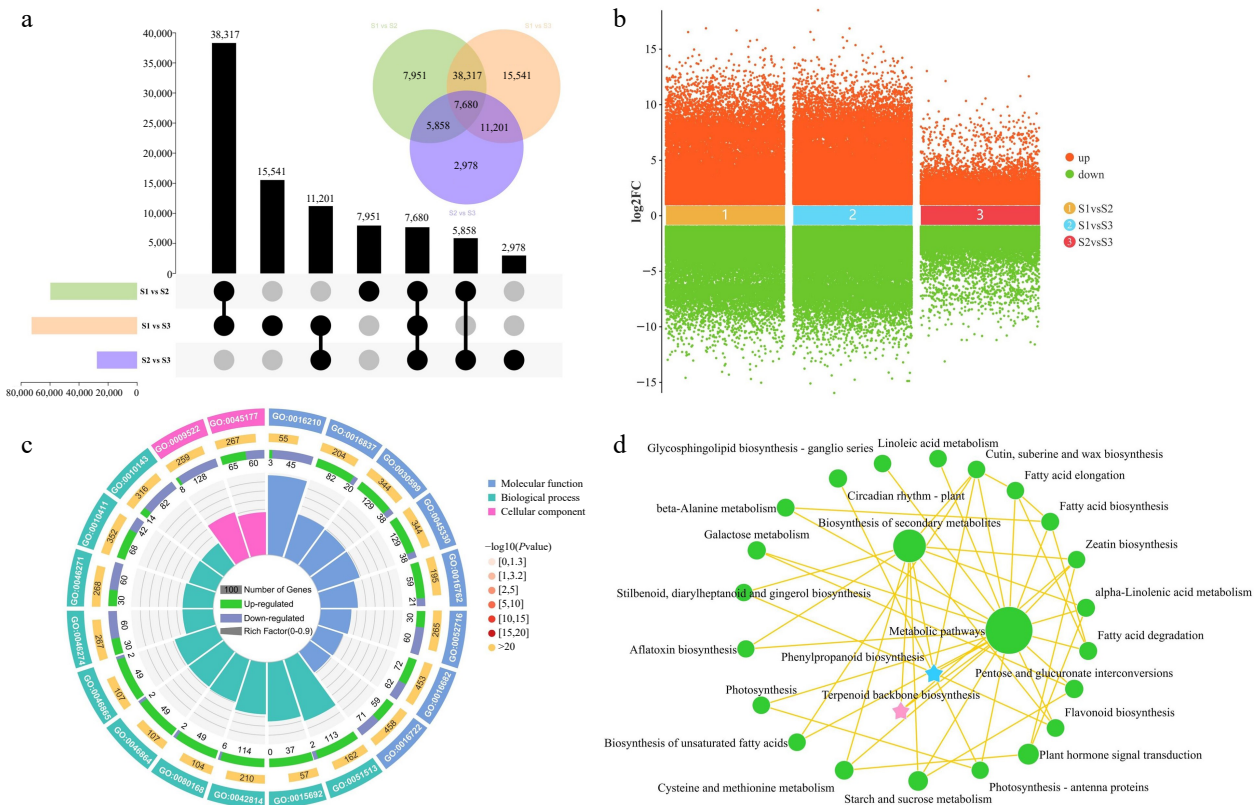
To elucidate the biological functions of DEGs identified in the comparative transcriptome analysis, functional annotations were assigned to the genes through the use of GO and KEGG analyses. The analysis of the top 20 GO terms revealed significant enrichment of DEGs across the categories of BP, CC, and MF. The main enriched functions consisted of "regulation of monopolar cell growth" (GO:0051513), "oxidoreductase activity, oxidizing metal ions" (GO:0016722), "phenylpropanoid catabolic process" (GO:0046271), and "terpenoid transport" (GO:0046865) (Fig. 4c). Subsequently, KEGG enrichment analysis was carried out to investigate the most highly represented metabolic and signal transduction pathways among the DEGs during the flower development of *T. amurensis*. The DEGs demonstrated notable enrichment in several KEGG pathways, encompassing "metabolic pathways", "biosynthesis of secondary metabolites", "plant hormone signal transduction", "phenylpropanoid biosynthesis", and "terpenoid backbone biosynthesis" (Fig. 4d). Significantly, the GO and KEGG enrichment analyses consistently unveiled significant enrichment of pathways related to terpenoid and phenylpropanoid biosynthesis. These findings provide compelling evidence for the significant involvement of genes associated with terpenoid and phenylpropanoid biosynthesis in the process of flower development.

### Differentially expressed metabolites (DEMs) analysis and metabolic pathway enrichment

To comprehensively analyze the differential expression and functional significance of metabolites across different stages of flowering, a comprehensive detection of metabolites was performed. A total of 664 metabolites were detected and identified throughout the various stages of flower development.



**Fig. 3** Identification and annotation analysis of unigenes of *T. amurensis* flowers. (a) Annotation information of unigenes in various public databases. (b) Statistical analysis of species distribution from the Nr database. (c) Function annotation of unigenes in KOG database. (d) The distribution across functional categories of unigenes in GO database.



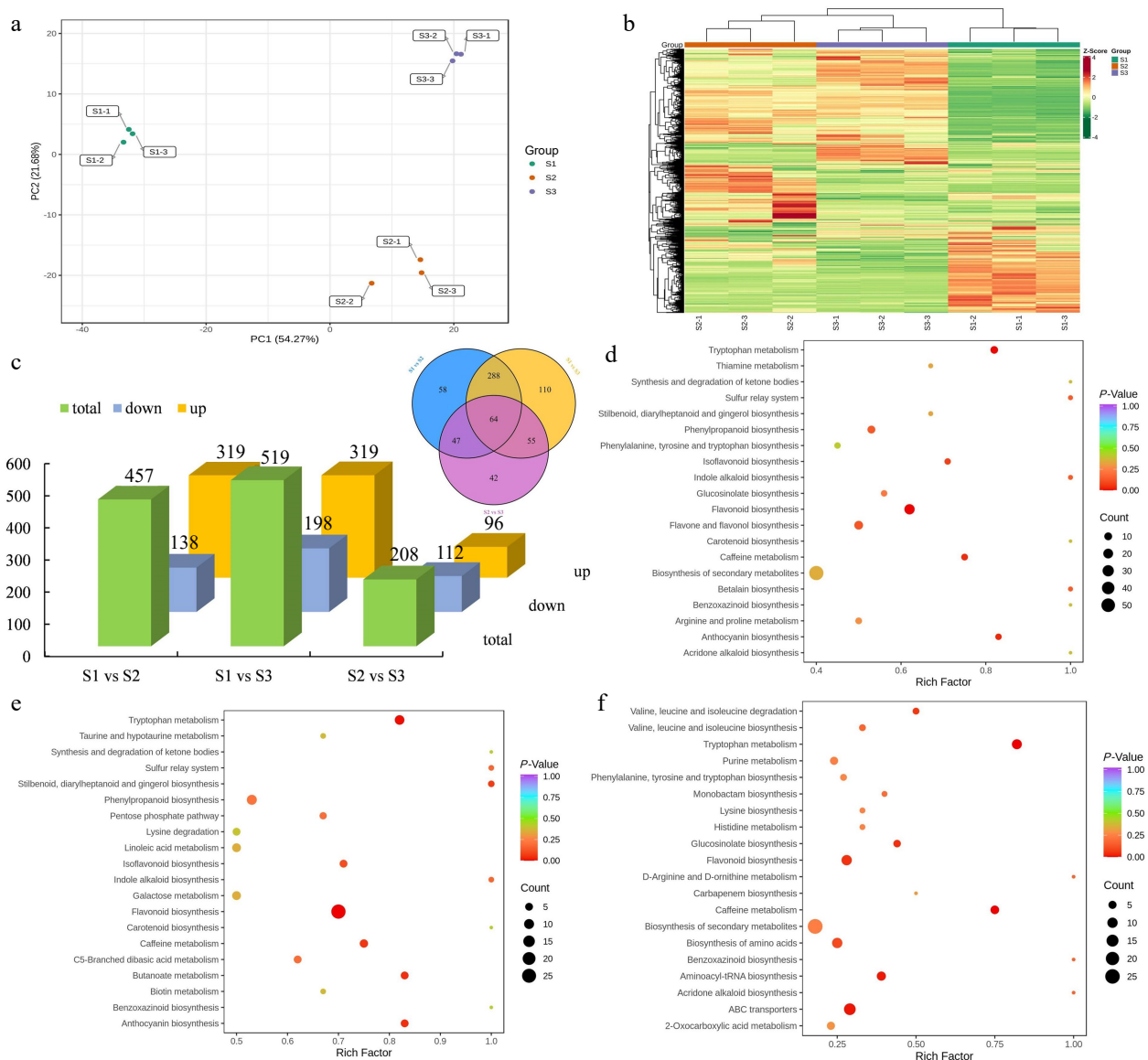
**Fig. 4** Identification and enrichment analysis of DEGs in three comparison groups. (a) UpSet and Venn diagram of DEGs in three comparison groups. (b) Volcano plot of upregulated and downregulated DEGs in three comparison groups. (c) Gene Ontology enrichment circle diagram of DEGs. (d) KEGG enrichment analysis network diagram of DEGs.

Aroma formation of flowers in nectar tree

These DEMs mainly comprising flavonoids (189), phenolic acids (105), lipids (74), amino acids and derivatives (59), organic acids (43), nucleotides and derivatives (39), alkaloids (36), lignans and coumarins (35), terpenoids (26), tannins (10) and other metabolites (48) (Supplemental Table S7). Principal component analysis (PCA) demonstrated distinct variations in the overall metabolite profiles among the S1, S2, and S3 groups, whereas the within-group sample variability was minimal (Fig. 5a). Furthermore, hierarchical clustering analysis was implemented to examine the patterns of metabolite accumulation across samples. The findings revealed consistent division of all samples into three distinct groups, aligning with the results obtained from PCA. Notably, specific metabolites exhibited higher concentrations during the S2 and S3 stages, suggesting their significant involvement in flower development (Fig. 5b). The comparative analysis between S1 and S2 revealed a total of

457 DEMs, with 319 upregulated and 138 downregulated. Similarly, in the S1 vs S3 and S2 vs S3 comparison groups, 519 and 208 DEMs were found, respectively. In addition, a shared group of 64 metabolites exhibited differential expression across all three comparison groups, indicating their consistent modulation throughout the various developmental stages (Fig. 5c).

Moreover, KEGG enrichment analysis was performed for DEMs. In the comparison between S1 and S2 groups, the DEMs were categorized into 74 KEGG-enriched pathways. Similarly, in the S1 vs S3 and S2 vs S3 comparisons, the overrepresented metabolic pathways consisted of 80 and 49 pathways, respectively. Notably, the KEGG pathways of "biosynthesis of secondary metabolites", "phenylpropanoid biosynthesis", and "flavonoid biosynthesis" were consistently identified as the shared and highly enriched pathways in the group comparisons of S1 vs S2 (Fig. 5d), S1 vs S3 (Fig. 5e), and S2 vs S3



**Fig. 5** Identification and enrichment analysis of DEMs. (a) PCA of the metabolome data from three stages. (b) Cluster heatmap of all DEMs in three stages. (c) Statistical analysis of upregulated and downregulated DEMs in three comparison groups. (d) KEGG enrichment analysis of the DEMs in the S1 vs S2 group. (e) KEGG enrichment analysis of the DEMs in the S1 vs S3 group. (f) KEGG enrichment analysis of the DEMs in the S2 vs S3 group.

(Fig. 5f). In addition, "ubiquinone and other terpenoid-quinone biosynthesis", "starch and sucrose metabolism", and "phenylalanine metabolism" were also overrepresented in the three comparison groups. According to the statistical analysis of enriched DEGs and DEMs, we observed significant enrichment of pathways related to terpenoid and phenylpropanoid biosynthesis. Overall, these results further confirmed that the DEMs involved in terpenoid and phenylpropanoid biosynthesis pathways were significantly altered during flower development in *T. amurensis*.

