

Article ID: 207185
DOI: 10.5586/asbp/207185

Publication History

Received: 2025-01-22
Accepted: 2025-06-15
Published: 2025-11-24

Handling Editor

Katarzyna Buczkowska; Adam Mickiewicz University, Poznań, Poland; <https://orcid.org/0000-0003-1491-8157>

Authors' Contributions

LW, WD: Research concept and design; LW: Collection and/or assembly of data; LW: Data analysis and interpretation; LW: Writing the article; LW, DW: Critical revision of the article; LW, DW: Final approval of the article

Funding

The research was supported by undergraduate education quality construction comprehensive reform project of Wuhan University (2025). This article is the research result of the 2025 Hubei Provincial College Students' Innovation and Entrepreneurship Training Program Project "Occurrence of Gall Formation on *Distylium racemosum* and Its Impact on Host Plants" (Project No. 202510486048).

Competing Interests

No competing interests have been declared.

Copyright Notice

© The Author(s) 2025. This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits redistribution, commercial and noncommercial, provided that the article is properly cited.

RESEARCH PAPER

Gall-Induced Trade-Offs in *Distylium racemosum*: Transcriptional Suppression of Photosynthesis and Activation of Terpenoid Metabolism Mechanisms

Leyang Wang, Wei Du 

College of Life Sciences, Wuhan University, China

* To whom correspondence should be addressed. Email: wdu@whu.edu.cn

Abstract

Galls are deformed manifestations of plant tissues stimulated by insects and other organisms that feed on or lay eggs on plants. *Distylium racemosum*, with the Chinese name "Wen-Mu-Shu", has established a relationship with *Neothoracaphis hangzhouensis* etc. over a long evolutionary period. However, the effects of galls on the transcriptional levels in *D. racemosum* leaves remain unknown. Using whole-transcriptome sequencing and de novo assembly, we generated a reference transcriptome (104, 537 unigenes). Notably, in galled leaves, transcripts encoding key terpenoid biosynthetic enzymes, including linalool 8-monooxygenase, were significantly up-regulated, indicating the preferential redirection of metabolic resources toward terpenoid metabolism compounds. Conversely, the reduced expression of transcripts associated with photosynthesis-related structural proteins or subunits (e.g., Lhca/Lhcb) was observed, suggesting compromised photosynthetic capacity. Overall, this study revealed the adaptive strategy of *D. racemosum* in coping with gall-induced stress through 'photosynthesis-terpenoid metabolism' resource reallocation. These findings provide a transcriptomic framework for gall formation in *D. racemosum*, offering insights into plant adaptive strategies under biotic stress.

Keywords

Distylium racemosum; Gall; Photosynthesis; Transcriptome; Terpenoid biosynthesis

1. Introduction

Gall formation is a manifestation of the reprogramming of plant cell growth and development stimulated by insects and other organisms. Approximately 132, 930 insect species worldwide can induce gall formation (Zhang et al., 2018). One of the oldest known fossil galls was produced on a tree fern in the central region of North America in the late Carboniferous era, approximately 302 Ma (Harris & Pitzschke, 2020). Insects produce the most diverse and complex galls in nature (Tooker & Helms, 2014), such as leaf, flower, fruit, stem, and root galls (Guedes et al., 2016; Mani, 1964; Miller, 1975). Many plants have gall leaves, such as *Robinia*

spp. (Molnár et al., 2009), *Prunus serotina* and *Rhus glabra* (Larson, 1998). *Obolodiplosis robiniae*, a pest gall midge, leads to delayed growth and development of *Robinia* spp. (Molnár et al., 2009). Two gallers, *Phytoptus cerasicrumena* and *Melaphis rhois*, affect the photosynthetic rate of leaves in *P. serotina* and *R. glabra* (Larson, 1998). Further, the expression levels of certain secondary metabolites (e.g., shikimate and phenylpropanoids) undergo significant changes during gall formation (Nabity et al., 2013).

The Hamamelidaceae family contains 28 genera (Xiang et al., 2019), nine of which have been reported to produce galls. Among these, seven genera (*Fothergilla*, *Parrotia*, *Distylium*, *Distyliopsis*, *Hamamelis*, *Sycopsis*, *Trichocladus*)

have been explicitly reported to form galls on leaves (Dedrich et al., 2022; Adroit et al., 2018; Zhou et al., 2025; Chen et al., 2011; Yang et al., 2014; Endress, 1993; Yeh & Ko, 2017; Conner et al., 2013; Lewis & Walton, 1958; Mizuki et al., 2024; Walker, 1944; Chetverikov et al., 2021; Gong et al., 1988). *Distylium racemosum*, the most widely distributed and hardiest species within its genus, was first described in 1835 and is distributed in western Japan, the Ryukyu Islands, Korea, and southern China naturally. In natural habitat, *D. racemosum* exhibits polymorphism, while mostly appearing as either a sprawling, upright tree (International Dendrology Society, 2024). To date, four species of *Nipponaphis* (Aoki et al., 2015; Aoki et al., 2019), *Monzenia minuta* (Yeh & Ko, 2017), *Tripartita formosana* (Yeh & Ko, 2017), and *Neothoracaphis hangzhouensis* (Shen et al., 1993) have been found to cause galls on the leaves of *D. racemosum*. In China, *N. hangzhouensis* migrates to *D. racemosum* in early November, lays eggs at the base of leaf bud scales in mid-January, and nymphs hatch in late March. By mid-May, the aphids emerge from galls and migrate to secondary host trees, including *Quercus alba*, *Q. robur*, and *Liriodendron tulipifera*, to spend the summer (Shen et al., 1993). Annually in spring, newly emerged leaves of these *D. racemosum* trees become covered with galls.

Differences in relative conductivity, malondialdehyde (MDA) content, protein content, and superoxide dismutase (SOD) enzyme activity for *D. racemosum* have been compared between galled and normal leaves (Ling, 2012). Although the alteration of plant gene expression under insect biotic stress is a widespread phenomenon (Giron et al., 2016), the effects of galler parasitism on the transcriptional levels of galled leaves of *D. racemosum* remain unknown. Omics technologies and systems biology have been widely applied in research related to insect gallers and their host plants (Oates et al., 2016). Previous studies on *D. racemosum* can no longer meet the data resource demands in this field, necessitating the implementation of transcriptomic studies. Although transcription factor families of model plants, such as *Arabidopsis*, have been thoroughly annotated (Gao & Dubos, 2024), and gene expression profiles during gall formation have been studied (Hirano et al., 2023), data for the vast majority of non-model plants remain scarce.

Transcriptomic technology has thus become indispensable for deciphering plant responses to biotic stress. Notable advancements include mechanistic insights into gall formation processes, as demonstrated by two researches (Takeda et al., 2024; Bailey et al., 2015). Comprehensive transcriptome profiling of *Artemisia indica* during *Rhopalomyia*-induced gall development revealed that transformation from source to sink organs occurs during gall development (Takeda et al., 2024). Systematic characterization of psyllid-triggered transcriptional dynamics in *Metrosideros polymorpha* revealed that the auxin response is involved in galling (Bailey et al., 2015). In addition, extracting biologically meaningful insights from extensive datasets has become a focal point in recent research, particularly for the annotation and functional prediction of unknown genes (Chen et al., 2024). For example, transcriptome annotation has been conducted to develop gene resources for the snail *Karatohelix adamsi* (Park et al., 2024).

In the present study, we used DNA nanoball (DNB) sequencing and Trinity assembly to draft a reference transcriptome for *D. racemosum*. We also conducted a differentially expressed gene (DEG) analysis between galled and normal leaves, providing data for further understanding the phenomenon of gall formation.

2. Materials and methods

2.1. Plant materials

Distylium racemosum, a small evergreen tree belonging to the Hamamelidaceae family, experiences annual infestations in spring with gall formation on newly developed leaves. *D. racemosum* living at Wuhan University in central China flowers from March to May, followed by fruiting from June to August. Based on our observations, the dominant aphid species infesting these trees is *Neothoracaphis hangzhouensis*. We collected both galled (three biological replicates: A1, A2, and A3) and healthy leaves (three biological replicates: B1, B2, and B3) at Wuhan University. After cleaning with ddH₂O, the leaves were immediately frozen in liquid nitrogen to prevent RNA degradation by endogenous and environmental RNases. RNA was then extracted from each sample and sequenced.

2.2. RNA extraction

For each individual sample, the appropriate amount of tissue was ground into powder under liquid nitrogen and transferred to 1.5 mL of preheated CTAB lysis reagent (supplemented with 2% β -mercaptoethanol) at 65°C and incubated at 65°C for 15 min. After incubation, the tube was cooled to 25°C and then centrifuged at 4°C with 12,000 $\times$ g for 5 min.

The supernatant was transferred to a 2.0 mL EP tube, and 200 μ L of chloroform/isoamyl alcohol (24:1) per milliliter of CTAB lysis buffer was added. The mixture was vortexed to mix and centrifuged at 4°C and 12,000 $\times$ g for 10 min.

Next, the clear upper aqueous layer was transferred to a 2.0 mL EP tube. An equal volume of phenol: chloroform: isoamyl alcohol (25:24:1) was added. The mixture was vortexed for mixing and centrifuged at 4°C and 12,000 $\times$ g for 10 min. The clear upper aqueous layer was then transferred to a 2.0 mL EP tube. An equal volume of chloroform: isoamyl alcohol (24:1) was added to the mixture. The mixture was vortexed for mixing and centrifuged at 4°C and 12,000 $\times$ g for 10 min. The final clear upper aqueous layer was transferred to a new 1.5 mL EP tube, and 2/3 of the aqueous layer volume of isopropanol was added. The mixture was gently inverted to mix well and placed at -20°C for 2 h. The RNA precipitate was collected by centrifugation at 4°C and 12,000 $\times$ g for 20 min. The supernatant was removed, and the precipitate was resuspended and washed with 1 mL of 75% ethanol. After washing, the precipitation was centrifuged again at 4°C and 17,500 $\times$ g for 3 min. The supernatant was discarded and the precipitate was dried in a biosafety cabinet for 4 min. Finally, 35 μ L of RNase-free water was added to dissolve the RNA.

2.3. mRNA library preparation

Notably, this study used a unique second-generation sequencing method, namely DNB sequencing technology (Drmanac et al., 2010) by BGI Genomics Co., Ltd. Compared with the commonly used Illumina bridge amplification, DNB sequencing employs rolling circle amplification, which prevents the exponential accumulation of error-containing strands during subsequent amplification (Li et al., 2019).

Library preparation was performed using the Optimal Dual-mode mRNA Library Prep Kit (BGI, Shenzhen, China). Briefly, 200 ng of total mRNA was denatured at a suitable temperature to open the secondary structure (set hot lid at 65°C and react in the following order to complete the combination of mRNA and mRNA Capture Beads: 65°C, 5 min; 25°C, 5 min; and 25°C, hold), and mRNA was enriched using oligo (dT)-attached magnetic beads. After reacting at 85°C for 8 min, RNA was fragmented using fragmentation reagents.