### DEGs are involved in hormone biosynthesis and signal transduction in the different flower development stages

Given the observed elevation in concentrations of IAA and BR throughout flower development, we proceeded with an in-depth examination of potential structural genes linked to the synthesis and signal transduction pathways of IAA and BR. Heat maps were generated to investigate the gene expression, incorporating the annotation information of the DEGs obtained from diverse public databases. IAA is the most abundant auxin in plants. By regulating the expression of structural genes in the auxin signaling pathway, it also indirectly affects plant growth and development<sup>[57]</sup>. Under the influence of different enzymes, IAA is synthesized through both the tryptamine (TAM) pathway and the indole pyruvic acid (IPA) pathway, starting from tryptophan (Trp). Moreover, IAA can also be synthesized *via* the indole-3-acetaldoxime (IAOx) pathway through the key intermediate indole acetonitrile (IAN) *via* the catalytic activity of indoleacetaldoxime dehydratase (CYP71A13) and nitrilase (NIT). In this study, 12 DEGs were identified, including the genes of *TAA1*, *trpB*, *DDC*, and *ALDH* family (Fig. 6a). The expression of *trpB* and *DDC* genes was upregulated in S2 and S3, suggesting their key role in auxin synthesis. The *ALDH* gene family had both up/downregulated genes at different flowering stages. Through differential gene expression screening, 67 DEGs, including the genes of *ARF*, *AUX/IAA*, *GH3*, *TIR1*, *SAUR*, and *AUX1* family, were detected in this pathway (Fig. 6a & Supplemental Table S8). In the examination of gene expression patterns, it was discovered that the *TIR1*, *GH3*, *ARF*, and *AUX/IAA* gene families exhibited a combination of up-regulated and down-regulated genes throughout all three flowering stages.

Brassinolide, a steroid hormone, acts as a plant growth regulator and is crucial for embryo, flower, and fruit development<sup>[58]</sup>. Regarding the BR synthesis pathway (Fig. 6b & Supplemental Table S9), we found two families (*CYP90B/724B* and *CYP85A1/2*) of genes that were differentially expressed at different stages of flower development. Significantly, within the *CYP85A1/2* family, all genes exhibited elevated expression levels during S2 and S3 stages, with the exception of the gene *Cluster-20864.178597*. Moreover, the transmission of BR signals from the receptor *BRI1* to the transcription factor *BZR1* involves a cascade of phosphatases and kinases, including *BAK1*, *BRI1*, *BSK*, *BIN2*, *TCH4*, *CYCD3*, *BSU1*, and *BZR1/2*<sup>[59]</sup>. The genes encoding *BAK1*, *BRI1*, *BSK*, *TCH4*, and *BSU1* showed both up/downregulation during flower development, while the expression levels of *CYCD3* and *BZR1/2* family genes were downregulated of S2 and S3 stages. Specifically, the expression profiles of genes belonging to the *BIN2* family were significantly elevated in the S2 stage, suggesting their potential role in promoting cell division during flower development.

### Analyses of the DEGs and DEMs involved in the terpenoid biosynthesis pathway during flowering

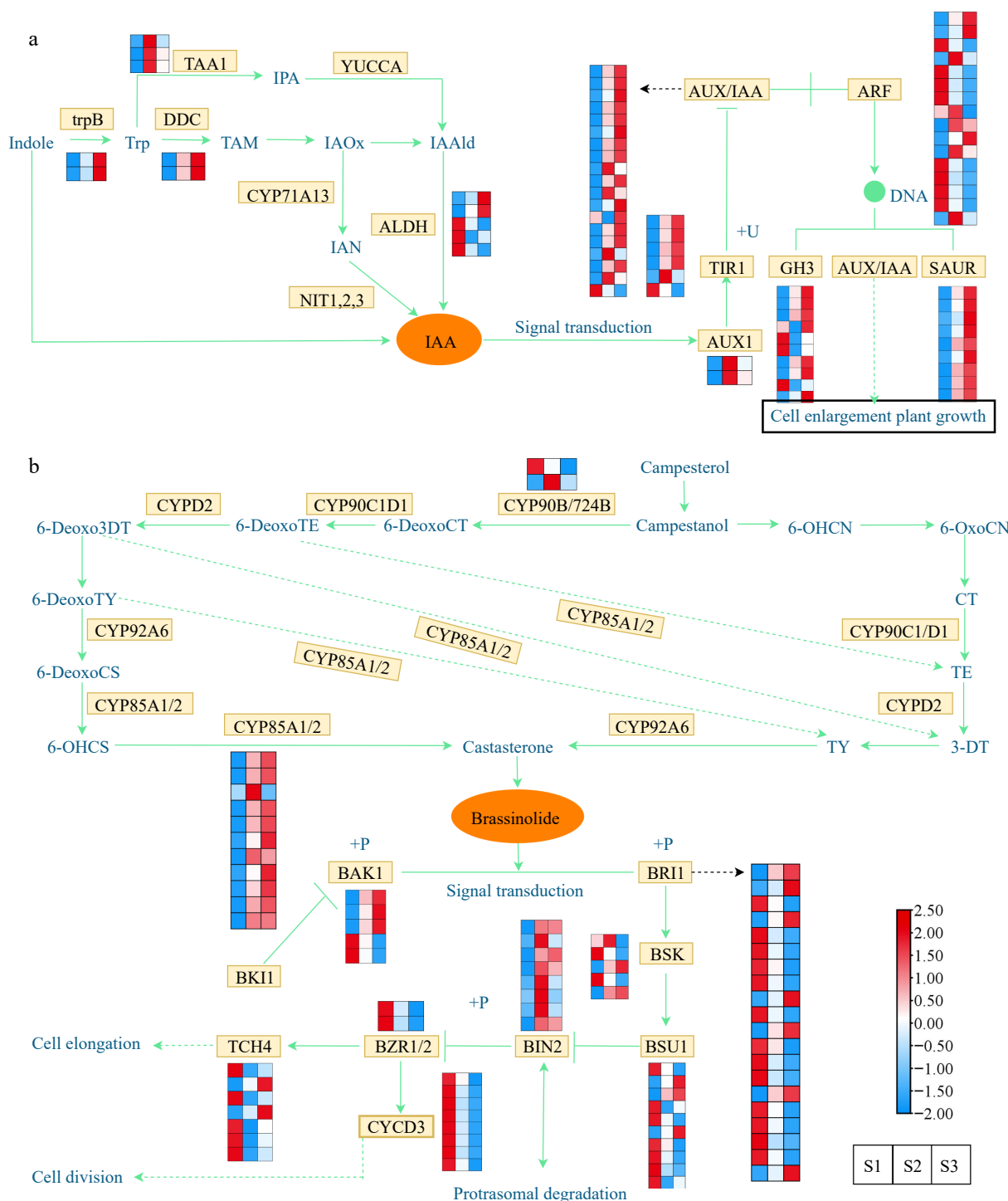
The biosynthesis pathway responsible for producing "terpenoid backbones" involves both the mevalonate (MVA) pathway and the 2-C-methyl-D-erythritol 4-phosphate (MEP) pathway. These pathways are responsible for the synthesis of monoterpenes and sesquiterpenes, respectively<sup>[60]</sup>. To enhance our comprehension of the molecular mechanisms behind the variations in aroma during *T. amurensis* flower development, the expression profiles of 165 DEGs associated with the MVA and MEP pathways, which are key pathways for terpene synthesis, were analyzed (Fig. 7a & Supplemental Table S10). Among them, only one DEG corresponded to *PMK* and *MCT*, while more than one DEG corresponded to *ACAT*, *HMGs*, *DXR*, and *TPS*. Moreover, 165 terpene biosynthesis pathway genes were significantly differentially expressed. Certain DEGs, such as genes encoding *LUP*, *DXS*, and *CHIP*, had low expression in S2 and S3 stages and high expression in the S1 stage, whereas others showed the opposite pattern (e.g., genes encoding *ACAT*, *NADPH*, and *HDR*). During flower development, particularly in the middle and late stages (S2 and S3), the expression profiles of the majority of DEGs related to terpene synthesis exhibited a substantial increase. This observation suggests that these stages are crucial for terpenoid production in flowers.

To gain a deeper understanding of the dynamics of DEMs accumulation, a K-means cluster analysis was conducted utilizing the relative contents of metabolites throughout the S1, S2, and S3 flowering stages. Throughout flowering, the accumulation patterns of the DEMs displayed various trends. Remarkably, during the S2 and S3 stages, DEMs assigned to cluster 5 and cluster 6, respectively, demonstrated considerably higher accumulation levels compared to the S1 stage (Fig. 7b). Intriguingly, the observed accumulation patterns of these DEMs closely mirrored the expression changes observed in the majority of genes associated with the terpenoid synthesis pathway during flower development. Terpenoids are key metabolites and important constituents of floral odors in many plants. Therefore, the DEMs in Cluster 5 and Cluster 6 were further investigated to identify the major terpenoids produced during the different flowering stages. 20 DEMs belonging to triterpenoids were found in the two clusters, including maslinic acid, corosolic acid, and jujubogenin, etc. They showed high accumulation levels in the S2 and S3 flowering stages (Fig. 7c & Supplemental Table S11). This suggests that the accumulation of these metabolites may directly affect the aroma production and profile in the flowers of *T. amurensis*.

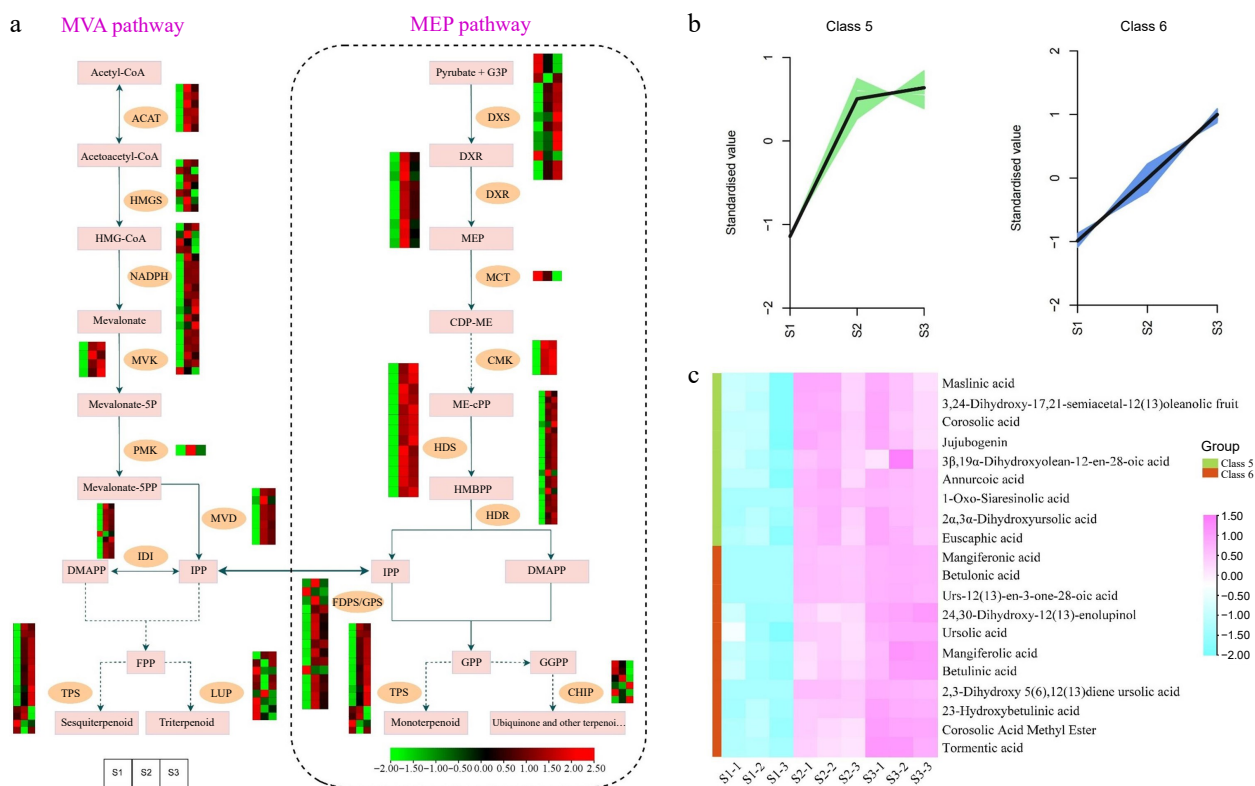
### Identification and expression analysis of DEGs and DEMs in the phenylpropanoid biosynthesis pathway during flowering

The phenylpropanoid biosynthesis pathway, accountable for generating phenylpropanolamine and various other compounds, serves crucial functions in plant growth, defense mechanisms, and acts as a primary pathway governing aroma production<sup>[30]</sup>. The KEGG enrichment analysis revealed an overrepresentation of the 'phenylpropanoid biosynthesis' pathway in three comparison groups, based on the analysis of DEGs and DEMs. We investigated the expression profiles of crucial genes and the accumulation of metabolites associated with the phenylpropanoid biosynthesis pathway. The analysis revealed that a combined total of ten DEMs (including sinapic acid,





**Fig. 6** Expression profiles of DEGs associated with phytohormone biosynthesis and signal transduction. (a) Expression profiles of DEGs involved in the auxin. (b) Expression profiles of DEGs involved in the brassinosteroid. The expression profiles of genes at the three stages are displayed in separate columns within the colored boxes, while each row represents a distinct gene. The color gradient indicates the expression levels, with red indicating upregulation and blue indicating downregulation. trpB, tryptophan synthase beta chain; DDC, tryptophan decarboxylase; TAA, tryptophan aminotransferase; YUCCA, indole-3-pyruvate monooxygenase; CYP71A13, indoleacetaldoxime dehydratase; ALDH, aldehyde dehydrogenase; NIT, nitrilase; AUX1, auxin influx carrier 1; TIR1, transport inhibitor response 1; AUX/IAA, auxin-responsive protein IAA; ARF, auxin response factor; GH3, gretchen hagen 3; SAUR, small auxin upregulated RNA; CYP90B/724B, steroid 22-alpha-hydroxylase; CYP92A6, typhasterol/6-deoxotyphasterol 2alpha-hydroxylase; CYP85A1/2, brassinosteroid-6-oxidase 1/2; BRI1; BRI1 kinase inhibitor 1; BAK1, brassinosteroid insensitive 1-associated receptor kinase 1; BSK, BR-signaling kinase; BSU1, serine/threonine-protein phosphatase; BIN2, brassinosteroid insensitive 2; BZR1/2, brassinosteroid resistant 1/2; TCH4, xyloglucan: xyloglucosyl transferase; CYCD3, cyclin D3.



**Fig. 7** Illustrations and expression patterns of DEGs and DEMs implicated in the biosynthesis pathway of terpenoids. (a) Expression levels of DEGs. (b) K-means cluster analysis of DEM accumulation. (c) Heatmap of DEMs in the terpenoid biosynthesis pathway. The expression profiles of genes at the three stages are displayed in separate columns within the colored boxes, while each row represents a distinct gene. The color gradient indicates the expression levels, with red indicating upregulation and green indicating downregulation. ACAT, acetyl-CoA C-acetyltransferase; HMGS, 3-hydroxy-3-methyl-glutaryl-CoA synthase; MVK, mevalonate kinase; PMK, phosphomevalonate kinase; MVD, mevalonate-5-pyrophosphate decarboxylase; IDI, isopentenyl-diphosphate delta-isomerase; LUP, lupeol synthase; DXS, 1-deoxy-D-xylulose 5-phosphate synthase; DXR, 1-deoxy-D-xylulose-5-phosphate reductoisomerase; MCT, 2-C-methyl-D-erythritol 4-phosphate cytidyltransferase; CMK, 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase; HDS, 1-hydroxy-2-methyl-2-butenyl 4-diphosphate synthase; HDR, 1-hydroxy-2-methyl-2-butenyl 4-diphosphate reductase; FDPS/GPS, farnesyl diphosphate synthase/geranyl diphosphate synthase; TPS, terpene synthases.

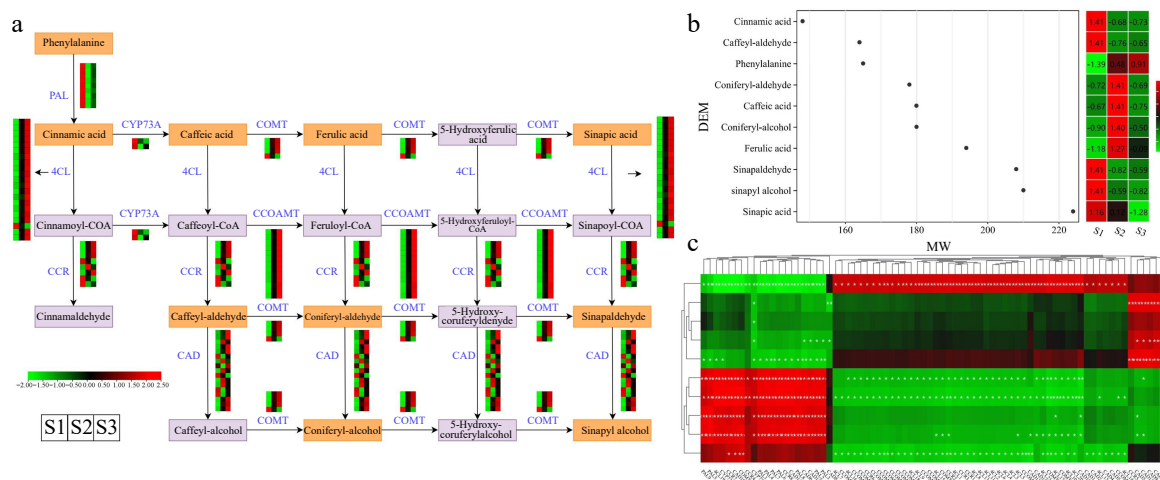
phenylalanine, and ferulic acid, etc.) and 74 DEGs (including nine *PALs*, eight *CCRs*, and four *COMTs*, etc.) exhibited distinct expression patterns throughout flower development (Fig. 8a & Supplemental Table S12). Heatmap analysis revealed that certain structural genes, such as one gene (*Cluster-20864.233536*) encoding for *CCR* and five genes (*Cluster-20864.112764*, *Cluster-20864.271610*, *Cluster-20864.253472*, *Cluster-20864.220953*, and *Cluster-20864.76579*) encoding for *CADs*, exhibited particularly elevated expression levels during the S2 stage in flowers. Moreover, apart from *Cluster-20864.148160*, the expression profiles of genes encoding *4CL* and *CCOMT* were elevated in the S2 and S3 developmental stages, in contrast to the S1 stage.

Based on the heatmap illustrating the accumulation levels of ten DEMs (with molecular weights between 145 and 225), the accumulation of ferulic acid, coniferyl-alcohol, caffeic acid, and coniferyl-aldehyde demonstrated a substantial elevation in the S2 stage, when compared to the S1 stage (Fig. 8b & Supplemental Table S13). Notably, phenylalanine accumulation exhibited a gradual upward trend, with the highest accumulation levels reaching during the S3 stage. In addition, we conducted a correlation analysis to explore potential regulatory relationships between the 10 DEMs and structural genes in the phenylpropanoid biosynthesis pathway (Fig. 8c & Supplemental Table S14). The results illustrated a noteworthy positive association

between the expression levels of *PALs*, *4CLs*, *CCRs*, and *CADs*, and the accumulation of cinnamic acid, sinapyl alcohol, sinapaldehyde, and caffeoyl-aldehyde. Notably, four DEMs (caffeic acid, coniferyl-aldehyde, coniferyl-alcohol, and ferulic acid) showed positive correlations only with *CAD*-related genes (including *CAD7*, *CAD9*, and *CAD12*). The findings suggested that the DEGs and DEMs identified have significant functions in the phenylpropanoid biosynthesis pathway and are probably key contributors to the flower's aroma accumulation in *T. amurensis*.

### Identification of transcription factors (TFs) and network analysis of structural genes involved in aroma formation

Apart from key structural genes, TFs play a vital role in directly modulating the expression profiles of multiple genes and metabolic pathways, consequently facilitating the generation and emission of floral fragrance. In current study, we identified 409 TFs from the transcriptome data, distributing across 45 families. Among these, the highest number of TFs were distributed in the *NAC* (37), *WRKY* (32), *C2H2* (29), and *bHLH* (28) families, among others (Fig. 9a & Supplemental Table S15). To explore the regulatory connections among the identified TFs, genes associated with terpenoid metabolism, and genes involved in phenylpropanoid biosynthesis, a co-expression



**Fig. 8** Illustrations and expression patterns of DEGs and DEMs involved in the phenylpropanoid biosynthesis pathway. (a) Expression levels of DEGs. (b) Heatmap of DEMs in the phenylpropanoid biosynthesis pathway. (c) Correlation analysis of phenylpropanoid biosynthesis-related structural gene expression profile and DEM accumulation. The expression profiles of genes at the three stages are displayed in separate columns within the colored boxes, while each row represents a distinct gene. The color gradient indicates the expression levels, with red indicating upregulation and green indicating downregulation. PAL, phenylalanine ammonia-lyase; 4CL, 4-coumarate-CoA ligase; CCR, cinnamoyl-CoA reductase; COMT, caffeic acid 3-O-methyltransferase; CCOAMT, caffeoyl-CoA O-methyltransferase; CAD, cinnamyl-alcohol dehydrogenase; CYP73A, trans-cinnamate 4-monoxygenase.

network was established. For this analysis, homologous proteins of *Arabidopsis* were sourced from STRING database.

The network depicted in Fig. 9b integrated genes related to terpenoid metabolism and the top five TF families among the DEGs. Notably, the analysis revealed robust interactions between these TFs and the genes involved in terpenoid metabolism. Notably, *MCT*, *MVD7*, *DXS11*, *HDR22*, and *LUP9* exhibited significant interaction with TFs, implying their potential roles in regulating aroma formation in *T. amurensis* flowers. Moreover, we discovered that the *C2H2-17* exhibited the most numerous expressions with other genes. This may indicate that it functions as the central regulator in the network. A second network (Fig. 9c) revealed the relationship between TFs and genes associated with phenylpropanoid biosynthesis. Among the genes present in the network, *CAD12*, *PAL9*, and *4CL13* exhibited notable interactions with TFs in comparison to the remaining genes. Unlike the terpenoid metabolism regulatory network, our findings indicated that *MYB-24* exhibited the strongest connectivity with phenylpropanoid biosynthesis-related genes, thus considered the hub regulator of phenylpropanoid biosynthesis. Interestingly, multiple shared transcription factors were present in both networks, including *WRKY-21*, *NAC-5*, *NAC-22*, and *WRKY-24*, etc, and strongly correlated with terpenoid metabolism and phenylpropanoid biosynthesis genes.

Sixteen TFs from the co-expression network were chosen for further investigation, and their expression patterns were assessed in the different stages (Fig. 9d). The expression levels of five shared TFs, namely *WRKY-26*, *NAC-11*, *C2H2-18*, *MYB-6*, and *WRKY-14* were notably high during the S2 and S3 stages of flower development. Furthermore, in the S2 and S3 stages, two central regulators, *C2H2-17* and *MYB-24*, displayed pronounced expression, further supporting their essential role in the modulation of floral fragrance development. Certain genes (including *WRKY-21*, *NAC-5*, *NAC-22*, and *WRKY-24*) declined their expression profiles in S2 and S3 stages relative to the S1 stage,

suggesting that they may be negative regulators of flower growth and development. These results indicate that transcription factors have strong interactions with genes related to floral scent metabolism and are an indispensable component of floral scent synthesis.