First-strand cDNA was generated by random hexamer-primed reverse transcription, followed by second-strand cDNA synthesis. The synthesized double-stranded cDNA was then subjected to an end-repair reaction. After cDNA end repair, a single 'A' nucleotide was added to the 3' end of the blunt fragments through an A tailing reaction. The reaction system for adaptor ligation was then configured to ligate adaptors with the cDNAs, and the library products were finally amplified using PCR and subjected to quality control.

Single-stranded library products were then generated by denaturation. The circularization reaction system was set up to obtain single-stranded cyclized DNA products. Uncyclized single-stranded linear DNA molecules were digested. The final single-strand circularized library was amplified using phi29 and rolling circle amplification to generate DNB, which carried more than 300 copies of the initial single-stranded circularized library molecule. DNBs were then loaded into a patterned nanoarray, and PE150 base reads were generated on the T7 platform (BGI-Shenzhen, China).

2.4. Data filtering

The raw data were filtered with SOAPnuke (Chen et al., 2018) (v1.6.5) by (I) removing reads containing adapters (adapter contamination), (II) removing reads whose unknown base ('N' base) ratio was more than 1%, and (III) removing reads whose low-quality base ratio (base quality less than or equal to 15) was more than 40%, after which clean reads were obtained and stored in the FASTQ format.

2.5. De novo assembly and quality assessment

Trinity (v2.13.2) (Grabherr et al., 2011) was used to assemble clean reads and BUSCO was used to assess assembly quality. Sequencing produced 39.50 Gb of high-quality data with Q20 (>98%) and Q30 (>93%) pass rates and >96% Clean Reads Ratio across all samples (Table S1). The assembly generated 104,537 unigenes, predominantly short sequences (10200 nt, > 20,000) decreasing with length but including a notable subset > 3000 nt (Figure S1; Table S2).

CD-HIT clustering was used to remove redundancies, yielding an All-Unigene dataset for downstream analysis.

2.6. Reference gene mapping

Clean data were mapped to the unique genes assembled using Bowtie2 (v2.4.5) (Langmead & Salzberg, 2012) and the expression levels of genes were calculated using RSEM (v1.3.1) (Li & Dewey, 2011).

2.7. Gene annotation

Unigenes were annotated using major functional databases (KEGG, GO, NR), and transcription factors were predicted. For plant transcription factors, getorf was used to determine the ORFs of the unigenes, followed by hmmsearch to align the ORFs with transcription factor protein domains. We then identified the functions of the unigenes based on the characteristics of the transcription factor families described in PlantTFDB (<https://planttfdb.gao-lab.org>).

2.8. Differential gene expression

Between-group differential gene analyses were performed using DESeq2 (Love et al., 2014) and DEGSeq (Wang et al., 2010). Poisson Dis was used for between-sample differential gene analysis. According to the GO and KEGG annotation results and official classifications, DEGs were functionally classified; phyper in R software was used for KEGG enrichment analysis, and the TermFinder package was used for GO Enrichment analysis. Criteria for selecting DEGs and parameters for enrichment analysis are listed in Table S3.

2.9. Simple sequence repeat (SSR) forecast

Unigenes were detected using MISA (v1.0) (Thiel et al., 2003), and Primer3 (v2.6.1) (Untergasser et al., 2012) was used to design primers for SSR detection.

2.10. RT-qPCR

RT-qPCR analysis was conducted for six highly expressed target genes, a selection strategy that minimizes stochastic amplification failures and enhances signal detection robustness in low-variability systems. To ensure rigorous experimental validation, we implemented a hierarchical replication approach: 1) Three independent biological replicates (derived from distinct experimental cohorts) were analyzed per target gene to account for inter-subject variability; 2) Each biological replicate underwent triplicate technical measurements (n = 3) using independently assembled reaction mixtures, thereby controlling for intra-assay variability. The GAPDH gene is a housekeeping gene that is highly expressed in almost all tissues, and its expression level is generally constant in the same type of cells or tissues. For a long time, the GAPDH gene has been widely used as an internal reference gene in qPCR (Zhang et al., 2011). The GAPDH gene was also selected as the internal reference gene in this study. Relative gene expression levels were calculated according to the $\Delta\Delta C_t$ method. Subsequently,

the correlation rate between the RT-qPCR results and transcriptome differential expression analysis was determined. We have detailed the technical details in [Appendix S1](#).

3. Results

3.1. Transcriptome gene annotation

We annotated the assembled unigenes, resulting in a Venn diagram of the gene annotations ([Figure 1](#)). Notably, 3,678 genes were common to GO, KEGG, and NR, but absent from the TF annotations. A substantial number of genes, 5,325, were shared between NR and KEGG alone, whereas fewer genes overlapped with the other combinations. There were 3,990 common unigenes in the four types of gene annotations. We annotated the unigenes aligned to the KEGG database and analyzed the gene classification ([Figure 2](#)). The KEGG functional distribution statistics revealed that the most enriched KEGG pathway was Carbohydrate metabolism, with 1,059 annotated genes. In addition, 229 unigenes were involved in the “Metabolism of terpenoids and polyketides” pathway. The subject IDs corresponding to genes in the “Metabolism of terpenoids and polyketides” pathway within the KEGG data base are shown in [Table S4](#), and the protein interaction network diagram for the metabolism of terpenoids and polyketides is shown in [Figure S2](#) and [Table S5](#). Our GO functional distribution statistics indicated that the most represented GO pathway was organelle, with 8,035 annotated genes ([Figure 3](#)). The subject IDs corresponding to the genes in the “organelle” pathway within the KEGG database are shown in [Table S6](#).

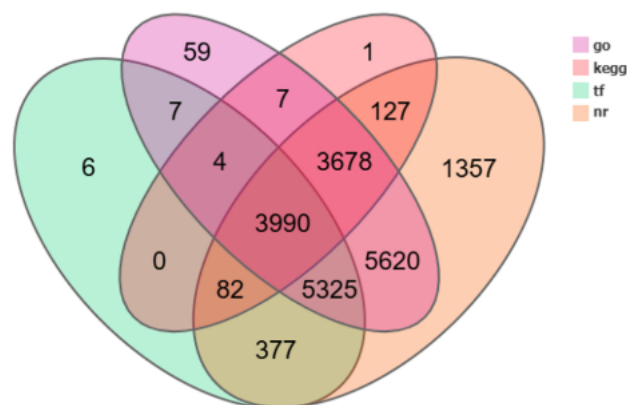


Figure 1 Venn diagram of gene annotations. The Venn diagram of gene annotations illustrates the overlap and uniqueness of gene annotations across four annotation databases: GO (Gene Ontology), KEGG (Kyoto Encyclopedia of Genes and Genomes), TF (Transcription Factor), and NR (Non-Redundant protein database). Pairwise and higher-order overlaps are also represented.

Functional annotation analysis revealed the key genes contributing to terpenoid metabolism and related biological processes. Unigene 10079-S3 was functionally annotated to the pathway involving key enzymes in terpenoid metabolism, including farnesyl-diphosphate farnesyltransferase (GO:0004310), squalene synthase (GO:0051996), and carotenoid biosynthesis (GO:0016117). Furthermore, Unigene-10091-S4 was functionally annotated to pathways involving

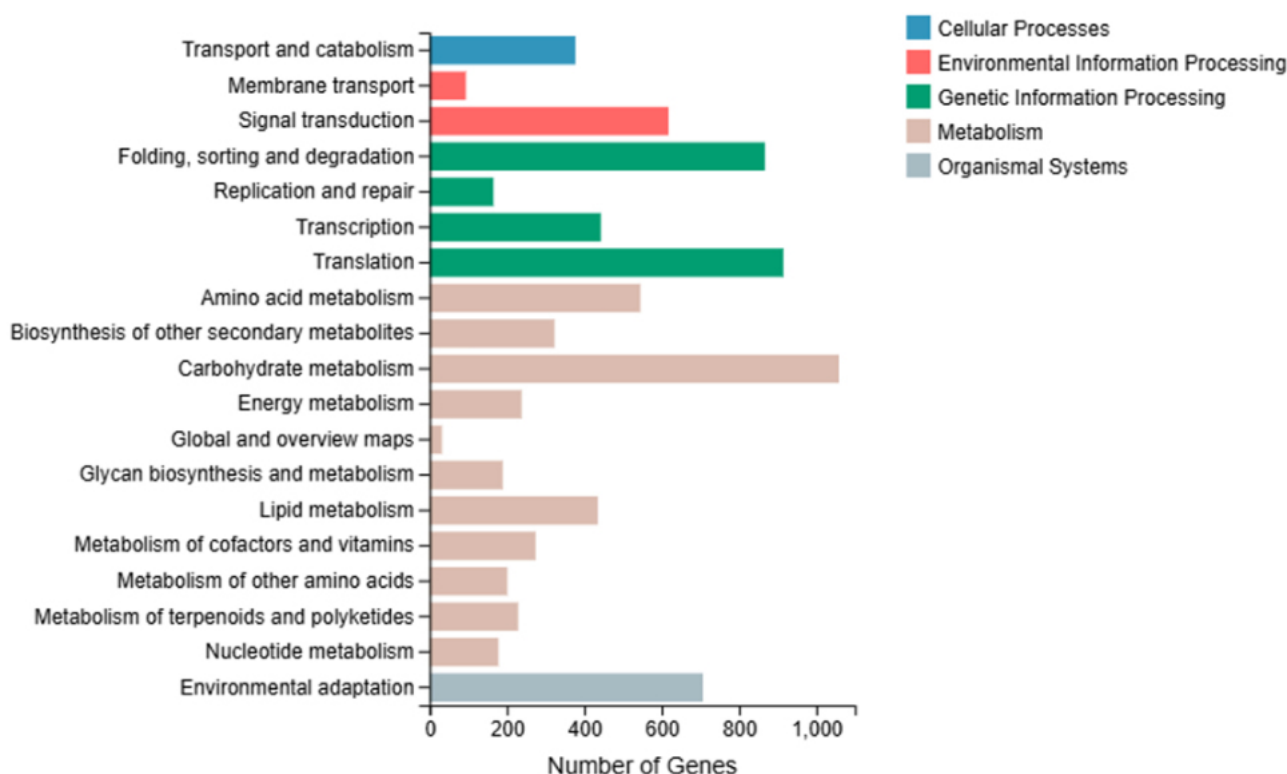


Figure 2 Functional distribution map and species distribution map. KEGG functional distribution statistics reveal that the most enriched KEGG pathway is Carbohydrate metabolism.