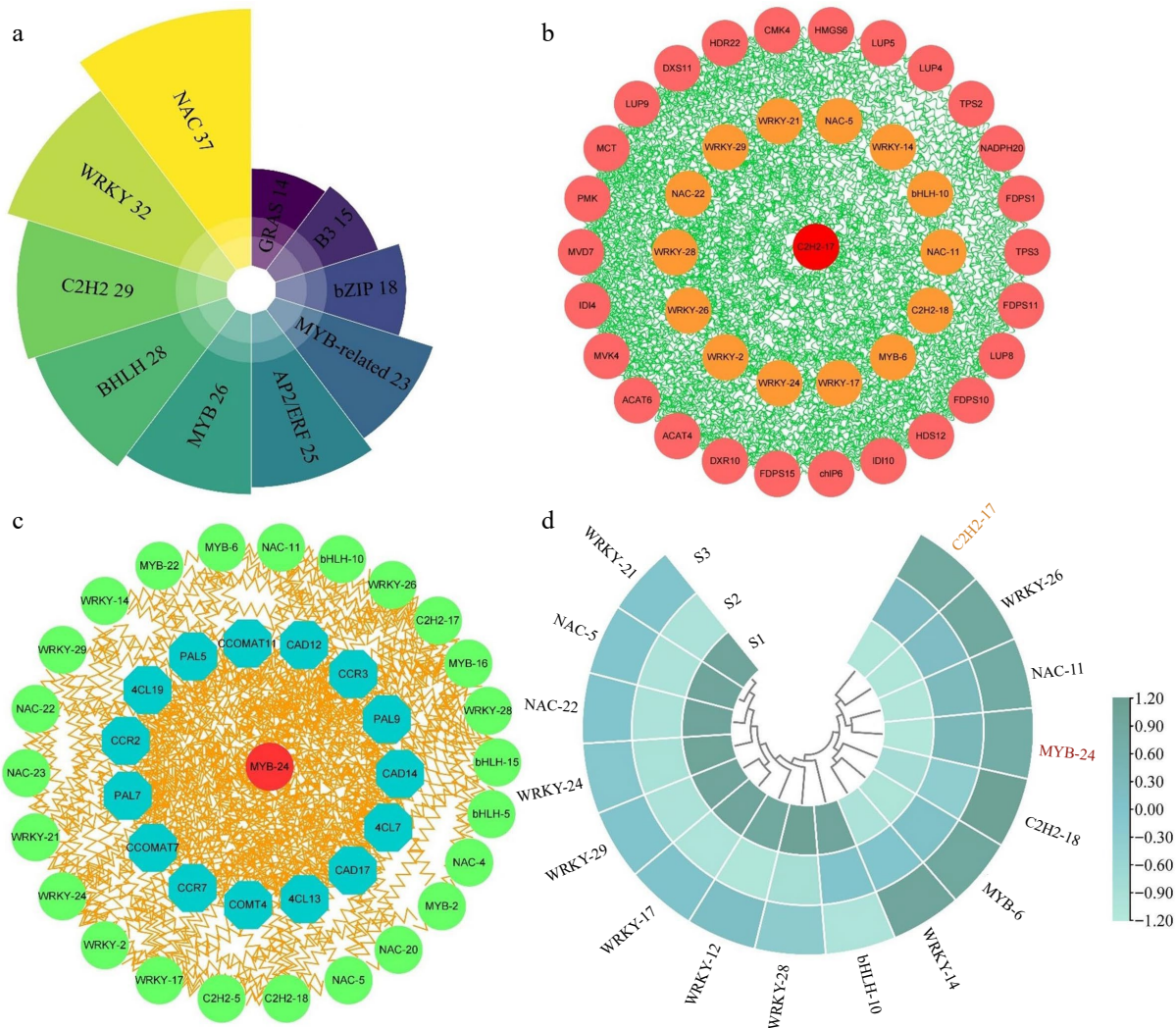
### Validation of RNA-Seq data by RT-qPCR

A total 14 DEGs were chosen for RT-qPCR assays to confirm the reliability of the RNA-seq data. Among them, five DEGs (*Cluster-20864.228631*, *Cluster-20864.206192*, *Cluster-20864.167641*, *C2H2-17*, and *MYB-24*) are transcription factors, three DEGs (*Cluster-20864.198501*, *Cluster-20864.140634*, *Cluster-20864.162042*) are associated with the terpenoid biosynthesis pathway, three DEGs (*Cluster-20864.210350*, *Cluster-20864.215451*, *Cluster-20864.197278*) are related to the phenylpropanoid biosynthesis pathway, while the rest are ordinary differentially expressed genes. The results demonstrated a high degree of agreement between the RNA-seq data and qPCR data in terms of gene expression patterns, exhibiting a correlation coefficient of 0.8584 (Supplemental Figs S1, S2, & S3). These results provide evidence of the dependable nature of the transcriptomic data.

## Discussion

### Alterations in phytohormone levels and their regulatory functions of flower development

Phytohormones assume a central role in governing plant development, orchestrating diverse processes including cell division, cell expansion, differentiation, as well as flower and fruit development<sup>[61]</sup>. Flower development involves various plant hormones, which, synergistically or antagonistically, influence flower size, shape, color, and fragrance<sup>[62,63]</sup>. Previous research has provided evidence supporting the role of auxin in promoting the growth and expansion of floral structures, as well as its ability to stimulate the synthesis of fragrance compounds<sup>[64,65]</sup>. A report indicates that the phytohormones



**Fig. 9** Analysis of correlation and expression patterns between structural genes and transcription factors. (a) Statistical analysis of the highest number of ten transcription factor families in three comparison groups. (b) Correlation network analysis of terpenoid biosynthesis-related structural genes and the highest number of five transcription factor families. (c) Correlation network analysis of phenylpropanoid biosynthesis-related structural genes and the top five transcription factor families. (d) The expression patterns of 16 selected transcription factors.

GA and JA act as growth regulators, accelerating flowering in *Arabidopsis thaliana*<sup>[66,67]</sup>. Furthermore, ABA metabolism also plays a crucial role in sweet cherry (*P. avium*) flower development and fruit setting process<sup>[68]</sup>. Our study revealed significant variations in the levels of eight endogenous hormones during different flower development stages, indicating their pivotal function in regulating flower development. Consistent findings have been documented in other plant species, including *Castanea henryi*<sup>[69]</sup> and *Gerbera hybrida*<sup>[70]</sup>, further supporting the results.

Plant hormones serve as signaling molecules and can significantly impact plant growth and metabolism<sup>[71,72]</sup>. Our observations align with the above results, as we also detected the enrichment of DEGs in the 'plant hormone signal transduction' pathway. Moreover, compared with other hormones, the concentrations of IAA and BR showed a progressive upward trend along with flower development, so we speculate that their signaling may play critical and flowering-specific regulatory roles. Previous studies showed that auxin response factors influence leaf and flower development. It has been found that

genes related to auxin biosynthesis and signal transduction contribute to the flowering of *Dendrobium officinale*<sup>[72,73]</sup>. Significantly, the majority of structural genes of auxin biosynthesis and signal transduction pathway, such as *AUX/IAA*, *GH3*, and *SAUR*, exhibited upregulation during the S3 stage in our investigation, highlighting their crucial involvement in stimulating flower growth. In research on flower development in sugar apples, several *Aux/IAA* genes were significantly differentially expressed, consistent with our findings<sup>[74]</sup>. Furthermore, in alfalfa, *MsGH3-5/6/11* acts synergistically during flower development and exhibits relatively high flower expression levels<sup>[75]</sup>. The *SAUR* gene family, responsible for transmitting early auxin responses, holds vital functions in plant development<sup>[76]</sup>. In the current study, it was evident that all DEGs affiliated with the *SAUR* gene family were specifically upregulated during the S3 stage. This observation suggests that the expression of the *SAUR* gene family potentially facilitates the developmental processes involved in flower growth. Similarly, the *SAUR* gene family could regulate pineapple's floral organ and fruit development<sup>[77]</sup>. Additionally, studies have shown that the

## Aroma formation of flowers in nectar tree

*CmSAUR13*, 28, 31, and 39 genes exhibited pronounced expression in female flowers and played a regulatory role in the process of flower development in *Cucumis melo*<sup>[78]</sup>.

Plant-specific steroid hormones known as brassinosteroids (BRs) are key regulators of various physiological processes ranging from seed development to flowering and senescence<sup>[79]</sup>. The observed fluctuations in the expression levels of structural genes (e.g., *CYP85A1/2* and *BAK1*) associated with the brassinosteroid signaling pathway during the S3 stage suggest their potential significance in the transcriptional regulation of flower development. Notably, overexpression of the *Arabidopsis* glycosyltransferase *UGT73C5*, which catalyzes the glycosylation of brassinosteroids, leads to significant modifications in flower development and flowering time<sup>[80]</sup>. Moreover, our findings are consistent with the notion that BR signaling serves as a promoter during vegetative and reproductive development, as demonstrated in *Nelumbo nucifera*<sup>[16]</sup>. Furthermore, a significant association has been established between plant hormone content, as well as genes involved in plant hormone synthesis and signal transduction, with the production of floral aroma<sup>[81,82]</sup>. Taken together, the collective action of multiple phytohormones is crucial for the precise regulation of flower development in *T. amurensis*.

### Potential roles of genes and metabolites engaged in the terpenoid biosynthesis pathway

Plant floral fragrance is a result of volatile aromatic compounds, including terpenes, coumarins, esters, and phenolic compounds<sup>[83]</sup>. Terpenoids, being highly abundant in plant volatiles, serve as a primary source of floral fragrance and play a crucial role as an indicator of fruit ripeness<sup>[84]</sup>. We discovered that the accumulation of 20 DEMs that were annotated as terpenoids, including maslinic acid, corosolic acid, jujubogenin, and tormentic acid, increased significantly in S2 and S3, potentially being the primary cause of aroma generation in *T. amurensis* flowers. Similar results were reported in *Camellia sinensis*<sup>[21]</sup> and *Vitis vinifera*<sup>[85]</sup>. In *Chrysanthemum indicum*, it was similarly demonstrated that the potential changes in floral scent are mainly caused by changes in volatile terpenoid metabolites during flower development<sup>[86]</sup>. According to the research conducted by Yue et al., the levels of monoterpenes present in grapes varied across their five stages of ripening, with higher concentrations observed at harvest than the initial ripening<sup>[87]</sup>. Furthermore, previous investigations have suggested that variations in the terpenoid components and relative terpenoid abundance were the primary factors contributing to the divergence in aroma between green and red prickly ash. These observations align with our findings in the flowers of *T. amurensis*<sup>[84]</sup>.

Extensive research has been conducted on the terpenoid biosynthesis pathway in various plant species, including *Jasminum sambac*<sup>[23]</sup>, *Camellia sinensis*<sup>[88]</sup>, and *Osmanthus fragrans*<sup>[89]</sup>. These studies have contributed to the identification of crucial genes connected to fragrance production. Nonetheless, the molecular mechanisms responsible for floral fragrance production in *T. amurensis* remain ambiguous due to the limited availability of genetic information. Remarkably, our analysis of the results unveiled a significant enrichment of DEGs in the terpenoid biosynthesis pathway, highlighting their indispensable role in the establishment of floral fragrance during development. Several key structural genes regulate the synthesis of terpenoid compounds. Recent studies identified, through

comparative transcriptome analysis, *DXS*, *DXR*, *MCT*, *CMT*, and *TPS* as critical structural genes of the terpenoid biosynthetic pathway for the production of floral fragrance<sup>[90,91]</sup>. Within our study, a total of 165 genes associated with terpenoid biosynthesis have been successfully identified. A significant number of structural genes, such as *ACAT*, *MVK*, *TPS*, *DXS*, *DXR*, *HDS*, *HDR*, and *GPS*, exhibited considerable differential expression levels during flower development and may promote the production of floral fragrance in *T. amurensis*.

Earlier research has indicated that heightened activity of upstream genes, like *ACAT* in the MVA pathway, and *DXS* in the MEP pathway, has the potential to enhance the metabolic flux towards terpenoid biosynthesis and the production of its building blocks<sup>[90,92]</sup>. Our findings corroborate previous research, demonstrating that the expression profiles of *HMGs* genes in the MVA pathway were notably higher in *Lilium* 'Siberia' (known for its strong fragrance emission) compared to *Lilium* 'Novano' (known for its very faint fragrance emission)<sup>[93]</sup>. Specifically, in our study, we observed high expression levels of *HMGs*-related genes during S2, implying their potential to enhance the production of aroma compounds. Moreover, in the MVA pathway, Yan et al. reported that suppressing the *IDI* genes reduced sesquiterpene content in *Pogostemon cablin*<sup>[94]</sup>. The second step in the MEP pathway is facilitated by *DXR*. A study has revealed that *DXR* overexpression in transgenic tobacco increases *DXR* activity, increasing the levels of photosynthetic pigments and volatile isoprenoid components<sup>[95]</sup>. Silencing of the *HDS* and *HDR* genes in *Nicotiana benthamiana* leaves. Inhibited the MEP pathway, indicating the involvement of *HDS* and *HDR* in terpenoid biosynthesis<sup>[96]</sup>. Upregulation of the downstream genes *GPS* and *TPS* was observed during the flowering process. In the previous study, the biosynthesis of monoterpenes involves the production of GPP, which is carried out by *GPS*. *GPS* has been studied and characterized in multiple plant species, including *Arabidopsis*<sup>[97]</sup>, *Phalaenopsis bellina*<sup>[98]</sup>, and *Jasminum sambac*<sup>[23]</sup>. *TPSs* constitute a highly diversified medium-sized gene family involved in plants' terpenoid biosynthesis. In *Arabidopsis*, it has been documented that the biosynthesis of nearly all 20 sesquiterpenes is attributed to the enzymes *TPS21* and *TPS11*<sup>[99]</sup>. Furthermore, studies have revealed that the *TPS* family genes are implicated in the emission of characteristic aroma terpenes in ripe kiwifruit and *Litchi chinensis*<sup>[100,101]</sup>. In our investigation, we observed a significant upregulation of 13 *TPS* genes during the S3 stage, suggesting a direct association between *TPS* expression and the generation of floral fragrance. These findings coincide with the outcomes reported in the aforementioned study, reinforcing the role of *TPS* in modulating the production of floral scent.

### The functions of genes and metabolites involved in phenylpropanoid biosynthesis in floral scent production

Phenylpropanoid compounds, following terpenes, represent the second most prevalent group of floral constituents in plants. They are synthesized through a complex series of branching pathway reactions starting from phenylalanine, the initial substrate<sup>[102]</sup>. The phenylpropanoid synthesis pathway, also known as the cinnamic acid pathway, is a component of the phenylpropionic acid synthesis pathway<sup>[103]</sup>. The crucial role of the phenylpropanoid biosynthetic pathway in the production of floral fragrance has been widely studied in plants

such as *Rhododendron fortunei*<sup>[103]</sup>, *Chimonanthus praecox*<sup>[104]</sup>, and *Chimonanthus praecox*<sup>[105]</sup>. Moreover, several essential structural genes, such as *PAL*, *C4H*, *4CL*, *SAMT*, *CCOAMT*, *CCR*, *CAD*, *EGS*, and *BALDH*, are involved in regulating phenylpropanoid biosynthesis<sup>[103]</sup>. In the current study, KEGG enrichment analysis identified a notable enrichment of DEGs in the phenylpropanoid biosynthesis pathway, consistent with previous findings in related studies. Moreover, 74 DEGs associated with phenylpropanoid biosynthesis displayed distinct expression patterns and could potentially enhance the production of floral fragrance in *T. amurensis*. Besides structural genes, 10 metabolites of the phenylpropanoid biosynthesis pathway also showed differential accumulation, with similar results observed in *Prunus mume*<sup>[106]</sup>.

The *PAL* gene assumes a critical role in governing the synthesis of phenylpropanoid and benzene ring compounds. The expression profiles of *PALs* were downregulated in the S3 stage and was significantly negatively correlated with the contents of phenylalanine, caffeic acid, and ferulic acid, which may indicate a critical negative regulatory role of *PALs* in aroma synthesis. In line with our research, a robust negative correlation was acknowledged between the downregulation of *PAL* expression and the levels of various benzene-like volatile organic compounds in *Rhododendron fortunei*, providing further support to our findings<sup>[103]</sup>. Using DNA microarray analysis, Verdonk and colleagues also identified that *PAL* controls the synthesis of aromatic volatile organic compounds in petunias<sup>[107]</sup>. Then, *4CL* can facilitate the synthesis of cinnamoyl-CoA from *t*-cinnamic acid, which in turn serves as a direct precursor for the production of benzaldehyde<sup>[108]</sup>. In our study, all *4CL* genes, except for *Cluster-20864.148160*, were highly expressed during the S3 stage, and many *4CL* genes were positively correlated with the contents of cinnamic acid and sinapic acid. This suggests that high expression of *4CL* may be crucial for the biosynthesis of volatile compounds responsible for the floral scent during the flowering period. Reports have indicated that two *Pm4CL* genes are significantly upregulated in *Prunus mume*, suggesting their involvement in the synthesis of volatile compounds, cinnamyl ester and eugenol<sup>[106]</sup>. Additionally, studies have shown that *4CL* exerts a crucial function in benzene/phenylpropane compound biosynthesis in *Ocimum sanctum*<sup>[109]</sup>, mulberry<sup>[110]</sup>, and *Agave amica*<sup>[111]</sup>, which aligns with our research findings. *CAD* participates in the oxidation reaction of cinnamyl alcohol to cinnamaldehyde in the phenylpropanoid metabolic pathway and is a key enzyme for aroma compound biosynthesis<sup>[112]</sup>. In the study, we identified the *CAD6*, *CAD10*, *CAD13*, and *CAD14* expression was positively correlated with the content of phenylpropanoid metabolites, indicating that the floral scent is affected by their expression and catalytic activities during flowering. In their extensive analysis of *Prunus mume*, Zhang et al. discovered that *PmCAD1* exerts a crucial function in the *in vitro* synthesis of cinnamyl alcohol, as evidenced by their comprehensive investigation of endogenous volatile compounds and transcriptome data<sup>[113]</sup>. Furthermore, transcriptomic differential expression analysis of petals from different varieties of *Rhododendron fortunei* revealed that 4 *CAD* genes were engaged in the synthesis of aromatic compounds in flowers<sup>[103]</sup>. Furthermore, transcriptomic differential expression analysis of petals from different varieties of *Rhododendron fortunei* revealed that four *CAD* genes were engaged in the synthesis of

aromatic compounds in flowers<sup>[114]</sup>. Moreover, the *CCOAMT* gene expression levels in the skin of pear fruit were associated with skin color fading, which in turn affected the aroma of the pear<sup>[115]</sup>. These studies indicate the potential significance of the phenylpropanoid biosynthesis pathway in the development of floral scent. However, further molecular experimental investigations are needed to explore its specific mechanisms in *T. amurensis* volatile compound production in the flowers.

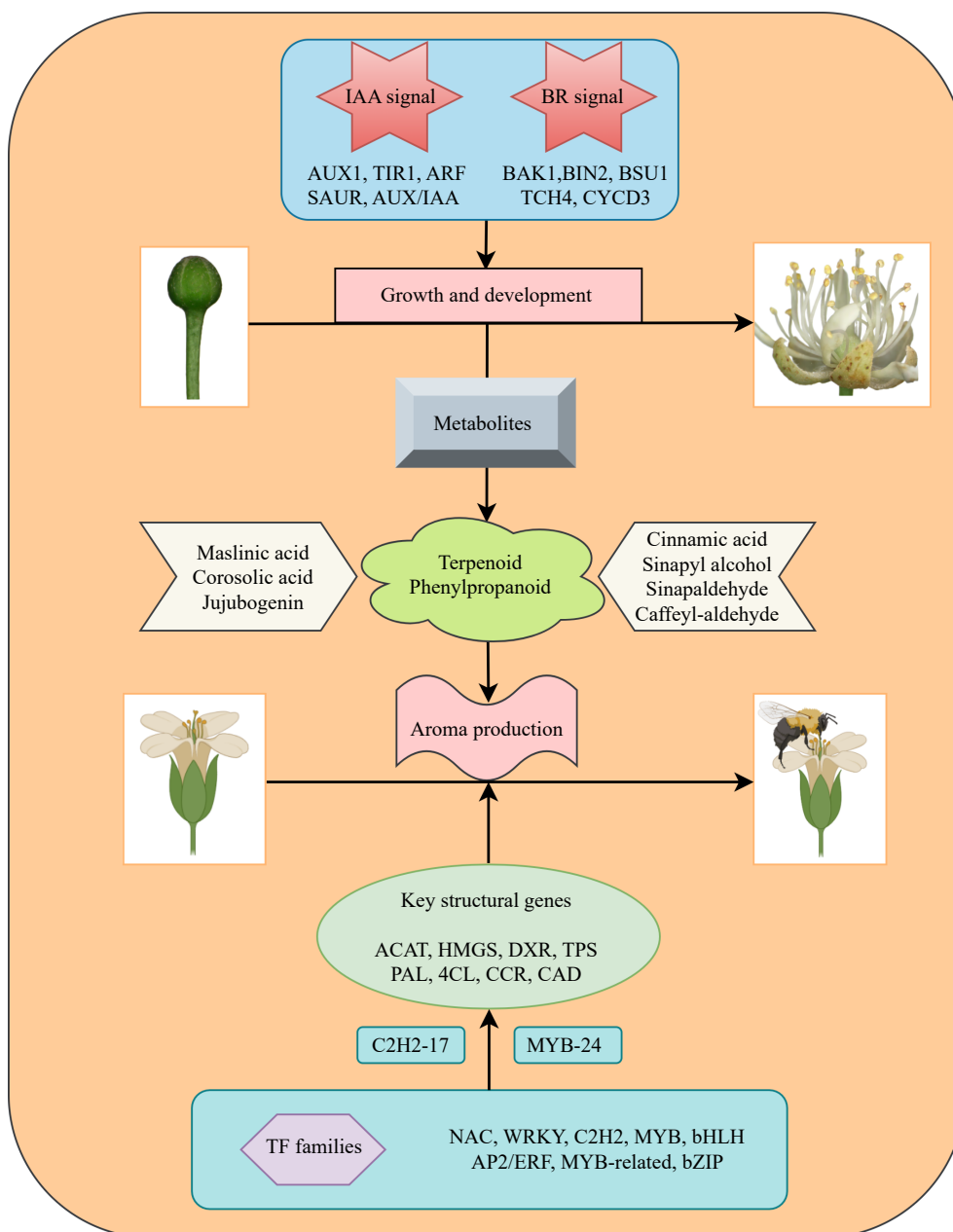
### TFs involved in the regulation of terpenoid and phenylpropanoid biosynthesis