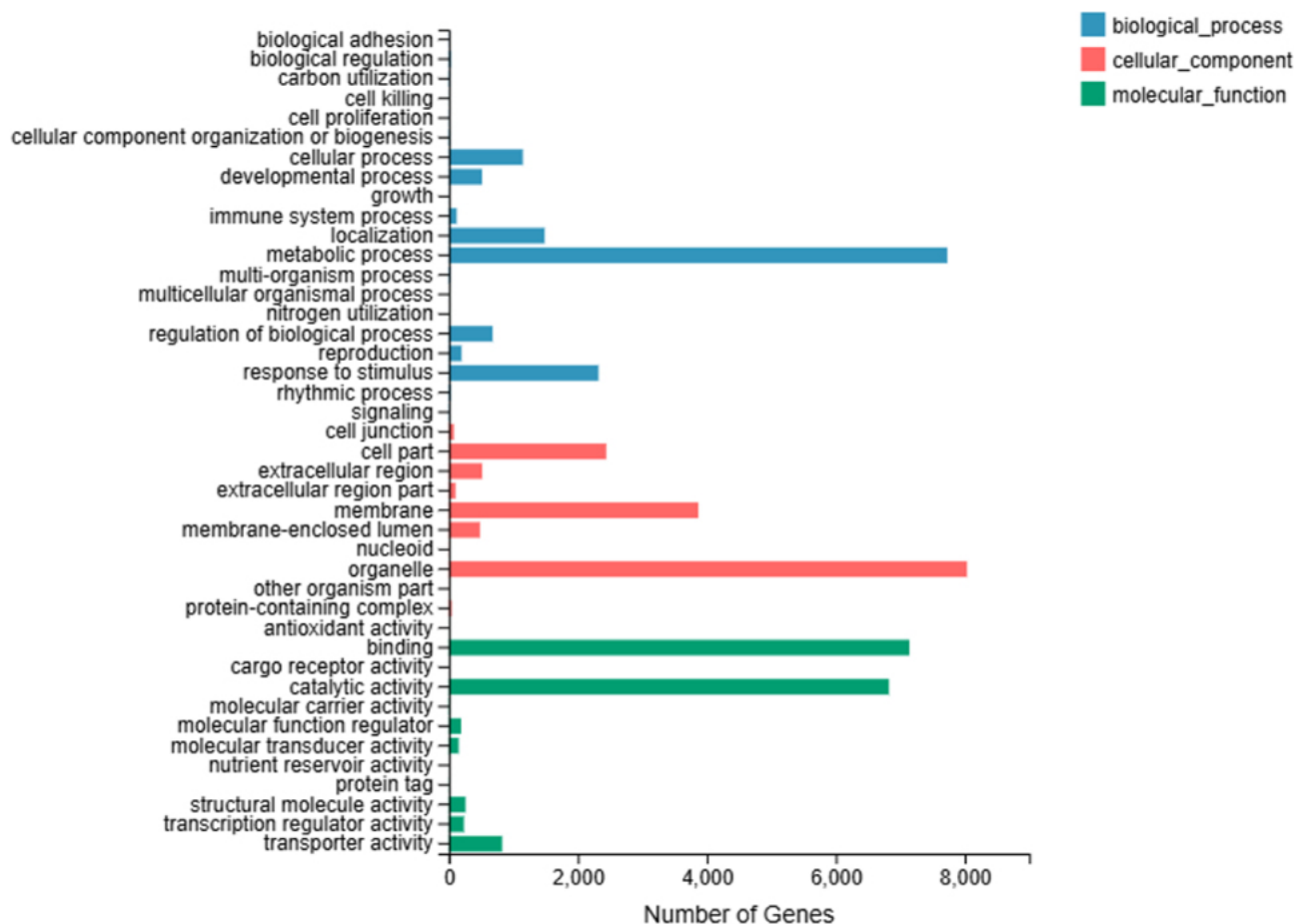


Figure 3 Functional distribution map and species distribution map. GO functional distribution statistics indicate that the most represented GO pathway is organelle.

key enzymes in terpenoid and sterol metabolism, including sterol 3-beta-glucosyltransferase activity (GO:0016906) and sterol biosynthetic processes (GO:0016126), highlighting its dual role in membrane lipid modification and secondary metabolite synthesis.

Based on NR annotation results, we calculated the proportion of annotations from different species and generated a species distribution chart. Our NR annotation species distribution indicated that *Nyssa sinensis* and *Vitis vinifera* had the highest annotation proportions (Figure 4). The transcription factor families to which the genes belonged were statistically classified as shown in Figure 5. The transcription factor family with the largest proportion of genes was the bHLH transcription factor family, with 943 annotated genes.

3.2. Transcriptome quantitative analysis

Through transcriptome-based gene quantification analysis, we identified 64,752 unigenes common to both the experimental and control groups, with 19,840 unigenes specific to the experimental group and 9,896 unigenes unique to the control group (Figure S3A). An expression level density plot was generated based on fragments per kilobase of transcripts per million mapped reads for each sample. The gene



Figure 4 Functional distribution map and species distribution map. NR annotation species distribution indicates that *Nyssa sinensis* and *Vitis vinifera* have the highest annotation proportions.



Figure 5 Classification of genes by transcription factor families. The transcription factor family with the largest proportion of genes is the bHLH transcription factor family.

expression levels roughly followed a normal distribution, indicating a typical gene expression pattern (Figure S3B). The results of this sequencing were satisfactory and met our expectations. Additionally, we calculated the Pearson correlation coefficients of gene expression levels between all pairs of samples and visualized the results as a heatmap. The heatmap showed similar gene expression profiles within the experimental and control groups, but significant differences in the overall gene expression patterns between groups (Figure S3C). The clustering and differential gene expression between the experimental and control groups were consistent with our expectations.

3.3. DEG analysis

We identified 16,498 upregulated unigenes and selected the top 20 pathways with the smallest Q values for Gene Ontology (GO) enrichment analysis (Figure 6, Table S7). GO analysis of the upregulated genes revealed significant enrichment of cellular components, biological processes, and extracellular regions. The cytoplasm (GO:0005737) showed the highest number of term candidate genes (806 out of 4003), with a rich ratio of 0.31 and a significant p-value of 6.01×10^{-11} . Other highly enriched terms included the extracellular space (GO:0005615, $p = 3.92 \times 10^{-10}$) and proteasome-mediated ubiquitin-dependent protein catabolic

processes (GO:0043161, $p = 3.44 \times 10^{-10}$). Enrichment analysis highlighted the proteasome complex and apoplast as key cellular components. Notably, although the “Monoterpenoid biosynthetic process” pathway was not among the top 20 pathways, it had a Q value < 0.1 and aligned with previous gene annotation results as well as subsequent KEGG enrichment findings. Consequently, it was included as the 21st enriched pathway in the GO-UP figure. For KEGG enrichment analysis of the upregulated genes, seven pathways with Q values < 0.05 were identified (Figure 7, Table S8). Pathway enrichment analysis using KEGG demonstrated a strong association with Proteasome-related processes (ko03050, $p = 8.97 \times 10^{-9}$) and Monoterpenoid biosynthesis (ko00902, $p = 6.45 \times 10^{-7}$). The Citrate cycle (ko00020) and amino sugar metabolism pathways were also significantly enriched, with rich ratios of 0.52 and 0.39, respectively. These findings suggested that the upregulated genes were primarily involved in protein degradation, carbohydrate metabolism, and amino acid biosynthesis. For the 9,412 downregulated unigenes, GO enrichment analysis revealed the top 20 pathways with the smallest Q values (Figure 8, Table S9). GO enrichment analysis of downregulated genes indicated significant enrichment in photosynthesis-related components, particularly the chloroplast thylakoid membrane (GO:0009535, $p = 1.12 \times 10^{-40}$, rich ratio = 0.50) and chloroplast stroma (GO:0009570, $p = 5.60 \times 10^{-26}$). Additionally, biological processes like photosynthesis (GO:0015979, $p = 5.10 \times 10^{-16}$) and cellular components, such as the thylakoid (GO:0009579), were significantly downregulated. These findings suggested impaired chloroplast activity and reduced photosynthetic function. KEGG analysis revealed eight enriched pathways with Q values < 0.05 (Figure 9, Table S10). KEGG pathway analysis of downregulated genes indicated significant enrichment in pathways associated with Photosynthesis (ko00195, $p = 5.81 \times 10^{-11}$) and Antenna proteins (ko00196, $p = 2.18 \times 10^{-10}$). Additional pathways, such as benzoxazinoid biosynthesis and glycosphingolipid biosynthesis, showed moderate enrichment with p-values of 1.85×10^{-6} and 2.59×10^{-4} , respectively. These results suggested suppression of energy metabolism in the downregulated genes.

The enzymes upregulated during monoterpenoid biosynthesis are shown in Figure 10. These are highlighted in blue boxes. The related unigenes and additional information for [EC 1.1.1208], [EC 1.1.1324], [EC 1.14.14.84], and [EC 4.2.3.111] are presented in Table S11. In the GO-DOWN and KEGG-DOWN DEG analyses, numerous pathways related to the downregulation of photosynthesis were identified. The relevant pathways obtained from the GO-DOWN DEG analysis included GO:0009535, GO:0009570, GO:0009534, GO:0015979, GO:0009579, GO:0009523, GO:0009543, GO:0031977, GO:0009768, GO:0009773, GO:0010287, GO:0031969, GO:0009654, GO:0010598, GO:0009522, GO:0009658, and GO:0016168. Among the top 20 enriched pathways, only two were not directly related to the downregulation of photosynthesis. In KEGG-DOWN DEG analysis, the top two enriched pathways were ko00195 and ko00196. The ko00195 pathway is involved in photosynthesis (Figure 11). The unigenes associated with the two downregulated enzymes and 18 structural proteins are presented in Table S12.