The generation of compounds accountable for floral fragrance requires the expression and function of genes encoding enzymes, which are governed by regulatory genes like transcription factors<sup>[104]</sup>. TFs can regulate the gene expression of the biosynthetic enzymes in the floral scent synthesis pathway, thereby affecting the metabolic pathway flow toward the synthesis of specific floral scent compounds<sup>[19]</sup>. The first identified regulatory transcription factor of floral scent synthesis in plants is *ODO1 (ODORANT1)* in petunias, which belongs to the *R2R3-MYB* TF family and is expressed explicitly in the petals, regulating the biosynthesis of phenylpropanoids/phenylalanine-derived compounds<sup>[36]</sup>. Similarly, in *Phalaenopsis orchids*, the overexpression of *bHLH4* or *bHLH6* results in elevated levels of volatile monoterpenes<sup>[116]</sup>. Furthermore, a recent study revealed the involvement of the *WRKY* gene in aroma synthesis through its regulation of the production of monoterpene volatiles. The expression of *OfWRKYs* is associated with monoterpene synthesis in *Osmanthus fragrans*<sup>[117]</sup>. So far, transcription factor family have been found to regulate terpenoid and phenylpropanoid compounds in plants, including *AP2/ERF*, *WRKY*, *NAC*, *bHLH*, *C2H2*, and *bZIP*, among others. However, there are no comprehensive reports on the transcription factors regulating floral scent synthesis in *T. amurensis*. In our investigation, we observed an overrepresentation of DEGs belonging to the *NAC*, *WRKY*, *C2H2*, *bHLH*, and *MYB* families, suggesting their potential involvement in the comprehensive regulation of flower development in *T. amurensis*. For example, *WRKY-14* the homolog of *AtWRKY71*, which has been reported to accelerate flowering by directly activating *FLOWING LOCUS T* and *LEAFY* in *Arabidopsis thaliana*, which is consistent with our findings<sup>[118]</sup>. Furthermore, *WRKY-21* displays significant expression during the S2 and S3 stages, exhibiting a homology with *Arabidopsis AtWRKY34*. This similarity suggests that *WRKY-21* likely plays a role in flowering, given that homologs in *Arabidopsis* serve as crucial controllers of pollen development<sup>[119]</sup>. Similarly, *MYB-6*, an orthologous counterpart of *AtMYB17*, exhibits notable expression in inflorescences and epidermal tissues, especially during the initial stages of flower development, highlighting the potential involvement of *MYB-6* in floral developmental processes<sup>[120]</sup>. In addition, network analysis showed significant interactions between members of different transcription factor families (such as *NAC-5*, *WRKY-14*, *bHLH-10*, and *MYB-6*) and structural genes involved in terpenoid and phenylpropanoid biosynthesis. It suggests that these TFs may regulate terpenoid and phenylpropanoid biosynthesis and accumulation, thereby affecting the production of floral fragrance. In *Citrus medica*, co-expression network analysis revealed that members of the *MYB*, *C2H2*, *NAC* TF families, and 13 *TPS* genes were associated with the synthesis of volatile terpenoid compounds<sup>[121]</sup>. *MYC* TFs are extensively studied in

## Aroma formation of flowers in nectar tree

the *bHLH* family and have been shown to regulate the biosynthesis of terpenoids in numerous plant species. The *AtMYC2* transcription factor was found to positively regulate *TPS11* and *TPS21* in *Arabidopsis* flowers, resulting in an elevation of sesquiterpene emission, in agreement with our research findings<sup>[122]</sup>. Furthermore, we found that *C2H2-17* and *MYB-24* showed the most numerous and significant interactions with other structural genes in the terpenoid and phenylpropanoid biosynthetic pathways, respectively, and exhibited high expression levels at the S3 stage, which may indicate their central regulatory roles on the synthesis of floral fragrance. Our research findings align with results obtained in winter sweet plants, as the regulation of floral scent compound biosynthesis in winter sweet involves the differential expression of 56 *C2H2* transcription factors<sup>[104]</sup>. In addition, in *Arabidopsis*, the *C2H2* TF

family member *AtZAT10* was shown to regulate the expression of ABA signaling pathway genes, indirectly affecting the synthesis of terpenoid compounds. Furthermore, four *MYB* members regulate phenylpropanoid volatiles in petunias<sup>[123]</sup>, overexpression of *Arabidopsis MYB* genes in petunia flowers demonstrated that regulating the flow of phenylpropanoid precursor metabolism affects floral scent synthesis<sup>[124]</sup>. Moreover, overexpression of the *R2R3-MYB* transcription factor *PtMYB14* from *Pinus taeda* leads to the accumulation of sesquiterpenes in transgenic white spruce (*Picea glauca*) plants<sup>[34]</sup>. The *PmMYB4* was shown to regulate downstream structural genes, leading to significant differences in aroma compounds among different cultivars, which is consistent with our research findings<sup>[106]</sup>. The findings above suggest that TFs assume a crucial role in regulating terpenoid and phenyl-



**Fig. 10** Diagram illustrating the process of flower development and aroma synthesis in the nectar tree (*Tilia amurensis*).

propanoid biosynthesis through the modulation of downstream structural genes, ultimately modulating floral scent production.

## Conclusions

In summary, we utilized metabolomics to quantify and characterize the secondary metabolites accumulated during the development of *T. amurensis* flowers. We further conducted a transcriptomic analysis of the flowers during different flowering stages. A total of 89,526 DEGs and 664 DEMs were discovered in the current study, and enrichment analysis indicated that the biosynthesis pathways of terpenoids and phenylpropanoids were commonly enriched in the different comparison groups. Moreover, the roles of two key plant hormones, IAA and BR, in flower development were evaluated based on their accumulation and the expression patterns of their signal transduction and biosynthesis pathways. Through the integration of transcriptomics and metabolomics techniques, we successfully determined the temporal accumulation patterns of essential metabolites, such as maslinic acid, jujubogenin, cinnamic acid, and sinapic acid. We identified that terpenoids and phenylpropanoids are the major metabolites involved in the production of floral fragrances. We also identified 165 and 74 DEGs associated with terpenoid and phenylpropanoid biosynthesis, respectively, and further predicted the underlying regulatory mechanisms involving structural genes such as *ACAT*, *HDS*, *TPS*, *4CL*, *CAD*, and *CCOAMT* in the synthesis of both compound classes. Furthermore, in this study, we uncovered the involvement of transcription factor families such as *NAC*, *WRKY*, and *MYB* in regulating terpenoid and phenylpropanoid biosynthesis pathways in *T. amurensis* flowers. We identified the transcription factors *C2H2-17* and *MYB-24* as central regulators with positive regulatory roles in the terpenoid and phenylpropanoid biosynthesis pathways, respectively. These findings lay the foundation for the generation of floral fragrance in *T. amurensis* (Fig. 10).

## Author contributions

The authors confirm contribution to the paper as follows: study conception and design: Zhang L, Zhao X; data collection: Li Y, Han R, Jiang T; analysis and interpretation of results: Li H, Zhang Q, Pei X; draft manuscript preparation: Cai K, Zhao Q.

## Data availability

RNA-Seq raw data from three samples were deposited in the National Center for Biotechnology Information (NCBI) under the accession number PRJNA962507.

## Acknowledgments

This research was supported by First-tier Distinguished Doctor Research Initiation Fund Project at Jilin Agricultural University (201020007). Thanks to Bullet Edits Limited for the linguistic editing and proofreading of the manuscript.

## Conflict of interest

The authors declare that they have no conflict of interest.

**Supplementary Information** accompanies this paper at (<https://www.maxapress.com/article/doi/10.48130/FR-2023-0024>)

## Dates

Received 29 June 2023; Accepted 21 September 2023; Published online 8 October 2023

## References

- Mu L, Liu Y. 2007. Genetic diversity of *Tilia amurensis* populations in different geographical distribution regions. *Chinese Journal of Plant Ecology* 31:1190–98
- Liu Y, Chang Q, Tang Z, Wu K, Abozeid A, et al. 2022. The interrelationship between latitudinal differences and metabolic differences in the natural distribution area of *Tilia amurensis* Rupr. *Forests* 13:1507
- Kwon EB, Kim YS, Han SM, Kim SG, Choi JG. 2022. The protective effect of *Tilia amurensis* honey on influenza A virus infection through stimulation of interferon-mediated IFITM3 signaling. *Biomedicine & Pharmacotherapy* 153:113259
- Chun SW, Lee JW, Ahn JY. 2022. Development and characterization of novel microsatellite markers in *Tilia amurensis* Rupr. using next-generation sequencing. *Molecular Biology Reports* 49:1637–41
- Dudareva N, Klempien A, Muhlemann JK, Kaplan I. 2013. Biosynthesis, function and metabolic engineering of plant volatile organic compounds. *New Phytologist* 198:16–32
- Abbas F, Zhou Y, He J, Ke Y, Qin W, et al. 2021. Metabolite and transcriptome profiling analysis revealed that melatonin positively regulates floral scent production in *Hedychium coronarium*. *Frontiers in Plant Science* 12:808899
- Li X, Cai K, Fan Z, Wang J, Wang L, et al. 2022. Dissection of transcriptome and metabolome insights into the isoquinoline alkaloid biosynthesis during stem development in *Phellodendron amurense* (Rupr.). *Plant Science* 325:111461
- Ali M, Hussain RM, Rehman NU, She G, Li P, et al. 2018. *De novo* transcriptome sequencing and metabolite profiling analyses reveal the complex metabolic genes involved in the terpenoid biosynthesis in Blue Anise Sage (*Salvia guaranitica* L.). *DNA Research* 25:597–617
- Mu L, Zheng J, Wang Y, Liu L. 2009. Chemical compositions and geographic variation in flowers and fruits of *Tilia amurensis*. *Scientia Silvae Sinicae* 45:46–52
- Izawa T. 2021. What is going on with the hormonal control of flowering in plants? *The Plant Journal* 105:431–45
- Janowska B, Andrzejak R. 2022. Cytokinins and gibberellins stimulate the flowering and post-harvest longevity of flowers and leaves of Calla lilies (*Zantedeschia Spreng.*) with colourful inflorescence spathes. *Agronomy* 12:1859
- Xu Y, Prunet N, Gan ES, Wang Y, Stewart D, et al. 2018. SUPERMAN regulates floral whorl boundaries through control of auxin biosynthesis. *The EMBO Journal* 37:e97499
- Han Y, Yang H, Jiao Y. 2014. Regulation of inflorescence architecture by cytokinins. *Frontiers in Plant Science* 5:669
- Dar RA, Nisar S, Tahir I. 2021. Ethylene: a key player in ethylene sensitive flower senescence: a review. *Scientia Horticulturae* 290:110491
- Zhang L, Song C, Guo D, Guo L, Hou X, et al. 2022. Identification of differentially expressed miRNAs and their target genes in response to brassinolide treatment on flowering of tree peony (*Paeonia ostii*). *Plant Signaling & Behavior* 17:2056364
- Sheng J, Li X, Zhang D. 2022. Gibberellins, brassinolide, and ethylene signaling were involved in flower differentiation and development in *Nelumbo nucifera*. *Horticultural Plant Journal* 8:243–50
- Ke Y, Abbas F, Zhou Y, Yu R, Fan Y. 2021. Auxin-responsive R2R3-MYB transcription factors HcMYB1 and HcMYB2 activate volatile