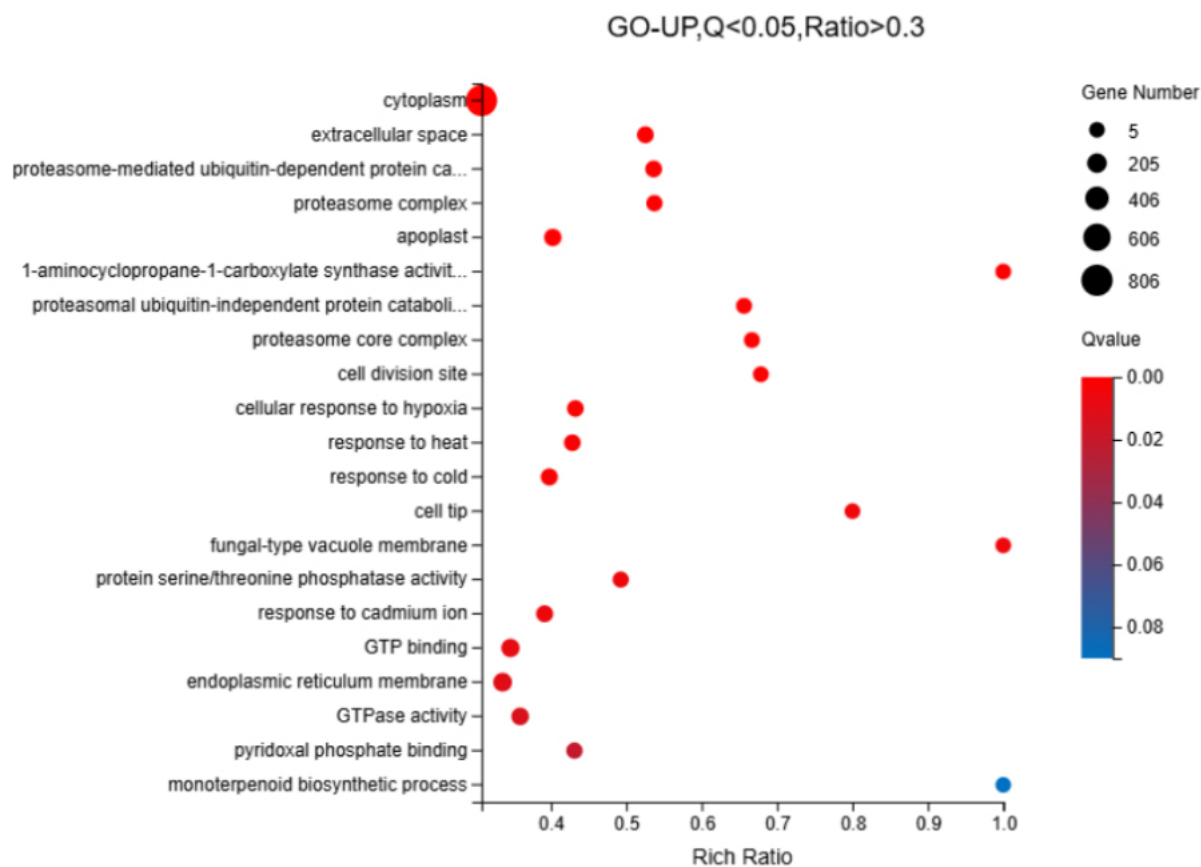


Figure 6 Differentially expressed gene analysis of GO-UP.

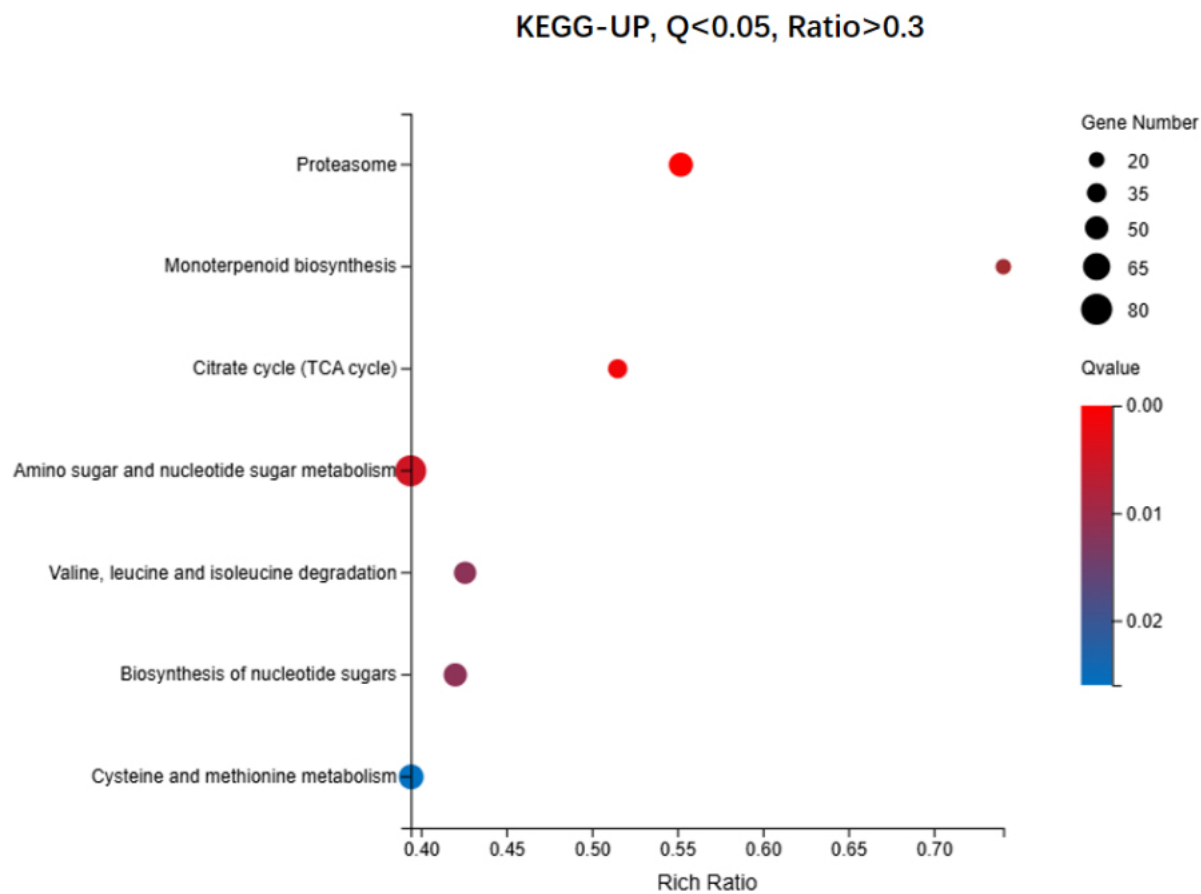


Figure 7 Differentially expressed gene analysis of KEGG-UP.

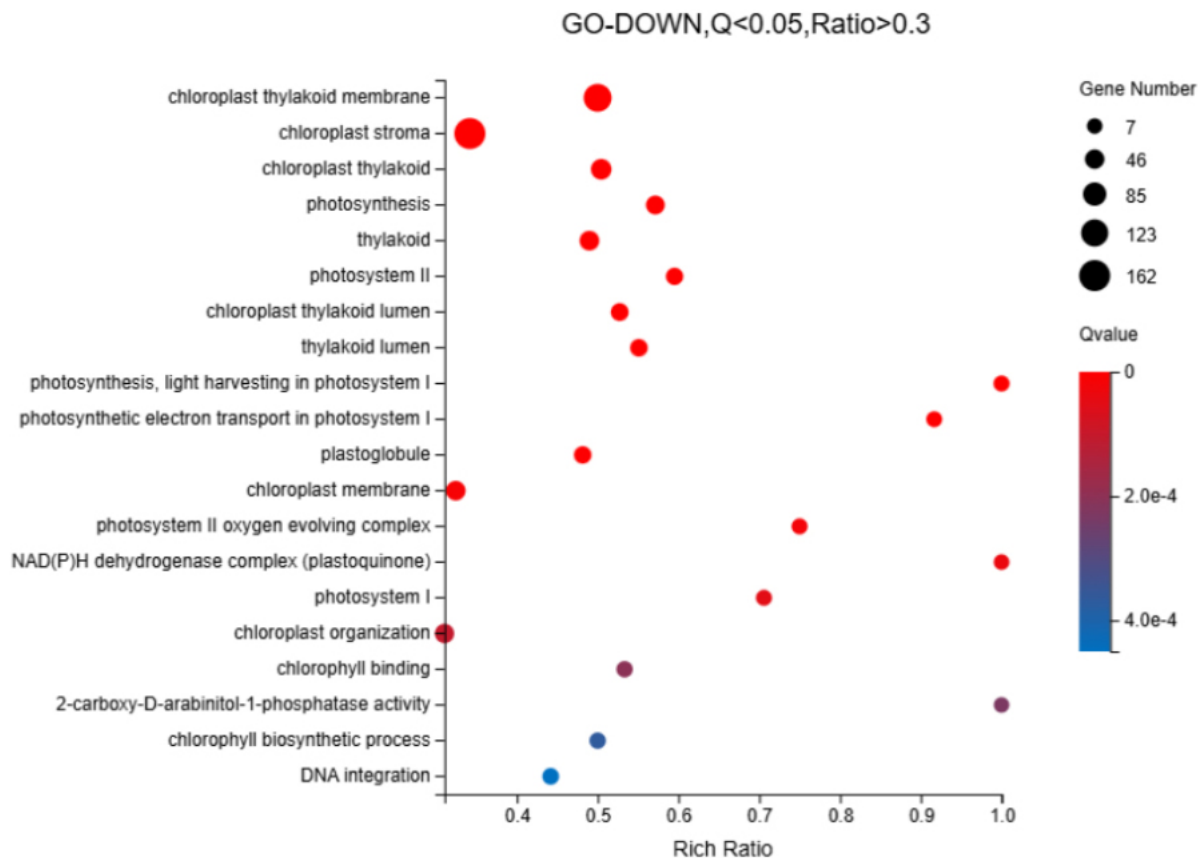


Figure 8 Differentially expressed gene analysis of GO-DOWN.

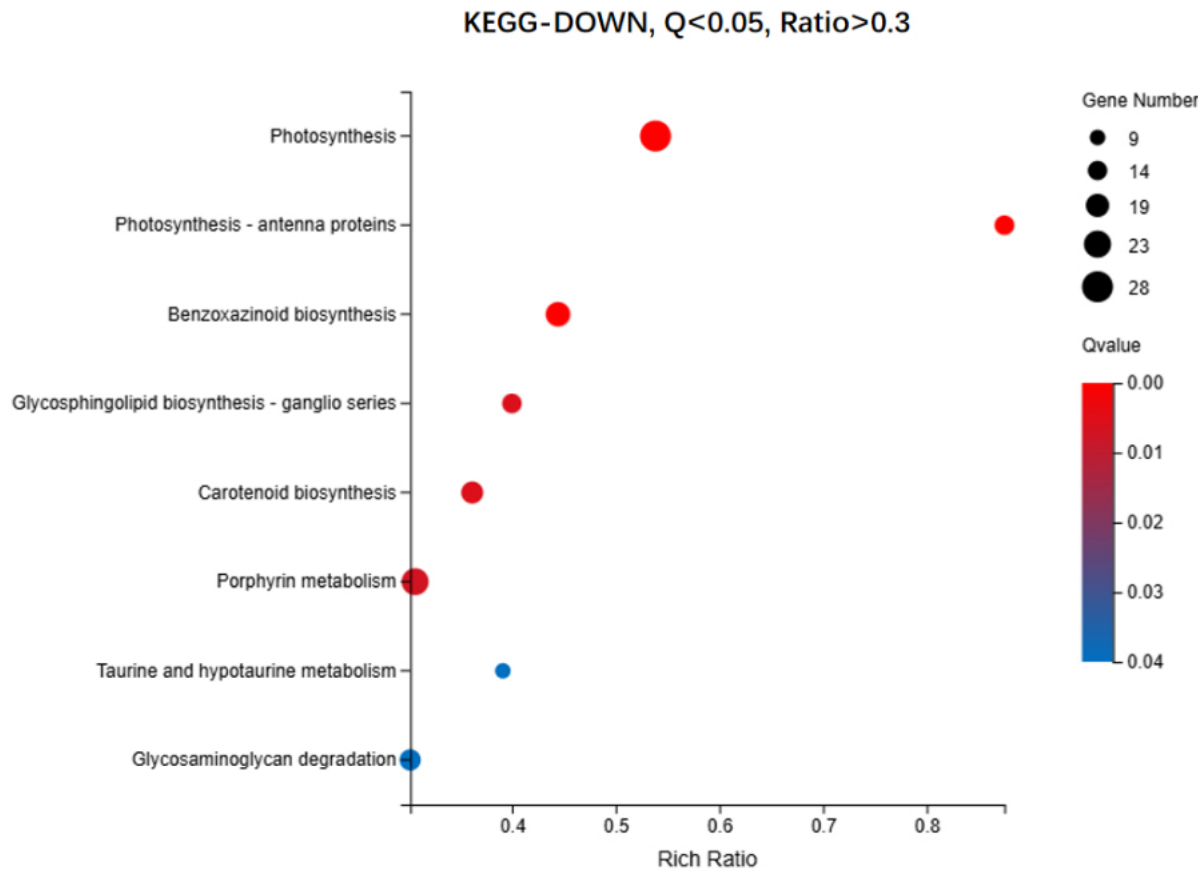
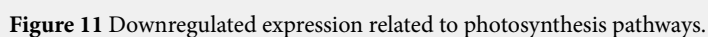
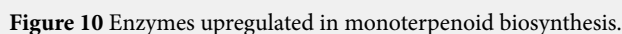


Figure 9 Differentially expressed gene analysis of KEGG-DOWN.



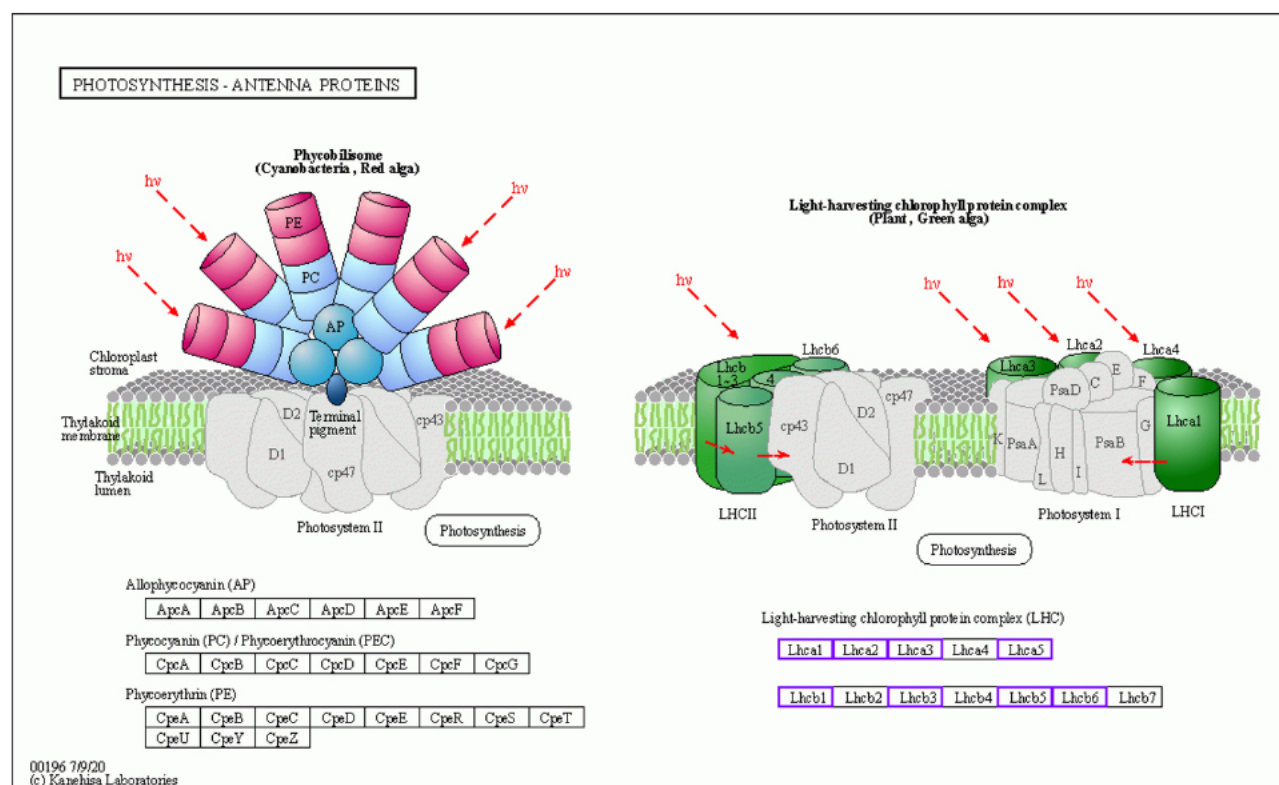


Figure 12 Downregulated expression related to photosynthesis-antenna protein pathways.

The ko00196 pathway corresponds to photosynthesis-antenna proteins (Figure 12), and its downregulation is attributed to the decreased expression of components of the light-harvesting chlorophyll protein complex, specifically Lhca1, Lhca2, Lhca3, Lhca5, Lhcb1, Lhcb3, Lhcb5, and Lhcb6. The downregulated expression of unigenes associated with each component of the light-harvesting chlorophyll protein complex, along with additional information, is presented in Table S13.

In this study, we annotated the DEGs transcripts. The up-regulated DEGs were enriched in pathways related to monoterpenoids, including the monoterpenoid biosynthetic process (GO:0016099) and Monoterpenoid biosynthesis (ko00902). In contrast, the downregulated DEGs were primarily associated with photosynthesis-related pathways, encompassing terms, such as chloroplast thylakoid membrane (GO:0009535), chloroplast stroma (GO:0009570), chloroplast thylakoid (GO:0009534), photosynthesis (GO:0015979), thylakoid (GO:0009579), photosystem II (GO:0009523), chloroplast thylakoid lumen (GO:0009543), thylakoid lumen (GO:0031977), photosynthesis, light harvesting in photosystem I (GO:0009768), photosynthetic electron transport in photosystem I (GO:0009773), plastoglobule (GO:0010287), chloroplast membrane (GO:0031969), photosystem II oxygen evolving complex (GO:0009654), NAD(P)H dehydrogenase complex (plastoquinone) (GO:0010598), photosystem I (GO:0009522), chloroplast organization (GO:0009658), chlorophyll binding (GO:0016168), chlorophyll biosynthetic

process (GO:0015995), as well as Photosynthesis (ko00195) and Photosynthesis - antenna proteins (ko00196).

3.4. SSR detection

We performed SSR detection of the unigenes and designed primers for each SSR. The length characteristics of the SSRs are shown in the SSR length statistics (Table S14), and the results of primer design are provided in the SSR primer design (Table S15). Combined with the previous work (Yin et al., 2018), this study, provides data to support SSR research on plants of the family Hamamelidaceae for assessing genetic diversity in *D. racemosum* and potentially numerous other Hamamelidaceae taxa. These loci will aid in future conservation genetics efforts across the family.

3.5. RT-qPCR

The results showed that the expression of these six genes, including Unigene15193-S3, Unigene16224-S4, Unigene51085-S3, Unigene2103-S3, Unigene20832-S2 and Unigene44219-S5, was consistent with the expression patterns of the RNA-seq data. The expression levels of these genes showed the same trend in the results of RNA-Seq and qPCR (Figure S4, Table S16). Further quantitative correlation analysis revealed that a strong positive correlation rate ($R^2 = 0.97$) was obtained using linear regression analysis (Figure S5), suggesting that the transcriptome data were reliable. The core finding of this article is based on the

RNA-Seq results. Verifying the reliability of RNA-Seq via qPCR offers additional assurance for subsequent analyses grounded in RNA-Seq data.

4. Discussion

Based on KEGG annotation, a protein interaction network for the metabolism of terpenoids and polyketides diagram was constructed (Figure S2). The core genes of the protein interaction network were Unigene41013-S3 and Unigene35068-S1, both of which were functionally annotated to KEGG entry: K00787, named FDPS, which is defined as farnesyl diphosphate synthase [EC 2.5.1.1 2.5.1.10]. It also plays an important role in terpenoid synthesis in jujubes (*Ziziphus jujuba*) (Wen et al., 2023). The FDPS gene can regulate plant development and various biological processes and promote the synthesis of secondary metabolites, such as terpenoids and sterols (Wen et al., 2023; Poulter, 2006).

Our NR annotation of species distribution indicated that *V. vinifera* had the high annotation proportions. Interestingly, terpenoids have also attracted considerable attention in *V. vinifera*. Terpenoids are among the most important constituents of grape flavor and wine bouquets (Data available: <http://urgv.evry.inra.fr/FLAGdb>) (Martin et al., 2010). Terpenoid pheromones play significant roles in insect pheromone blends (Boncan et al., 2020; Rebholz et al., 2023; Yang et al., 2025). During gall formation in *D. racemosum*, the plant may regulate insect behavior through terpenoid-derived volatile compounds.

In plants, bHLH is the second-largest transcription factor family following MYB, with diverse biological functions, including the regulation of plant growth and development, responses to abiotic stress, and metabolic networks (Feller et al., 2011). In our study, the transcription factor family with the largest proportion of genes was the bHLH family with 943 annotated genes (Figure 5). The bHLH transcription factor family in the model plant *Arabidopsis* has been relatively well studied (Gao & Dubos, 2024). These studies established that the bHLH transcription factor family regulates stomatal development in *Arabidopsis*. From bryophytes to dicots, the stomatal developmental pathway in land plants relies on a conserved bHLH-mediated regulatory module (Chang et al., 2023). We inferred that during gall formation in *D. racemosum*, the bHLH transcription factor family may also influence metabolic homeostasis by modulating stomatal development at gall sites.

Elevated expression of secondary metabolites is a prominent feature of galls compared to that in normal leaves. For instance, the phenolic content was significantly higher in the galls than in the healthy leaves of *Hamamelis japonica* (Mizuki et al., 2024), while benzofuran-type and dibenzofuran-type phytoalexins were highly accumulated in the galls of *Parrotia subaequalis* (Zhou et al., 2025). By regulating the levels of plant secondary metabolites or antagonizing the gene expression of other products involved in gall formation, it may be possible to effectively inhibit gall development and control insect pests (Desnitskiy et al., 2023; Nabity et al., 2013; Pang et al., 2021; Yamaguchi et al., 2012).