- biosynthesis in *Hedychium coronarium* flowers. *Frontiers in Plant Science* 12:710826
18. Qiao Z, Hu H, Shi S, Yuan X, Yan B, et al. 2021. An update on the function, biosynthesis and regulation of floral volatile terpenoids. *Horticulturae* 7:451
  19. Colquhoun TA, Clark DG. 2011. Unraveling the regulation of floral fragrance biosynthesis. *Plant Signaling & Behavior* 6:378–81
  20. Sapir-Mir M, Mett A, Belausov E, Tal-Meshulam S, Frydman A, et al. 2008. Peroxisomal localization of Arabidopsis isopentenyl diphosphate isomerases suggests that part of the plant isoprenoid mevalonic acid pathway is compartmentalized to peroxisomes. *Plant Physiology* 148:1219–28
  21. Xu Q, He Y, Yan X, Zhao S, Zhu J, et al. 2018. Unraveling a crosstalk regulatory network of temporal aroma accumulation in tea plant (*Camellia sinensis*) leaves by integration of metabolomics and transcriptomics. *Environmental and Experimental Botany* 149:81–94
  22. Wei G, Tian P, Zhang F, Qin H, Miao H, et al. 2016. Integrative analyses of nontargeted volatile profiling and transcriptome data provide molecular insight into VOC diversity in cucumber plants (*Cucumis sativus*). *Plant Physiology* 172:603–18
  23. Zhou C, Zhu C, Tian C, Xu K, Huang L, et al. 2022. Integrated volatile metabolome, multi-flux full-length sequencing, and transcriptome analyses provide insights into the aroma formation of postharvest jasmine (*Jasminum sambac*) during flowering. *Postharvest Biology and Technology* 183:111726
  24. Yan J, Wang H, Wang Y, Xu S, Wan Y, et al. 2021. Integrated metabolome and transcriptome analysis reveals candidate genes involved in metabolism of terpenoids and pthalides in celery seeds. *Industrial Crops and Products* 172:114011
  25. Cheng S, Fu X, Mei X, Zhou Y, Du B, et al. 2016. Regulation of biosynthesis and emission of volatile phenylpropanoids/benzenoids in *petunia* *hybrida* flowers by multi-factors of circadian clock, light, and temperature. *Plant Physiology and Biochemistry* 107:1–8
  26. Mei X, Wan S, Lin C, Zhou C, Hu L, et al. 2021. Integration of metabolome and transcriptome reveals the relationship of benzenoid–phenylpropanoid pigment and aroma in purple tea flowers. *Frontiers in Plant Science* 12:762330
  27. Dong N, Lin H. 2021. Contribution of phenylpropanoid metabolism to plant development and plant–environment interactions. *Journal of Integrative Plant Biology* 63:180–209
  28. Tohge T, de Souza LP, Fernie AR. 2017. Current understanding of the pathways of flavonoid biosynthesis in model and crop plants. *Journal of Experimental Botany* 68:4013–28
  29. Wang S, Alseekh S, Fernie AR, Luo J. 2019. The structure and function of major plant metabolite modifications. *Molecular Plant* 12:899–919
  30. Li W, Lu X, Li J. 2022. The effect of organic nutrient solution on flavor in ripe cherry tomato fruit—Transcriptome and metabolomic analyses. *Environmental and Experimental Botany* 194:104721
  31. Pan Q, Wang Q, Yuan F, Xing S, Zhao J, et al. 2012. Overexpression of ORCA3 and G10H in *Catharanthus roseus* plants regulated alkaloid biosynthesis and metabolism revealed by NMR-metabolomics. *PLoS ONE* 7:e43038
  32. Xu Q, Wang S, Hong H, Zhou Y. 2019. Transcriptomic profiling of the flower scent biosynthesis pathway of *Cymbidium faberi* Rolfe and functional characterization of its jasmonic acid carboxyl methyltransferase gene. *BMC Genomics* 20:125
  33. Ding K, Pei T, Bai Z, Jia Y, Ma P, et al. 2017. SmMYB36, a novel R2R3-MYB transcription factor, enhances tanshinone accumulation and decreases phenolic acid content in *Salvia miltiorrhiza* hairy roots. *Scientific Reports* 7:5104
  34. Bedon F, Bomal C, Caron S, Levasseur C, Boyle B, et al. 2010. Subgroup 4 R2R3-MYBs in conifer trees: gene family expansion and contribution to the isoprenoid- and flavonoid-oriented responses. *Journal of Experimental Botany* 61:3847–64
  35. Colquhoun TA, Kim JY, Wedde AE, Levin LA, Schmitt KC, et al. 2011. *PhMYB4* fine-tunes the floral volatile signature of *Petunia* *hybrida* through *PhC4H*. *Journal of Experimental Botany* 62:1133–43
  36. Verdonk JC, Haring MA, Van Tunen AJ, Schuurink RC. 2005. *ODORANT1* regulates fragrance biosynthesis in petunia flowers. *The Plant Cell* 17:1612–24
  37. Yang J, Zhang J, Wang Z, Zhu Q, Wang W. 2001. Hormonal changes in the grains of rice subjected to water stress during grain filling. *Plant Physiology* 127:315–23
  38. Shi W, Li H, Liu T, Polle A, Peng C, et al. 2015. Exogenous abscisic acid alleviates zinc uptake and accumulation in *Populus* *× canescens* exposed to excess zinc. *Plant, Cell & Environment* 38:207–23
  39. Cai K, Zhang D, Li X, Zhang Q, Jiang L, et al. 2023. Exogenous phytohormone application and transcriptome analysis provides insights for adventitious root formation in *Taxus cuspidata* S. et Z. *Plant Growth Regulation* 100:33–53
  40. Zou S, Wu J, Shahid MQ, He Y, Lin S, et al. 2020. Identification of key taste components in loquat using widely targeted metabolomics. *Food Chemistry* 323:126822
  41. Li X, Li Y, Zhao M, Hu Y, Meng F, et al. 2021. Molecular and metabolic insights into anthocyanin biosynthesis for leaf color change in chokecherry (*Padus virginiana*). *International Journal of Molecular Sciences* 22:10697
  42. Yue Y, Liu J, Shi T, Chen M, Li Y, et al. 2019. Integrating transcriptomic and GC-MS metabolomic analysis to characterize color and aroma formation during tepal development in *Lycoris longituba*. *Plants* 8:53
  43. Ibragic S, Barbini S, Oberlerchner JT, Potthast A, Rosenau T, et al. 2021. Antioxidant properties and qualitative analysis of phenolic constituents in *Ephedra* spp. by HPTLC together with injection port derivatization GC-MS. *Journal of Chromatography B* 1180:122877
  44. Qiu W, Su W, Cai Z, Dong L, Li C, et al. 2020. Combined analysis of transcriptome and metabolome reveals the potential mechanism of coloration and fruit quality in yellow and purple *Passiflora edulis* Sims. *Journal of Agricultural and Food Chemistry* 68:12096–106
  45. Xie C, Mao X, Huang J, Ding Y, Wu J, et al. 2011. KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Research* 39:W316–W322
  46. Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, et al. 2020. TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Molecular Plant* 13:1194–202
  47. Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34:i884–i890
  48. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, et al. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* 29:644–52
  49. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, et al. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* 25:3389–402
  50. Li B, Dewey CN. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12:323
  51. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, et al. 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature Biotechnology* 28:511–15
  52. Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* 15:550
  53. Du Z, Zhou X, Ling Y, Zhang Z, Su Z. 2010. agriGO: a GO analysis toolkit for the agricultural community. *Nucleic Acids Research* 38:W64–W70
  54. Zheng Y, Jiao C, Sun H, Rosli HG, Pombo MA, et al. 2016. iTAK: a program for genome-wide prediction and classification of plant transcription factors, transcriptional regulators, and protein kinases. *Molecular Plant* 9:1667–70
  55. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, et al. 2003. Cytoscape: a software environment for integrated models of

- biomolecular interaction networks. *Genome Research* 13:2498–504
56. Schmittgen TD, Livak KJ. 2008. Analyzing real-time PCR data by the comparative  $C_T$  method. *Nature Protocols* 3:1101–8
57. Yu D, Qanmber G, Lu L, Wang L, Li J, et al. 2018. Genome-wide analysis of cotton GH3 subfamily II reveals functional divergence in fiber development, hormone response and plant architecture. *BMC Plant Biology* 18:350
58. Bögre L, Magyar Z, López-Juez E. 2008. New clues to organ size control in plants. *Genome Biology* 9:226
59. Chen Z, Lu X, Gao J, Xuan Y, Ren J. 2020. Integrating transcriptomic and metabolomic analysis of hormone pathways in *Acer rubrum* during developmental leaf senescence. *BMC Plant Biology* 20:410
60. Hampel D, Mosandl A, Wüst M. 2005. Induction of de novo volatile terpene biosynthesis via cytosolic and plastidial pathways by methyl jasmonate in foliage of *Vitis vinifera* L. *Journal of Agricultural and Food Chemistry* 53:2652–57
61. Perilli S, Di Mambro R, Sabatini S. 2012. Growth and development of the root apical meristem. *Current Opinion in Plant Biology* 15:17–23
62. Frick EM, Strader LC. 2018. Roles for IBA-derived auxin in plant development. *Journal of Experimental Botany* 69:169–77
63. Hu S, Zhang M, Yang Y, Xuan W, Zou Z, et al. 2020. A novel insight into nitrogen and auxin signaling in lateral root formation in tea plant [*Camellia sinensis* (L.) O. Kuntze]. *BMC Plant Biology* 20:232
64. Cucinotta M, Cavalleri A, Chandler JW, Colombo L. 2021. Auxin and flower development: a blossoming field. *Cold Spring Harbor Perspectives in Biology* 13:a039974
65. Zhang Q, Gong M, Xu X, Li H, Deng W. 2022. Roles of auxin in the growth, development, and stress tolerance of horticultural plants. *Cells* 11:2761
66. Srikanth A, Schmid M. 2011. Regulation of flowering time: all roads lead to Rome. *Cellular and Molecular Life Sciences* 68:2013–37
67. Brioudes F, Joly C, Szécsi J, Varaud E, Leroux J, et al. 2009. Jasmonate controls late development stages of petal growth in *Arabidopsis thaliana*. *The Plant Journal* 60:1070–80
68. Leng P, Zhang Y, Du Y, Wang J, Jiang L, et al. 2018. Expression pattern of ABA metabolic and signalling genes during floral development and fruit set in sweet cherry. *Plant Growth Regulation* 84:71–80
69. Wu G, Zhu Z, Qiu Q, Fan X, Yuan D. 2022. Transcriptome Analysis Reveals the Regulatory Networks of Cytokinin in Promoting Floral Feminization in *Castanea henryi*. *International Journal of Molecular Sciences* 23:6389
70. Huang G, Han M, Yao W, Wang Y. 2017. Transcriptome analysis reveals the regulation of brassinosteroids on petal growth in *Gerbera hybrida*. *PeerJ* 5:e3382
71. Cheng Q, Bai S, Ge G, Li P, Liu L, et al. 2018. Study on differentially expressed genes related to defoliation traits in two alfalfa varieties based on RNA-Seq. *BMC Genomics* 19:807
72. Kou X, Zhao X, Wu B, Wang C, Wu C, et al. 2022. Auxin response factors are ubiquitous in plant growth and development, and involved in crosstalk between plant hormones: a review. *Applied Sciences* 12:1360
73. Si C, Zeng D, da Silva JAT, Qiu S, Duan J, et al. 2023. Genome-wide identification of Aux/IAA and ARF gene families reveal their potential roles in flower opening of *Dendrobium officinale*. *BMC Genomics* 24:199
74. Liu K, Feng S, Pan Y, Zhong J, Chen Y, et al. 2016. Transcriptome analysis and identification of genes associated with floral transition and flower development in sugar apple (*Annona squamosa* L.). *Frontiers in Plant Science* 7:1695
75. Li J, Min X, Luo K, Abdoulaye AH, Zhang X, et al. 2023. Molecular characterization of the GH3 family in alfalfa under abiotic stress. *Gene* 851:146982
76. Ren H, Gray WM. 2015. SAUR proteins as effectors of hormonal and environmental signals in plant growth. *Molecular Plant* 8:1153–64
77. Zhang Y, Ye T, She Z, Huang S, Wang L, et al. 2023. Small Auxin Up RNA (SAUR) gene family identification and functional genes exploration during the floral organ and fruit developmental stages in pineapple (*Ananas comosus* L.) and its response to salinity and drought stresses. *International Journal of Biological Macromolecules* 237:124061
78. Tian Z, Han J, Che G, Hasi A. 2022. Genome-wide characterization and expression analysis of SAUR gene family in Melon (*Cucumis melo* L.). *Planta* 255:123
79. Srivastava M, Srivastava AK, Orosa-Puente B, Campanaro A, Zhang C, et al. 2021. SUMO conjugation to BZR1 enables brassinosteroid signaling to integrate environmental cues to shape plant growth. *Current Biology* 31:668–69
80. Poppenberger B, Fujioka S, Soeno K, George GL, Vaistij FE, et al. 2005. The UGT73C5 of *Arabidopsis thaliana* glucosylates brassinosteroids. *Proceedings of the National Academy of Sciences of the United States of America* 102:15253–58
81. Di T, Zhao L, Chen H, Qian W, Wang P, et al. 2019. Transcriptomic and metabolic insights into the distinctive effects of exogenous melatonin and gibberellin on terpenoid synthesis and plant hormone signal transduction pathway in *Camellia sinensis*. *Journal of Agricultural and Food Chemistry* 67:4689–99
82. Liu W, Zhang Y, Wang L, Ahmad B, Shi X, et al. 2023. Integrated transcriptome and metabolome analysis unveiled the mechanisms of xenia effect and the role of different pollens on aroma formation in 'Yali' pear (*Pyrus bretschneideri* Rehd). *Scientia Horticulturae* 307:111503
83. Urrutia M, Rambla JL, Alexiou KG, Granel A, Monfort A. 2017. Genetic analysis of the wild strawberry (*Fragaria vesca*) volatile composition. *Plant Physiology and Biochemistry* 121:99–117
84. Fei X, Qi Y, Lei Y, Wang S, Hu H, et al. 2021. Transcriptome and metabolome dynamics explain aroma differences between green and red prickly ash fruit. *Foods* 10:391
85. Li H, Li J, Dong Y, Hao H, Ling Z, et al. 2019. Time-series transcriptome provides insights into the gene regulation network involved in the volatile terpenoid metabolism during the flower development of lavender. *BMC Plant Biology* 19:313
86. Zhu L, Liao J, Liu Y, Zhou C, Wang X, et al. 2022. Integrative metabolome and transcriptome analyses reveal the molecular mechanism underlying variation in floral scent during flower development of *Chrysanthemum indicum* var. *aromaticum*. *Frontiers in Plant Science* 13:919151
87. Yue X, Ren R, Ma X, Fang Y, Zhang Z, et al. 2020. Dynamic changes in monoterpene accumulation and biosynthesis during grape ripening in three *Vitis vinifera* L. cultivars. *Food Research International* 137:109736
88. Tang D, Shen Y, Li F, Yue R, Duan J, et al. 2022. Integrating metabolite and transcriptome analysis revealed the different mechanisms of characteristic compound biosynthesis and transcriptional regulation in tea flowers. *Frontiers in Plant Science* 13:1016692
89. Yang X, Yue Y, Li H, Ding W, Chen G, et al. 2018. The chromosome-level quality genome provides insights into the evolution of the biosynthesis genes for aroma compounds of *Osmanthus fragrans*. *Horticulture Research* 5:72
90. Wang W, Feng J, Wei L, Khalil-Ur-Rehman M, Nieuwenhuizen NJ, et al. 2021. Transcriptomics integrated with free and bound terpenoid aroma profiling during "shine muscat" (*Vitis labrusca* × *V. vinifera*) grape berry development reveals coordinate regulation of MEP pathway and terpene synthase gene expression. *Journal of Agricultural and Food Chemistry* 69:1413–29
91. Zhang C, Liu H, Hu S, Zong Y, Xia H, et al. 2022. Transcriptomic profiling of the floral fragrance biosynthesis pathway of *Liriodendron* and functional characterization of the *LtuDXR* gene. *Plant Science* 314:111124
92. Zhu C, Zhang S, Fu H, Zhou C, Chen L, 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. *Frontiers in Plant Science* 10:1638