In the GO-UP DEG analysis, the “monoterpenoid biosynthetic process” pathway was identified, whereas in the KEGG-UP DEG analysis, the “monoterpenoid biosynthesis” pathway was enriched. The Rich Ratios for these monoterpenoid-related pathways were both close to 1 (100%). Plants are the direct source of most terpene compounds. Monoterpenoids have significant ecological value, aiding in plant defense against animals (Fan et al., 2023). In the present study, the differential expression of terpenoid compounds at the transcriptional level may have resulted from the stress response of leaves to biological stimulation by gallers. In *Semiliquidambar cathayensis* (Hamamelidaceae), a phylogenetically related species to *D. racemosum*, transcriptomic analysis revealed constitutive expression of monoterpenoid biosynthesis genes across leaf, stem, and root tissues (Tian et al., 2022). While this study focused on basal metabolic regulation rather than stress-induced responses, the widespread expression of monoterpenoid-related genes suggests that Hamamelidaceae plants may inherently possess a robust genetic framework for monoterpenoid production. Of particular interest is [EC 4.2.3.111], (-)- α -terpineol synthase, which plays a critical role in terpenoid biosynthesis (Lücker et al., 2004; Martin & Bohlmann, 2004). In our study, we identified the DEGs associated with [EC 4.2.3.111], namely Unigene33316-S5, Unigene33318-S5, and Unigene33320-S5.

Based on the differential expression annotation, we identified the upregulation of core metabolic pathways related to monoterpenoid biosynthesis in galled leaves (Figure 13). Notably, we observed the upregulation of linalool 8-monooxygenase. Linalool 8-monooxygenase catalyzes two critical consecutive enzymatic reactions during monoterpenoid biosynthesis. These reactions convert 8-hydroxylinalool into 8-oxolinalool, thereby exerting cascading effects on the entire metabolic network. However, no studies have specifically investigated linalool 8-monooxygenase in plant biotic stress responses. Additionally, three other enzymes involved in monoterpenoid biosynthesis, specifically (+)-neomenthol dehydrogenase, 8-hydroxygeraniol dehydrogenase, and (-)- α -terpineol synthase, exhibited upregulated expression levels. These three enzymes utilize substrates derived from downstream metabolites of geranyl-PP. Geranyl-PP is a pivotal product in terpenoid backbone biosynthesis (Kanehisa Laboratories, 2023).

Through DEG analysis, we observed downregulation of gene expression in *D. racemosum* associated with numerous pathways involved in photosynthesis. Downregulation of photosynthetic capacity in galled leaves is a relatively common phenomenon in *Litsea acuminata*, *Glochidion obovatum*, *Eurya japonica*, and *Artemisia montana* (Huang et al., 2015). The downregulation of photosynthesis-related pathways is not associated only with nutrient availability; under biotic stress, downregulation of photosynthesis-related genes is part of the plant defense response (Bilgin et al., 2010). Significantly downregulated photosynthesis-related pathways, along with the expression of terpenoid-related compounds, may constitute an integrated aspect of the complex defense strategies in plants. Further, the activation of floral organ genes by galls suggests that gall

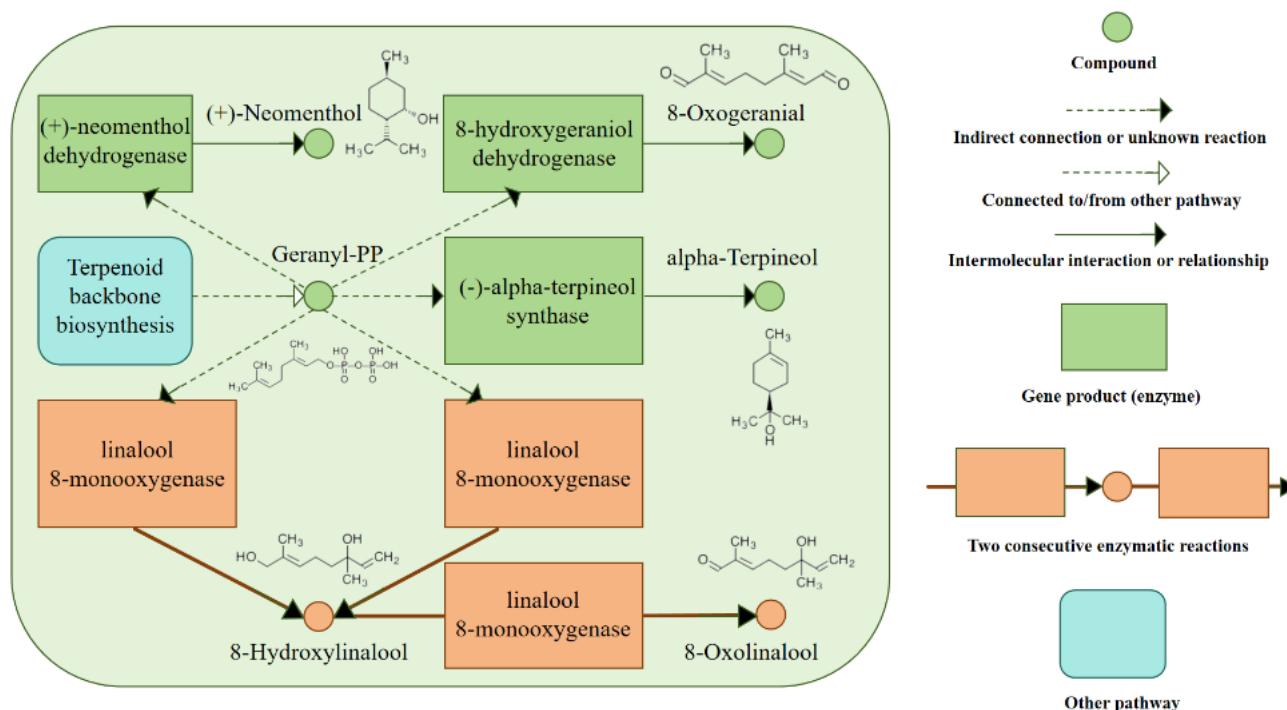


Figure 13 Upregulation of core metabolic pathways in the monoterpenoid biosynthesis of galled leaves.

formation may co-opt genetic programs from plant reproductive development to promote cell proliferation or redirect metabolic pathways, thereby establishing passive sink organs that facilitate nutrient storage and provide shelter for parasites (Takeda et al., 2024). Based on our findings, we speculate that the transcriptional downregulation of photosynthesis-related pathways in galls results from multiple factors, including the transformation of galls into sink organs and the reallocation of carbon resources to secondary metabolites.

Thus, our study indicated that under insect-induced stress, *D. racemosum* allocated carbon resources to production of secondary metabolites rather than maintaining photosynthesis. The carbon metabolism in *D. racemosum* leaves undergoes significant alterations during gall formation. Carbon resources are preferentially allocated for synthesizing enzymes involved in terpenoid metabolism, such as (+)-neomenthol dehydrogenase, 8-hydroxygeraniol dehydrogenase, (-)-alpha-terpineol synthase, and linalool 8-monooxygenase, leading to marked upregulation of the entire terpenoid biosynthesis pathway. Conversely, fewer resources are directed toward photosynthesis-related structural proteins or subunits, including Lhca (Lhca 1, 2, 3, 5), Lhcb (Lhcb 1, 3, 5, 6), and the structural proteins of photosystems I and II, resulting in significant downregulation of photosynthetic pathways.

5. Conclusions

This study provides insights into the transcriptional responses of *D. racemosum* leaves to gall formation. We identified 104,537 unigenes and functionally annotated them using multiple databases. Transcriptome sequencing and

DEG analyses revealed significant downregulation of photosynthesis-related genes in galled leaves, indicating a reduction in photosynthetic activity. In contrast, genes associated with terpenoid biosynthesis were upregulated, particularly within pathways related to monoterpenoid biosynthesis. These results highlight the integration of energy production and defense strategies in *D. racemosum* during gall formation. Metabolomic studies reveal that *Distylium buxifolium* (Hamamelidaceae) produces high levels of jasmonic acid, which provides resistance against gall parasitism, whereas *D. racemosum* (Hamamelidaceae) does not exhibit such a significant level of elevated jasmonic acid production (Zhang et al., 2024). Why has *D. racemosum* evolved to prioritize terpenoid biosynthesis over the more potent jasmonic acid pathway as a response to gall parasitism? A deeper exploration of terpenoid-mediated modulation of plant-galler interspecific relationships will be a critical research focus moving forward.

6. Supplementary material

The following supplementary material is available for this article:

Appendix S1. The steps used for cDNA synthesis and q-PCR.

Figure S1. Length distribution of unigenes.

Figure S2. Protein interaction network diagram for the metabolism of terpenoids and polyketides. The core genes of the protein interaction network are Unigene41013-S3 and Unigene35068-S1, both of which were functionally annotated to KEGG Orthology Entry: K00787, named FDPS, defined as farnesyl diphosphate synthase [EC 2.5.1.1 2.5.1.10]

Figure S3. (A) Venn diagram showing the differentially expressed genes between the experimental and control groups; (B) density plot of gene expression levels based on Fragments Per Kilobase of transcript per Million mapped reads (FPKM) for each sample. The X-axis represents $\log_2$ (FPKM+1), and the Y-axis represents gene density, which is the proportion of genes at each expression level relative to the total number of expressed genes; (C) heatmap of gene expression correlations between samples.

Figure S4. Amplification fold-change of six genes.

Figure S5. Correlation rate between relative expression in the transcriptome and RT-qPCR.

Table S1. Quality statistics of filtered reads.

Table S2. Quality metrics of the clustered unigenes.

Table S3. Criteria for Differentially Expressed Genes DEGs Selection and Enrichment Analysis Parameters.

Table S4. Subject IDs corresponding to genes in the Metabolism of terpenoids and polyketides pathway within the KEGG data base.

Table S5. Protein interaction network diagram for the metabolism of terpenoids and polyketides.

Table S6. Subject IDs corresponding to genes in organelle pathway within the GO data base.

Table S7. Differentially expressed gene analysis of GO-UP.

Table S8. Differentially expressed gene analysis of KEGG-UP.

Table S9. Differentially expressed gene analysis of GO-DOWN.

Table S10. Differentially expressed gene analysis of KEGG-DOWN.

Table S11. Enzymes upregulated in monoterpenoid biosynthesis KEGG-UP.

Table S12. Downregulated expression related to photosynthesis pathways KEGG-DOWN.

Table S13. Downregulated expression related to photosynthesis – antenna proteins pathways KEGG-DOWN.

Table S14. SSR length statistics.

Table S15. SSR primer design table.

Table S16. Comparison of RT-qPCR and RNA-Seq.

References

- Adroit, B., Malekhosseini, M., Girard, V., Abedi, M., Rajaei, H., Terral, J. F., & Wappler, T. (2018). Changes in pattern of plant-insect interactions on the *Persian ironwood* (*Parrotia persica*, Hamamelidaceae) over the last 3 million years, *Review of Palaeobotany and Palynology*, 258, 22–35. <https://doi.org/10.1016/j.revpalbo.2018.06.007>
- Aoki, S., Kurosu, U., Uematsu, K., Fukatsu, T., & Kutsukake, M. (2019). Two closely related species of *Nipponaphis* (Hemiptera: Aphididae) that migrate between *Distylium racemosum* and *Machilus* trees in Japan. *Entomological Science*, 22(2), 220–229. <https://doi.org/10.1111/ens.12358>
- Aoki, S., Kutsukake, M., Kurosu, U., Yeh, H. T., Sano, M., & Fukatsu, T. (2015). *Nipponaphis* species (Aphididae: Hormaphidinae) that form green galls on *Distylium racemosum* in Japan. *Entomological Science*, 18(4), 420–434. <https://doi.org/10.1111/ens.12144>
- Bailey, S., Percy, D. M., Hefer, C. A., & Cronk, Q. C. (2015). The transcriptional landscape of insect galls: psyllid (Hemiptera) gall formation in Hawaiian *Metrosideros polymorpha* (Myrtaceae). *BMC Genomics*, 16, article 943. <https://doi.org/10.1186/s12864-015-2109-9>
- Bilgin, D. D., Zavala, J. A., Zhu, J., Clough, S. J., Ort, D. R., & DeLucia, E. H. (2010). Biotic stress globally downregulates photosynthesis genes. *Plant, Cell and Environment*, 33(10), 1597–1613. <https://doi.org/10.1111/j.1365-3040.2010.02167.x>
- Boncan, D. A. T., Tsang, S. S. K., Li, C., Lee, I. H. T., Lam, H. M., Chan, T. F., & Hui, J. H. L. (2020). Terpenes and terpenoids in plants: Interactions with environment and insects. *International Journal of Molecular Sciences*, 21(19), article 7382. <https://doi.org/10.3390/ijms21197382>
- Chang, G., Ma, J., Wang, S., Tang, M., Zhang, B., Ma, Y., Li, L., Sun, G., Dong, S., Liu, Y., Zhou, Y., Hu, X., Song, C. P., & Huang, J. (2023). Liverwort bHLH transcription factors and the origin of stomata in plants. *Current Biology*, 33(13), 2806–2813. <https://doi.org/10.1016/j.cub.2023.05.050>
- Chen, J., Sorin, M., & Qiao, G. X. (2011). *Asiphonipponaphis*, a new genus of Hormaphidinae (Hemiptera, Aphididae) causing galls on *Distylium chinense* from China, with description of a new species. *ZooKeys*, 111, 1–10. <https://doi.org/10.3897/zookeys.111.1283>
- Chen, Y., Chen, Y., Shi, C., Huang, Z., Zhang, Y., Li, S., Li, Y., Ye, J., Yu, C., Li, Z., Zhang, X., Wang, J., Yang, H., Fang, L., & Chen, Q. (2018). SOAPnuke: a MapReduce acceleration-supported software for integrated quality control and preprocessing of high-throughput sequencing data. *GigaScience*, 7(1), 1–6. <https://doi.org/10.1093/gigascience/gix120>
- Chen, Z., Ain, N. U., Zhao, Q., & Zhang, X. (2024). From tradition to innovation: conventional and deep learning frameworks in genome annotation. *Briefings in Bioinformatics*, 25(3), article bbae138. <https://doi.org/10.1093/bib/bbae138>
- Chetverikov, P. E., Craemer, C., Cvrković, T., Klimov, P. B., Petanović, R. U., Romanovich, A. E., Sukhareva, S. I., Zukoff, S. N., Bolton, S., & Amrine, J. (2021). Molecular phylogeny of the phytoparasitic mite family Phytoptidae (Acariformes: Eriophyoidea) identified the female genitalic anatomy as a major macroevolutionary factor and revealed multiple origins of gall induction. *Experimental and Applied Acarology*, 83, 31–68. <https://doi.org/10.1007/s10493-020-00571-6>
- Conner, K., Olive, J., Zhang, L., Jacobi, J., & Putnam, M. L. (2013). First report of bacterial gall on *Loropetalum chinense* caused by *Pseudomonas savastanoi* in the United States. *Plant Disease*, 97(6), 835–835. <https://doi.org/10.1094/PDIS-11-12-1011-PDN>
- Dederich, A. E., Halbert, S. E., & VON Dohlen, C. D. (2022). Description of a new species of *Hamamelistes* forming galls on *Fothergilla* spp. (Hamamelidaceae) and the generic limits of *Hormaphidini* (Sternorrhyncha: Aphididae: Hormaphidinae). *Zootaxa*, 5183(1), 203–219. <https://doi.org/10.11646/zootaxa.5183.1.16>
- Desnitskiy, A. G., Chetverikov, P. E., Ivanova, L. A., Kuzmin, I. V., Ozman-Sullivan, S. K., & Sukhareva, S. I. (2023). Molecular Aspects of Gall Formation Induced by Mites and Insects. *Life*, 13(6), 1347. <https://doi.org/10.3390/life13061347>
- Drmanac, R., Sparks, A. B., Callow, M. J., Halpern, A. L., Burns, N. L., Kermani, B. G., Carnevali, P., Nazarenko, I.,

- Nilsen, G. B., Yeung, G., Dahl, F., Fernandez, A., Staker, B., Pant, K. P., Baccash, J., Borcharding, A. P., Brownley, A., Cedeno, R., Chen, L., Chernikoff, D., ... & Reid, C. A. (2010). Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays. *Science*, 327(5961), 78–81. <https://doi.org/10.1126/science.1181498>
- Endress, P. (1993). Hamamelidaceae. In K. Kubitzki, J. G. Rohwer, & V. Bittrich (Eds.), *Flowering plants, Dicotyledons: Magnoliid, Hamamelid and Caryophyllid families* (pp. 322–331). Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-662-02899-5>
- Fan, M., Yuan, S., Li, L., Zheng, J., Zhao, D., Wang, C., Wang, H., Liu, X., & Liu, J. (2023). Application of Terpenoid Compounds in Food and Pharmaceutical Products. *Fermentation*, 9(2), 119. <https://doi.org/10.3390/fermentation9020119>
- Feller, A., Machemer, K., Braun, E. L., & Grotewold, E. (2011). Evolutionary and comparative analysis of MYB and bHLH plant transcription factors. *The Plant Journal*, 66(1), 94–116. <https://doi.org/10.1111/j.1365-3113.2010.04459.x>
- Gao, F., & Dubos, C. (2024). The arabidopsis bHLH transcription factor family. *Trends in Plant Science*, 29(6), 668–680. <https://doi.org/10.1016/j.tplants.2023.11.022>
- Giron, D., Huguet, E., Stone, G. N., & Body, M. (2016). Insect-induced effects on plants and possible effectors used by galling and leaf-mining insects to manipulate their host-plant. *Journal of Insect Physiology*, 84, 70–89. <https://doi.org/10.1016/j.jinsphys.2015.12.009>
- Gong, M., Ke, J., Yan, C., & Wang, X. (1988). A preliminary study on the regularity of occurrence of the disease on *Mytilaria laosensis* caused by gall mites and its control. *Forest Research*, 1, 325–330. <https://www.cabidigitallibrary.org/doi/full/10.5555/19910649010>
- Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., & Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, 29, 644–652. <https://doi.org/10.1038/nbt.1883>
- Guedes, L. M., Aguilera, N., Becerra, J., Hernández, V., & dos Santos Isaías, R. M., (2016). Leaf and stem galls of *Schinus polygamus* (Cav.) Cabr (Anacardiaceae): Anatomical and chemical implications. *Biochemical Systematics and Ecology*, 69, 266–273. <https://doi.org/10.1016/j.bse.2016.10.012>
- Harris, M. O., & Pitzschke, A. (2020). Plants make galls to accommodate foreigners: some are friends, most are foes. *The New phytologist*, 225(5), 1852–1872. <https://doi.org/10.1111/nph.16340>
- Hirano, T., Okamoto, A., Oda, Y., Sakamoto, T., Takeda, S., Matsuura, T., Ikeda, Y., Higaki, T., Kimura, S., & Sato, M. H. (2023). Ab-GALFA, a bioassay for insect gall formation using the model plant *Arabidopsis thaliana*. *Scientific Reports*, 13, article 2554. <https://doi.org/10.1038/s41598-023-29302-8>
- Huang, M. Y., Huang, W. D., Chou, H. M., Chen, C. C., Chen, P. J., Chang, Y. T., & Yang, C. M. (2015). Structural, biochemical, and physiological characterization of photosynthesis in leaf-derived cup-shaped galls on *Litsea acuminata*. *BMC Plant Biology*, 15, article 61. <https://doi.org/10.1186/s12870-015-0446-0>
- International Dendrology Society (2024, April 12). '*Distylium racemosum*' from the website *Trees and Shrubs Online*. <https://www.treesandshrubsonline.org/articles/distylium/distylium-racemosum/>
- Kanehisa Laboratories (2023, June 3). *Monoterpenoid biosynthesis*. <https://www.kegg.jp/pathway/ko00902>
- Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, 9, 357–359. <https://doi.org/10.1038/nmeth.1923>
- Larson, K. (1998). The impact of two gall-forming arthropods on the photosynthetic rates of their hosts. *Oecologia*, 115, 161–166. <https://doi.org/10.1007/s004420050503>
- Lewis, I. F., & Walton, L. (1958). Gall-formation on *Hamamelis virginiana* resulting from material injected by the aphid *Hormaphis hamamelidis*. *Transactions of the American Microscopical Society*, 77(2), 146–200. <https://doi.org/10.2307/3224114>
- Li, B., & Dewey, C. N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*, 12, article 323. <https://doi.org/10.1186/1471-2105-12-323>
- Li, Q., Zhao, X., Zhang, W., Wang, L., Wang, J., Xu, D., Mei, Z., Liu, Q., Du, S., Li, Z., Liang, X., Wang, X., Wei, H., Liu, P., Zou, J., Shen, H., Chen, A., Drmanac, S., Liu, J. S., Li, L., Jiang, H., Zhang, Y., Wang, J., Yang, H., Xu, X., Drmanac, R., & Jiang, Y. (2019). Reliable multiplex sequencing with rare index mis-assignment on DNB-based NGS platform. *BMC Genomics*, 20, article 215. <https://doi.org/10.1186/s12864-019-5569-5>
- Ling, Z. (2012). Physiological and biochemical responses of *Distylium racemosum* to gall's formation. *Journal of Fujian College of Forestry*, 32, 10–12. Corpus ID: 88276467
- Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, 15, article 550. <https://doi.org/10.1186/s13059-014-0550-8>
- Lücker, J., Bowen, P., & Bohlmann, J. (2004). *Vitis vinifera* terpenoid cyclases: functional identification of two sesquiterpene synthase cDNAs encoding (+)-valencene synthase and (–)-germacrene D synthase and expression of mono- and sesquiterpene synthases in grapevine flowers and berries. *Phytochemistry*, 65(19), 2649–2659. <https://doi.org/10.1016/j.phytochem.2004.08.017>
- Mani, M. S. (1964). Galls on flower and fruit. *Ecology of plant galls* (pp. 136–148). Springer. https://doi.org/10.1007/978-94-017-6230-4_8
- Martin, D. M., Aubourg, S., Schouwey, M. B., Daviet, L., Schalk, M., Toub, O., Lund, S. T., & Bohlmann, J. (2010). Functional annotation, genome organization and phylogeny of the grapevine (*Vitis vinifera*) terpene synthase gene family based on genome assembly, FLcDNA cloning, and enzyme assays. *BMC Plant Biology*, 10, article 226. <https://doi.org/10.1186/1471-2229-10-226>
- Martin, D. M., & Bohlmann, J. (2004). Identification of *Vitis vinifera* (–)-alpha-terpineol synthase by in silico screening of full-length cDNA ESTs and functional characterization of recombinant terpene synthase. *Phytochemistry*, 65(9), 1223–1229. <https://doi.org/10.1016/j.phytochem.2004.03.018>
- Miller, H. N. (1975). Leaf, stem, crown, and root galls induced in chrysanthemum by *Agrobacterium tumefaciens*. *Phytopathology*, 65, 805–811. <https://doi.org/10.1094/Phyto-65-805>
- Mizuki, M., Kaneko, Y., Yukie, Y., Suyama, Y., Hirota, S. K., Sawa, S., Kubo, M., Yamawo, A., Sasabe, M., & Ikeda, H. (2024). Evolution of secondary metabolites, morphological structures and associated gene expression