## Aroma formation of flowers in nectar tree

93. Hu Z, Tang B, Wu Q, Zheng J, Leng P, et al. 2017. Transcriptome sequencing analysis reveals a difference in monoterpene biosynthesis between scented *Lilium* 'Siberia' and unscented *Lilium* 'Novano'. *Frontiers in Plant Science* 8:1351
94. Yan W, Yang Y, Wu Y, Yu J, Zhang J, et al. 2021. Isopentenyl diphosphate isomerase (*IP1*) gene silencing negatively affects patchouli alcohol biosynthesis in *Pogostemon cablin*. *Plant Molecular Biology Reporter* 39:557–65
95. Yang J, Adhikari MN, Liu H, Xu H, He G, et al. 2012. Characterization and functional analysis of the genes encoding 1-deoxy-D-xylulose-5-phosphate reductoisomerase and 1-deoxy-D-xylulose-5-phosphate synthase, the two enzymes in the MEP pathway, from *Amomum villosum* Lour. *Molecular Biology Reports* 39:8287–96
96. Page JE, Hause G, Raschke M, Gao W, Schmidt J, et al. 2004. Functional analysis of the final steps of the 1-deoxy-D-xylulose 5-phosphate (DXP) pathway to isoprenoids in plants using virus-induced gene silencing. *Plant Physiology* 134:1401–13
97. Bouvier F, Suire C, d'Harlingue A, Backhaus RA, Camara B. 2000. Molecular cloning of geranyl diphosphate synthase and compartmentation of monoterpene synthesis in plant cells. *The Plant Journal* 24:241–52
98. Hsiao YY, Jeng MF, Tsai WC, Chuang YC, Li CY, et al. 2008. A novel homodimeric geranyl diphosphate synthase from the orchid *Phalaenopsis bellina* lacking a DD (X)<sub>2-4</sub>D motif. *The Plant Journal* 55:719–33
99. Tholl D, Chen F, Petri J, Gershenzon J, Pichersky E. 2005. Two sesquiterpene synthases are responsible for the complex mixture of sesquiterpenes emitted from *Arabidopsis* flowers. *The Plant Journal* 42:757–71
100. Wang X, Zeng Y, Nieuwenhuizen NJ, Atkinson RG. 2021. TPS-b family genes involved in signature aroma terpenes emission in ripe kiwifruit. *Plant Signaling & Behavior* 16:1962657
101. Abbas F, Guo S, Zhou Y, Wu J, Amanullah S, et al. 2022. Metabolome and transcriptome analysis of terpene synthase genes and their putative role in floral aroma production in *Litchi chinensis*. *Physiologia Plantarum* 174:e13796
102. Mostafa S, Wang Y, Zeng W, Jin B. 2022. Floral scents and fruit aromas: Functions, compositions, biosynthesis, and regulation. *Frontiers in Plant Science* 13:860157
103. Yang G, Qin Y, Jia Y, Xie X, Li D, et al. 2023. Transcriptomic and metabolomic data reveal key genes that are involved in the phenylpropanoid pathway and regulate the floral fragrance of *Rhododendron fortunei*. *BMC Plant Biology* 23:8
104. Shang J, Tian J, Cheng H, Yan Q, Li L, et al. 2020. The chromosome-level wintersweet (*Chimonanthus praecox*) genome provides insights into floral scent biosynthesis and flowering in winter. *Genome Biology* 21:200
105. Fu J, Huang S, Qian J, Qing H, Wan Z, et al. 2022. Genome-wide identification of petunia *HSF* genes and potential function of *PhHSF19* in benzenoid/phenylpropanoid biosynthesis. *International Journal of Molecular Sciences* 23:2974
106. Wang X, Song Z, Ti Y, Ma K, Li Q. 2022. Comparative transcriptome analysis linked to key volatiles reveals molecular mechanisms of aroma compound biosynthesis in *Prunus mume*. *BMC Plant Biology* 22:395
107. Verdonk JC, De Vos CHR, Verhoeven HA, Haring MA, Van Tunen AJ, et al. 2003. Regulation of floral scent production in petunia revealed by targeted metabolomics. *Phytochemistry* 62:997–1008
108. Deng Y, Sun X, Gu C, Jia X, Liang L, et al. 2017. Identification of pre-fertilization reproductive barriers and the underlying cytological mechanism in crosses among three petal-types of *Jasminum sambac* and their relevance to phylogenetic relationships. *PLoS ONE* 12:e0176026
109. Rastogi S, Kumar R, Chanotiya CS, Shanker K, Gupta MM, et al. 2013. 4-coumarate: CoA ligase partitions metabolites for eugenol biosynthesis. *Plant and Cell Physiology* 54:1238–52
110. Wang C, Yu J, Cai Y, Zhu P, Liu C, et al. 2016. Characterization and functional analysis of 4-coumarate: CoA ligase genes in mulberry. *PLoS ONE* 11:e0155814
111. Kutty NN, Ghissing U, Mitra A. 2021. Revealing floral metabolite network in tuberose that underpins scent volatiles synthesis, storage and emission. *Plant Molecular Biology* 106:533–54
112. Tong Z, Li H, Zhang R, Ma L, Dong J, et al. 2015. Co-downregulation of the hydroxycinnamoyl-CoA: shikimate hydroxycinnamoyl transferase and coumarate 3-hydroxylase significantly increases cellulose content in transgenic alfalfa (*Medicago sativa* L.). *Plant Science* 239:230–37
113. Zhang T, Bao F, Ding A, Yang Y, Cheng T, et al. 2022. Comprehensive analysis of endogenous volatile compounds, transcriptome, and enzyme activity reveals *PmCAD1* involved in cinnamyl alcohol synthesis in *Prunus mume*. *Frontiers in Plant Science* 13:820742
114. Liu Z, Mohsin A, Wang Z, Zhu X, Zhuang Y, et al. 2021. Enhanced biosynthesis of chlorogenic acid and its derivatives in methyl-jasmonate-treated *Gardenia jasminoides* cells: a study on metabolic and transcriptional responses of cells. *Frontiers in Bioengineering and Biotechnology* 8:604957
115. Wang Z, Du H, Zhai R, Song L, Ma F, Xu L. 2017. Transcriptome analysis reveals candidate genes related to color fading of 'Red Bartlett' (*Pyrus communis* L.). *Frontiers in Plant Science* 8:455
116. Chuang YC, Hung YC, Tsai WC, Chen WH, Chen HH. 2018. PbbHLH4 regulates floral monoterpene biosynthesis in *Phalaenopsis* orchids. *Journal of Experimental Botany* 69:4363–77
117. Ding W, Ouyang Q, Li Y, Shi T, Li L, et al. 2020. Genome-wide investigation of WRKY transcription factors in sweet osmanthus and their potential regulation of aroma synthesis. *Tree Physiology* 40:557–72
118. Yu Y, Liu Z, Wang L, Kim SG, Seo PJ, et al. 2016. WRKY71 accelerates flowering via the direct activation of *FLOWERING LOCUS T* and *LEAFY* in *Arabidopsis thaliana*. *The Plant Journal* 85:96–106
119. Lei R, Li X, Ma Z, Lv Y, Hu Y, et al. 2017. Arabidopsis WRKY2 and WRKY34 transcription factors interact with VQ20 protein to modulate pollen development and function. *The Plant Journal* 91:962–76
120. Zhang Y, Cao G, Qu L, Gu H. 2009. Characterization of *Arabidopsis* MYB transcription factor gene *AtMYB17* and its possible regulation by *LEAFY* and *AGL15*. *Journal of Genetics and Genomics* 36:99–107
121. Xu Y, Zhu C, Xu C, Sun J, Grierson D, et al. 2019. Integration of metabolite profiling and transcriptome analysis reveals genes related to volatile terpenoid metabolism in finger citron (*C. medica* var. *sarcodactylis*). *Molecules* 24:2564
122. Hong G, Xue X, Mao Y, Wang L, Chen X. 2012. *Arabidopsis* MYC2 interacts with DELLA proteins in regulating sesquiterpene synthase gene expression. *The Plant Cell* 24:2635–48
123. Muhlemann JK, Klempien A, Dudareva N. 2014. Floral volatiles: from biosynthesis to function. *Plant, Cell & Environment* 37:1936–49
124. Ben Zvi MM, Negre-Zakharov F, Masci T, Ovadis M, Shklarman E, et al. 2008. Interlinking showy traits: co-engineering of scent and colour biosynthesis in flowers. *Plant Biotechnology Journal* 6:403–15



Copyright: © 2023 by the author(s). Published by Maximum Academic Press, Fayetteville, GA. This article is an open access article distributed under Creative Commons Attribution License (CC BY 4.0), visit <https://creativecommons.org/licenses/by/4.0/>.