- patterns in galls induced by four closely related aphid species on a host plant species. *Molecular Ecology*, 33(16), article e17466. <https://doi.org/10.1111/mec.17466>
- Molnár, B., Boddum, T., Szócs, G., & Hillbur, Y. (2009). Occurrence of two pest gall midges, *Obolodiplosis robiniae* (Haldeman) and *Dasineura gleditchiae* (Osten Sacken) (Diptera: Cecidomyiidae) on ornamental trees in Sweden. *Entomologisk Tidskrift*, 130, 113–120. Corpus ID: 55458859
- Nabity, P. D., Haus, M. J., Berenbaum, M. R., & DeLucia, E. H. (2013). Leaf-galling phylloxera on grapes reprograms host metabolism and morphology. *Proceedings of the National Academy of Sciences*, 110 (41), 16663–16668. <https://doi.org/10.1073/pnas.1220219110>
- Oates, C. N., Denby, K. J., Myburg, A. A., Slippers, B., & Naidoo, S. (2016). Insect gallers and their plant hosts: From omics data to systems biology. *International Journal of Molecular Sciences*, 17(11), 1891. <https://doi.org/10.3390/ijms17111891>
- Pang, Z., Chen, J., Wang, T., Gao, C., Li, Z., Guo, L., Xu, J., & Cheng, Y. (2021). Linking Plant Secondary Metabolites and Plant Microbiomes: A Review. *Frontiers in Plant Science*, 12, article 621276. <https://doi.org/10.3389/fpls.2021.621276>
- Park, J. E., Patnaik, B. B., Sang, M. K., Song, D. K., Jeong, J. Y., Hong, C. E., Kim, Y. T., Shin, H. J., Ziwei, L., Patnaik, H. H., et al. (2024). Transcriptome sequencing of the endangered land snail *Karafutohelix adamsi* from the Island Ulleung: De novo assembly, annotation, valuation of fitness genes and SSR markers. *Genes & Genomics*, 46, 851–870. <https://doi.org/10.1007/s13258-024-01511-z>
- Poulter, C. D. (2006). Farnesyl diphosphate synthase: A paradigm for understanding structure and function relationships in E-polyprenyl diphosphate synthases. *Phytochemistry Reviews*, 5, 17–26. <https://doi.org/10.1007/s11101-005-4887-1>
- Rebholz, Z., Shewade, L., Kaler, K., Larose, H., Schubot, F., Tholl, D., Morozov, A. V., & O'Maille, P. E. (2023). Emergence of terpene chemical communication in insects: Evolutionary recruitment of isoprenoid metabolism. *Protein Science*, 32(5), article e4634. <https://doi.org/10.1002/pro.4634>
- Shen, B. Y., Jing, J. L., & Zhu, J. G. (1993). Preliminary study on the biology of *Neothoracaphis hangzhouensis* [杭州新胸蚜生物学的初步研究]. *Plant Protection*, 2, 5–7. https://kns.cnki.net/kcms2/article/abstract?v=KlmjsyJjhUSQzVoCL4jgnAqsLhAc4iPPGYyip0RtryMuMquTZ7NfPRMGm7xUIA0pFVnBuFdjeMtvEzy9syfaTWL4pOh_MR6y2D5_hdVVcGn1hq8aFmsEtX8i7w9LiI-BV1S1JkVrvGI8dpY2eLJLwIqEbmQUVnkYCbUWs_QFFqyKoDiPOdjZPXKXKYUnyvi&uniplatform=NZKPT&language=CHS
- Takeda, S., Yoza, M., Ueda, S., Takeuchi, S., Maeno, A., Sakamoto, T., & Kimura, S. (2024). Exploring the diversity of galls on *Artemisia indica* induced by *Rhopalomyia* species through morphological and transcriptome analyses. *Plant Direct*, 8(7), article e619. <https://doi.org/10.1002/pld3.619>
- Thiel, T., Michalek, W., Varshney, R. K., & Graner, A. (2003). Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *Theoretical and Applied Genetics*, 106, 411–422. <https://doi.org/10.1007/s00122-002-1031-0>
- Tian, X., Yan, L., Jiang, L., Xiang, G., Li, G., Zhu, L., & Wu, J. (2022). Comparative transcriptome analysis of leaf, stem, and root tissues of *Semiliquidambar cathayensis* reveals candidate genes involved in terpenoid biosynthesis. *Molecular Biology Reports*, 49(6), 5585–5593. <https://doi.org/10.1007/s11033-022-07492-0>
- Tooker, J. F., & Helms, A. M. (2014). Phytohormone dynamics associated with gall insects, and their potential role in the evolution of the gall-inducing habit. *Journal of chemical ecology*, 40, 742–753. <https://doi.org/10.1007/s10886-014-0457-6>
- Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M., & Rozen, S. G. (2012). Primer3: New capabilities and interfaces. *Nucleic Acids Research*, 40(15), article e115. <https://doi.org/10.1093/nar/gks596>
- Walker, E. H. (1944). A revision of *Distylium* and *Sycopsis* (Hamamelidaceae). *Journal of the Arnold Arboretum*, 25(3), 319–341. <https://www.jstor.org/stable/43782434>
- Wang, L., Feng, Z., Wang, X., Wang, X., & Zhang, X. (2010). DEGseq: an R package for identifying differentially expressed genes from RNA-seq data. *Bioinformatics*, 26(1), 136–138. <https://doi.org/10.1093/bioinformatics/btp612>
- Wen, C., Zhang, Z., Shi, Q., Niu, R., Duan, X., Shen, B., & Li, X. (2023). Transcription Factors ZjMYB39 and ZjMYB4 Regulate Farnesyl Diphosphate Synthase- and Squalene Synthase-Mediated Triterpenoid Biosynthesis in *Jujube*. *Journal of Agricultural and Food Chemistry*, 71(11), 4599–4614. <https://doi.org/10.1021/acs.jafc.2c08679>
- Xiang, X., Xiang, K., Ortiz, R. D. C., Jabbour, F., & Wang, W. (2019). Integrating palaeontological and molecular data uncovers multiple ancient and recent dispersals in the pantropical Hamamelidaceae. *Journal of Biogeography*, 46(11), 2622–2631. <https://doi.org/10.1111/jbi.13690>
- Yamaguchi, H., Tanaka, H., Hasegawa, M., Tokuda, M., Asami, T., & Suzuki, Y. (2012). Phytohormones and willow gall induction by a gall-inducing sawfly. *The New Phytologist*, 196(2), 586–595. <https://doi.org/10.1111/j.1469-8137.2012.04264.x>
- Yang, F., Wang, Y., & Chan, Z. (2014). Perspectives on screening winter-flood-tolerant woody species in the riparian protection forests of the three gorges reservoir. *PloS one*, 9, article e108725. <https://doi.org/10.1371/journal.pone.0108725>
- Yang, J. C., Zhang, J. P., Wu, C. Y., Bai, Y., Guedes, R. N. C., Dewar, Y., Li, F. Q., & Zang, L. S. (2025). Diversity and role of volatile terpene and terpenoid pheromones in insects. *Journal of Economic Entomology*, 118(1), 9–18. <https://doi.org/10.1093/jee/ent071>
- Yeh, H. T., Ko, C. C. (2017). A new genus and two new species of galling aphids (Hemiptera: Aphididae) from *Distylium* (Hamamelidaceae) in Taiwan. *Zootaxa*, 4323(2), 209–218. <https://doi.org/10.11646/zootaxa.4323.2.4>
- Yin, Q., Huang, C., Huang, Y., Chen, S., Ye, H., Fan, Q., & Liao, W. (2018). Identification and development of microsatellite markers in *Hamamelis mollis* (Hamamelidaceae). *Applications in Plant Sciences*, 6(10), article e01189. <https://doi.org/10.1002/aps3.1189>
- Zhang, M., Zhou, C., Song, Z., Weng, Q., Li, M., Ji, H., Mo, X., Huang, H., Lu, W., Luo, J., Li, F., & Gan, S. (2018). The first identification of genomic loci in plants associated with resistance to galling insects: a case study in *Eucalyptus* L'Hér. (Myrtaceae). *Scientific reports*, 8, article 2319. <https://doi.org/10.1038/s41598-018-20780-9>
- Zhang, W. X., Lian, F. L., Liang, Y. Q., Zhang, S. T., Chen, S. T., Liu, Y. J., Hong, Z., Gong, F. Y., Sun, J., & Wang, Z. A. (2024). Analysis of secondary metabolites in *Distylium racemosum* and *Distylium buxifolium* based on broadly targeted metabolomics [基于广泛靶向代谢组学的蚊母

- 树和小叶蚊母树次生代谢产物分析]. *Journal of Zhejiang Forestry Science and Technology*, 44(6), 30–39.
<https://doi.org/10.3969/j.issn.1001-3776.2024.06.005>
- Zhang, X. H., Rao, X. L., Shi, H. T., et al. (2011). Overexpression of a cytosolic glyceraldehyde-3-phosphate dehydrogenase gene OsGAPC3 confers salt tolerance in rice. *Plant Cell, Tissue and Organ Culture*, 107(1), 1–11.
<https://doi.org/10.1007/s11240-011-9950-6>
- Zhou, P. J., Wu, X. Y., Zhao, Z. Y., Zang, Y., Sun, Z. S., Li, Y. L., Li, N., Xiong, J., Choo, Y. M., Jin, Z. X., Li, J., & Hu, J. F. (2025). Benzofurans and dibenzofurans from galls on twigs of the endangered Chinese endemic tree *Parrotia subaequalis* and their inhibitory properties against *Staphylococcus aureus* and ATP-citrate lyase. *Phytochemistry*, 229, article 114309.
<https://doi.org/10.1016/j.phytochem.2024.114309>